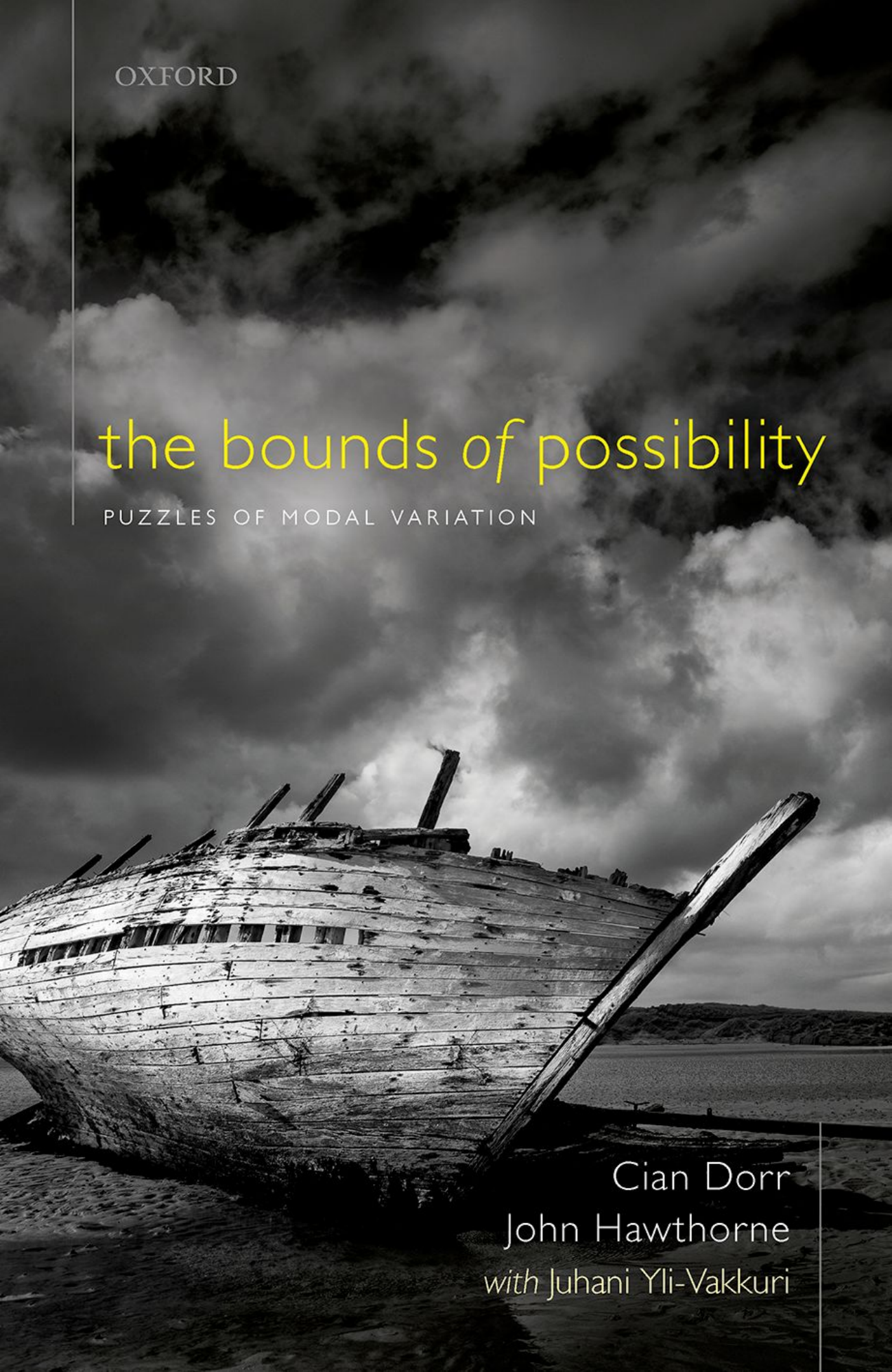
OXFORD
the bounds of possibility
PUZZLES OF MODAL VARIATION
Cian Dorr
John Hawthorne
with Juhani Yli-Vakkuri

The Bounds of Possibility

The Bounds of Possibility

Puzzles of Modal Variation

Cian Dorr and John Hawthorne

with Juhani Yli-Vakkuri

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

© Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri 2021

The moral rights of the authors have been asserted

First Edition published in 2021

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in
a retrieval system, or transmitted, in any form or by any means, without the
prior permission in writing of Oxford University Press, or as expressly permitted
by law, by licence or under terms agreed with the appropriate reprographics
rights organization. Enquiries concerning reproduction outside the scope of the
above should be sent to the Rights Department, Oxford University Press, at the
address above

You must not circulate this work in any other form
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2021944899

ISBN 978–0–19–284665–5

DOI: 10.1093/oso/9780192846655.001.0001

Printed and bound in Great Britain by
Clays Ltd, Elcograf S.p.A.

Links to third party websites are provided by Oxford in good faith and
for information only. Oxford disclaims any responsibility for the materials
contained in any third party website referenced in this work.

Contents

Foreword ix
List of Figures xi

Introduction 1

1. Logical Tools 14
   1.1 Classical Propositional Logic 14
   1.2 Higher-Order Languages 17
   1.3 Higher-Order Logic 23
   1.4 Modal Logic 32
   1.5 Rigidity 41
   1.6 Worlds 46

2. Tolerance Puzzles 53
   2.1 A Schema for Tolerance Arguments 53
   2.2 Quantified Tolerance Arguments 61
   2.3 Varying the Modality 65
   2.4 Originating Matter 69

3. Motivating Non-Contingency 79
   3.1 Tolerance Arguments and the Sorites 79
   3.2 Non-Soritical Motivations for Non-Contingency 85
   3.3 The Security Argument 89

4. Coincidence Puzzles 100
   4.1 Coincidence Puzzles 102
   4.2 Denying Distinctness 108
   4.3 Denying Non-Coincidence 114
   4.4 Denying Resiliency 119

5. Accepting Hypertolerance 122
   5.1 Chisholm 122
   5.2 From Sufficiency to Overlap Essentialism 129
   5.3 Doubts about Overlap Essentialism 139

6. Hypertolerance and Supervenience 146
   6.1 Supervenience Principles 146
   6.2 Supervenience and Tolerance Puzzles 151
   6.3 Coarse-Grained Hypertolerance 158
   6.4 Denying Microphysical Supervenience 165

7. Rejecting Iteration 171
7.1 Iteration and Tense Operators 172
7.2 Williamson's Earrings 174
7.3 Ancestral Hypertolerance and Higher-Order Identity 178
7.4 Structure 187

8. Iteration for Metaphysical Necessity 199
8.1 The Broadness of Metaphysical Necessity 200
8.2 The Broadest Necessity in the Setting of Classicism 211
8.3 Is There Only One Metaphysically Necessary Truth? 217
8.4 The Broadest Necessity in a Grain-Neutral Setting 221

9. Tolerance and Chance 227
9.1 Chance and Iteration 227
9.2 Chance Fixity, Humeanism, and Approximation 234
9.3 Chance and Robustness 240

10. Tolerance and Counterpart Theory 246
10.1 Counterpart Theory 246
10.2 Counterpart Theory and the Basic Modal Logic 250
10.3 Logically Orthodox Counterpart Theory 257

11. Resolving the puzzles 262
11.1 Security and Semantic Plasticity 262
11.2 Plenitude Principles 267
11.3 The Case for Plenitude 274
11.4 Plenitude and Non-Contingency 279
11.5 Plenitude and Common Nouns 282

12. Refinements and Choice Points 291
12.1 Plenitude, Plasticity, and Hypertolerance 291
12.2 Plenitude, Plasticity, and Coincidence Puzzles 294
12.3 Three Choice Points Concerning Tablehood 296
12.4 Which Properties Are Discriminating? 303

13. Alternatives and Challenges 307
13.1 Alternative Sources of Plasticity 307
13.2 The Challenge from Speech Reports 315
13.3 People 318

14. Indiscernible Tolerance Arguments 327
14.1 Qualitativeness 327
14.2 Indiscernible Tolerance Arguments 331
14.3 Rejecting Indiscernible Tolerance 339
14.4 Indiscernible Hypertolerance and Microphysical Supervenience 351

15. Non-Qualitativeness and Aboutness 361
15.1 The Non-Qualitativeness of Tablehood 361
15.2 From Non-Qualitativeness to Plenitude 364
15.3 Aboutness 366
15.4 What Is Tablehood About? 370
15.5 New Objects 373

Appendix A Modal Logics with and without Necessitation 385
Appendix B Rigidity and Ancestral Iteration 390
Appendix C Consequences of the Rigidity Axioms 395
Appendix D Narrower Modalities in Higher-Order S4 400

References 411
Index of Names 427
Index of Topics and Named Principles 430

Foreword

This book had an unusual history. In late 2018, Juhani and John drafted a short monograph entitled 'Living on the Edge', on puzzles of modal variation, developing a treatment of those puzzles based on a "plenitudinous" ontology of material objects and widespread semantic shiftiness in demonstratives and common nouns. Cian read the draft and had ideas about how to develop it further. Optimistically, we then planned a short*ish* monograph that would be completed by the three of us by the summer of 2019. Things did not go according to plan: we had not properly controlled for Cian's completist tendencies, or for John's habit of indulging them, which led to a book that exemplifies its own subject matter. Its Ship-of-Theseus-like transformation was largely overseen by almost daily morning online meetings between John and Cian, since we quickly realized that regular three-way meetings were too hard to co-ordinate. (Cian bears sole responsibility for the appendices.) The resulting book develops the original strategy for dealing with the puzzles, but also uses them as an entry point into a wide swath of other metaphysical questions about modality. It also introduces and deploys some tools of higher order modal logic that we think deserve a more central place within metaphysics (and indeed within philosophy more generally). We hope that the added heft won't be too daunting; for those too short on time to read the book from cover to cover, the Introduction describes some shorter paths.

We have benefited from extremely helpful conversations with many philosophers about these ideas, including Andrew Bacon, Mike Caie, Maegan Fairchild, Kit Fine, Peter Fritz, Jeremy Goodman, Yoaav Isaacs, Harvey Lederman, Sarah-Jane Leslie, Ofra Magidor, Ted Sider, Jack Spencer, Timothy Williamson, Stephen Yablo, and Dean Zimmerman. Jessica Moss and Jin Zeng provided useful feedback on parts of the manuscript, and Ofra Magidor gave insightful comments on an early draft. We owe special thanks to Ted Sider , who read a complete draft and provided many detailed and incisive comments, and to two anonymous referees for OUP. We also got useful feedback from a series of discussion sessions on the book that were organized at the Dianoia Institute at the Australian Catholic University, from the Grain Exchange reading group, from a seminar at NYU, organized by Paul Horwich and Crispin Wright, and from a seminar that Cian taught at NYU in Fall 2020: we thank all the participants in these events. We are also grateful to the organizers of and audiences at colloquia and conferences where one or other of us presented material from the book, including events at the University of Pittsburgh, the University of Oxford, Madrid University, and the "Metaphysics on the Mountain" online conference organized by Boise University. Thanks to

Foreword

This book had an unusual history. In late 2018, Juhani and John drafted a short monograph entitled 'Living on the Edge', on puzzles of modal variation, developing a treatment of those puzzles based on a "plenitudinous" ontology of material objects and widespread semantic shiftiness in demonstratives and common nouns. Cian read the draft and had ideas about how to develop it further. Optimistically, we then planned a short*ish* monograph that would be completed by the three of us by the summer of 2019. Things did not go according to plan: we had not properly controlled for Cian's completist tendencies, or for John's habit of indulging them, which led to a book that exemplifies its own subject matter. Its Ship-of-Theseus-like transformation was largely overseen by almost daily morning online meetings between John and Cian, since we quickly realized that regular three-way meetings were too hard to co-ordinate. (Cian bears sole responsibility for the appendices.) The resulting book develops the original strategy for dealing with the puzzles, but also uses them as an entry point into a wide swath of other metaphysical questions about modality. It also introduces and deploys some tools of higher order modal logic that we think deserve a more central place within metaphysics (and indeed within philosophy more generally). We hope that the added heft won't be too daunting; for those too short on time to read the book from cover to cover, the Introduction describes some shorter paths.

We have benefited from extremely helpful conversations with many philosophers about these ideas, including Andrew Bacon, Mike Caie, Maegan Fairchild, Kit Fine, Peter Fritz, Jeremy Goodman, Yoaav Isaacs, Harvey Lederman, Sarah-Jane Leslie, Ofra Magidor, Ted Sider, Jack Spencer, Timothy Williamson, Stephen Yablo, and Dean Zimmerman. Jessica Moss and Jin Zeng provided useful feedback on parts of the manuscript, and Ofra Magidor gave insightful comments on an early draft. We owe special thanks to Ted Sider , who read a complete draft and provided many detailed and incisive comments, and to two anonymous referees for OUP. We also got useful feedback from a series of discussion sessions on the book that were organized at the Dianoia Institute at the Australian Catholic University, from the Grain Exchange reading group, from a seminar at NYU, organized by Paul Horwich and Crispin Wright, and from a seminar that Cian taught at NYU in Fall 2020: we thank all the participants in these events. We are also grateful to the organizers of and audiences at colloquia and conferences where one or other of us presented material from the book, including events at the University of Pittsburgh, the University of Oxford, Madrid University, and the "Metaphysics on the Mountain" online conference organized by Boise University. Thanks to

Una Dorr for helping with the illustrations. Finally we would like to thank Peter Momtchiloff at OUP for his patience and support throughout this project.

Cian would like to thank the other participants in the "Grain exchange" reading group—Andrew Bacon, Mike Caie, Peter Fritz, Jeremy Goodman, and Harvey Lederman—for years of discussion which have had a major influence on the more logical parts of the book. Closer to home, Oscar and Una Dorr have provided a steady supply of occasions for paternal pride, and helped keep life bubbling merrily along when 2020 got weird. Cian's contribution is dedicated to Jessica Moss: companion, ally, interlocutor, role model, lodestar.

John would like to thank his friends Juhani and Cian who helped make this project such a fun and rewarding journey, and especially Christina Dietz whose love, encouragement and support kept him going throughout this project, large portions of which were written during the challenging times of the pandemic. Little Marlowe has been an inspiration and John's contribution is dedicated to her.

Juhani would like to especially thank the other participants in the Dianoia Higher-Order Logic Reading Group who read the penultimate draft of the manuscript—Andrew Bacon, Nevin Climenhaga, Steve Finlay, Peter Fritz, Dmitri Gallow, Amir Kiani, Gillian Russell, and Jin Zeng—as well as Catharine Diehl, Beau Madison Mount, and Chunxiao Qu for helpful discussions.

List of Figures

1.1. Logical constants 22
1.2. Notational shorthands 22
1.3. Axioms and rules of H_0 25
1.4. The basic modal logic H_{KA} 36
1.5. Axioms of H_{KR} 42
2.1. Tolerance Argument (*de re* version) 57
2.2. The buckets 60
2.3. Tolerance Argument (quantified version) 63
2.4. The Borja *Ecce Homo* 69
3.1. Security Argument 90
3.2. Chance Security Argument 95
4.1. Coincidence Argument 104
6.1. The earrings about to be made 154
9.1. Chance Tolerance Argument 228
9.2. Approximate Chance Tolerance Argument 239
9.3. Robustness Argument 242
12.1. Three models of the plate-stamping case 302
14.1. The Basic Theory of Qualitativeness 329
14.2. Indiscernible Table Argument 338
15.1. A nanospoon and its environment 365
15.2. The Basic Theory of Aboutness 368

Introduction

This book is long. It could have been somewhat shorter. But it couldn't have been just one sentence long.

This is a recurring pattern: for almost any familiar object, one can find respects in which it could have been somewhat different, but could not have been radically different. For example:

(i) The Great Pyramid could have been a little bit smaller. But it couldn't have been as small as a thimble.
(ii) The *Mona Lisa* could have been slightly different as regards its spatial distribution of colours. But it could not have had the spatial distribution of colours that we find in Edvard Munch's *The Scream*.
(iii) The Vienna Circle could have had somewhat different members. But it couldn't have had Sigmund Freud, Arnold Schoenberg, and Franz Kafka as its only members.
(iv) The game of chess could have been played according to somewhat different rules: for example, it could have been played without the *en passant* rule, or without the rule whereby a thrice-repeated position leads to a draw. But it could not have been played according to the rules of Twister.
(v) The table before us could have been originally made using somewhat different parts: for example, any one of its legs could have been different. But this very table could not have been originally made of a completely different collection of parts.

These are what tradition calls '*de re*' modal claims: they have to do with what *specific* things could have been like, not just with what general *sorts* of thing there could have been. For example, we are claiming that the Great Pyramid *itself* could have been somewhat smaller, not just that there could have been *some pyramid or other* (in the same place, made by the same people …) that is somewhat smaller than the Great Pyramid actually is. Unlike those other claims, ours implies that *there is something* that could have been somewhat smaller than it in fact is. Likewise, when we say that the Great Pyramid couldn't have been thimble-sized we mean that *that very pyramid* couldn't have been thimble-sized. The ancient Egyptians could of course have placed *a* thimble-sized pyramid in the Valley of the Kings right around where the Great Pyramid actually stands; but it would not have been the Great Pyramid itself.

The Bounds of Possibility: Puzzles of Modal Variation. Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri, Oxford University Press. © Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri 2021.
DOI: 10.1093/oso/9780192846655.003.0001

Our primary concern in this book will be with some arguments that threaten to undermine these obvious-looking judgements. These arguments purport to show that if the objects in question are *tolerant*—capable of being somewhat different in the relevant respects from the way they in fact are—they are also *hypertolerant*—capable of being vastly different in these respects from the way they in fact are.

The motivating thought in these arguments is that the fact that the relevant objects are tolerant—assuming it *is* a fact—doesn't seem to be a mere accident. It is hard to believe that it is even *possible* for there to be a pyramid for which a certain slightly smaller size was impossible, or a table that couldn't have been made of certain slightly different parts. There is thus pressure to think that it is *necessary* that the objects in question are tolerant, in the relevant respect. But since modest differences can add up to vast differences, being necessarily tolerant seems to entail being hypertolerant. For example, suppose the Great Pyramid is necessarily tolerant in the following sense: necessarily, however tall it is, it could have been 10 per cent shorter. Since it is in fact approximately 134 m tall, it could have been approximately 121 m tall. But if it were 121 m tall, it would then be possible for it to be 10 per cent shorter than *that*, i.e. approximately 109 m tall. So it is possible for it to be possible for it to be 109 m tall; but from this, it seems to follow that it is is just *possible* for it to be 109 m tall. Multiple applications of this mode of reasoning will force us to say that the Great Pyramid could, after all, have been thimble-sized.

These arguments can be fitted to a general schema. Each begins with a *Tolerance* premise that says that a certain thing or certain things could have been slightly different in a certain way or range of ways. Each then adds a *Non-contingency* premise according to which the Tolerance premise is necessarily true if true at all. With the necessitation of the Tolerance premise in place, one can use certain principles of modal logic—most interestingly, an *Iteration* premise that says that what is possibly possible is possible—to derive a *Hypertolerance* conclusion, which says that the thing or things could have been *arbitrarily* different in the given respect. Many arguments of this form are puzzling in the way that many well-known philosophical arguments are: the premises are rather plausible, but the conclusion seems unacceptable. While the label 'paradox' may be a bit strong, these arguments at least present us with an intriguing class of *puzzles*, since it is not obvious what to say about them. Should we accept the conclusion that the objects in question are hypertolerant? And if not, which premise should we deny?

While these "Tolerance Puzzles" may initially seem like mere curiosities, they offer an entry point into an array of deep and challenging metaphysical questions having to do with modality. This book will be structured as an exploration of various options for addressing Tolerance Puzzles. But along the way, we will be investigating most of the questions about modality that have taken centre stage in metaphysics since around 1970.

In the last couple of decades many metaphysicians seem to have been losing interest in these questions. Some herald an era of "post-modal" metaphysics, in which modal concepts will no longer play a central role in setting the metaphysical agenda.[1] Sometimes, the thought is that modal notions are too *superficial* for debates involving them to be worth spending much time on.[2] Sometimes, the complaint is that the modal questions are too *coarse*, so that the answers to them fail to settle the answers to the really important "hyperintensional" questions in the vicinity. In some respects, then, we are swimming against the tide. But we make no apologies. The focus on questions involving modality brought a lot of clarity and discipline to metaphysics, a field notorious for its tendency to degenerate into obscurity and unruliness. This flowering was made possible by the wide uptake of modal logic, especially in the wake of Kripke (1959, 1963), which created a field rich enough to be of interest not just within metaphysics but across a wide range of fields within and far beyond philosophy. Whether or not one finds the modal questions interesting enough in themselves to deserve the central place in metaphysics they enjoyed in the late twentieth century, the new-found rigour that they made available should make them an obligatory part of any reasonable metaphysical education. And for what it's worth, we think that the modal questions that will come up in the course of our exploration of Tolerance Puzzles are among the most interesting questions in metaphysics, neither too superficial nor too coarse to be set aside as somehow second rate.

Although we are thus continuing a tradition, there is one important way in which our work departs from its twentieth-century precursors: the formal language we use in regimenting claims and arguments is that of *higher-order* modal logic, which goes beyond the *first-order* modal logic that is central to Kripke's work. Higher-order languages are characterized by the availability of quantification into a wide range of different grammatical categories, not just into the category of 'singular terms' characteristic of first-order quantifiers. These languages provide a simple and rigorous way of regimenting informal talk of properties, relations, conditions, operations, and propositions or states of affairs, and thus serve as a helpful vehicle for expressing claims at the level of generality to which metaphysics tends to aspire. We don't mean to suggest that higher-order resources are indispensible for articulating Tolerance Puzzles; but as we will see, they are useful for capturing the general structure which the puzzles have in common, and in our discussions of particular strategies for solving Tolerance Puzzles, higher-order regimentations will play an increasingly important role.

Higher-order languages are far from new: indeed, higher-order quantification is present already in Frege's *Begriffschrift* (1879), perhaps the founding document

[1] See Sider 2020: ch. 1. [2] See Sider 2011: ch. 11.

of analytic philosophy. But thanks especially to Quine (1948, 1970), higher-order formalization fell out of favour within metaphysics in the latter half of the last century, only quite recently making something of a comeback. Like first-order logic (modal and nonmodal), higher-order logic (modal and nonmodal) is a rich field, whose interest extends far beyond metaphysics, and indeed far beyond philosophy (e.g. into mathematics and computer science). We are thus hopeful that its broader uptake within metaphysics will bring further clarity and discipline to that field.

The Case for Tolerance

The chapters that follow offer a wide-ranging exploration of the main strategies for solving Tolerance Puzzles. After mapping out the argument more carefully and relating it to some other puzzles, we will spend several chapters each on the options of accepting Hypertolerance, denying Iteration, and denying Non-contingency. Of course, Tolerance Arguments can also be blocked by rejecting the Tolerance premises: for example, one might claim that the Great Pyramid couldn't have been even a little bit shorter than it in fact is, that the table couldn't have been made of even slightly different originating matter, etc. Logically speaking, it would have made sense to have included a chapter devoted to Tolerance-denial. But, while there is room for resistance to certain specific versions of Tolerance as they occur in certain puzzles, we in general found the Tolerance premises so overwhelmingly plausible that we didn't have a chapter's worth of things to say about them. So in place of such a chapter, we will now briefly present our reasons for being sceptical of Tolerance-denial as a general strategy for handling Tolerance Puzzles.

The plausibility of Tolerance claims does not depend on any particularly philosophical sensibility. In many cases, there is a direct route to a Tolerance premise from certain very mundane beliefs that we could easily find ourselves expressing in everyday life. A carpenter, finding to her chagrin that a table she made doesn't quite fit through a doorway, might say: 'If only I had made the planks a little narrower, this table would have fit through here.' On the face of it, she is making a claim that entails that there is a table such that if she had made the planks a little narrower, that very table would have fit through the doorway. Similarly, consider 'If Neurath hadn't joined the Vienna Circle, it would have been much less influential.' Of course, these are counterfactual conditionals, not possibility claims. But in combination with other bits of mundane background knowledge they entail corresponding possibility claims: for example, if the table *would* have been small enough to fit through the doorway if the carpenter had made the planks a little narrower, then it certainly *could* have been small enough to fit through the doorway. And there is no need to appeal to counterfactuals, since possibility claims about particular objects are also common in our ordinary discourse. For example,

a carpenter might say: 'I could have made this small enough to fit through the door, if only you had sent me the details', or a historian might say: 'Many versions of the castling rule in chess were proposed, any of which could easily have become standard for the game.'

Of course, these ordinary possibility judgements involve notions of possibility much more demanding than the very permissive status of *metaphysical* possibility that features in many Tolerance Arguments.[3] But the inference from the claim that something could *easily* have been the case to the claim that it is metaphysically possible for it to be the case seems unproblematic.[4]

One might try to resist the case for Tolerance from ordinary practice by allowing, for example, that 'This table could have been less than four feet wide but isn't' has a true reading, while denying that this reading is *de re* in the sense of entailing 'There is something that could have been less than four feet wide but isn't.'[5] Analogously, the most salient reading of 'The president could have been Hillary Clinton but isn't' is true but doesn't entail 'There is someone who could have been Hillary Clinton but isn't.' It is a familiar observation that some expressions—such as 'someone', 'every boy', and 'the president of the US'—generate systematic ambiguity when they share clauses with modal words like 'could', 'possible', and 'likely'. For example, 'Someone was likely to come in' can mean 'It was likely that someone would come in' or 'There was someone who was likely to come in.'[6] However, "narrow scope" readings—those which block existential generalization—seem absent, or at least hard to access, for names and demonstratives. 'John is in the car, and John could have broken his leg' and 'That guy in the car could have broken his leg' seem unambiguously to imply 'There is someone in the car who could have broken his leg.'[7] Perhaps there are a few special cases where demonstratives

[3] For the standard way of introducing that status see Kripke 1972. We will have much more to say about what metaphysical necessity and possibility are, and how they relate to other kinds of necessity and possibility, in Chapter 8.

[4] And even if one somehow rejected that inference, one can, as we shall see, generate gripping Tolerance Puzzles using certain other modalities, such as nomic possibility and *having nonzero objective chance*, for which the inference from the ordinary 'could's to the ones that generate the puzzle is, if anything, even more straightforward.

[5] Leibniz suggests that his Tolerance-unfriendly brand of essentialism can be reconciled with various ordinary claims which seem to conflict with it by treating certain uses of proper names as definite descriptions: 'as, for instance, when we mean by Adam the first man, whom God puts in a pleasure garden, which he leaves through sin, and from whose side God makes a woman' (Leibniz 1956: 515f.). Elsewhere he routinely uses proper names as predicates: 'I will now show you some [worlds], wherein shall be found, not absolutely the same Sextus as you have seen . . . but several Sextuses resembling him' (Leibniz 1710: §414).

[6] The most familiar (though not the only) approach to such ambiguities explains them using the same notion of "scope" that also covers the ambiguities that arise when multiple quantificational expressions share clauses with one another. For example, 'At least one person reads every paper published in *Mind*' can mean either 'There is at least one person who reads every paper published in *Mind*' or 'Every paper published in *Mind* is read by at least one person'.

[7] The idea that proper names are just as accessible to existential generalization when they occur in the scope of modals as they are anywhere else is one of the central themes in Kripke 1972. Kripke adopts a theoretical framework (originated by Carnap 1947) that uses the ideology of "designation relative to

in modal sentences admit readings which don't allow existential generalization: consider 'That could have been me' (pointing to the lottery-winner on television). But the target sentences seem manifestly not special in that way. On the intended reading of 'This table could have been narrower than it is', the inference to 'There is a table that could have been narrower than it is' is completely straightforward. In any case, there is no real need to conduct the discussion using names and demonstratives at all: we can just as well get the puzzles going using quantified judgements like: 'Every table you made today could easily have been made a bit smoother by sanding it more carefully.'

In making such appeals to "how we talk", we are not putting forward arguments of the form 'The folk say that *P*; the folk are to be trusted as regards whether *P*; therefore *P*.' We don't need to conceive of ourselves as philosophical anthropologists, documenting the practices of the folk and worrying about whether it would be in some sense chauvinistic or presumptuous to override them. We are in the relevant sense among the folk; the relevant ordinary claims are ones that we are highly confident in, and take ourselves to know using the same methods that non-philosophers use. Appeals in this book to "ordinary practice" should be taken in this spirit.

One might be tempted to contrast our "conservative" methodology that places a premium on these ordinary beliefs with some imagined "revolutionary" alternative that sets ordinary beliefs aside and approaches the subject without prejudice. But it is hard to imagine a sensible version of what such a methodology could actually look like. If at the outset of philosophical theorizing we set aside our confidence in the things we ordinarily take ourselves to know about the topic at hand, where is our confidence in the premises of philosophical arguments going to come from? In practice, an attempt to follow the path of the revolutionary seems likely to amount to a credulous embrace of whatever collection of philosophical arguments happens to first catch your eye. We don't want to go so far as to put forth some general doctrine that ordinary platitudes can never be overturned by philosophy. But the more carefully one looks at the bits of philosophy that would be required to turn our puzzles into an argument against some ordinary Tolerance belief, the harder it is to imagine how one could reasonably regard the conjunction of the arguments' premises as more compelling than the Tolerance claims they attempt to undermine.[8]

Tolerance Puzzles arise from certain tensions between ordinary Tolerance beliefs and certain other prima facie plausible claims, among which are the claims that the relevant objects are not *hypertolerant* in the relevant respects: it seems to

a world". In this setting, the claim can be put by saying that names designate the same thing relative to every world: they are "rigid designators". But for those who, like us, are not particularly fond of this semantic ideology, it is important to see that the central thought can be stated in a far more neutral way. (See Yli-Vakkuri, unpublished *b* for a critical discussion of Kripke's semantic ideology.)

[8] Our attitude here is similar to Moore's (1959) attitude to arguments for scepticism.

many of us that the Great Pyramid could not have been thimble-sized, that this very table could not have been made of some completely non-overlapping piece of wood, etc. But there is a salient contrast in strength between the pressure to accept Tolerance and the pressure to reject Hypertolerance, at least when both are interpreted as claims about metaphysical possibility. The tendencies of thought that make Hypertolerance strike us as odd do *not* seem particularly firmly embedded in ordinary practice. Of course, with a little Socratic questioning, one can elicit judgements from non-philosophers about all sorts of philosophical issues, including questions of Hypertolerance. But such dispositions are notoriously weak and variable, and it would be disastrous to lump them together with the more spontaneous and mundane dispositions that underlie our Tolerance beliefs.[9]

True, we have many mundane pre-philosophical views to the effect that particular objects *would* not have been created under various counterfactual circumstances ('This table wouldn't have been made if I hadn't come across that particularly beautiful slab of oak'), or to the effect that they *could* not have been created without certain circumstances obtaining, in some quite demanding sense of 'could' ('The Great Pyramid could never have been constructed without access to an abundant supply of slave labour'). But whereas there is a plausible route from ordinary 'would have' and 'could feasibly have' claims to claims of metaphysical possibility, there is no straightforward route from 'wouldn't have' or 'couldn't feasibly have' claims to claims of metaphysical impossibility. Giving up Tolerance would require positing pervasive error right at the heart of our everyday practices, whereas accepting Hypertolerance for metaphysical possibility would make barely any difference. So, if we were convinced that there was no way to avoid the choice between giving up Tolerance and embracing Hypertolerance, we are inclined to think that accepting Hypertolerance would quite obviously be the way to go.[10]

Despite this asymmetry in the initial state of play, when philosophical neophytes are confronted with the puzzles, their first impulse is often to reject the Tolerance premise in favour of some quite radical view on which the objects in question couldn't exist at all without being exactly the way they actually are in the relevant respect. Perhaps this impulse isn't so surprising: as everyone who has taught philosophy knows, many philosophical neophytes start out with an affinity for radically revisionary views about all sorts of questions. But the impulse to reject Tolerance doesn't just spring from a taste for denying the obvious. It is also due,

[9] Moreover, with artful Socratic questioning one can also warm people up to many initially odd-sounding Hypertolerance claims. For example, we might get someone to concede that an unfinished book is still a book, and that a book that has only just been begun is an unfinished book. At that point, the claim that this book could have been just one sentence long may seem much more tempting.

[10] As we will see, Tolerance Puzzles can be raised using many different interpretations of possibility, including closer-to-home statuses like easy or feasible possibility as well as the broad status of metaphysical possibility. Hypertolerance claims involving the narrower modalities would be much harder to reconcile with our ordinary practices. However, the main motivations for thinking that Iteration holds for metaphysical possibility do not carry over to these narrower modalities.

in part, to a bad argument which purports to derive intolerance from the logic of identity. The argument goes something like this: 'The table that would have been made if the carpenter had made the boards a little narrower is not exactly the same size as this table. So, the table that would have been made if the carpenter had made the boards a little narrower is not identical to this table. So, if the carpenter had made the boards a little narrower, this table would not have been made.'[11]

The problem here is with the starting premise, not with the subsequent inferences. Since everything is exactly the same size as itself, and since this table is the one that would have been made if the carpenter had made the boards a little narrower, the table that would have been made *is* exactly the same size as this table. The truth in the vicinity of the first premise is that the table that would have been made if the carpenter had made the boards a little narrower *would not have been* exactly the same size that this table *actually* is. But that truth does nothing to support the claim that this table wouldn't have been made if the carpenter had made the boards a little narrower, just as the fact that if you had drunk three cups of coffee you would have drunk more coffee than you did drink does nothing to support the claim that you couldn't have drunk three cups of coffee.

Note that the contrast between qualitative identity and numerical identity plays no role in this diagnosis: contrary to what is often taught to students, the problem is not that the first premise is true only if 'identical' means 'qualitatively identical' while the second step is true only if 'identical' means 'numerically identical'. Whatever qualitative identity is, everything is certainly qualitatively identical to itself, and so 'x is not qualitatively identical to y' implies 'x is not numerically identical to y'.[12]

A slightly different way of generating a spurious conflict between Tolerance claims and the logic of identity avoids the grammatical blunder of ignoring mood by instead appealing to things like possible worlds or situations: 'The table in this situation is four feet wide, but the table in that situation is not four feet wide; so, the table in this situation is not the table in that situation.' The argument has a superficial form that resembles good arguments, like 'The table that John thumped is big, but the table Cian thumped is not big; so the table that

[11] Look (2013) seems to make something like this argument, while also apparently conflating the principle we call Leibniz's Law (that identity requires having the same properties) with the Identity of Indiscernibles (its converse): 'This is indeed an important consequence of [Leibniz's] Principle of the Identity of Indiscernibles. When we speak of the Adam who brought sin into the world and an Adam who did not, we cannot, strictly speaking, be referring to the same individual.' Arguments of this kind are quite widely discussed and taken seriously in the temporal setting, as arguments for the claim that nothing can persist through change: see e.g. Gallois 2016.

[12] For evidence that some students at least are getting this faulty training, see https://quizlet.com/71822821/personal-identity-concepts-and-theories-flash-cards/: a flash card on 'Numerical vs. Qualitative Identity' reads: in part: 'Two things can be numerically identical but qualitatively non-identical (a penny minted in 1914 and that same penny found in 2014).' Similarly, in the context of discussing change, Gallois (2016) says that 'it seems that *a* and *b* can be numerically identical without being qualitatively identical by having different qualities at different times.'

John thumped is not the table Cian thumped.' But it also has a superficial form that resembles bad arguments, like 'In Boston, Van is admired, but in Oxford, Quine is not admired, so Van is not Quine', and 'The apartment in the brochure is luxurious, and the apartment in the video isn't luxurious, so the apartment in the brochure is not the apartment in the video.' The last of these arguments is the best guide to what is going on with the argument about situations. Being four feet wide is compatible with not being four feet wide in some non-obtaining situation, just as not being luxurious is compatible with being falsely characterized as luxurious by some misleading brochure.

In fact, the very concept of identity is something of a red herring as far as questions of tolerance are concerned. We don't need to mention identity at all to raise the questions. For emphasis, we may ask: 'Could a somewhat smaller table be identical to this one?' or 'Could this table have been somewhat smaller while still being one and the same table?'; but nothing important is lost if we stick to 'Could this table have been somewhat smaller?'.[13] If we do drag in identity, we will need to be careful not to put things in confusing ways that make some invalid arguments look like mere applications of Leibniz's Law. By sticking to identity-free formulations, we can avoid such dangers without effort.

Apart from these altogether confused arguments, various philosophers have given other, more sophisticated reasons for rejecting Tolerance. Leibniz was led by his "conceptual containment theory of truth" to the view that Adam 'would not have been our Adam, but another Adam, had other events happened to him' (Leibniz 1989: 73). Quine (1953*b*,*c*) is sceptical of the very *intelligibility* of all claims about *de re* modality, motivated apparently by the thought that such claims would have to be explained somehow in terms of analyticity (a property of sentences), where such an explanation would require making arbitrary choices amongst the various expressions that denote some object. And Lewis (1968) is sometimes (erroneously) interpreted as holding that literally speaking, no ordinary thing could have different properties from those it in fact has, on the grounds that this would require being part of many different "worlds" whereas in fact every ordinary thing is part of only one "world".[14] While interesting, these arguments are too idiosyncratic, or otherwise alien, for us to engage with here.

Of course, one could motivate the rejection of Tolerance simply from the Tolerance Puzzles themselves. If one found Hypertolerance repugnant enough, and the cases for Iteration and Non-contingency compelling enough, then one might feel inclined to reject Tolerance, even if one was initially well disposed towards it. However such a reaction seems wrongheaded to us, since, as we have already said, the case for Tolerance looks dramatically stronger than the case against

[13] Lewis (1986*a*: 193) makes the same point: 'We *do* state plenty of genuine problems in terms of identity. But we needn't state them so.'

[14] See Chapter 10 for a discussion of what we take to be Lewis's actual views.

Hypertolerance—and far stronger still than the combined cases for Iteration, Non-contingency, and the denial of Hypertolerance. And the more time we have spent in the company of the puzzles, doing our best to articulate interesting arguments for elements of that trio, the more we have been confirmed in our sense that none of the arguments in the offing have any prospect of rising to the heady level that would be required to justify anything as revolutionary as a wholescale abandonment of Tolerance.

Philosophers who are more open than we are to the denial of Tolerance often try to make it seem more liveable by appealing to the distinction between strict and loose talk. The thought is that even if one holds that *strictly speaking* a certain object couldn't have existed without having all the properties it in fact has, one can still grant that various sentences whose *literal* truth would conflict with that claim—e.g. 'Each of these tables could have been less than four feet wide'—are true in some loose, nonliteral sense.[15] If such a view could be sustained, it would indeed take some of the sting out of the denial of Tolerance. However, it is quite difficult to sustain. The challenge that must be met is not just that of describing a certain mapping from sentences to the propositions that they "loosely express". One must also explain what stops the propositions in question from counting as being *literally* expressed by the sentences in question. And in the case at hand, it is quite hard to imagine how this could go. To illustrate the difficulty, imagine that the mapping in question is specified in part by associating words like 'could' and 'possible' with a certain property of propositions which the proposition that this table is less than four feet wide does have, despite the fact that it is (strictly speaking) not possible for it to be true. What grounds could there be for denying that the use of 'could' to attribute this status counts as a literal one? There is already good reason to think that 'could' admits a wide range of literal interpretations. For example, the 'could' in 'I couldn't have come to your party, since I had to visit my

[15] A canonical source for such contrasts is the discussion of change over time in Butler 1736. According to Butler, remarks like 'The same tree has stood fifty years in this place', while false 'in a strict and philosophical manner of speech', are nevertheless appropriate 'in a loose and popular sense'. Similar ideas are in play in Chisholm (1976). In the modal case, in discussing Leibniz's view that any object having a certain property could not have existed without that property, Mondadori (1975) says that 'clearly, if one is a super-essentialist, one will not be in a position to interpret de re modal predications and most counterfactual conditionals in a literal way . . . [they] should accordingly be interpreted, and made sense of, in such a way that they turn out not to mean what they actually say'. Mondadori suggests that counterpart theory could be used to provide an account of the relevant non literal use of these sentences (see also Cover and O'Leary-Hawthorne 1999: 115ff.).

The idea of rejecting the literal truth of Tolerance while employing counterpart theory to account for its nonliteral truth is also suggested by a discussion in Kripke 1972 (51, n. 18). Kripke suggests that there is some kind of puzzle arising from the fact that certain Tolerance claims about originating matter seem true, while corresponding Hypertolerance claims seem false, and entertains a response on which "strict identity" applies only to fundamental objects, with "some sort of 'counterpart' notion" taking its place when the modal properties of objects like tables are in question. In Chapter 10 we will discuss the potential relevance of counterpart theory to Tolerance Puzzles, though it seems to us more promising to combine it with the denial of Iteration than with the denial of the literal truth of Tolerance.

aunt' is naturally interpreted differently from the one in 'I could have come to your party, but only at the cost of letting down my aunt.'[16] It seems prima facie arbitrary and uninteresting to decree that among all these ordinary uses of 'could', some wide range of them don't count as "literal".[17]

In any case, an appeal to loose talk does not really make the puzzle go away, since the relevant mode of loose interpretation can be applied uniformly to the whole argument. Loose talk doesn't prohibit deductive argumentation; standard deductive argument forms like Modus Ponens remain equally compelling when they are applied to derive loosely used conclusions from loosely used premises, so long as the same modes of loose interpretation remain in play throughout. So if we can validly deduce a Hypertolerance claim from loosely true premises, we would expect the Hypertolerance claim to be loosely true. But the repugnance of Hypertolerance is by no means driven by a sudden insistence on a strict and literal interpretation. 'Loosely speaking, the Great Pyramid could have been thimble-sized' is no more appealing than the plain 'The Great Pyramid could have been thimble-sized.' Of course, one could block such a deduction by claiming that some of the inference rules employed are not such as to preserve loose truth: but such logical revisionism will have similar costs whether we say that the propositions in play are literally or merely loosely expressed. In what follows, we won't have much more to say about the concept of loose use, which we have not found helpful in sharpening any of the interesting questions raised by Tolerance Puzzles.

Despite the appeal that Tolerance-denial holds for many of those coming across Tolerance Puzzles for the first time, the kind of super-essentialism that would be required to justify Tolerance-denial as a general strategy has very little going for it. As we will see, other approaches to the puzzles offer far richer philosophical rewards.

Guided Tour

Our goal is to provide a careful articulation of the whole family of Tolerance Puzzles, with an eye to their general underlying form, and to examine in detail various strategies for resolving them. As we mentioned above, we have found the level of rigour afforded by formalization in higher-order logic to be helpful in navigating this difficult terrain. So we will begin, in Chapter 1, with a systematic

[16] The canonical defence of this view is Kratzer 1977.

[17] If you were focused on Tolerance claims expressed using words like 'identical' and 'same', you might think that you could characterize the loose use simply by positing a special loose interpretation for those words (following Butler 1736). By contrast with the case of 'could', the challenge to explain why a use of 'identical' to express a relation less demanding than the one philosophers normally use it to express would count as a nonliteral use does not seem so daunting. The answer might be analogous to the reasons why ordinary uses of 'They are exactly the same height' are plausibly classified as nonliteral. But as we emphasized above, there is no need to mention identity at all in stating Tolerance.

presentation of the formal tools that we will be relying on throughout the rest of the book. Some readers may prefer to skip this chapter on a first reading; they may come back to it later if it at some point they find themselves concerned by a potential ambiguity in one of our informal statements, or drawn to a position that accepts the premises but not the conclusion of some argument we are treating as valid. We hope that Chapter 1 will also be able to stand independently of the rest of the book as a primer on higher-order modal logic, written with an eye to using it in metaphysics, as opposed to targeting it as an object of mathematical study.

Chapter 2 lays out in detail the structure of the Tolerance Arguments that are our central topic, presenting a general schema under which various examples can be subsumed. The schema can be instantiated not only with many different parameters of variation but also with many different interpretations of the relevant notion of possibility. Interpretations of interest range from mere metaphysical possibility to more demanding statuses like *having a nonzero objective chance* and *being true at some time*. Chapter 3 says more about the motivations for the crucial Non-contingency premises, contrasting certain forceful motivations with others which, while tempting, are shown to be untenable by the Sorites Paradox. Chapter 4 lays out a different family of "Coincidence Puzzles", which have sometimes been discussed alongside Tolerance Puzzles. We believe it is helpful to consider the two families of puzzles side by side, since some strategies for resolving Tolerance Puzzles generalize fairly easily to Coincidence Puzzles, while others do not.

Our exploration of the strategies for solving Tolerance Puzzles begins in Chapters 5 and 6, which explore the option of accepting Hypertolerance in many or all of the puzzles. Chapter 5 discusses some arguments against Hypertolerance which have been influential in the literature but which we do not find compelling, while Chapter 6 develops what seems to us to be the most forceful challenge to Hypertolerance, based on certain physicalistic supervenience principles. Chapters 7 and 8 turn to the strategy of denying the Iteration premise. Chapter 7 argues that denying Iteration for metaphysical possibility does not provide a sufficiently general solution to the full range of Tolerance Puzzles; Chapter 8 argues that because metaphysical possibility is the *broadest* form of possibility, Iteration holds for it. Chapter 9 focuses on one particular kind of Tolerance Puzzle where there are special barriers to both Iteration-denial and Hypertolerance, namely those in which the operative modality is defined in terms of *objective chance*. Chapter 10 takes up a cluster of ideas from the literature that goes under the name of "counterpart theory" and has been widely thought of as offering some distinctive help with Tolerance Puzzles (often via Iteration-denial). Our view is that counterpart theory is a red herring.

Chapter 11 finally unveils what we take to be the most promising general strategy for resolving Tolerance Puzzles (as well as Coincidence Puzzles)—one that combines a plenitudinous ontology of material objects with semantic shiftiness in

various expressions that figure in the puzzles (e.g. 'The Great Pyramid', 'chess', and 'table'). We explain how this picture undermines the central motivation for Non-contingency that emerged from Chapter 3. Chapter 12 refines the strategy, arguing that a mixed approach that sometimes accepts Hypertolerance and sometimes rejects Non-contingency is better than a uniform treatment; we also survey some interesting further choice points. Chapter 13 considers some nearby alternatives that appeal to shiftiness in other relevant expressions (e.g. 'possible'), and also discusses what we see as the most pressing challenges to our favoured approach.

Finally, Chapters 14 and 15 focus on a special class of Tolerance Puzzles which are new to the literature and raise special challenges that are not straightforwardly addressed by the previous discussion. These puzzles turn on a narrow notion of "indiscernible possibility" on which *qualitative* truths are automatically necessary; this makes for a distinctive new motivation for Non-contingency, based on the qualitativeness of certain properties (like *being a table*). Chapter 14 lays out these puzzles and explores the options left open when the relevant qualitativeness premises are accepted, and Chapter 15 presents our favoured approach to the new puzzles, which involves denying that properties like *being a table* are qualitative, and explores some further questions for it using the ideology of "plural aboutness".

There are several paths through this book in addition to the most obvious approach of reading it from cover to cover. Chapters 2, 3, and 11 are the core chapters: readers who just want to know what the central puzzles are and how we propose to solve them can just read these three chapters. We should warn such readers that, as with all workable strategies for resolving these puzzles, our approach has some uncomfortable and disconcerting features. To fully appreciate why we have our made our peace with these surprises, one will need to absorb the costs of alternative strategies. So we recommend that readers not initially taken with our positive view supplement the three core chapters with those chapters that engage with whichever alternative approach they find most attractive. On the other hand, those who are inclined to regard our view as obviously correct should have a look at the discussion of objections in Chapter 13. Other readers may be particularly interested in our background agenda of demonstrating the utility of higher-order logic for bringing rigour to metaphysics. These readers should certainly not skip Chapter 1; in addition to the core chapters, they will find material of special interest in Chapters 7 and 8 (dealing with Iteration) and in Chapters 14 and 15 (on qualitativeness).

1
Logical Tools

Deductive argumentation will play an important role in this book, as a way of setting up puzzles and mapping out the space of options for resolving them. Many of the arguments we discuss will be intuitive and easily grasped. Nevertheless, there will be a useful role for formal regimentation in providing additional clarity and rigour. Since the formalism required goes somewhat beyond the content of an introductory logic class, this chapter will present all of the required logical tools, in a way that we hope will be accessible to non-specialists.

We will be laying out a certain formal system relative to which certain sentences can be said to be "theorems"; relative to which some can be said to be "logical consequences" of others; relative to which certain collections of sentences can be said to be "logically inconsistent"; and so forth. In endorsing it, we are taking on some controversial commitments, namely commitments to the truth of its theorems. But we stress that we do not intend our endorsement of it to commit us to such further theses as that its theorems are *analytic* or *conceptual* truths; that it would be *incoherent* or inherently irrational to deny them; or that they are truths of *logic* in some sense that goes importantly beyond their being true. We are in fact quite sceptical about the philosophical usefulness of these categories. And even setting this scepticism aside, we think that they are more distracting than helpful when one is concerned with questions of metaphysics. We are willing to debate views that reject even our most central logical principles, and would not wish to resort to appeals to analyticity or accusations of incoherence in doing so. In setting out the principles, we will discuss some of the most interesting ways in which they might be disputed, although a full-dress defence of the principles would be a book in itself. The rest of this book will be almost entirely concerned with views compatible with the logical principles expounded in this chapter.

1.1 Classical Propositional Logic

Our formal language will contain, inter alia, the usual operators $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, and $\neg$. We will be taking for granted that this language obeys classical propositional logic: i.e. that the sentences certified as true by the standard method of truth tables do in fact express only truths, and that the arguments certified as valid by that method are in fact truth-preserving. We will moreover be taking for granted the adequacy

The Bounds of Possibility: Puzzles of Modal Variation. Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri, Oxford University Press. © Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri 2021.
DOI: 10.1093/oso/9780192846655.003.0002

of the standard ways of translating between English sentences involving the words 'and', 'or', and 'not' and such a formal language (at least for the kinds of English sentences we will be considering).[1]

These commitments bear special emphasis in an area of study where vagueness is an important theme. One distinctive thing about vague words is that they can be used to construct so-called "borderline" sentences and questions, which induce a distinctive kind of perplexity. For example, there are some people with middling numbers of hairs on their heads for whom we would be reluctant to give a flat answer of 'yes' or 'no' to the question 'Is so-and-so bald?', even if we were very familiar with the number and arrangement of hairs on their heads. We have an impulse to set aside the original yes/no question and instead say something more informative, like 'Well, he has quite a lot of hair around the sides of his head but he is completely bald on top', or 'He is a bit bald but not very bald.' When we force people back to the original question—'Yes, but is he *bald*?'—they tend to start coming out with speeches that would, if taken literally, require giving up classical propositional logic. For example, someone might say, 'He is not bald, but he is also not *not* bald.' Taken at face value, this is of the form '*P* but not *P*', which plausibly implies the outright contradiction '*P* and not *P*'. In classical propositional logic, everything can be validly derived from a contradiction, since '*P* or *Q*' follows from *P*, while *Q* follows from '*P* or *Q*' and 'Not *P*'. So proponents of classical logic can't take 'He is not bald, but he is also not *not* bald' as literally true. But that seems fine: it is plausible that there is some kind of nonliteral speech going on when we say such things. Besides, it would be little wonder if we got logically confused when being forced to override our natural impulse to respond to 'Is he bald?' by saying something informative that doesn't strictly speaking count as an answer.

Most philosophers (with the exception of the school of Priest (1987)) are not up for accepting outright contradictions. But many have a tendency to say things about borderline cases which suggest routes to a contradiction, sometimes leaving it unclear how these routes are supposed to be blocked. For example, some will say, 'It is neither true nor false that he is bald.' It is very natural to treat 'It is true that *P*' as interchangeable with *P*, and 'It is false that *P*' as interchangeable with 'Not *P*'. But if we do this, then 'It is not true that *P* and it is not false that *P*' will be interchangeable with the outright contradiction 'Not *P* and not not *P*'. Some philosophers, aware of this danger, explicitly reject the natural equivalences, instead treating 'It is true that *P*' and 'It is false that *P*' as strengthenings of *P* and 'Not *P*', respectively. But this is a hard view to live with, since it is hard to generate natural-sounding argumentative prose without relying on these equivalences: to avoid writing premises out multiple

[1] We take no stand on the contentious question how the word 'if' in English relates to →, although we will occasionally use 'if' stipulatively as a material conditional.

times, it is extremely convenient to be able to make arguments along the lines of '*P*. But if that's true, then *Q*. So, *Q*.' or '*P*. But if *Q* were false, *P* would be false. So *Q*.'

Another turn of phrase often invoked in borderline cases is 'There is no fact of the matter as regards whether . . .'. If one is willing to treat this as just another sentential operator, perhaps with a logic similar to that of 'it is contingent whether . . .', it poses no threat here to classical propositional logic. But it is tempting to treat 'He is bald' as interchangeable with 'He is bald and there is a fact of the matter as regards whether he is bald' and 'He is not bald' as interchangeable with 'He is not bald and there is a fact of the matter as regards whether he is bald.' And if we succumb to this temptation, we will be led to the view that 'There is no fact of the matter as regards whether he is bald' implies the outright contradiction 'It is not the case that he is bald, and it is not the case that he is not bald.'

Being committed to classical propositional logic means accepting 'Either he is bald or he is not bald', even when the person in question is one of those perplexing, middlingly hairy people. It's not a comfortable speech: it feels in tension with our impulse to say 'Forget about the word "bald", and let's just talk about how the hair is distributed on the guy's head.' But it is hard to turn this discomfort into an argument, since the obvious things to say can so easily be turned into arguments for outright contradictions.

There are many different accounts of the nature of vagueness which are friendly to classical logic, and they offer many different kinds of therapy that might help us overcome the initial repugnance that such instances of the Law of the Excluded Middle tend to inspire.[2] We will not need to decide on any particular view about the nature of vagueness, or any particular school of therapy. If you are a reader who has not found a therapy you like, you will likely often be tempted to respond to things we say by accusing us of ignoring vagueness, or tendentiously assuming that there is a fact of the matter about some question. In no case is this correct. If it ever should *sound* like we are ignoring vagueness, what is really happening is that we are saying things that, as proponents of classical logic, we endorse even in the presence of vagueness.

Why do we assume classical propositional logic? There are several reasons, First, we think it enjoys by far the greatest abductive support among its competitors both in philosophy and other areas of inquiry (see Williamson 2018). Second, the specific weakenings of classical propositional logic that have been urged as responses to the puzzles of vagueness are generally *so* weak that it is a bit hard to imagine what it would look like to conduct any extended piece of informal reasoning without implicitly invoking any rules that those weakenings reject. For example, they generally give up the rule of *reductio ad absurdum*, which lets us conclude 'Not *P*' if we have shown how to derive a contradiction from *P* together

[2] Examples include the views of Lewis (1973, 1993), Fine (1975), Sorensen (1988), Williamson (1994), Keefe (2000), Graff (2000), Dorr (2003), and Bacon (2018*b*).

with other things we have already established.[3] But this is an extremely central way of arguing for negative conclusions. Trying to engage in systematic informal reasoning about some complex puzzle without ever engaging in a *reductio* is very difficult—rather like writing a novel without using the letter 'e'. Perhaps it can be done; but it is probably best to try doing it the ordinary way first.[4] And third, as a practical matter, we already need to keep track of a large number of consistent packages of views; if we also had to keep track of all of the further possible positions that open up for those who are willing to add some logical heresy to the mix, the book would have to be a lot longer. We hope that the heretics will nevertheless be able to extract some value from our discussion.

1.2 Higher-Order Languages

This is a book about the metaphysics of *objects*—mostly, physical objects one could pick up or bump into. But we are also going to need to talk about properties and relations. For example, the general schema for Tolerance Puzzles we will present in Chapter 2 will be talking about certain "closeness" relations among properties (e.g. the relation *being properties F and G such that, for some n, F is* being n millimetres tall *and G is* being $n + 1$ millimetres tall). Here, as in so much of metaphysics, the point of talking about properties is to achieve certain kinds of generality. But to do this work, we need to be able to assume that, for example, 'x has the property of being n millimetres tall' can be treated as interchangeable with 'x is n millimetres tall', at least in the contexts we are concerned with. It would defeat the purpose of talking in terms of properties if we had to start engaging seriously with views like 'Although there are many different numbers n such that the Great Pyramid could have been n millimetres tall, there is only one number n such that the Great Pyramid could have had the property of being n millimetres tall.'

These are dangerous waters, however; one can easily land in a contradiction if one generalizes the thinking behind such interchangeability in a natural way. The obvious generalization is the following schema:

[3] See Field 2003, Fine 2017*b*, and Weber 2010.

[4] Forbes (1984, 1985) advocates a unified approach to Tolerance paradoxes and standard Sorites Paradoxes, based on a very radical weakening of classical logic. Modus Ponens—which he dubs 'the fallacy of detachment'—is invalid in his logic, as are a host of other familiar rules like modus tollens, disjunctive syllogism, conjunctive syllogism, hypothetical syllogism, etc. (The logic is the one generated by Łukasiewicz's model theory where the compositional semantic values of sentences are numbers in the real interval $[0, 1]$ (Łukasiewicz and Tarski 1930; Goguen 1969), when one identifies the valid arguments with those whose conclusions never have a value less than the minimum of the values of the premises.) While one could *perhaps* learn to live with the idea that Modus Ponens itself is invalid for certain complex 'if'-sentences in English (McGee 1985), Forbes's rejection is vastly more far-reaching, since it extends to almost all multi-premise reasoning that uses vague vocabulary. Given how pervasive vagueness is, we find it hard to imagine what it would even look like to do philosophy or any other reasoning without implicitly relying on some or other of the rules that Forbes rejects.

Naïve Property Conversion For all x, x has the property of being an x such that P iff P

Here P can be any formula: in the interesting cases, it will be a formula in which the variable x occurs free. The contradiction is that of Russell's Paradox (Russell 1903). Abbreviate 'the property of being an x such that x does not have x' as r. Then plugging in 'x does not have x' for P in Naïve Property Conversion yields 'For all x, x has r iff x does not have x.' By universal instantiation, this implies 'r has r iff r does not have r', which is of the form 'P iff not P'. This is inconsistent (in classical propositional logic).

Our strategy for legitimizing and making rigorous the kinds of property-talk that we will need to engage in will be to treat talk of properties, relations, and propositions as a mere manner of speaking, to be officially cashed out in the formalism of higher-order logic. Higher-order logic is a generalization of first-order logic where quantification into the position of any type of expression is allowed—for example, quantification into predicate position and quantification into sentence position.

To clarify what is meant by "quantifying into" different kinds of syntactic positions, consider the following simple sentence, which is well formed in familiar first-order languages:

(1) $\neg\text{Foolish}(\text{Socrates})$

In first-order logic, there is only one way to existential generalization in this sentence, namely to (validly) infer

(2) $\exists x(\neg\, \text{Foolish}(x))$

But in higher-order logic, any of the syntactic constituents of (1) can be replaced by a variable of the same syntactic category as that constituent, bound by an existential quantifier. So we get four more sentences out of (1):

(3) $\exists X(\neg X(\text{Socrates}))$

(4) $\exists O(O(\text{Foolish}(\text{Socrates})))$

(5) $\exists p(\neg p)$

(6) $\exists p(p)$

The variable X in (3) is of the same syntactic category as an ordinary monadic predicate (like 'Foolish'); the variable O in (4) is of the same syntactic category as a monadic sentential operator (like ¬); the variable p in (5) and (6) is of the same

syntactic category as a sentence. All of these sentences will follow from (1) by the higher-order version of the existential generalization rule (which we will present in the next section).

As Prior (1971) points out, English seems to exhibit some of the quantificational flexibility that these formal languages display. For example, we could try rendering (3) as 'There is something that Socrates isn't' (which seems to have a reading where it could be completed by 'Namely: foolish' rather than by 'Namely: the Parthenon'). We could render (5) as 'There is something that is not the case.' However, we would be hard-pressed to provide helpful English glosses on most higher-order sentences if we had to limit ourselves systematically to these resources. So, when we are speaking informally, we will use words like 'property', 'relation', and 'proposition' in glossing higher-order quantification. For example, we will pronounce (3) as 'There is a property that Socrates does not have', (4) as 'There is a property that the proposition that Socrates is foolish has', (5) as 'There is a true proposition', and (6) as 'There is a false proposition.'[5] In adopting this way of speaking, we do not mean to be taking any stand on how expressions like 'property' work in ordinary English, and a fortiori do not want to regard the formal sentences as being introduced by the stipulation that they should be equivalent to the English sentences we use to gloss them. ('Socrates instantiates himself' is a grammatical English sentence, but there is no natural mapping from English into the formal language that is both defined on that sentence and maps 'There is a property that Socrates does not instantiate' to $\exists X(\neg X(\text{Socrates}))$.) Rather—like Williamson (2003: 416–17, 459)—we think that the language of higher-order logic can be learnt "by the direct method", without taking a stand on how it relates to ordinary language. And we think that it is a very good language for students of metaphysics to learn, since it can be used to raise many interesting and precise questions while bypassing a panoply of puzzles, such as the one about Naïve Property Conversion, that ordinary English talk of "properties" presents us with when taken at face value. We hope that this book can serve as something of an advertisement for its usefulness as a tool for doing metaphysics.

To specify the syntax of a higher-order language rigorously, we need to begin by specifying the syntactic categories to which its expressions can belong. The standard way of doing this is to think of the language as given not just by the *set* of

[5] This way of using 'proposition' is a bit risky, since some philosophers treat 'proposition' as tightly tied to propositional attitudes like belief, while simultaneously adopting an account of propositional attitude ascriptions on which they generate exceptions to otherwise valid schemas about identity. Certain such philosophers might want to say, 'The proposition that someone is looking at Hesperus is distinct from the proposition that someone is looking at Phosphorus, since someone could believe the former but not the latter; but the property of looking at Hesperus is identical to the property of looking at Phosphorus, since Hesperus is Phosphorus.' We invite those who think that propositional attitude contexts (in English) are logically exceptional in this way, and use 'proposition' in a way that implicates it in these exceptions, to substitute 'state of affairs' whenever we have 'proposition'. What's wanted is some word that can stand to 0 as 'property' stands to 1 and 'binary relation' stands to 2.

all its well-formed expressions, or "terms", but by a *relation* which associates each term with a unique item called a "type", which fixes its syntactic category. It doesn't really matter what the types *are*, since their job is just to serve as markers, but we will take them to be strings of symbols (just as terms are also standardly taken to be strings of symbols). In the so-called "relational" variety of higher-order language we will use, the set of types is the smallest set T such that:

(i) T contains the letter 'e'.

(ii) Whenever T contains $\sigma_1, \ldots,$ and σ_n, it also contains $\langle \sigma_1, \ldots, \sigma_n \rangle$ (i.e. the result of concatenating $\sigma_1 \ldots \sigma_n$ with commas and two angle brackets).

Type e will be the type associated with singular terms. Type $\langle \sigma_1, \ldots, \sigma_n \rangle$ will be the type associated with n-place predicates that require arguments of types $\sigma_1, \ldots, \sigma_n$. For example, a term of type $\langle e \rangle$ is an an ordinary monadic predicate taking a singular term as an argument; a term of type $\langle e, e \rangle$ is an ordinary dyadic predicate; a term of type $\langle \langle e \rangle \rangle$ is a monadic predicate that takes a term of type $\langle e \rangle$ rather than an ordinary singular term as its argument; a term of type $\langle e, \langle e, e \rangle \rangle$ is a dyadic predicate that takes one singular term argument and one type-$\langle e, e \rangle$ argument. We understand rule (ii) to include the case where $n = 0$, so as a special case we have the type $\langle \rangle$. Terms of this type are called *formulae*, or *sentences* if they don't contain any free occurrences of variables: this is analogous to the common practice in first-order languages of treating sentence letters as 0-place predicates.[6]

The terms of each type include an infinite supply of variables of that type, and perhaps also certain constants. We can make complex formulae out of other terms by combining a predicate with arguments of the right types: when A is a term of type $\langle \sigma_1, \ldots, \sigma_n \rangle$ (for some $n > 0$) and $B_1, \ldots, B_n$ are terms of types $\sigma_1 \ldots, \sigma_n$ respectively, $A(B_1, \ldots, B_n)$ is a formula. (We sometimes omit the parentheses and commas.) We also have a way of making complex terms that are *not* formulae out of variables and formulae: when P is a formula and $v_1 \ldots, v_n$ (for some $n > 0$) are variables of types $\sigma_1, \ldots, \sigma_n$, $(\lambda v_1 \ldots v_n.P)$ is a term of type $\langle \sigma_1, \ldots, \sigma_n \rangle$. Lambda-terms are the formal analogue of complex predicates in English and other natural languages. For example, we can read '$\lambda x.F(x) \wedge G(x)$' as 'is such that it is F and it is G' (or just 'is F and G'); more generally, we can read $(\lambda v_1 \ldots v_n.P)$

[6] Many other treatments of higher-order logic use a different, "functional", notation for types, where each complex type $\sigma_1 \rightarrow \sigma_2$ is generated from two simpler types, and the basic types are e and t. So long as we ban types of the form $\sigma \rightarrow e$, this system is equivalent to ours, where t corresponds to $\langle \rangle$, $e \rightarrow t$ corresponds to $\langle e \rangle$, $t \rightarrow t$ corresponds to $\langle \langle \rangle \rangle$, $(e \rightarrow t) \rightarrow (e \rightarrow t)$ corresponds to $\langle \langle e \rangle, e \rangle$, etc. (See Dorr 2016*b*: App. A for a rigorous definition.) We have no deep objection to the intelligibility of terms of types ending in e, but they do raise some further interpretative challenges, and depending on how these are resolved, they may add complications to the logic.

as 'be (a) $v_1 \ldots v_n$ such that P'. Sometimes it is grammatically smoother to use gerunds, as in 'being F and G' or 'being $v_1 \ldots v_n$ such that P'.[7]

Predication and lambda-abstraction are the only two ways of constructing complex expressions. By contrast with standard presentations of first-order syntax, where symbols like $\neg$ and $\wedge$ are treated as punctuation marks corresponding to *sui generis* ways of constructing complex expressions, for us these symbols are just terms of certain types, so that the relevant constructions are subsumed under the general category of higher-order predication. For example, the conjunction symbol $\wedge$ is a term of type $\langle\langle\rangle, \langle\rangle\rangle$, which can thus combine with any two formulae P and Q to make a third, $\wedge(P, Q)$. We treat the usual notation $P \wedge Q$ as a shorthand for this.[8] The treatment of quantification relies on Frege's epochal insight (1879: §9–11, 1892: 173) that quantifiers can be understood as predicates that take other predicates as arguments: for example, 'Everything perishes' is treated like 'Perishing is universal.' Thus for each type σ, the quantifiers $\forall_\sigma$ and $\exists_\sigma$ are terms of type $\langle\langle\sigma\rangle\rangle$: predicates whose arguments are predicates whose arguments are terms of type σ.[9] The usual notation for quantifiers is treated as an abbreviation: when v is a variable of type σ, $\forall vP$ abbreviates $\forall_\sigma(\lambda v.P)$. Finally, for each type σ, we have an identity predicate $=_\sigma$. $=_e$ is just the familiar relation of numerical identity of objects. But predications of higher-order identity are not so unfamiliar: in English, we arguably can express some of them using sentences of the form 'To be F is to be G', or 'For it to be the case that P is for it to be the case that Q': see Dorr 2016*b* (§§1–4).[10]

[7] There is a different dialect of higher-order language where, given a term R of type $\langle e, e\rangle$ and a term a of type e, one can directly form a term (Ra) of type $\langle e\rangle$, analogous to our $(\lambda x.Rax)$. Typically this approach where terms are fed their arguments one at a time is the only allowed way of combining terms with arguments, so the closest analogue to our $R(a, b)$ would be $((Ra)b)$. This alternate syntax is usually combined with the alternative notation for types discussed in the previous footnote, but it can also be used with our notation for types as in Muskens 2007. For natural translations between the two dialects, and some potential wrinkles, see Dorr 2016*b* (App. B).

[8] By contrast, the letter λ is, like the parentheses, commas, and period, not a meaningful unit in its own right.

[9] We intend these to correspond to *unrestricted* quantification in English, rather than allowing our formal language to take over the characteristic kind of context-sensitivity generated by the phenomenon of "tacit restriction" in English. This isn't the right place to engage with philosophers who deny the very intelligibility of unrestricted quantification.

[10] Many presentations of logical systems like ours make do with a smaller list of logical constants, treating the rest as mere abbreviations. For example, one might stipulate that any string of symbols of the form $P \vee Q$ where P and Q are formulae is an abbreviation of $\neg(\neg P \wedge \neg Q)$. The cost of this is that the usual translations from English into the formal language become more tendentious than we would like. The standard translation of 'For it to be either raining or snowing is for it not to be both raining and snowing' would be '(Raining $\vee$ Snowing) = $\neg$($\neg$Raining $\wedge$ $\neg$Snowing)'. If we treat 'Raining $\vee$ Snowing' as an abbreviation along the lines considered, this string will abbreviate a sentence of the form $P = P$, a theorem of our logic. But the English sentence we started with is much more debatable, and would be rejected by proponents of certain relatively "fine-grained" theories of higher-order identity. So we prefer a somewhat longer list of logical constants that makes our regimentations of English somewhat less hostage to grain-theoretic assumptions. Related doubts may, nevertheless, arise as regards the adequacy of some of the standard translations we will be working with. For example, we will go along with the familiar rendition of English binary quantifications using singulary quantifiers, e.g. translating 'Every dog barks' as $\forall x(\mathrm{Dog}(x) \rightarrow \mathrm{Bark}(x))$, although again some fine-grained theorists may regard this as a case of mere necessary equivalence rather than identity.

$\neg : \langle\langle\rangle\rangle$	$\forall_\sigma, \exists_\sigma : \langle\langle\sigma\rangle\rangle$ for each type σ
$\wedge, \vee, \rightarrow, \leftrightarrow : \langle\langle\rangle, \langle\rangle\rangle$	$=_\sigma : \langle\langle\sigma, \sigma\rangle\rangle$ for each type σ

Fig. 1.1 Logical constants.

$$P \wedge Q \coloneqq \wedge(P, Q) \qquad P \vee Q \coloneqq \vee(P, Q)$$
$$P \rightarrow Q \coloneqq \rightarrow(P, Q) \qquad P \leftrightarrow Q \coloneqq \leftrightarrow(P, Q)$$
$$a = b \coloneqq =_\sigma(a, b) \quad \textit{where } a \text{ and } b \text{ are of type } \sigma$$
$$\forall \upsilon P \coloneqq \forall_\sigma(\lambda \upsilon.P) \qquad \exists \upsilon P \coloneqq \exists_\sigma(\lambda \upsilon.P) \qquad \textit{where } v \text{ is of type } \sigma$$
$$P_1 \wedge P_2 \wedge \ldots \wedge P_n \coloneqq (P_1 \wedge (P_2 \wedge (\ldots \wedge (P_{n-1} \wedge P_n) \ldots)))$$
$$P_1 \vee P_2 \vee \ldots \vee P_n \coloneqq (P_1 \vee (P_2 \vee (\ldots \vee (P_{n-1} \vee P_n) \ldots)))$$
$$P_1 \rightarrow P_2 \rightarrow \ldots \rightarrow P_n \coloneqq (P_1 \rightarrow (P_2 \rightarrow (\ldots \rightarrow (P_{n-1} \rightarrow P_n) \ldots)))$$

Fig. 1.2 Notational shorthands.

Figure 1.1 lists our basic logical constants and their types, and Figure 1.2 lists the notational shorthands.

The use of higher-order languages as a vehicle for theory-building has a long and distinguished history, including such luminaries as Frege (1879), Russell (1903, 1908), Ramsey (1925), Carnap (1934), Church (1940), Prior (1971), and Montague (1974). And this tradition has recently undergone a resurgence.[11] Nevertheless, the formalism is methodologically controversial. There is a mainstream tradition, spearheaded by Quine (1948), that insists that serious theorizing should be conducted in, or at least canonically translatable into, a first-order language in which all variables are syntactically interchangeable. But as we have seen, one has to work very hard to find a way of regimenting talk of properties and relations within such a language, especially if we want propositions and relations to be abundant enough for quantification over them to play its trademark theoretical role as a device of generalization, but not so abundant as to generate logical disaster.[12] By our lights, even the best first-order theories of properties and relations look inelegant and artificial in a way that should raise warning flags even in those not antecedently committed to the good standing of the higher-order approach. But this book need not be lost on proponents of first-order theories of properties and relations. Any

[11] See, e.g., Williamson 2003, 2013*b*; Dorr 2016*b*; Bacon 2018*a*, 2019, 2020; Bacon, Hawthorne, and Uzquiano 2016; Bacon and Uzquiano 2018; Bacon and Dorr forthcoming; Fritz 2017*b*; Bacon and Russell 2019; Caie, Goodman, and Lederman 2020; and the papers in Fritz and Jones forthcoming.

[12] Quine himself combined a methodological commitment to first-order languages with eliminativism about a wide range of entities including properties, relations, and propositions; but the tendency of more recent proponents of "methodological first-orderism" is to assume that lots of instances of Naïve Property Conversion are true, though normally without saying much about how the true instances are to be distinguished from the false ones.

such theory that is worth its salt should put you in a position to give a first-order "translation" of the sorts of claims we are formulating in a higher-order language, where each higher-order quantifier is mapped to some appropriately restricted first-order quantifier. We invite those who favour such a theory to reinterpret all of our main claims along these lines.

1.3 Higher-Order Logic

So far we have just described the *syntax* of our higher-order language and attempted to convey a feeling for what its sentences are supposed to mean. Let's now turn to the basic logical system we will be presupposing in subsequent chapters when reasoning in this language, which we will call H_0.

Most of the principles of H_0 are just straightforward generalizations of principles already familiar from classical first-order logic. For example, in standard systems of first-order logic, every instance of the following schema is a theorem:

UI$_{\text{sub}}$ $\qquad \forall \nu P \rightarrow P[a/\nu]$

Here, P is any formula, ν is any variable, a is any term, and $P[a/\nu]$ stands for the result of replacing every free occurrence of ν in P with an occurrence of a, so long as this can be done without any of these occurrences of a becoming bound. In H_0 this schema is extended the obvious way so that a can be a term of any type, so long as ν is a variable of that same type—as it must be for the operation of substituting a for ν even to make syntactic sense.

The one axiom of H_0 that *isn't* just a generalization of something from first-order logic is the *Extensional Beta-Conversion* schema:

Eβ $\qquad (\lambda \nu_1 \ldots \nu_n.P)(a_1, \ldots, a_n) \leftrightarrow P[a_1/\nu_1, \ldots, a_n/\nu_n]$

Here, P is a formula; $\nu_1, \ldots, \nu_n$ are one or more variables of any types; $a_1, \ldots, a_n$ are terms of the same types; and $P[a_1/\nu_1, \ldots, a_n/\nu_n]$ is the result of substituting occurrences of $a_1, \ldots, a_n$ for those of $\nu_1, \ldots, \nu_n$ in P (assuming this can be done without any variable free in $a_1, \ldots, a_n$ becoming bound). Eβ is the closest analogue in H_0 of the Naïve Property Conversion schema discussed in the previous section. But it is better to think of it not as some special technical trick for getting the advantages of that schema while avoiding inconsistency, but simply as the analogue of various biconditionals in English that let us go back and forth between subject-predicate sentences with complex predicates and more complex sentences involving simpler predicates—e.g. 'Zuckerberg is rich and famous if and only if Zuckerberg is rich and Zuckerberg is famous.'[13]

[13] Eβ is a restriction of the stronger and and more controversial *β-conversion* schema, whose instances are all biconditionals $Q \leftrightarrow Q'$ where Q' comes from Q by replacing an occurrence in Q

Just as in the first-order case, there are many mathematically equivalent ways of rigorously specifying the logic. In introductory classes, systems of first-order logic are often specified by means of a *natural deduction proof system*. In such a system, a proof of an instance of $\mathrm{UI}_{\mathrm{sub}}$ might look something like this:

1	$\forall vP$	Assumption
2	$P[a/v]$	$\forall$-Elim (1)
3	$\forall vP \rightarrow P[a/v]$	$\rightarrow$-Intro (1–2)

Such systems can be extended straightforwardly to higher-order languages; we just have to remember to keep the types of variables straight when applying the rules for the quantifiers and for identity.[14] But since the details of natural deduction systems get fiddly, and we are not actually planning to write down lots of complete proofs, for our purposes it is more convenient to use a "Hilbert-style" proof system. In such systems one specifies a set of formulae called 'axioms', and then characterizes the set of theorems as the smallest set containing all the axioms which satisfies certain closure conditions, or 'inference rules'. A specification of H_0 in this format is given in Figure 1.3. In each schema, the schematic letters may be replaced with terms of any types that make the whole formula well formed, and likewise for the subscripts on quantifiers and the identity symbol.

Note that the official axiom-scheme UI is a bit different from the $\mathrm{UI}_{\mathrm{sub}}$ schema discussed above. We can prove any instance of the latter schema using an instance of UI and one of Extensional Beta:

1.	$\forall vP \rightarrow (\lambda v.P)a$	UI
2.	$(\lambda v.P)a \leftrightarrow P[a/v]$	Eβ
3.	$((\lambda v.P)a \leftrightarrow P[a/v]) \rightarrow (\forall vP \rightarrow (\lambda v.P)a) \rightarrow \forall vP \rightarrow P[a/v]$	PC
4.	$(\forall vP \rightarrow (\lambda v.P)a) \rightarrow \forall vP \rightarrow P[a/v]$	3, 2, MP
5.	$\forall vP \rightarrow P[a/v]$	4, 1, MP

(Recall that when v is of type σ, $\forall vP$ is short for $\forall_\sigma(\lambda v.P)$.) Similarly, by combining our official axiom-schemes EG and LL with Eβ, we can recover versions of

of a subformula of the form $(\lambda v_1 \ldots v_n.P)(a_1, \ldots, a_n)$ with $P[a_1/v_1, \ldots, a_n/v_n]$. All instances of Eβ are also instances of β-conversion, but β-conversion goes much further: for example, using the β-conversion instance $((\lambda y.Fy)x = Fx) \leftrightarrow (Fx = Fx)$ and the Ref-instance $Fx = Fx$, we can prove $(\lambda y.Fy)x = Fx$, and more generally we can strengthen any instance of Eβ from a biconditional to a claim of propositional identity.

[14] To make it possible to derive every instance of Eβ, we would just add two new basic rules, one ('Reduction') corresponding to the left-to-right direction of Eβ, and the other ('Abstraction') corresponding to the right-to-left direction.

Axiom schemes:

PC $\vdash P$ whenever P is a theorem of classical propositional logic

Eβ $\vdash (\lambda v_1 \ldots v_n.P)(a_1, \ldots, a_n) \leftrightarrow P[a_1/v_1, \ldots, a_n/v_n]$

UI $\vdash \forall F \rightarrow Fa$

EG $\vdash Fa \rightarrow \exists F$

UD $\vdash \forall v(P \rightarrow Q) \rightarrow P \rightarrow \forall vQ$, v not free in P

ED $\vdash \forall v(P \rightarrow Q) \rightarrow \exists vP \rightarrow Q$, v not free in Q

Ref $\vdash a = a$

LL $\vdash a = b \rightarrow Fa \rightarrow Fb$

Rules:

MP If $\vdash P \rightarrow Q$ and $\vdash P$ then $\vdash Q$.

GEN If $\vdash P$ then $\vdash \forall vP$.

Fig. 1.3 Axioms and rules of H_0.

existential generalization and Leibniz's Law which correspond more directly to those encountered in first-order logic:

$\mathbf{EG_{sub}}$ $\quad P[a/v] \rightarrow \exists vP$

$\mathbf{LL_{sub}}$ $\quad a = b \rightarrow P[a/v] \rightarrow P[b/v]$

H_0 is in many ways a rather weak system. Even in the pure language with only logical constants, there are many questions that can be raised that H_0 doesn't settle either way.[15] Some of the most metaphysically interesting of these questions have to do with how "fine-grained" the higher-order realm is. For example, it is compatible with H_0 that (7), (8), or neither is true:

(7) $\quad \forall p(p = (p \wedge p))$

(8) $\quad \forall p(p \neq (p \wedge p))$

[15] By contrast, the only questions that can be raised in the pure language of *first*-order logic whose answers aren't settled by standard classical first-order logic are equivalent, in that logic, to cardinality-theoretic sentences like $\exists x \exists y(x \neq y)$. And those questions aren't all *that* interesting, since there is such a strong case for the view that there are infinitely many things, which settles them all (since distinctions between infinite cardinalities can't be expressed in the pure first-order language).

For all H_0 tells us, conjoining a proposition with itself might always, never, or only sometimes yield the same proposition. The question which of these is the case seems pretty interesting from a metaphysical point of view. It is not obvious what the answer is.[16] Moreover, all three of the answers compatible with H_0 can be derived from certain strong and systematic theories which deliver answers to many other questions of a similar sort, capturing distinctive visions of the "grain of reality".[17] For those with verificationist leanings, the lack of consensus as regards a specific procedure for settling the truth values of (7) and (8) might suggest that they are somehow unintelligible, or perhaps ambiguous between different meanings corresponding to different verification procedures. But verificationism is false, since already in natural languages we find many sentences that are intelligible and not relevantly ambiguous despite the lack of such consensus. We, at least, are unperturbed, and regard "grain-theoretic" questions in the vein of (7) and (8) as theoretically central questions in metaphysics.[18]

[16] Some champions of "post-modal metaphysics" write as if the sorts of "coarse-grained" theories that might motivate (7) were obviously false, or could only be motivated by some sort of prejudice. But this seems quite wrong to us. There is an influential theory of *grounding* due to Fine (2012) that implies (8) (given LL), but this is *very* far from providing a knockdown objection to (7), especially since the obvious higher-order generalization of Fine's theory is inconsistent with H_0 (Krämer 2013; Fritz 2020). O In fact, our view is that Classicism—a theory that implies (7), which we will discuss in Chapter 8—has sufficient advantages in terms of simplicity and strength to be regarded as the *default* view, which sets the standard against which competing views should be measured.

[17] Theories that imply (7) include Classicism (see Chapter 8). The 'Only Logical Circles' principle defended in Dorr 2016*b* implies (8). And Ditter (2020) explores a view, motivated by considerations having to do with "essence", that motivates the rejection of both (7) and (8).

[18] Mathematical work on and in higher-order logic typically focuses on systems *much* stronger than H_0. The most influential such system, which Henkin (1950) dubbed "The Simple Theory of Types", is roughly equivalent—bracketing various syntactic differences, including those discussed in notes 6 and 7—to the result of adding six further axiom-schemes to H_0:

β-conversion (See note 13 above.)

η-conversion $P \leftrightarrow Q$, where Q comes from P by replacing an occurrence of a predicate of the form $(\lambda v_1 \dots v_n.Rv_1 \dots v_n)$ with R (where $v_1 \dots v_n$ are not free in R).

Fregean Axiom $(p \leftrightarrow q) \rightarrow (p = q)$

Functionality $\forall z_1 \dots \forall z_n(P = Q) \rightarrow (\lambda z_1 \dots z_n.P) = (\lambda z_1 \dots z_n.Q)$

Infinity $\exists R^{\langle e,e\rangle}(\forall x \exists y Rxy \wedge \exists y \forall x \neg Rxy \wedge \forall x \forall y \forall z \forall w((Rxy \wedge Rzw) \rightarrow (x = y \leftrightarrow z = w)))$

Choice $\exists S \forall x(\exists y Rxy \rightarrow \exists y(Rxy \wedge \forall z(Sxz \leftrightarrow z = y)))$

Church's original (1940) "Simple Theory of Types" was a bit weaker, essentially equivalent to the result of dropping the Fregean Axiom from this list but adding the "Plenitude" axiom discussed in Dorr 2016*b* (§6).

All six of the additional axioms amount to deep and non-obvious metaphysical theses. The only one we have a firm shared opinion about is the Fregean Axiom, which we reject. For example, despite the fact that snow is white if and only if grass is green, we deny that for snow to be white is for grass to be green; it follows by EG that $\exists p \exists q(p \leftrightarrow q \wedge p \neq q)$. Deciding which of the other five to accept will require a systematic exploration of the theoretical merits of larger theories which include them or rule them out. In the case of Choice, we take the integral role it plays in much of real analysis (and in physical theories that presuppose real analysis) to add up to a powerful abductive case for its truth. η-conversion also offers a lot of desirable strength at little cost in simplicity. Infinity seems well supported both on empirical and theoretical grounds. For some discussion of β-conversion, and a defence of a restricted version (where all of $v_1, \dots, v_n$ are required to have at least one free occurrence

The system H_0 specified by the axioms and rules in Figure 1.3 isn't strictly speaking a single set of formulae: rather, it is a function that yields a set of formulae for any given choice of nonlogical constants. In endorsing H_0, we claim that all the closed theorems with *no* nonlogical constants are true. That is already a strong commitment, since it implies that for every closed theorem containing arbitrary nonlogical constants, the result of replacing those constants with variables bound by initial universal quantifiers is true, so that the theorem itself is true under every interpretation of the constants as denoting entities of the relevant type.[19] We also think that there are many straightforward ways of introducing nonlogical constants on which they do come to denote entities of the appropriate type, so that the closed theorems of H_0 for the expanded language will still be true (on all uniform resolutions of any ambiguity or context-sensitivity that the new vocabulary may have introduced). In applying H_0, we will often need to employ various nonlogical constants related to the subject matter at hand, which we take to have this straightforward sort of meaning. But we do not maintain that all closed theorems will be true no matter how one expands the language. That would be absurd! One is free, if one wishes, to introduce a new symbol by stipulating whatever equivalences one pleases between new sentences involving it and old sentences not involving it. Such stipulations might even make certain H_0-theorems involving the new symbol synonymous with contradictions not involving it. This would be a perverse thing to do, since the usefulness of formal languages turns largely on their having simple and uniform logical behaviour. But if you want to throw those advantages away, nothing can stop you.

One common procedure for expanding a formal language is to find some natural-language expression and import it as a constant, of a type corresponding in the most obvious way to its English syntactic category, while taking the truth-preservingness of the natural translations from the expanded language into English for granted. Such "naïve formalization" can be fruitful. But bearing in mind the above disclaimer, it would be foolish to assume that all closed theorems of H_0 will be true in an expanded language obtained by naïve formalization. Thanks to Russell (1905), definite descriptions are the classic example of how naïve formalization can disrupt logical simplicity. In English, 'the King of France' seems (roughly) syntactically interchangeable with ordinary pronouns like 'he', 'her', and 'it', so a naïve formalization would render it using a new type-*e* constant, say 'KoF'. But when the obvious translations are held fixed, this will disrupt H_0: in particular, some of the instances of Eβ in the expanded language seem not to express truths. For in English, 'It is not the case that Russell met the King of France' seems to be

in P'), see Dorr 2016*b* (§5–6). In Chapter 8 we say a little about the pros and cons of Functionality, against the background of certain other controversial principles (weaker than the above six even when Functionality is added).

[19] This status corresponds to "logical truth" according to the seminal account in Tarski 1936*b*. Williamson (2013*b*: §3.3) calls it 'metaphysical universality'.

true, whereas 'The King of France is such that it is not the case that Russell met him' seems not to be. If these appearances are correct, then in a formal language with constants 'KoF', 'Russell', 'Met' modelled on their English analogues, the Eβ-instance

$$\neg \text{Met}(\text{Russell}, \text{KoF}) \leftrightarrow (\lambda x.\neg \text{Met}(\text{Russell}, x))(\text{KoF})$$

will not be true.

This sort of naïve formalization might nevertheless be useful for some purposes, e.g. if we want to use a formal language as a simple model for studying the workings of natural language. But the whole point of using a formal language in theorizing about non-linguistic matters is that the language should lack many of the complexities that make the attempt to formulate exceptionless logical laws directly in natural language so difficult. For such theoretical purposes, we want languages that are as simple as they can be, compatible with being capable of saying the things we want to say about the world. In English, not only names and definite descriptions, but also pronouns ('it'), demonstratives ('that dog'), indefinite descriptions ('a dog'), and quantificational expressions ('every dog', 'no dog') all fall into the same basic semantical category, that of determiner phrases (DPs). This syntactic uniformity makes for a lack of simplicity in the quantificational principles of English. For example, the simple schema 'If DP VP then something VP' suggested by 'If your dog bit me then something bit me' founders on the fact that 'No dog' is a DP. The characteristic clarity and freedom from structural ambiguity achieved by standard formal languages is largely due to their *not* following the lead of natural languages in this respect. And as Russell (1905) showed, this applies just as much to definite descriptions like 'the King of France' as to 'every dog' and 'no dog'. Since the standard apparatus of quantifiers and identity can plausibly be used to say everything about the world we might have wanted to express using definite descriptions in English, there is no important loss in expressive power, and plenty to gain in terms of simplicity, by choosing to use a formal language that does not include type-*e* expressions modelled after English definite descriptions.[20, 21]

If you have qualms about H_0, it is a good idea to carefully distinguish between views that merely cast doubt on theorems of the logic for some particular expansion of the language with nonlogical vocabulary, on the one hand, and views which

[20] This is expressed a little tendentiously. On some very fine-grained views about propositional identity, the formal "translations" of English sentences do not strictly speaking preserve the proposition expressed, precisely because of the ways in which they depart from the English syntactic structures. We are wary of such fine-grained views, for reasons we will explain in §7.4. But even if they were true, the moral would be that when one's interest is in theorizing about the non-linguistic world, the capacity to express *exactly* the propositions expressible in English is not an important desideratum for a formal language.

[21] For one potential complication to Russell's method of eliminating definite descriptions see Kripke 2005.

cast doubt on some of its theorems containing only our logical constants, on the other. Doubts of the first kind will only be relevant to applications of the logic essentially involving the nonlogical vocabulary in question. For example, if we can express some metaphysically interesting claim as a theorem of H_0 without using additions to the language modelled on English definite descriptions, the question how H_0 would need to be modified to accommodate such additions is neither here nor there.[22] For a more controversial example, suppose we naïvely formalize the English verb 'believe' as a constant of type $\langle\langle\rangle, e\rangle$. Suppose moreover that we endorse the widespread—though highly controversial—view that (9) is unequivocally true in English:

(9) Hesperus is Phosphorus, and someone believes that Hesperus is Hesperus but does not believe that Hesperus is Phosphorus.

Then, if we also naïvely formalize the names 'Hesperus' and 'Phosphorus' as type-e constants h and p, we should not accept H_0 for the expanded language. For these stipulations about the new constants require (9) to have the same truth value as (10), whose negation is a theorem of H_0:

$$\text{(10)} \qquad \mathrm{h} = \mathrm{p} \wedge \exists x(\mathrm{Believe}(\mathrm{h} = \mathrm{h}, x) \wedge \neg\, \mathrm{Believe}(\mathrm{h} = \mathrm{p}, x))$$

In endorsing H_0, we are not taking a stand on whether English sentences like (9) are unequivocally true. If they are, that doesn't mean it is *impossible* to translate English sentences involving 'believe' into a higher-order language in which all H_0-theorems are true. But it does mean that any such translation will have to depart from the naïve one—in which 'believe' is translated uniformly by a type-$\langle\langle\rangle, e\rangle$ constant and 'Hesperus' and 'Phosphorus' are translated uniformly by type-e constants—in some way that is bound to be controversial.

That is not to say that debates concerning (9) are completely irrelevant to the commitments that come with our endorsement of H_0. There is one possible way of developing the view that (9) is true that will require rejecting even certain theorems of H_0 that are entirely free of nonlogical constants. If (9) is true in English, so that (10) is true in the naïvely expanded language, *some* axiom or rule of H_0—most plausibly, either Eβ or LL—must fail there. One option here is to pin the blame entirely on LL. But if all the axioms and rules other than LL hold in the expanded

[22] For a contrasting attitude, see Robertson Ishii and Salmón 2020. Having pointed out that (the substitutional version of) Universal Instantiation fails in certain languages expanded with analogues of English definite descriptions, they take this to undermine the schema even within the pure higher-order language. If one were thinking of lambda-terms as stipulative equivalents of English expressions of the form 'the property of being an x such that ...', one might naturally expect them to display the same kinds of logical irregularities that characterize other English expressions beginning with 'the'. But there is far less basis for such suspicion when lambda-terms are thought of as analogous in the first instance to complex English noun and verb phrases, like 'female fox' and 'loves someone'.

language, LL must already fail in the language with only logical constants. For using just PC, EG, and Eβ, one can get from (10) to (11), a sentence involving no nonlogical constants:

(11) $\exists y \exists z \exists B(y = z \wedge \exists x(B(y = y, x) \wedge \neg B(y = z, x)))$ [23]

(11) involves only logical constants, and is inconsistent in H_0.

The radical view that endorses this argument as sound has recently been defended by Caie, Goodman, and Lederman (2020). The "Classical Opacity" package they defend accepts the fragment of H_0 not involving the = symbol, but rejects many (universal generalizations of) instances of LL. On their picture, it is just false that identical objects share all their properties, since properties like *being believed to be identical to Hesperus* can distinguish an object from some objects identical to it.

It is is worth noting that if we simply redefine = as $\lambda xy.\forall Z(Z(x) \leftrightarrow Z(y))$ ("Leibniz equivalence"), Ref and LL both become provable from the remaining axioms and rules of H_0. So, on this interpretation of =, Caie, Goodman, and Lederman accept the entirety of H_0. However, the interpretation of = they are interested in is one that ties it to the English word 'identical', and related expressions in English such as counting words. They think that 'There is only one planet orbiting second from the Sun' is true in English, even though a planet orbiting second from the Sun is such that there is a planet orbiting second from the Sun that is not Leibniz-equivalent to it.

This is not the time for a full-dress discussion of this radical package. The idea that having different properties suffices for non-identity seems so central to ordinary reasoning about identity that surviving without it is a disorienting and daunting prospect. We sympathize with Bacon and Russell's (2019) remark that views where Leibniz equivalence is more demanding than identity make that relation seem 'even more identity-like than identity'. And we are especially perturbed by the way in which (as Caie, Goodman, and Lederman admit) Classical Opacity requires invalidating ordinary English inference-patterns involving counting words, like 'Some *F* is *G*; some *F* is not *G*; so there are at least two *F*s', and

[23] The derivation depends on sequential applications of Eβ and EG:

1.	$\mathrm{h} = \mathrm{p} \wedge \exists x(\text{Believe}(\mathrm{h} = \mathrm{h}, x) \wedge \neg\, \text{Believe}(\mathrm{h} = \mathrm{p}, x))$	Premise
2.	$(\lambda B.\mathrm{h} = \mathrm{p} \wedge \exists x(B(\mathrm{h} = \mathrm{h}, x) \wedge \neg B(\mathrm{h} = \mathrm{p}, x)))(\text{Believe})$	1, Eβ
3.	$\exists B(\mathrm{h} = \mathrm{p} \wedge \exists x(B(\mathrm{h} = \mathrm{h}, x) \wedge \neg B(\mathrm{h} = \mathrm{p}, x)))$	2, EG
4.	$(\lambda z.\exists B(\mathrm{h} = z \wedge \exists x(B(\mathrm{h} = \mathrm{h}, x) \wedge \neg B(\mathrm{h} = z, x))))(\mathrm{p})$	3, Eβ
5.	$\exists z \exists B(\mathrm{h} = z \wedge \exists x(B(\mathrm{h} = \mathrm{h}, x) \wedge \neg B(\mathrm{h} = z, x)))$	4, EG
6.	$(\lambda y.\exists z \exists B(y = z \wedge \exists x(B(y = y, x) \wedge \neg B(y = z, x))))(\mathrm{h})$	5, Eβ
7.	$\exists y \exists z \exists B(y = z \wedge \exists x(B(y = y, x) \wedge \neg B(y = z, x)))$	6, EG

'Some *F* is *G*; there is at most one *F*; so every *F* is *G*.'[24] The package also seems to require rejecting the validity of compelling inference-patterns involving definite descriptions (which they do not discuss), like 'Some *F* is *G*; there is at most one *F*; so the *F* is *G*.'[25] As these inferences show, Classical Opacity rejects deeply rooted patterns of thought, ones that we are not willing to relinquish.[26]

There is another kind of departure from H_0, which (unlike Classical Opacity) rejects the quantifier axiom-schemas UI and EG, or at least $\mathrm{UI}_{\mathrm{sub}}$ and $\mathrm{EG}_{\mathrm{sub}}$. Logics without these schemas are known as 'free logics'. The most familiar motivations for them involve languages with constants corresponding to non-referring English words such as (arguably) 'Pegasus' and 'phlogisticated'.[27] They also are sometimes motivated as part of a package that also rejects LL-instances involving words like 'believe'.[28] We are open to the view that naïve formalization will in these cases lead to failures of UI, EG, and/or Eβ. But this has little bearing on the propriety of the instances of these schemas we will be appealing to, which will not involve such expressions. However, there is a more radical vision that is outright incompatible with our commitments, on which UI and EG fail even in the pure language with only logical constants. Consider the following theorem of H_0, where *X* is of the type of sentential operators ($\langle\langle\rangle\rangle$):

(12) $$\forall X(X(\exists pXp) \rightarrow \exists pXp)$$

[24] These patterns don't just sound good in the abstract: even when we plug in something like 'planet orbiting second from the sun' for *F* and 'believed to be visible in the evening' for *G*, the results still have an air of validity that seems hard to explain away.

[25] Two applications of this rule will get us from uncontroversial premises to 'The planet orbiting second from the sun is believed to be visible in the evening, and the planet orbiting second from the sun is not believed to be visible in the evening.' Whether or not this is semantically contradictory, it is evidently bad in some way, so it is natural in the context of Classical Opacity to lay the blame on the inference rule.

[26] The view also seems to require a rather mysterious and demanding theory of *reference*, since referring to a thing takes a lot more than referring to something identical to it. Suppose you have gone by the name 'Scott Soames' all your life, and you overhear your best friend saying 'Scott Soames is a philosopher.' You might assume that when they said 'Scott Soames' they were referring to you, and that they were saying that you were a philosopher. But on the Classical Opacity approach, such thoughts seem problematic. After all, if you have ever been even slightly less confident that you are Scott Soames than that Scott Soames is Scott Soames, you are not Leibniz-equivalent to Scott Soames, in which case you can't infer that your friend was referring to you from the fact that your friend was referring to Scott Soames. Indeed, given the ways in which "identity confusion" can arise even for repeated uses of the same name, there is a threat that proponents of Classical Opacity will be driven to conclude that it rarely if ever happens that distinct uses of the same expression are such that there is something to which both uses refer.

[27] See Lambert 1963 and Bacon 2013. Note that insofar as these words motivate any logical revisionism, they motivate a rejection of UI, not just $\mathrm{UI}_{\mathrm{sub}}$: proponents will not want to infer $(\lambda x.\exists y(y = x))(\mathrm{Pegasus})$ from $\forall x\exists y(y = x)$. It is not obvious that such examples pose any challenge to Eβ.

[28] The thought is that even though there is something identical to Jack the Ripper, one can, e.g., believe [fail to believe] that Jack the Ripper is a murderer without there being any *x* such that one believes [fails to believe] that *x* is a murderer. The most obvious way of blocking this inference would already reject the Eβ inference to $(\lambda x.\,\mathrm{Believe}(\mathrm{Murderer}(x), \mathrm{a})(\mathrm{JtR}))$; Bacon and Russell 2019 and Bacon 2019 explore an alternative which preserves Eβ and pins the blame on UI and EG proper.

If a property of propositions is had by the proposition that some proposition has it, then some proposition has it. Sounds obvious! But there is a tradition, going back to Poincaré (1905–6) and Russell (1906), of wanting to weaken the quantifier schemas to disallow what are called "impredicative" instances, in such a way as to block the derivation of theorems like (12).[29] There seem to be two basic motivations behind this resistance. One is the sense that there is something problematically 'circular' about a proposition specified by applying a quantifier to a "totality" (property of propositions) that includes (is instantiated by) that very proposition. Another is the desire to block the derivation of certain surprising theorems of H_0, including Prior's theorem (see Bacon, Hawthorne, and Uzquiano 2016) and the Russell-Myhill theorem (which we will discuss in Chapter 7). We don't see what's bad about the supposed "circularity", and regard the theorems as important discoveries. We also find the original, un-restricted quantifier rules so simple and powerful by comparison with the proposed replacements for them, and so much more compelling than the negations of the putatively 'paradoxical' conclusions of the aforementioned theorems, that one should really think twice about giving them up.[30]

Our basic logic H_0 is not beyond dispute; the task of defending it against the myriad alternatives that have been proposed would require a book in itself. Nevertheless, we think it is in excellent standing. And as we will see, the space of treatments of Tolerance Puzzles that it leaves open is plenty rich enough to keep us busy.

1.4 Modal Logic

Given the nature of our topic, questions about necessity and possibility will loom large. Questions about *metaphysical* necessity and possibility will be especially

[29] Those who propose restricting UI and EG to avoid 'impredicative' instances like (12) often throw a bone to those of us to whom they seem obviously true, by multiplying senses of the quantifier. True, (12) may have to be rejected, but we still get to assert $\forall X(X(\exists pXp) \rightarrow \exists^{+}pXp)$, where $\exists^{+}$ is a broader, "higher-level" propositional quantifier than our original $\exists$. The analogue of EG still fails for $\exists^{+}$, since (e.g.) we don't have $\forall X(X(\exists^{+}pXp) \rightarrow \exists^{+}pXp)$; but we do get $\forall X(X(\exists^{+}Xp) \rightarrow \exists^{++}pXp)$ as a consolation prize, where $\exists^{++}$ is a still higher-level propositional quantifier. And we can keep going forever. (This kind of "ramified" system is generally implemented, following Russell and Whitehead, by complicating the syntax so that every term has a "level" specified as part of its syntactic category, although we can get essentially the same effect by simply replacing each of our quantifiers with an infinite hierarchy of quantifier-constants.)

The proposed substitutes do not strike us as adequate replacements for the compelling thought behind (12).

[30] Robertson Ishii and Salmón (2020) argue that a certain comprehension principle which is a theorem of H_0, but not of its free weakenings, needs to be given up in order to preserve the principle that for every property of objects F, there is a *singleton set* $\{F\}$, itself an object, that has F and nothing else as a member. Given their presupposition that sets can have both individuals and properties as members, they seem to be interpreting the higher-order formalism in a very different way from us.

central, but we will also be concerned with a range of other non-epistemic statuses that are analogous, in ways to be explained, to metaphysical necessity and possibility. To raise these questions in our formal language, we will need to add some new operators corresponding to the relevant uses of words like 'necessarily' and 'possibly' in English, in addition to the logical constants employed in the previous section. Syntactically, these operators are constants $\Box$ and $\Diamond$ of type $\langle\langle\rangle\rangle$ (the same type as negation): they "denote properties of propositions". Our most basic logical commitment as regards these operators is just that H_0 remains correct when they are added to the language. This doesn't go without saying. As we said in §1.3, we are open to the idea that introducing constants like 'Believe' in the obvious way will disrupt H_0; and if this is right, it is plausible that *epistemic* interpretations of modal operators (corresponding to English 'must' and 'might') will be similarly ill-behaved. But following Kripke (1972), we see a deep contrast between epistemic and non-epistemic uses of modal words, and thus see no reason to expect whatever logical idiosyncracies are associated with the epistemic uses to carry over to the others.[31]

H_0 already has some interesting consequences for modal logic, such as the following qualified version of the necessity of identity:

NI$^-$ $\qquad\qquad \forall x \forall y(x = y \rightarrow \Box x = x \rightarrow \Box x = y)$

This is the result of applying GEN twice to an instance of $\mathrm{LL}_{\mathrm{sub}}$. x and y may be of any type. In the case where they are of type e, this claim was once seen as deeply paradoxical, but has been mainstream since its defence by Kripke (1971).[32]

But to get anywhere interesting in our reasoning about possibility and necessity, we will need to employ some laws specific to the modal operators, in addition to the general-purpose laws of higher-order logic. We will not assume very *much* modal logic. For example, we will not assume the Iteration principle $\forall p(\Diamond\Diamond p \rightarrow \Diamond p)$ or the closely related $\forall p(\Box p \rightarrow \Box\Box p)$, since some important strategies for escaping Tolerance Puzzles turn on rejecting these principles. But we will be taking for granted a certain more minimal modal logic for the range of operators we are

[31] Given H_0 for the modal language, we can argue compellingly against the Fregean Axiom that there are only two propositions (see note 18). For example, we can argue that since it is necessary that if snow is white then snow is white, but not necessary that if snow is white then grass is green, by LL it is false that for snow to be white is for grass to be green. The best bet for proponents of the Fregean Axiom who don't want to give up H_0 for the modality-free language is to claim, Kripke notwithstanding, that non-epistemic uses of modal words introduce the same kinds of failures of UI and EG that have often been posited for propositional attitude verbs. This package is not incoherent, but the implausibility of the higher-order identity claims that go with it puts a heavy argumentative burden on its proponents.

[32] Much of the controversy focused not on the quantified sentence NI$^-$, but on its instantiation with type-e constants, "naïvely formalizing" English proper names like 'Hesperus' and 'Phosphorus', in place of the variables x and y. Kripke *also* defends that schema. But if one wanted to disagree with him about this, it seems more promising to pin the logical ill-behavedness on the introduction of the names while keeping H_0 for the language with just $\Box$, $\Diamond$, and the logical constants.

concerned with, partly because it strikes us as clearly correct, and partly because we don't see that rejecting it opens up any distinctive and interesting strategies for resolving the paradoxes.

Let's start with two particularly compelling principles:

K $\forall p \forall q(\Box(p \to q) \to \Box p \to \Box q)$

Duality $\forall p(\Box \neg p \leftrightarrow \neg \Diamond p)$[33]

These principles, or their schematic analogues, are common to almost all systems studied in the literature on modal logic, and will be part of our basic modal logic. But by themselves they don't get us all that we need. For example, the following claims seem similarly compelling, but do not follow from K and Duality in H_0:

$$\Box \forall p(p \to p)$$
$$\forall p \Box(p \to p)$$
$$\forall X \forall y \Box\Box(Xy \vee \neg Xy)$$
$$\neg \exists x \Diamond(x \neq x)$$

The standard method of specifying a theory of modality strong enough to imply claims like these is not to add further axioms like K and Duality above, or even further axiom-schemes, but rather to specify the theory as the result of adding the "necessitation" rule $P/\Box P$ to the axioms and rules of some other theory, e.g. H_0 together with K and Duality.

We do in fact accept all the theorems of the theory H_K that results when we add K and Duality to H_0 and close under MP, GEN, and necessitation. But we have two reasons for being dissatisfied. The first is that reliance on theories specified with the aid of necessitation tends to (though admittedly doesn't have to) encourage the idea that there is some important status of "logical truth", stronger than mere truth, which one is claiming on behalf of the theorems of the theory in question. For one can very easily find oneself coming out with sentences like 'The result of applying $\Box$ to a logical truth is another logical truth' when one is trying to expound or justify the commitments involved in accepting a theory of this sort. But as we have already explained, we do not wish to suppose there is any such important status, and we would be especially wary if it is supposed to be independently plausible that the status is closed under necessitation.

[33] If we treated $\Diamond$ as mere shorthand for $\lambda p.\neg\Box\neg p$, Duality would become a theorem of H_0. Alternatively one could treat $\Box$ as mere shorthand for $\lambda p.\neg\Diamond\neg p$; in that case Duality is not a theorem of H_0, but does follow from $\forall p(\Diamond\neg\neg p \leftrightarrow \Diamond p)$. But we won't adopt either of these definitions, for the same reason that we didn't define $\forall$ in terms of $\exists$ or vice versa; and neither $\Box = (\lambda p.\neg\Diamond\neg p)$ nor $\Diamond = (\lambda p.\neg\Box\neg p)$ will be a theorem of our basic modal logic.

Our second reason for dissatisfaction is more clear-cut: $\mathsf{H_K}$ is oddly and artificially weak, failing to imply certain things that seem very much of a piece with things it does imply. For example, it implies each of the following plausible claims:

$$\forall x(x = x)$$
$$\Box\forall x(x = x)$$
$$\Box\Box\forall x(x = x)$$
$$\vdots$$

But (in the absence of Iteration) it fails to imply a certain generalization which subsumes these claims. We will write this generalization as

$$\Box^*\forall x(x = x)$$

and pronounce it as 'It is ancestrally necessary that everything is self-identical.' It stands to the above sequence of sentences as 'No number is its own successor' stands to the sequence '0 is not its own successor', '1 is not its own successor', '2 is not its own successor' .... This is straightforward to formulate using the resources of higher-order logic.[34]

From a very abstract perspective, this kind of weakness is an inevitable feature of any consistent, recursively axiomatizable theory meeting a certain minimal standard of complexity. For any such theory T, Gödel's second incompleteness theorem means that we will be able to state and prove theorems tantamount to '0 isn't the code-number of a proof of a contradiction in T', '1 isn't the code-number of a proof of a contradiction in T', etc.; but we won't be able to prove 'No number is the code-number of a proof of a contradiction in T.' Still, such failures of generality can be more or less blatant. And in the particular case of $\mathsf{H_K}$, it is sufficiently blatant to seem like a real flaw. If we can find a simple theory strong enough to imply the generalization as well as its instances, that seems like it would be a significant improvement, analogous to the move from some minimal number theory that can only prove that each specific number is not its own successor to one that proves the generalization.

[34] To define $\Box^*$ we use the same sort of technique that Frege used to define *being an ancestor of* in terms of *being a parent of*. First, we define the property of being a *finite iteration of* $\Box$: an operation (property of propositions) X is a finite iteration of $\Box$ iff it has every property of operations that's had by $(\lambda p.p)$ and preserved by composition with $\Box$:

$$\mathbf{F}_\Box := \lambda X.\forall Z((Z(\lambda p.p) \wedge \forall X(ZX \rightarrow Z(\lambda p.X\Box p))) \rightarrow ZX)$$

We then define being ancestrally necessary as having every finite iteration of necessity:

$$\Box^* := \lambda p.\forall X(\mathbf{F}_\Box X \rightarrow Xp)$$

In Appendix A, Lemma A1f, we confirm that $\Box^*P$ does, as one would expect, imply P, $\Box P$, $\Box\Box P$, etc.

$\Box^*$**N**	$\Box^*\forall v_1 \ldots \forall v_n \Box P$, where P is a theorem of H_0, and $v_1 \ldots v_n$ are any (zero or more) variables.
$\Box^*$**K**	$\Box^*\forall p \forall q(\Box(p \to q) \to (\Box p \to \Box q))$.
$\Box^*$**Duality**	$\Box^*\forall p(\Diamond\neg p \leftrightarrow \Box\neg p)$.

Fig. 1.4 The basic modal logic H_{KA}.

Happily, there is a reasonably simple theory that implies the missing generalizations and does not require any extra rules like necessitation. This theory, which we will call H_{KA}, is the smallest extension of H_0 which contains all of the additional axioms in Figure 1.4 and is closed under MP and GEN. In Appendix A we will show that H_{KA} includes H_{K}. For any theorem P of H_0, e.g. $\forall x(x = x)$, H_{KA} implies $\Box^*P$, since $\Box^*\Box P$ is an instance of $\Box^*$N (with $n = 0$), and $\forall q(q \wedge \Box^*\Box q) \to \Box^* q$ is a straightforward consequence (in H_0) of the definition of $\Box^*$.[35]

Whereas $\Box^*$K and $\Box^*$Duality are just two individual sentences, $\Box^*$N is a schema. Indeed, it's a schema of a somewhat unusual sort, since figuring out whether a particular formula is one of its instances requires figuring out whether a certain *other* formula is a theorem of H_0, a question which is undecidable (by Church's essential undecidability theorem). But this isn't essential: we could if we wished make do with a much less expansive, and decidable, schema whose instances are just those instances of $\Box^*$N where P results from prefixing universal quantifiers to any of the axioms for H_0 listed in Figure 1.3.[36]

H_{KA} is, intuitively, only a very little bit stronger than H_{K}, which as modal logics go is very weak. But neither H_{KA} nor H_{K} is so weak as to be uncontroversial for any of the modalities of interest. Restricting our attention to the language where the only constants are $\Box$ and $\Diamond$ and the logical constants of §1.2, and setting aside the more radical objections that require rejecting even H_0 for this language, the central controversies arise from the fact that necessitation can be applied to *open* as well as closed theorems of H_0 (in H_{KA}: such theorems can play the role of P in instances of $\Box^*$N). Consider for example $(\forall x Yx) \to Yx$, an open formula with two free variables, x of some type σ and Y of type $\langle\sigma\rangle$. Since it is a theorem of H_0—an instance of $\mathrm{UI}_{\mathrm{sub}}$—its necessitation $\Box(\forall x Yx \to Yx)$ is a theorem of H_{K}. By K and MP, so is $\Box\forall x Yx \to \Box Yx$. Applying GEN to this gives $\forall x(\Box\forall x Yx \to \Box Yx)$. Since x isn't free in the antecedent of the conditional, we can apply UD and MP to get

[35] See Lemma A3 in Appendix A.

[36] By Craig's (1953) "reaxiomatization lemma", any theory that can be axiomatized by a recursively enumerable axiom-scheme like $\Box^*$N can be axiomatized by a decidable set of axioms, so nothing deep turns on this choice.

$\Box\forall xYx \rightarrow \forall x\Box Yx$. A final application of GEN to Y then gives us a version of the "Converse Barcan Formula" (Barcan 1946):

CBF $\forall Y(\Box\forall xYx \rightarrow \forall x\Box Yx)$[37]

If it is necessary that everything has a property, then each thing necessarily has that property. This sounds intuitive enough! But its controversial character can be seen when we consider the result of applying $\mathrm{UI}_{\mathrm{sub}}$ to this formula, substituting $\lambda x.\exists y(y = x)$ for Y. Then the antecedent, $\Box\forall x(\lambda x.\exists y(y = x))x$, is the necessitation of a theorem of H_0, so the consequent, $\forall x\Box(\lambda x.\exists y(y = x))x$, is a theorem of H_{K} too. By K and the necessitation of the Eβ-instance $(\lambda x.\exists y(y = x))x \rightarrow \exists y(y = x)$, this can be simplified to the claim that everything is necessarily identical to something:

NE $\forall x\Box\exists y(y = x)$

In the special case where $\Box$ is metaphysical necessity and the variable x is of type e, the necessitation of NE (which is also a theorem, since H_{K} is closed under necessitation) is the doctrine of *necessitism*, recently defended at book length by Williamson (2013*b*).

Many philosophers accept *contingentism*, the negation of necessitism; and more generally, they reject NE for a wide variety of necessity operators $\Box$, including many that are much narrower than metaphysical necessity. Most of them think that there in fact are things that could rather easily have failed to be identical to anything, and are prepared to point to examples—e.g. they think that Saul Kripke would have failed to be identical to anything if Myer and Dorothy Kripke had never met. Necessitists can agree with contingentists that Kripke would never have been conceived in that case, and that as a result, he would never have had a shape or size or mass, never have borne any spatial relation to anything else, never have had any experiences or beliefs, and so on—in short, he would never have been *concrete*. But they think that despite his non-concreteness, he would still have been identical to himself, and therefore identical to *something* (when the quantifier is used unrestrictedly).[38]

[37] It would be more customary to use "CBF" as a label for the schema $\forall x\Box P \rightarrow \Box\forall xP$. Instances of our version result from applying GEN to instances of this schema with $P := Yx$. In the other direction, an application of $\mathrm{UI}_{\mathrm{sub}}$ to our CBF, substituting $\lambda x.P$ for Y, gives $\Box\forall x(\lambda x.P)x \rightarrow \forall x\Box(\lambda x.P)x$. We can get from that to the usual version of CBF by applying K to the necessitations of the H_0-theorems $\forall xP \rightarrow \forall x(\lambda x.P)x$ and $(\lambda x.P)x \rightarrow P$.

[38] We will not need to pick a once-and-for-all definition of 'concrete'. In the particular settings where we use the word, there will always be some more specific and more familiar property that could be substituted in: for example, any thesis about what needs to be the case for a certain table to be concrete could be replaced by a thesis about what needs to be the case for that table to have a shape, or for it to have a spatial location.

As it happens, we are necessitists. But necessitism is not crucial to the most central arguments that we will be discussing in this book. That would be unfortunate, given that many philosophers—perhaps most of those who have considered the question—regard necessitism as patently false. We believe that in most cases, the arguments we will be talking about can be modified in such a way that contingentists will also accept them as valid, without very substantially diminishing the plausibility of any premise. The reason we don't actually do this is that contingentists are divided among themselves as regards how H_K should be weakened to block the derivation of NNE, so that the weakened premises in the modified arguments would require several different kinds of qualification, making them unpleasantly complicated. For example, some contingentists accept, while others reject, the necessity of identity:

NI $\forall x \forall y (x = y \rightarrow \Box x = y)$

Given that NI^- is already a theorem of H_0, NI is equivalent in that setting to the simpler thesis that everything is necessarily identical to itself:

□Ref $\forall x \Box x = x$[39]

Contingentists who accept NI think that (e.g.) Kripke could have been identical to himself without being identical to anything; those who reject it are, by contrast, free to accept the principle that nothing could be identical to itself without being identical to anything ($\forall x \Box (x = x \rightarrow \exists y (y = x))$).[40]

Many contingentists—motivated by the desire to make sense of ordinary speeches like 'There are sixty-three possible knives that could be made from these seven handles and nine blades'—have embraced the project of introducing a manner of speaking such that when they adopt it, they will accept all the sentences that necessitists accept.[41] Contingentists who have, or aspire to have, some such systematic reinterpretation up their sleeves won't need to fuss around making

[39] NI and □Ref are valid on the very influential model theory of Kripke 1963, although Kripke later voices some doubts (Kripke 1971: 164). Stalnaker (1994) rejects both.

[40] Two related choice points for contingentists are whether to accept or reject □Eβ and □EG, the necessitations of the Eβ and EG schemas. No contingentist who accepts H_0 and K can accept both, since from the □Eβ-instance $\Box(\neg\exists y(y = x) \rightarrow (\lambda x.\neg\exists y(y = x))x)$ and the □EG-instance $\Box(\lambda x.\neg\exists y (y = x))x \rightarrow \exists x\neg\exists y(y = x)$ one can (with K and MP) derive $\Box\exists y(y = x)$.

A further option for contingentists like Stalnaker (1994), who reject both □Eβ and □Ref, is to accept the following schema, which Williamson (2013*b*: ch. 4) dubs the "Being Constraint":

$$\Box(R(a_1, \ldots, a_n) \rightarrow \exists x(x = a_i))$$

In our terminology: nothing can instantiate a property or relation without being identical to something. Note, however, that if this principle is applied in type ⟨⟩ and NE is rejected in that type, it will require rejecting even very basic principles like $\forall p \Box(p \rightarrow p)$.

[41] See Fine 1977*a*, 2005*c*, 2016*b*; Williamson 2013*b* (ch. 7); and Fritz and Goodman 2017.

piecemeal modifications to the premises and conclusions of the arguments we will be discussing to render them valid by their lights: they can simply apply their reinterpretation across the board. We do not expect such reinterpretation to substantially affect the defensibility of any of the claims we will be concerned with.

Despite being controversial in these ways, H_{KA} is still very weak. For example, it doesn't pass judgement on any of the following claims:

4	$\forall p(\Box p \rightarrow \Box\Box p)$	or equivalently, $\forall p(\Diamond\Diamond p \rightarrow \Diamond p)$
T	$\forall p(\Box p \rightarrow p)$	or equivalently, $\forall p(p \rightarrow \Diamond p)$
B	$\forall p(p \rightarrow \Box\Diamond p)$	or equivalently, $\forall p(\Diamond\Box p \rightarrow p)$
5	$\forall p(\Diamond p \rightarrow \Box\Diamond p)$	or equivalently, $\forall p(\Diamond\Box p \rightarrow \Box p)$
ND	$\forall x\forall y(x \neq y \rightarrow \Box x \neq y)$	or equivalently, $\forall x\forall y(\Diamond x = y \rightarrow x = y)$
BF	$\forall Y(\forall x\Box Yx \rightarrow \Box\forall xYx)$	or equivalently, $\forall Y(\Diamond\exists xYx \rightarrow \exists x\Diamond Yx)$

4 (a.k.a. Iteration), T, B, and 5 are individual sentences, while ND and BF are schemas with one instance per type.[42] Since we will be seriously engaging with solutions to Tolerance Puzzles that turn on the denial of Iteration for various modalities, it would be a bad idea to build it into our background logic. Assuming 5, B, ND, or BF would be similarly premature.[43] T, by contrast, is obviously true on many of the interpretations of $\Box$ we are interested in, including metaphysical necessity. However, it is implausible for certain other interpretations of $\Box$ for which H_{KA} remains plausible, and which turn out to feature in some interesting Tolerance Arguments. For example, Chapter 9 will be concerned with Tolerance Arguments where $\Box$ is interpreted as 'There is an objective chance of 1 at *t* that . . . '. Here T faces potential counterexamples: for example, a dart might land on a particular point on a dartboard even though at some earlier time the chance of its not landing on that very point was 1.

Since the non-theorems we have just listed will all have interesting roles to play later on, it is worth pointing out some implications among them in H_{KA} (and H_{K}):

[42] BF is a version of the Barcan Formula, which Barcan (1946) proposed adding as an additional axiom to a system of quantified modal logic (which already proved its converse for the the same reason that H_{K} does). It would be more customary to use the label 'BF' for the schema $\forall x\Box P \rightarrow \Box\forall xP$, which is equivalent to our quantified version in H_{K}.

[43] On the usual way of developing the Iteration-denying approach to Tolerance Puzzles (Salmon 1986*a*), failures of 5 go hand in hand with failures of Iteration. Salmón (Salmon 1989) additionally denies that the B axiom is a 'logical (or analytic)' truth, but expresses agnosticism about whether it is a truth. As a reminder, our purposes do not require saying anything about which truths are logical truths. In Chapter 4 we will discuss a puzzle which could be used to motivate the rejection of B.

- B follows from T and 5.
- 4 follows from B or $\Box$B and $\Box$5.[44]
- 5 follows from B and $\Box$4.[45]
- ND follows from B.[46]
- BF follows from B and $\Box$B.[47]

We will refer to the result of adding 4 and its necessitation $\Box 4$ to H_{KA} as H_{K4}; H_{S4} is the result of adding 4 and T and their necessitations, and H_{S5} is the result of adding 4, T, and B—or equivalently, T and 5—and their necessitations. (In Appendix A we show that each of these theories is closed under necessitation, and that it would make no difference if we started with H_{K} rather than H_{KA}, since the extra strength of H_{KA} over H_{K} is washed away by the addition of 4.)

It is worth noting one more claim which is not a theorem of H_{KA} or even of H_{S5}, but which is often assumed in work in higher-order modal logic (e.g. Gallin 1975 and Williamson 2013*b*), namely that necessary coextensiveness suffices for identity:

Intensionalism $\forall X \forall Y(\Box \forall z_1 \ldots \forall z_n(Xz_1 \ldots z_n \leftrightarrow Yz_1 \ldots z_n) \rightarrow X = Y)$.[48]

Intensionalism implies both Iteration and T in H_{K}. For suppose $\Box p$. Since $\Box(p \rightarrow (p \leftrightarrow (p \vee \neg p)))$, we have $\Box(p \leftrightarrow (p \vee \neg p))$ by K, and thus $p = (p \vee \neg p)$ by Intensionalism. But since $p \vee \neg p$ and $\Box\Box(p \vee \neg p)$, we can infer both p and $\Box\Box p$ by LL. Since we need to stay neutral about Iteration, it would be a bad idea to build Intensionalism into the background logic. Moreover, Intensionalism—and even the weaker "Ancestral Intensionalism" principle with a $\Box^*$ in place of the $\Box$—has many controversial consequences about higher-order identity that don't contain $\Box$ or $\Diamond$ at all, and which we also won't be assuming. For example, it implies that $\forall p(p = \neg\neg p)$ and that $\forall p \forall q(p = (p \vee (p \wedge q)))$.[49] By contrast, H_{KA} and even

[44] Suppose $\Box p$; then $\Box\Diamond\Box p$ by B or $\Box$B; but by $\Box$5 we have $\Box(\Diamond\Box p \rightarrow \Box p)$; hence $\Box\Box p$ by K.

[45] Suppose $\Diamond p$; then $\Box\Diamond\Diamond p$ by B; but $\Box(\Diamond\Diamond p \rightarrow \Diamond p)$ by $\Box$4, so $\Box\Diamond p$ by K.

[46] This result is due to Prior (1963: 206–7). $\Box(x = y \rightarrow \Box x = x \rightarrow \Box x = y)$ is the necessitation of an $\mathrm{LL}_{\mathrm{sub}}$-instance and $\Box\Box x = x$ is the double necessitation of a Ref-instance, so by K and Duality we have $\Diamond x = y \rightarrow \Diamond\Box x = y$, which by B implies $\Diamond x = y \rightarrow x = y$ and hence $x \neq y \rightarrow \Box x \neq y$.

[47] This result is also due to Prior (1957). Note first that every instance of the following schema is a theorem of H_{K}:

CBF* $\Diamond \forall x P \rightarrow \forall x \Diamond P$

(This follows from $\forall x \Box(\forall x P \rightarrow P)$, which entails $\forall x(\Diamond \forall x P \rightarrow \Diamond P)$ by K and Duality.) Now, suppose $\forall x \Box Yx$. Then by B, $\Box\Diamond\forall x \Box Yx$, so by the necessitation of CBF*, $\Box\forall x\Diamond\Box Yx$; so by $\Box(\Diamond\Box Yx \rightarrow Yx)$ (which follows from $\Box$B and a necessitated UI-instance), we have $\Box\forall x Yx$.

[48] Here, $n \geq 0$; $z_1, \ldots, z_n$ can be variables of any types $\sigma_1, \ldots, \sigma_n$, in which case X and Y are of type $\langle \sigma_1, \ldots, \sigma_n \rangle$.

[49] More generally, Intensionalism entails the validity of "Logical Equivalence", the schema whose instances are $(\lambda v_1 \ldots v_n.P) = (\lambda v_1 \ldots v_n.Q)$ whenever $P \leftrightarrow Q$ is a theorem of H_0. We call the view that we get by adding this schema to H_0 'Classicism'; it is explored in more detail in Chapter 8 and in Bacon and Dorrforthcoming.

the much stronger H_{S5} have no theorems not involving $\Box$ or $\Diamond$ that aren't already theorems of H_0, since if we substitute the truth operator $\lambda p.p$ everywhere for $\Box$ and $\Diamond$, all the axioms of H_{S5} become theorems of H_0.

1.5 Rigidity

There are some occasions when it would be expressively useful to be able to make sense of talk about *sets* (or collections, or pluralities) of properties, relations, or propositions. For example, consider the following sentence:

(13) Mary could have had all of John's favourite properties.

One might be tempted to render (13) in our higher language as (14) or (15):

(14) $\forall X(\text{Favourite}(X, \text{John}) \rightarrow \Diamond X(\text{Mary}))$

(15) $\Diamond\forall X(\text{Favourite}(X, \text{John}) \rightarrow X(\text{Mary}))$

(14) says that all of John's favourite properties are such that Mary could have had them; (15) says that it could have happened that Mary had every favourite property of John. Both of these seem to be possible readings of (13). But neither captures its most natural reading, which is one that has to be false if *being tall* and *not being tall* are two of John's favourite properties; by contrast, both (14) and (15) could still be true under this supposition. To express the intended reading unambiguously, one wants to reach for something along the lines of 'There is a set that contains all and only John's favourite properties and is such that possibly, Mary instantiates every member of it.' However, insofar as we are treating property-talk as a scheme for pronouncing higher-order quantification, we can't take such talk entirely at face value: if we take 'is a member of' to be an expression of type $\langle e, e\rangle$, it can't meaningfully take predicates as arguments. Moreover, it would defeat the purpose if we treated talk of 'sets of properties' as simply a way of pronouncing quantification into type $\langle\langle e\rangle\rangle$ (i.e. as interchangeable with talk of properties of properties). For the sentence

(16) $\exists Z(\forall Y(ZY \leftrightarrow \text{Favourite}(Y, \text{John})) \wedge \Diamond\forall Y(ZY \rightarrow Y(\text{Mary})))$

(where Z is a variable of type $\langle\langle e\rangle\rangle$) certainly does not capture the intended meaning of (13), since it is automatically true if there is any contingent truth: when p is contingently true, $\lambda Y.\,\text{Favourite}(Y, \text{John}) \wedge p$ will serve as a witness.

The communicative utility of set-talk in this setting depends crucially on certain assumptions about the *modal* behaviour of sets, notably the assumption that when something belongs to a set it does so necessarily. By contrast, it is false (on any

$\Box^*$**Rigid Comprehension** $\Box^*\forall Y\exists C(\mathsf{R}C \wedge \forall z_1 \ldots \forall z_n(Cz_1 \ldots z_n \leftrightarrow Yz_1 \ldots z_n))$

$\Box^*$**Persistent Rigidity** $\Box^*\forall C(\mathsf{R}C \rightarrow \Box\mathsf{R}C)$

$\Box^*$**Persistence** $\Box^*\forall C\forall z_1 \ldots \forall z_n(\mathsf{R}C \rightarrow Cz_1 \ldots z_n \rightarrow \Box Cz_1 \ldots z_n)$

$\Box^*$**Inextensibility** $\Box^*\forall C\forall Y(\mathsf{R}C \rightarrow (\forall z_1 \ldots \forall z_n(Cz_1 \ldots z_n \rightarrow \Box Yz_1 \ldots z_n) \rightarrow \Box\forall z_1 \ldots \forall z_n(Xz_1 \ldots z_n \rightarrow Yz_1 \ldots z_n)))$.

Fig. 1.5 Axioms of H_{KR}.

interesting notion of necessity) that whenever a property instantiates a property it does so necessarily. It is only for some special Xs that $\forall y(Xy \rightarrow \Box Xy)$ is true. If we want a way of expressing in our formal language the kinds of thoughts that we were trying to convey by invoking sets of properties, we will need some way of singling out these special properties whose pattern of instantiation is "modally constant" or "rigid". Since we will have recurrent need for claims of this kind, we will help ourselves to a notion of *rigidity*: for any types $\sigma_1, \ldots, \sigma_n$, we will have a predicate $\mathsf{R}_{\sigma_1,\ldots,\sigma_n}$ of type $\langle\langle\sigma_1, \ldots, \sigma_n\rangle\rangle$. And we will reason about rigidity according to a "logic of rigidity" given by the axiom-schemas in Figure 1.5. We assume, at least, that these axioms are true when $\Box$ expresses metaphysical necessity; we will occasionally be discussing other interpretations of $\Box$ for which Inextensibility is less plausible.[50] We will refer to the logic that results when the rigidity axioms are added to H_{KA} as H_{KR}. (Similarly $\mathsf{H}_{\mathsf{K4R}}$, $\mathsf{H}_{\mathsf{S4R}}$, and $\mathsf{H}_{\mathsf{S5R}}$ are the logics that result from adding the axioms to H_{K4}, H_{S4}, and H_{S5}.)

Rigid Comprehension says that every property (or relation) is coextensive with a rigid one. Persistent Rigidity says that every rigid property (or relation) is necessarily rigid. Persistence says that whatever has a rigid property (or relation) has it necessarily. And Inextensibility says that a rigid property necessitates any property (or relation) that each of its instances has necessarily. The axioms in Figure 1.5 assert the ancestral necessity of these four claims.[51]

[50] For example, Inextensibility is implausible when we interpret $\Box$ as expressing the property of having objective chance 1. Each point in the infinite collection of all points on a certain dartboard might have chance 1 of not being hit by a dart, without there being chance 1 that no member of that collection is hit by a dart. More controversially, Lewis (1973) endorses a theory of counterfactuals that allows for failures of what he calls "the limit assumption": for example, he suggests that the collection of all properties of the form *shorter than l* where *l* is longer than one inch could be such that each property in the collection would have been instantiated by a certain line if it had been longer than an inch, although it is false that if the line had been longer than an inch it would have instantiated every property in the collection. Thus Inextensibility fails when $\Box$ is interpreted as 'would have been true if the line had been longer than an inch'. These potential failures of Inextensibility help illuminate an important way in which the operators in question could fail to be "fully necessity-like" even if they conform to the basic modal logic.

[51] Inextensibility corresponds to an axiom of the plural modal logic developed by Linnebo (2013: 211). It says, in effect, that BF holds for quantifiers restricted to the instances of any rigid property. Thus insofar as BF can be thought of as saying there couldn't be new things, Inextensibility can be thought

With these principles in place, talk of rigid properties can take over the theoretical utility of set-talk, since they are analogues of the modal principles about sets that underwrite the expressive usefulness of set-talk in modal settings. For example, we can capture the intended reading of (13) using a variant of (16) restricted to rigid properties of properties, as follows:

(17) $\exists Z(\mathsf{R}Z \wedge \forall Y(ZY \leftrightarrow \text{Favourite}(Y, \text{John})) \wedge \Diamond\forall Y(ZY \rightarrow Y(\text{Mary})))$

Given the further premise Favourite(Tall,John) ∧ Favourite($\lambda x.\neg$Tall(x),John), we can now, as one would wish, derive the falsity of (17). For given Persistence, $\mathsf{R}Z \wedge Z(\text{Tall}) \wedge Z(\lambda x.\neg \text{Tall}(x))$ implies $\Box(Z(\text{Tall}) \wedge Z(\lambda x.\neg \text{Tall}(x)))$. In $\mathsf{H_K}$, the combination of this with $\Diamond\forall Y(ZY \rightarrow Y(\text{Mary}))$ implies the inconsistent $\Diamond(\text{Tall}(\text{Mary}) \wedge \neg \text{Tall}(\text{Mary}))$. Similarly, we can deduce the *truth* of (17) from the claim that tallness is John's only favourite property ($\forall Y(\text{Favourite}(Y, \text{John}) \rightarrow Y = \text{Tall})$ and the claim that Mary could be tall. Suppose $\mathsf{R}Z \wedge \forall Y(Z(Y) \rightarrow Y = \text{Tall})$; then by the necessity of identity, we have $\mathsf{R}Z \wedge \forall Y(Z(Y) \rightarrow \Box Y = \text{Tall})$, so by Inextensibility, $\Box\forall Y(Z(Y) \rightarrow Y = \text{Tall})$. In $\mathsf{H_K}$, the combination of this with $\Diamond$Tall(Mary) implies $\Diamond\forall Y(Z(Y) \rightarrow Y(\text{Mary}))$.

In our informal English renditions of higher-order sentences involving 'Rigid' we will either use the word 'collection' or the resources of plural quantification, rather than saying 'rigid property'. For example, we would pronounce (17) as 'there is a collection of properties that contains all and only John's favourite properties, and is such that possibly, Mary has every member of it', or just 'John's favourite properties are such that possibly, Mary has each of them.'[52]

Since it seems weird to think that there could be two distinct collections with the same members, this way of pronouncing quantification over rigid properties

of as saying that there couldn't be new instances of any rigid property. Proposition C4 in Appendix C shows that Inextensibility becomes redundant if we add the ancestral necessity of both ND and BF to the logic. Given the other rigidity axioms, ND implies a "Negative Persistence" claim, according to which anything that *isn't* an instance of a rigid property is necessarily not an instance, while BF rules out a rigid property gaining new instances by there coming to be new things. But we will not be taking ND or BF for granted.

[52] An alternative approach to regimenting this kind of talk is to complicate our basic higher-order syntax. Loosely following Fine (1977*b*), one might introduce a new rule for forming types: when $\sigma_1, \ldots, \sigma_n$ are types, $[\sigma_1, \ldots, \sigma_n]$ is a type, thought of as the type of sets/collections/pluralities of n-tuples of things of type $\sigma_1, \ldots, \sigma_n$. There is just one new rule for forming terms: When F, $a_1, \ldots, a_n$ are of types $[\sigma_1, \ldots, \sigma_n], \sigma_1, \ldots, \sigma_n$, respectively, $Fa_1 \ldots a_n$ is of type $\langle\rangle$; we also have variables in all the new types, and quantifiers subscripted with new types. One can translate from our language into the language with the extra types by translating $\mathsf{R}_{\langle\sigma_1,\ldots,\sigma_n\rangle}$ as $\lambda C.\exists X^{[\sigma_1,\ldots,\sigma_n]}(C = \lambda z_1 \ldots z_n.Xz_1 \ldots z_n)$. There are various possible strategies for translating from the language with the extra types back into ours by mapping variables of new types to variables of the corresponding old type (replacing square brackets with angle brackets), and mapping quantifiers subscripted with new types to quantifiers with appropriate restrictions. For example, $\forall_{[e]}$ will become $\lambda X^{\langle\langle e\rangle\rangle}.\forall Y^{\langle e\rangle}(XY \rightarrow \mathsf{R}Y)$; it is less obvious what, e.g., $\forall_{\langle[e]\rangle}$ should map to.

suggests adding a further axiom according to which no two rigid properties or relations are coextensive:

Rigid Extensionality

$$\forall C \forall D((\mathsf{R}C \wedge \mathsf{R}D \wedge \forall z_1 \ldots z_n(Cz_1 \ldots z_n \leftrightarrow Dz_1 \ldots z_n)) \rightarrow C = D)$$

However, since we will never need to appeal to Rigid Extensionality, we will not officially add it to the logic: this means that the uniqueness presupposition of the definite article in expressions like 'the collection of all *x*'s qualitative properties' is not part of what we are officially committed to in using such locutions.[53]

Against the background of H_{KA}, the rigidity axioms have several noteworthy consequences that do not themselves involve the special rigidity predicates. One important consequence, proved in Appendix B, is that the ancestral necessity operator $\Box^*$ obeys Iteration, i.e. $\forall p(\Box^* p \rightarrow \Box^* \Box^* p)$.[54] This will play an major role in Chapter 7, which will consider Tolerance Puzzles stated in terms of ancestral modality, for which the result rules out an Iteration-denying solution.

Another consequence is that for any property of propositions, there is *a greatest lower bound* of the instances of that property: a proposition that ancestrally necessitates all of them, and is ancestrally necessitated by any other proposition that ancestrally necessitates all of them.

Boolean Completeness $\Box^* \forall X \exists p(\forall q(Xq \rightarrow \Box^*(p \rightarrow q)) \wedge \forall r(\forall q(Xq \rightarrow \Box^*(r \rightarrow q)) \rightarrow \Box^*(r \rightarrow q)))$[55]

Intuitively, for a given X, we can choose a rigid C coextensive with X and take p to be $\forall q(Cq \rightarrow q)$. In the case where X has only two instances, their conjunction is also a witness to Boolean Completeness; more generally, a witness to Boolean Completeness behaves, as far as $\Box$ is concerned, like the conjunction of all the X propositions, even if there are infinitely many of them. Parallel reasoning gives us a *least upper bound* of the X propositions, which behaves like their disjunction. In what follows we will occasionally find it useful to talk about the conjunctions and disjunctions of perhaps-infinite collections of propositions: these can be identified respectively with the propositions that every member of the collection is true,

[53] In a very fine-grained theory, one might not be able to maintain both Rigid Extensionality and natural principles like 'The conjunction of any two rigid properties is rigid'—for if X, Y, and Z are rigid, the latter principle will entail that $\lambda x.Xx \wedge (Yx \wedge Zx)$ and $\lambda x.(Xx \wedge Yx) \wedge Zx$ are both rigid, which given Rigid Extensionality would entail that they are identical; but a very fine-grained theory that attempts to stay relatively close to Structure (without coming so close as to fall into inconsistency) might entail that this is not the case when X, Y, and Z are all distinct.

[54] This follows from the stronger result (Proposition B12) that $\Box^*$ conforms to the logic $\mathsf{H}_{\mathsf{S4R}}$, i.e. that replacing $\Box$ uniformly with $\Box^*$ in any theorem of $\mathsf{H}_{\mathsf{S4R}}$ yields a theorem of H_{KR}.

[55] See Proposition C6 in Appendix C.

and that some member of the collection is true.[56] A third noteworthy R-free consequence of the rigidity axioms is that there is (with ancestral necessity) a truth that ancestrally necessitates all truths:

Actuality $\quad \Box^*\exists p(p \wedge \forall q(q \rightarrow \Box^*(p \rightarrow q)))$

The existential quantification will be witnessed by the proposition that $\forall p(Cp \rightarrow p)$, where C is a rigid property coextensive with truth.[57] Note that Actuality does not follow obviously from Boolean Completeness: the latter tells us that there is a proposition that necessitates all truths and is necessitated by every other proposition that does so, but leaves it open that the only such propositions might be necessarily false.[58]

The fact that the rigidity axioms have these attractive R-free consequences naturally raises the question whether we can find a simple axiomatization of the R-free theorems of H_{KR}. The obvious way to do this is to find some definition of each rigidity predicate $\mathsf{R}_{\sigma_1 \ldots \sigma_n}$ in terms of modal operators and logical constants, such that substituting these definitions for R maps theorems to other theorems. We can then easily axiomatize the R-free fragment by substituting the definitions into the axioms in Figure 1.5. Appendix C shows that the following definition does the job:

$$\mathsf{R} := \lambda C.\ \text{Persistent } C \wedge \text{Inextensible } C$$

where

$$\begin{aligned} \text{Persistent} &:= \lambda C.\Box^*\forall z_1 \ldots z_n(Cz_1 \ldots z_n \rightarrow \Box Cz_1 \ldots z_n) \\ \text{Inextensible} &:= \lambda C.\Box^*\forall Y(\forall z_1 \ldots z_n(Cz_1 \ldots z_n \rightarrow \Box Yz_1 \ldots z_n) \\ &\qquad \rightarrow \Box\forall z_1 \ldots z_n(Cz_1 \ldots z_n \rightarrow Yz_1 \ldots z_n)) \end{aligned}$$

That is: a property or relation is persistent iff it is ancestrally necessary that it is necessary to all its instances, and inextensible iff it is ancestrally necessary that it necessitates every property necessary to all its instances. Since this definition turns $\Box^*$Persistence and $\Box^*$Inextensibility into trivial theorems of H_0, it lets us axiomatize the R-free fragment just by the results of substituting the definitions into $\Box^*$Rigid Comprehension and $\Box^*$Persistent Rigidity.[59] Nevertheless, we prefer not to commit ourselves officially to this or any other definition of rigidity in

[56] Boolean Completeness generalizes from propositions to any type $\langle \sigma_1, \ldots, \sigma_n \rangle$:

$$\exists R(\forall S(FS \rightarrow R \leq S) \wedge \forall T(\forall S(FS \rightarrow T \leq S)) \rightarrow T \leq S)$$

where $R \leq S := \Box^*\forall x_1 \ldots \forall x_n(Rx_1 \ldots x_n \rightarrow Sx_1 \ldots x_n)$.

[57] See Proposition C7 in Appendix C.

[58] We believe that Boolean Completeness and Actuality are independent in H_{KA}: see Bacon and Dorr forthcoming for further discussion.

[59] We can simplify this axiomatization even further by getting rid of Persistent Rigidity and adding the single axiom $\Box^*\forall p(\Box^* p \rightarrow \Box^*\Box^* p)$: see Proposition C2.

modal terms, since we feel we have a pretty clear grip on the notion of rigidity independent of its relation to any particular modality.

1.6 Worlds

A central tool in studying the properties of various logical systems is *model theory*, or as it is sometimes (misleadingly) known, *model-theoretic semantics*.[60] In this discipline, one typically proceeds by giving a set-theoretic definition of something called a "model", along with a definition of a "true in" relation that holds between sentences in the relevant language and these models. A sentence is *valid relative to* a class of models iff it is true in every member of the class. The central results are *soundness* theorems, to the effect that every theorem of some axiomatically specified system is valid in a certain class of models, and *completeness* theorems, to the effect that every sentence valid in the class of models is a theorem of the theory. Soundness theorems are especially important for the purposes of showing that certain sentences *aren't* theorems of a certain system: armed with a soundness theorem, we can show that something isn't a theorem by finding a model in which it isn't true. Completeness theorems are often useful as tools of discovery, since in many cases, finding a set-theoretic proof that a sentence is true in all models is easier than coming up with a derivation of that sentence in the relevant axiomatic system; they also motivate the search for countermodels in cases where one suspects but has not yet established non-theoremhood. Completeness and soundness results can also be used to probe other technical properties like decidability.

The best-known classes of models for higher-order languages are sound for our systems—and thus potentially useful for establishing claims of non-theoremhood—but not complete for those systems, or indeed for any recursively axiomatizable systems.[61] But there are broader definitions of "model" for which one can prove soundness and completeness theorems for our basic logic H_0, and these definitions can be extended in such a way as to do the same for modal logics like H_{KA} and H_{KR}.[62] Our exposition in this chapter has deliberately been entirely free of model theory, since the validities that will matter to the metaphysical debates we will be focusing on in the rest of the book generally have fairly short and easy-to-find derivations, and there will not be a major role for claims which are in

[60] Misleading because it is quite opaque what the whole project has to do with explaining, or theorizing about, the meanings of words in any language (formal or natural). We have no sympathy with any suggestion to the effect that fully understanding what our formal language means would require knowing any model theory. For more on the unfortunate connotations of Tarski's decision to use the word 'semantics' for model theory, see Burgess 2008.

[61] We are thinking of, e.g., the "standard models" of Gallin 1975.

[62] See Muskens 2007. Some philosophers might wish to enter into a debate as regards which, if any, of these different classes of models is such that the sentences valid in them are exactly the logical truths; but as we have explained, we have not found the ideology of logical truth helpful.

fact consistent, but for which there would be a serious reason to worry about their consistency in the absence of a model-theoretic proof.

In general, while model theory is vitally important for theorists who want to *study* rigorously specifiable languages and theories stated in them, it is only an occasionally useful tool for theorists who want to *use* such languages in order to communicate and argue for views about the world. For example, the development of model theory for first-order logic did not have much of an effect on the way that philosophers expressed the kinds of claims and arguments that are readily formalized in first-order languages. But in the case of modal logic, the sociological situation is a bit different: a certain model theory, namely the "possible worlds" model theory introduced by Kripke (1959, 1963), has had a pervasive effect on the ways in which philosophers engage in object-level modal reasoning. The distinctive feature of Kripke-style models is the central role played in them by so-called "Kripke frames": pairs $\langle W, R \rangle$, where W is a non-empty set and R a set of ordered pairs of members of that set; the models come with a definition of a relation 'S is true at w', where S is a sentence of the relevant language and $w \in W$. A sentence of the form $\Box P$ is true at w iff P is true at every v such that Rwv; $\Diamond P$ is true at w iff P is true at some v such that Rwv. While W is mnemonically called the "set of possible worlds" of the model, the definition of model typically places no constraints on what objects might go into W. The heuristic virtues of model theory as a tool for discovery are particularly evident in the case of this kind of model, which in turn may explain its pervasive influence in metaphysics. We seem to be a lot better at reasoning about points and binary relations than reasoning directly using modal operators, so it is often helpful to exploit a completeness theorem together with some relatively easy fact about Kripke frames to convince oneself that there is a proof of something in some given system of modal logic, even if one hasn't found such a proof yet, or to exploit the soundness theorem together with some particular easy-to-construct Kripke frame to convince oneself that there isn't a proof of something.

The success on these terms of the possible-worlds model theory has encouraged many metaphysicians to incorporate some of the apparatus of the model theory into the object language in which they theorize about modal questions. The picture is that there really *are* things called 'worlds', one of which is 'the actual world', and for each interpretation of 'necessarily' and 'possibly', a corresponding relation of 'accessibility' among worlds, such that something is necessarily the case iff it is the case at every world accessible from the actual world, and possibly the case iff it is the case at some world accessible from the actual world. This is a convenient thing to think, since it lets us *directly* deploy our facility with reasoning in extensional fragments of our language to probe questions expressed using modal operators, rather than having to make a detour through set-theoretic models based on arbitrary Kripke frames. But it also raises some challenging foundational questions:

What *are* worlds? What is it for one of them to be accessible from another? What is it for one to be actual? What is it for a proposition to be true at one of them?

Our higher-order system naturally suggests some ways of making sense of the language of worlds which settle questions like these, and provide interpretations of the claim that there are possible worlds on which it already follows from the logical principles we have already canvassed rather than being some risky new posit. The approaches involve reconstructing quantification over worlds as a certain restricted kind of higher-order quantification: as we might put it, worlds aren't objects.[63] One natural option—following Plantinga (1974), Adams (1974), and Salmón (Salmon 1989*b*: 6–7)—takes worlds to be certain *collections* (i.e. rigid properties) of propositions; a proposition is true at a world just in case it is a member of (i.e. instantiates) it. In conformity with the idea that worlds are "maximally specific", we require them to answer every yes/no question: for every proposition and every world, either the proposition is true at the world or its negation is true at the world (or both):

$$\begin{aligned} \text{World}(W) &:= \mathsf{R}\,W \wedge \forall p(Wp \vee W\neg p) \\ \text{At}(p, W) &:= Wp \end{aligned}$$

An *actual* (or *actualized*) world is one such that everything true at it is true. A *possible* world (in some given sense of 'possible') is one that could (in that sense) be actual:

$$\begin{aligned} \text{Actual}(W) &:= \forall p(\text{At}(p, W) \rightarrow p) \\ \text{Possible}(W) &:= \Diamond\, \text{Actual}(W) \end{aligned}$$

Finally, a world V *accesses* a world W ('V sees W'; 'W is accessible from V') just in case the W is possible at V:

$$\text{Accesses}(V, W) := \text{World}(V) \wedge \text{World}(W) \wedge \text{At}(V, \text{Possible}(W))$$

Persistence guarantees that any proposition that is true at some possible world is possibly true. And Rigid Comprehension guarantees that there *is* an actual world, at which all and only the true propositions are true, and from which all and only the possible worlds are accessible.[64]

[63] Talk about *sets* of worlds will likewise be reconstructed as shorthand for talk about collections, i.e. rigid properties, of entities of the relevant type.

[64] Remember again that the actual world is not an object or collection of objects: '[t]he world is the totality of facts, not of things' (Wittgenstein 1921: §1). If we understand Wittgenstein's claim as one about what we are calling 'the actual world', it will be natural for someone who accepts that claim and also is willing to identify 'facts' with true propositions (or obtaining states of affairs) to think that there are other "totalities" in the same general category as the one Wittgenstein calls 'the world', namely the other collections we are calling 'worlds'. It might seem more natural to use 'world' in a way

Note that on this definition there are plenty of worlds—for example, the collection of all propositions—that are impossible not just in some particular sense of interest, but in every sense of 'possible' that conforms to our basic modal logic. This consequence seems unproblematic, indeed welcome; given that we want to leave open that there are impossible but possibly possible worlds, it would be a mistake to build possibility in some sense into the notion of a world.[65]

This way of glossing world-talk fits pretty well with the role that worlds are typically taken to play in philosophical theorizing. But the fit may still be quite imperfect, unless the modality we are interested in obeys certain principles that are not guaranteed our basic modal logic. Some surprises are apt to arise if there are failures of the Necessity of Distinctness or of the Barcan Formula (ND and BF: see §1.4). If BF fails for type $\langle\rangle$, that means the collection of all propositions is such that there could be propositions not belonging to it—*new* propositions. If so, then very plausibly there could be a new proposition whose negation was also new.[66] If there could be such a proposition then each world is only contingently a world, since if there was a new proposition with a new negation, it would no longer be such as to contain every proposition or its negation. Indeed, there may be possible worlds that are not worlds at themselves, since they necessitate that there are new propositions with new negations (cf. Stalnaker 2012: ch. 2). Meanwhile, if ND fails for type $\langle\rangle$, so that certain distinct propositions could be identical, then some non-worlds may only contingently be non-worlds, since all the propositions such that

that corresponds to our 'actual world', thinking of the role of the word 'possible' in 'possible world' as analogous to its role in 'possible winner' rather than to the role of 'white' in 'white cat' (see Fine 2005*c*). However, this doesn't fit the way of talking that is inspired by the study of Kripke models.

[65] Another approach, inspired by Prior and Fine (1977), Fine (1977*a*), and Chisholm (1981: 129), identifies worlds with *propositions* rather than collections of propositions. On the natural way of developing this approach, one must use some modal operator already in defining 'true at' and 'world'. A proposition is a world iff it necessitates (in the chosen sense) every proposition or its negation; p is true at w iff it is (in the chosen sense) necessary that if w than p:

$$\mathrm{World}(p) \coloneqq \forall q(\Box(p \to q) \vee \Box(p \to \neg q))$$
$$\mathrm{At}(p, q) \coloneqq \Box(p \to q)$$

The definitions of 'Actual', 'Possible', and 'Accesses' can remain as before, modulo the change in type. This approach works less smoothly than the worlds-as-collections approach unless the chosen notion of necessity is something (e.g. ancestral metaphysical necessity) for which we are willing to presuppose the truth of Iteration. Without Iteration, the fact that a certain proposition is true at a certain world might be a contingent one. We can allow for impossible but possibly possible worlds; but when a world is impossible, all propositions—even contradictions—are true at it. Thus, the fact that p is true at a world two accessibility-steps away from the actual world will not guarantee the possible possibility of p. By contrast, this intuitive implication does hold on the worlds-as-collections approach. (Suppose A is actual, W is accessible from A, W' is accessible from W, and p is true at W'. Then $\Box W'p$ by Persistence, and $\Box \mathsf{R}\, W$ by Persistent Rigidity, so $\Box\Box W'p$ by $\Box$Persistence, so $\Box\Box(\forall q(W'q \to q) \to p)$ by the closure of $\Box\Box$ under logical consequence. But $\Box W \Diamond \forall q(W'q \to q)$ by Persistence and the fact that W' is accessible from W, and $\Diamond \forall r(Wr \to r)$ by the fact that W is accessible from A. Hence $\Diamond\Diamond\forall q(W'q \to q)$, and thus $\Diamond\Diamond p$.)

[66] If it is necessary that every proposition is its negation's negation—$\Box\forall p(p = \neg\neg p)$—then any new proposition would have to have a new negation. For let C be the collection of all propositions: then since $\forall p(Cp \to C\neg p)$, $\forall p(Cp \to \Box C\neg p)$ by Persistence, so $\Box\forall p(Cp \to C\neg p)$ by Inextensibility, so $\Box\forall p(C\neg p \to C\neg\neg p)$ by necessitated UI, and thus $\Box\forall p(C\neg p \to Cp)$ by $\Box\forall p(p = \neg\neg p)$.

neither they nor their negations belong to them are such that, possibly, each is identical to a proposition such that either it or its negation does belong to them. ND-failures may also lead to cases where a proposition contingently fails to be true at some world, because it is contingently distinct from some proposition that is true at that world.

These surprises might not be so hard to learn to live with, but there is another potential source of surprise that goes more to the heart of the standard practice of talking about worlds. This practice takes for granted that a proposition is possible iff true at some possible world, and necessary iff true at every possible world:

Leibnizian Possibility $\Diamond p \leftrightarrow \exists w(\text{World}(w) \wedge \text{Possible}(w) \wedge \text{At}(w, p))$

Leibnizian Necessity $\Box p \leftrightarrow \forall w(\text{World}(w) \rightarrow \text{Possible}(w) \rightarrow \text{At}(w, p))$

We have seen how our definitions guarantee (in H_{KR}) the right-to-left direction of Leibnizian Possibility; similar reasoning gives us the left-to-right direction of Leibnizian Necessity. But the other directions are not automatic. Given our definitions, they are equivalent to one another, and to the principle that every consistent collection of propositions can be extended to a consistent collection that includes every proposition or its negation:

MaxCon $(\mathsf{R}X \wedge \Diamond\forall p(Xp \rightarrow p)) \rightarrow$
$\exists X'(\mathsf{R}X' \wedge \Diamond\forall p(X'p \rightarrow p) \wedge \forall p(Xp \rightarrow X'p) \wedge \forall q(X'q \vee X'\neg q))$

MaxCon is equivalent in H_{KR} to the perhaps simpler principle that every possible proposition is necessitated by an *atomic* proposition, i.e. one which is possible and necessitates every proposition or its negation:

Atomicity $\Diamond p \rightarrow \exists p'(\Diamond p' \wedge \Box(p' \rightarrow p) \wedge \forall q(\Box(p' \rightarrow q) \vee \Box(p' \rightarrow \neg q)))$[67]

But Atomicity is not a theorem of H_{KR}.[68] In that logic we can prove Actuality, and hence the existence of a *true* atomic proposition. But we cannot rule out a scenario where a certain contingently false proposition is such that for every possible proposition that necessitates it, there is a stronger proposition that is also possible.[69]

[67] See Proposition C10 in Appendix C.

[68] This follows from model-theoretic results in Bacon and Dorr forthcoming.

[69] This is not just an artefact of our proposed definitions of 'World', 'Possible', and 'At'. Even if we take these as primitive, we can prove MaxCon from Leibnizian Possibility or Leibnizian Necessity along with the claims (a) that if a proposition isn't true at a world its negation is, and (b) that if a world is possible, the collection of all propositions true at it is such that it is possible for all its members to be true. Thus for example the situation is not changed by switching to the alternative worlds-as-propositions approach discussed in note 65: indeed, that approach makes Leibnizian Possibility immediately equivalent to Atomicity (since a possible proposition not necessitated by any atomic proposition is ipso facto not true at any world).

Interestingly, MaxCon/Atomicity, ND, and BF are all provable in $\mathsf{H}_{\mathsf{S5R}}$ (the result of adding the rigidity axioms to H_{S5}). §1.4 already discussed Prior's results that ND follows from the B axiom and BF from B and its necessitation. Meanwhile, given the necessitation of Actuality—which as noted in §1.5 is a theorem of H_{KR}—5 and BF together imply Atomicity.[70] But 5 and B are at least as controversial as Iteration, and it would be a mistake to take either of them for granted.[71]

One could add any or all of ND, BF, and Atomicity to H_{KR} without adding B or 5. A package containing this trio, or at least Atomicity, would doubtless be convenient to philosophers with an ingrained habit of dragging in possible worlds even when what they are trying to get across could be expressed perfectly well using ordinary modal operators. But that in itself is not much of an argument for the principles. And it is far from obvious which if any of the trio can be motivated independently of B and 5. In Chapter 4 we will discuss some possible motivations for ND, the most promising of which, due to Williamson (1996), involves certain principles involving an 'actually' operator (intended to regiment certain natural-language sentences involving the phenomenon of "modal anaphora"). A somewhat similar argument can be given for BF, as we will discuss in Chapter 8, and there may be a way of doing something similar for Atomicity/MaxCon.[72] Without such a motivation, one will need to be more cautious about world-talk than is customary.[73]

In unpublished work, Peter Fritz gives a simple axiomatization of a theory of possible worlds in a propositionally quantified language; shows that this theory follows in S5 from Atomicity and (the propositional special case of) Intensionalism; and argues that it "doesn't leave anything out", by showing that the only questions in its language that it does not settle are questions about how many possible worlds there are.

[70] This result is due to Gallin (1975: 85). By the necessitation of Actuality, we have $\Diamond p \rightarrow \Diamond(p \wedge \exists p'(p' \wedge \forall q(q \rightarrow \Box(p \rightarrow q))))$, hence $\Diamond p \rightarrow \Diamond\exists p'(p' \wedge \Box(p' \rightarrow p) \wedge \forall q(\Box(p' \rightarrow q) \vee \Box(p' \rightarrow \neg q)))$. By BF this yields $\Diamond p \rightarrow \exists p'\Diamond(p' \wedge \Box(p' \rightarrow p) \wedge \forall q(\Box(p' \rightarrow q) \vee \Box(p' \rightarrow \neg q)))$, and hence $\Diamond p \rightarrow \exists p'(\Diamond p' \wedge \Diamond\Box(p' \rightarrow p) \wedge \Diamond\forall q(\Box(p' \rightarrow q) \vee \Box(p' \rightarrow \neg q)))$. By CBF* (see note 47), this implies $\Diamond p \rightarrow \exists p'(\Diamond p' \wedge \Diamond\Box(p' \rightarrow p) \wedge \forall q\Diamond(\Box(p' \rightarrow q) \vee \Box(p' \rightarrow \neg q)))$. So by 5 (in the dual form $\Diamond\Box P \rightarrow \Box P$), we get $\Diamond p \rightarrow \exists p'(\Diamond p \wedge \Box(p' \rightarrow p) \wedge \forall q(\Box(p' \rightarrow q) \vee \Box(p' \rightarrow \neg q)))$, i.e. Atomicity.

There is also a sort of converse of this fact: in H_{S5}, Atomicity implies the necessitation of Actuality. For in H_{S5} we can show that any atomic proposition p' is necessarily atomic: if $\Diamond p'$ then $\Box\Diamond p'$, and if $\forall q(\Box(p' \rightarrow q) \vee \Box(p' \rightarrow \neg q))$ then $\forall q(\Box\Box(p' \rightarrow q) \vee \Box\Box(p' \rightarrow \neg q))$ by Iteration, so $\forall q\Box(\Box(p' \rightarrow q) \vee \Box(p' \rightarrow \neg q))$ by K, so $\Box\forall q(\Box(p' \rightarrow q) \vee \Box(p' \rightarrow \neg q))$ by BF. But any true atomic proposition witnesses Actuality; so every atom necessitates Actuality, which by Atomicity implies that Actuality is necessary. When we use the definition of 'rigid' as 'persistent and inextensible' mooted in §1.5, Actuality is equivalent to Rigid Comprehension (see Proposition C9 in Appendix C); thus in the setting of H_{S5} Atomicity (or MaxCon) can serve as a replacement for the rigidity axioms.

[71] Indeed, as noted in §1.4 (note 44) B and $\Box 5$ together imply Iteration.

[72] One question we have not so far been able to settle is whether ND and BF imply Atomicity (in H_{KR}). Of course, if Atomicity follows from ND and BF, then insofar as considerations related to modal anaphora support ND and BF, they support Atomicity. If not, then one might get somewhere be appealing to more expressive varieties of modal anaphora such as those explored by Vlach (1973) and Cresswell (1990: ch. 3).

[73] A very different possible motivation for MaxCon comes from the logic of counterfactuals, where the controversial Conditional Excluded Middle principle (Stalnaker 1980) can be combined with the principle that true counterfactuals with contradictory consequents always have impossible antecedents to argue that for each proposition p, the collection of all propositions that *would* be true if p were true is a maximal consistent collection containing p.

So, unless one is willing to presuppose principles that go well beyond our base logic, it is hard to find an interpretation of world-talk that renders unproblematic the things that philosophers tend to take for granted when they connect truth at worlds and accessibility to possibility and necessity. This is not a problem for possible-worlds model theory understood as a tool for determining which modal claims are derivable from which, but it does mean we need to be cautious about relying on world-talk to articulate claims about the structure of modal reality.[74] Fortunately, the kinds of things that metaphysicians typically want to say in the contexts where they are presupposing Leibnizian Possibility and Leibnizian Necessity very often are straightforwardly equivalent (given those presuppositions) to purely modal claims which also make perfect sense even without those presuppositions. We think it is good discipline, whenever possible, to write down claims "in the language of boxes and diamonds" without bringing in worlds at all. But since, as far as we can see, Atomicity is orthogonal to the central puzzles we will be discussing, readers who find the habit of world-talk hard to shake off won't be missing out on anything too important as far as these puzzles are concerned.

This concludes the presentation of the basic framework of higher-order modal logic that will be in the background of our discussion of Tolerance Puzzles, and that we more generally propose as a good framework for metaphysical theorizing. Some aspects, such as the theory of rigidity and the invocation of ancestral necessity, are novel; others are common to pretty much all of the systems of higher-order logic that have been explored in the literature. By contrast with most presentations in that literature, we have said almost nothing about "metatheory", i.e. the mathematical study of the formal properties of our system. That is because our aim here is to theorize *in* a higher-order language, a project which doesn't require doing very much theorizing *about* such languages. Most analytic philosophers are comfortable using first-order languages to regiment claims and arguments even though far fewer of them are well equipped to engage in metatheory concerning such languages. We see no reason why the situation with higher-order languages should be any different. And we hope that our exposition in this chapter will set to rest any nagging concern that one somehow needs to learn a lot of metatheory in order to do higher-order metaphysics.

[74] "Standard" possible-worlds models for higher-order logic (as in Gallin 1975) do validate Atomicity, but that doesn't interfere with the use of the soundness theorem for such models to establish claims of non-derivability. As it happens, though, it is also possible to introduce a class of "possible worlds models" for which our logic is sound and complete, and so for example where Atomicity won't be true in all models. In these models, each world is associated with a "domain of propositions", and any member of the domain of propositions of any world can be evaluated as true or false at any other world, but there is no guarantee that the domain of one world contains any proposition that is true only at exactly one other world. For more see Bacon and Dorr, forthcoming.

2
Tolerance Puzzles

Armed with our logical toolkit, we can now get down to the main business of the book. We have already, in the Introduction, given a range of examples of the puzzles of interest and gone some way towards isolating the central moving parts. In the present chapter, we will articulate the puzzles in a more systematic way, by presenting two general argument-schemas under which they can be subsumed. The first schema is tailored to puzzles that concern some specific object (such as the Great Pyramid); the second schema covers puzzles that generalize over all objects of a certain kind (e.g. tables). Once the schemas are on the table, we will be able to see how puzzles having to do with change over time, like the celebrated Ship of Theseus puzzle, can be set up in such a way as to instantiate them. We will also give special attention to a subfamily of puzzles which play a prominent role in the literature (and in some later chapters), where the kind of tolerance at issue concerns originating matter.

2.1 A Schema for Tolerance Arguments

Every Tolerance Argument is concerned with some family of properties—e.g. height-properties, properties that specify originating matter, properties that specify the rules of a game, and so on. (In most of our examples, the properties in the family are pairwise incompatible, although this is not necessary to the logic of the argument.) Instances of our first schema have a particular object in view, and the central 'Tolerance' premise says that this object could have had any property that is "close" to one that it in fact has. Here, 'close' is just a placeholder for some binary relation among the properties in the family. But in the instantiations of interest, the operative closeness relation can reasonably be understood as a form of similarity. For example, the sorts of definitions of '*G* is close to *F*' which will generate interesting arguments include the following:

> *F* and *G* are heights whose ratio is between 9/10 and 10/9 (for short: '*G* is a close height to *F*')
>
> *F* and *G* are fully specific ways that paint-colours could be distributed over a 4×3-foot rectangular object, and *F* differs from *G* only as regards the colour-pattern in one square inch ('*G* is a close paint-distribution to *F*')

The Bounds of Possibility: Puzzles of Modal Variation. Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri, Oxford University Press. © Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri 2021.
DOI: 10.1093/oso/9780192846655.003.0003

> For some finite collections of atoms C and D of atoms such that for every atomic number, C and D contain equally many atoms with that atomic number, and at least 90 per cent of the atoms in C are also in D, F is *being originally composed by* C and G is *being originally composed by* D ('G is origin-close to F')

In symbolic settings, we will write '$G \sim F$' for 'G is close to F'. The choice of the symbol $\sim$ and the word 'close' suggest symmetry, and indeed many of the closeness relations we will be working with are symmetric; however, the assumption that closeness is symmetric is not required for an argument to instantiate the schema.

One kind of definition of 'close', which corresponds to a common way of setting up Tolerance Puzzles in the literature, involves simply writing down a list of properties $H_1, \ldots, H_n$ and defining 'G is close to F' to mean that G immediately follows F on this list—i.e. '$F = H_1$ and $G = H_2$ or … or $F = H_{n-1}$ and $G = H_n$'. Working with closeness relations of this sort simplifies the argument in some ways, by dispensing with the need for higher-order quantification (which can be replaced by suitable conjunctions). However, arguments where the closeness relation is given by a finite list also have disadvantages. Even when a Tolerance or Hypertolerance claim only mentions a specific list of properties, the appeal of the former and the repugnance of the latter are often motivated by some more general thought. It is methodologically advantageous to use a schema that allows those general thoughts to be fully on display.

With an interpretation of 'close' in hand, we define several auxiliary predicates. First, we have a property of *tolerance*: being such that you could instantiate any property close to one you in fact instantiate.

> x is *tolerant* := for any F and G such that x instantiates F and G is close to F, it is possible for x to instantiate G.

For example, when 'close' stands for height-closeness, x is tolerant iff either x doesn't have a height (in which case it is *vacuously* tolerant, since no property it instantiates is close to any property), or every height whose ratio with x's height is between 9/10 and 10/9 is a height x could have had. Likewise, with paint-distribution closeness, x is tolerant iff either x is not a 4×3-foot object covered in paint, or x is such an object, and could have still been such an object while having any of the paint-distributions differing only in one square inch from its actual paint-distribution.

In generating an instance of our schema, we will have to not only fix on a definition of 'close', but also choose a *singular term* such as a name or demonstrative:

'the Great Pyramid'; 'the *Mona Lisa*'; 'the Vienna Circle'; 'this table'.[1] Using '*a*' to stand in for this singular term, and defining 'tolerant' as above, we can state the Tolerance premise very simply as follows:

Tolerance *a* is tolerant.

We may sometimes lapse into informally characterizing Tolerance as the claim that *a* could have been somewhat different in the relevant respect. But this is potentially misleading: for example, for *a* to be tolerant with respect to height, *every* height close to the height it has must be possible for it. If—bizarrely—the Great Pyramid could have been 9/10 of its actual height but couldn't have been 19/20 of its actual height, then while it would be true that it could have been somewhat different in height, the Tolerance premise based on height-closeness would be false.

The mere truth of Tolerance doesn't get one very far. To get to any potentially puzzling conclusions, we need its necessary truth. Our second premise bridges this gap:

Non-contingency If *a* is tolerant, it is necessary that *a* is tolerant.

Understand this as a material conditional, to avoid distracting issues related to the meaning of 'if' in English.

The third premise makes no special mention of the chosen closeness relation or object:

Iteration Whatever is possibly possible is possible.

As we will see, one of the most widely discussed reactions to Tolerance Arguments in the literature treats them as a *reductio* of Iteration.

Finally we will need one more premise, rarely highlighted in the literature:

Persistent Closeness When properties are close, they are necessarily close.

For example, in the case of height-closeness, this amounts to the claim that when two heights stand in a ratio between 9/10 and 10/9, it is necessary that they do so. This may seem so obvious as to go without saying, but it needs to be included for the argument to be valid, and as we will see, it is more tendentious in certain other instantiations. (Note that if one is working with a 'list'-style definition of 'close', Persistent Closeness becomes a triviality: by the necessity of identity, it can't

[1] An ordinary definite description like 'the table before us' will also be fine, so long as we settle on a policy of resolving all of the *de re*/*de dicto* ambiguities that may arise from the interaction of such expressions with modal operators in the *de re* way (i.e. the way that makes existential generalization legitimate).

be contingently true that $F = H_1$ and $G = H_2$, or …, or $F = H_{n-1}$ and $G = H_n$. The need for this premise thus comes from our desire for the generality that quantification over properties affords.)

The conclusion of each argument fitting the schema will be a claim which we can write as follows:

Hypertolerance *a* is hypertolerant.

To say what this means, we will need to consider the *ancestral*, or transitive closure, of our chosen closeness relation. *G* is *ancestrally close* to *F*—in symbols, $G \sim^* F$—iff there is a finite sequence of two or more properties beginning with *G* and ending with *F*, in which each element after the first is close to its predecessor.[2] Hypertolerance is the analogue of tolerance but with ancestral closeness in place of closeness:

> *x* is *hypertolerant* ≔ for any *F* and *G* such that *x* instantiates *F* and *G* is ancestrally close to *F*, it is possible for *x* to instantiate *G*.

For many of the families of properties of interest, every property in the family is ancestrally close to every other, so that an object that instantiates a member of the family will be hypertolerant just in case every member of the family is possible for it. For example, every height can be connected to every other height by a finite sequence of heights in which neighbouring elements have a ratio between 9/10 and 10/9, so when the operative closeness relation is height-closeness, *x* is hypertolerant iff *x* either doesn't have a height, or *x* could have had any height.

The final argument is given in Figure 2.1. The argument is valid, given our minimal modal logic (§1.4). We can show this in four steps.

Step One: Apply Modus Ponens to Tolerance and Non-contingency to derive:

[2] Quantification over finite sequences of properties is a bit cumbersome to regiment in the higher-order language introduced in Chapter 1. Fortunately, we can avoid the need to mention sequences by instead quantifying over properties of properties, using a technique invented by Frege (1879). Say that a property of properties is *hereditary* iff it is had by every property that is close to a property that has it. And say that a property of properties *encompasses* a property iff it is had by every property that is close to that property. Then we can define '*G* is ancestrally close to *F*' as '*G* has every hereditary property of properties that encompasses *F*'. Formally (letting *Z* be a variable of type $\langle\langle e\rangle\rangle$):

$$G \sim^* F :=_f \forall Z(\forall U((\exists V(ZV \wedge U \sim V) \rightarrow ZU) \wedge (U \sim F \rightarrow ZU)) \rightarrow ZG)$$

Compare the Fregean definition of '*y* is an ancestor of *x*' as '*y* has every property of people that all *x*'s parents have, and that is had by the parents of anything that has it'.

It obviously follows from this definition that every property that is close to a given property is ancestrally close to it. Slightly less obviously, it follows that ancestral closeness is transitive: if *H* is ancestrally close to *G* and *G* is ancestrally close to *F*, then *H* is ancestrally close to *F*. It suffices to show that if *G* is ancestrally close to *F*, *being ancestrally close to F* is hereditary and encompasses *G*, since in that case, it is instantiated by every property that has all such properties (i.e. is ancestrally close to *G*). Since, by hypothesis, *G* is ancestrally close to *F*, this follows from the claim that *being ancestrally close to F* is hereditary. But this is obviously true, since when *Y* is close to *X* and *X* has every hereditary property that encompasses *F*, *Y* does too.

Tolerance *a* is tolerant.

$$\forall F \forall G(G \sim F \rightarrow Fa \rightarrow \Diamond Ga)$$

Non-contingency If *a* is tolerant, then it is necessary that *a* is tolerant.

$$\forall F \forall G(G \sim F \rightarrow Fa \rightarrow \Diamond Ga) \rightarrow \Box \forall F \forall G(G \sim F \rightarrow Fa \rightarrow \Diamond Ga)$$

Iteration Whatever is possibly possible is possible.

$$\forall p(\Diamond \Diamond p \rightarrow \Diamond p)$$

Persistent Closeness When properties are close, they are necessarily close.

$$\forall F \forall G(G \sim F \rightarrow \Box G \sim F)$$

Hypertolerance *a* is hypertolerant.

$$\forall F \forall G(G \sim^* F \rightarrow Fa \rightarrow \Diamond Ga)$$

Fig. 2.1 Tolerance Argument (*de re* version).

Necessitated Tolerance It is necessary that *a* is tolerant.

$$\Box \forall F \forall G(G \sim F \rightarrow Fa \rightarrow \Diamond Ga)$$

Step Two: Using Necessitated Tolerance and Persistent Closeness, derive:

Stepwise Necessitated Tolerance Whenever *G* is close to *F*, it is necessary that if *a* is *F*, then it is possible for *a* to be *G*.

$$\forall F \forall G(G \sim F \rightarrow \Box(Fa \rightarrow \Diamond Ga))$$

Proof: Suppose *G* is close to *F*. By Persistent Closeness, it is necessary that *G* is close to *F*. But by Necessitated Tolerance, it is necessary that if *a* is *F* and *G* is close to *F*, it is possible for *a* to be *G*. Hence it is necessary that if *a* is *F*, it is possible for *a* to be *G*.[3]

Step Three: From Stepwise Necessitated Tolerance and Iteration, derive:

[3] Formally, we apply CBF to Necessitated Tolerance to derive $\forall F \forall G \Box(G \sim F \rightarrow Fa \rightarrow \Diamond Ga)$, and combine this with Persistent Closeness (using the K axiom) to derive Stepwise Necessitated Tolerance. This part of the argument is invalid in some of the weakenings of our background logic that have been advocated by contingentists, who reject CBF. In particular, "positive" higher-order contingentists who accept the necessitation of Eβ (see note 40 in Chapter 1) will reject the inference from 'Necessarily, *G* is close to *F* and for any *X* and *Y*, if *a* is *Y* and *X* is close to *Y*, it is possible that *a* is *X*' to 'Necessarily, if *a* is *F* and *G* is close to *F*, it is possible that *a* is *G*', for the same reason that they would reject the inference from 'It is possible that Socrates is nonexistent and for any *x*, if *x* is nonexistent, *x* is a round square' to 'It is possible that Socrates is a round square.' We can fix up the argument so that even these contingentists will accept it as valid by strengthening Persistent Closeness as follows:

Persistent Closeness+ Whenever *G* is close to *F*, it is necessary that: *G* is close to *F*, and *F* is identical to some property, and *G* is identical to some property.

If 'necessary' is interpreted as metaphysical necessity, this will be rejected as false by some contingentists for many of the closeness relations we have been considering. For example, some higher-order contingentists will think that when *C* is a collection of atoms, it is not metaphysically necessary that

Possibility Transfer For any F and G such that G is close to F, if it is possible for a to be F, then it is possible for a to be G.

$$\forall F \forall G(G \sim F \rightarrow \Diamond Fa \rightarrow \Diamond Ga)$$

Proof: Stepwise Necessitated Tolerance implies that whenever G is close to F, if it is possible for a to be F, then it is *possible* for it to be possible for a to be G. But by Iteration, if it is possible for it to be possible for a to be G, it is possible for a to be G.

Step Four: All that remains is to derive Hypertolerance from the combination of Tolerance and Possibility Transfer. The validity of this step should be clear: if every property close to one that a does or could instantiate is one a could instantiate, then every property that can be reached in finitely many steps from one that a instantiates is one a could instantiate.[4]

Of course, many Tolerance Arguments are not at all puzzling. Some have obviously true conclusions; some have an obviously false premise; many have both. But there are many instantiations of the pattern which present us with a non-obvious choice as to whether to accept the conclusion or give up a premise. We will speak in these cases of "Tolerance Puzzles", which we can think of as consisting of five individually plausible but jointly incompatible claims: Tolerance, Non-contingency, Iteration, Persistent Closeness, and Non-hypertolerance (the negation of Hypertolerance). And of course one can run a valid argument from any four of the five to the negation of the fifth; we will call these 'Tolerance Arguments' too.

Unlike many presentations of Tolerance Puzzles in the literature, the above argument-scheme does not use the expression 'possible world' or any related

there is a property identical to *being originally composed by C*. But for this closeness relation, even the original Persistent Closeness is highly controversial, as we will discuss in §2.4. We will suggest controlling for this by running the argument using some narrower modality for which Persistent Closeness is uncontroversially true; a parallel move can be employed to get around contingentist concerns about Persistent Closeness+.

[4] To spell this out, let Z be *being a property a could have* ($\lambda X.\Diamond Xa$), and let F be some property that a has. By Tolerance, every property close to F is Z. And by Possibility Transfer, every property close to a Z property is itself Z. But by the definition of "ancestrally close", every property ancestrally close to F has every property that is had by every property close to F and by every property close to a property that instantiates it. Hence, every property ancestrally close to F has Z. By Eβ, that means every property ancestrally close to F is one that a could instantiate. Since F was arbitrary, we can conclude that all the properties ancestrally close to properties of F are properties a could instantiate.

Step Four is the only part of the argument that turns on higher-order quantification other than wide-scope universal quantification, which could, if we wished, be understood schematically. For those wary of higher-order resources, it is worth noting that even without them, we can already derive $\forall F \forall G(G \sim^n F \rightarrow Fa \rightarrow \Diamond Ga)$ for each number $n > 0$, where $G \sim^n F$ means 'G is n closeness steps away from F'. (That is: $G \sim^0 F :=: G = F$; $G \sim^{n+1} F :=: \exists H(H \sim^n F \wedge G \sim H)$.) This can be shown by a straightforward induction, using Tolerance for the $n = 1$ case and Possibility Transfer for the induction step. For the closeness relations of interest, we can generally find some large n for which this weaker variant of Hypertolerance is just as implausible as the official one with ancestral closeness.

vocabulary. One can of course substitute all occurrences of 'necessarily' and 'possibly' with 'At every possible world' and 'At some possible world', respectively, but nothing interesting is added by doing this. By contrast, starting with Chisholm (1967), Tolerance Puzzles have sometimes been presented as if they were primarily challenges to the standard principles that would license such substitutions (e.g. Leibnizian Possibility and Leibnizian Necessity from §1.6).[5] But this is a mistake: Tolerance Puzzles are a lot more interesting than they would be if they merely challenged some piece of philosophical lore about possible worlds!

Some Tolerance Puzzles are somewhat reminiscent of the notorious Sorites Paradox, since they involve a closeness relation that holds only between pairs of *very similar* properties.[6] But one can easily construct Tolerance Puzzles using much less demanding closeness relations. To illustrate the flexibility of the schema, imagine that a certain bucket, Flimsy, is made of a piece of plastic, *A*, while another bucket is made simultaneously by the same machine out of a different piece of plastic, *B* (see Figure 2.2). It seems possible that Flimsy could still have been made while *B* was discarded rather than being made into a bucket. But if things were like *that*, it seems like it would have been possible for Flimsy to have been made as a strong bucket, twice as thick as it actually is, by adding *B* to the melting pot along with *A*. And if things were like *that*, it seems like it would have been possible for Flimsy to have been made as a flimsy bucket by discarding *A* rather than adding it to the melting pot, so that Flimsy was made of just *B*. And finally, if things were like *that*, it seems like it would have been possible for Flimsy to still be a flimsy bucket made of *B* while *A* was made into a different bucket, simultaneously by the same machine. Given Iteration, it follows that it is in fact possible for Flimsy to be a bucket made of *B* while a different bucket is made out of *A*, simultaneously by the same machine. And that seems puzzling. What—we might wonder—could make the difference between this allegedly possible state of affairs and the actual configuration, given that both involve *A* and *B* being simultaneously made into buckets by the same machine?[7]

[5] For example, Mackie and Jago (2017) characterize the combination of Necessitated Tolerance and Hypertolerance as constituting 'a difficulty for the paraphrase of *de re* modal claims … in terms of transworld identity', a problem which would not arise at all for counterpart theorists like Lewis (1968), who deny transworld identity. Regimenting the argument in modal terms makes it clear that if counterpart theory helped at all, it would have to be by undermining the plausibility of one of the five putatively incompatible claims, or denying their incompatibility. In Chapter 10 we will argue that counterpart theory does not help in either of these ways.

[6] We will discuss this connection further in §3.1.

[7] One might be tempted to explain the difference in terms of people's different decisions as regards to how to use the *name* 'Flimsy'; but this seems misguided, given that it also seems clearly possible for the bucket actually made of *B* to be called 'Flimsy'. In any case we don't have to imagine the story in such a way that people are in the practice of naming buckets at all—the name is one we are using in telling the story, not part of the story.

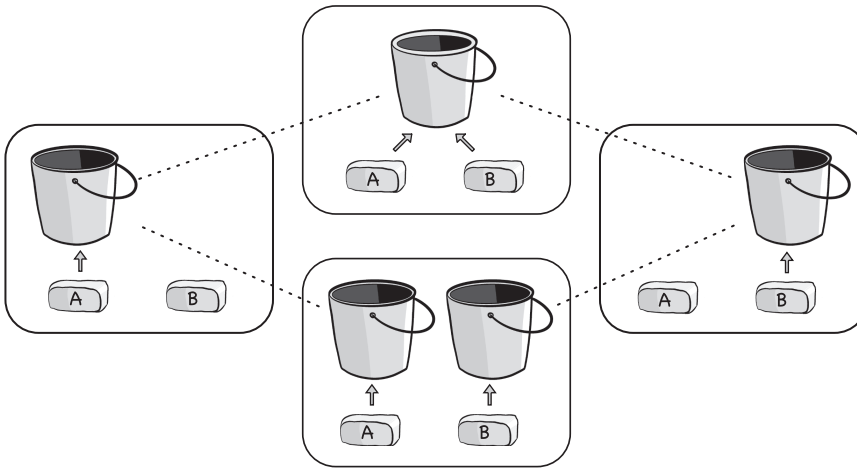

Fig. 2.2 The buckets.

To regiment this as a Tolerance Puzzle, define 'G is close to F' to mean that G follows F in the following list of five properties:[8]

H_1: *being a bucket made of A while a different bucket is made of B*

H_2: *being a bucket made of A while B is discarded*

H_3: *being a fat bucket made of both A and B*

H_4: *being a bucket made of B while A is discarded*

H_5: *being a bucket made of B while a different bucket is made of A*

Let's call the Tolerance Argument based on this definition of 'close' and with the name 'Flimsy' in the role of *a* the 'Bucket Argument'; call its premises 'Bucket Tolerance', 'Bucket Non-contingency', etc.

Bucket Tolerance is equivalent to the conjunction of four material conditionals: if Flimsy has H_1 it could have H_2; if it has H_2 it could have H_3; if it has H_3 it could have H_4; if it has H_4 it could have H_5. Even if we didn't know that any of these conditionals had a false antecedent, each of them would seem rather plausible. There are a couple of different thoughts in the background. In the case of the H_1-to-H_2 and H_4-to-H_5 conditionals, it is tempting to reach for some general principle to the effect that whether or not a given object is formed from certain matter cannot depend on what is happening at the same time to some completely separate piece of matter. It is not obvious how to say this precisely. But whatever we might think

[8] Of course, the appeal to lists is inessential here; what we mean is just the disjunction of '$G = H_2$ and $F = H_1$, or $G = H_3$ and $F = H_2$, or $G = H_4$ and $F = H_3$, or $G = H_5$ and $F = H_4$'.

of the generalization (in Chapter 11, we will see some reasons to be suspicious), as with many other Tolerance claims, the conditionals in question are well supported by various specific ordinary judgements about close-to-home possibilities. These judgements include counterfactuals ('If that piece of plastic had also been made into a bucket, this would have been one of two similar buckets'), claims of easy possibility, safety, and danger ('There was a danger that this would be the only bucket we made today'), chance claims, and so on. In the case of the H_2-to-H_3 and H_3-to-H_4 conditionals, it is even less clear that there is a plausible general principle from which they could be derived, but again there is no need to appeal to such a principle since we can instead appeal to a range of specific, ordinary judgements: 'This bucket would have been much more durable if its walls had been twice as thick'; 'There was a chance that this bucket would be much less durable, since we nearly threw away one of the two lumps of plastic that went into it'.

In fact, it is stipulated as part of the setup of the case that Flimsy has H_1, so motivating Bucket Tolerance only requires arguing that it could have had H_2. But motivating Bucket Non-contingency requires thinking about all four conditionals, and noticing that the reasons for thinking them true which we just surveyed are independent of the falsity of their antecedents. (It isn't obvious how to turn this reflection into anything like an argument for Bucket Non-contingency—in Chapter 3 we will make an effort to pin down what the case for Non-contingency might be in arguments like this one.) As with any closeness relation defined from a finite list, Bucket Persistent Closeness follows immediately from the necessity of identity for properties. Bucket Iteration is just the standard Iteration principle for metaphysical possibility. Bucket Hypertolerance, finally, is equivalent to the claim that if Flimsy has any of the five properties on the list, it could have had all of the subsequent ones. And that seems prima facie implausible, for reasons we have noted. We are given that Flimsy has H_1; but there is something mystifying about the suggestion that it could have had H_5, since it is hard to see what could make the difference between a possible situation in which Flimsy was formed from *B* and a different bucket was formed from *A* and the actual situation in which Flimsy is formed from *A* and a different bucket is formed from *B*.[9]

2.2 Quantified Tolerance Arguments

To construct a Tolerance Argument according to the schema in the previous section, we need to pick a singular term to substitute for the schematic letter '*a*'. But not all of the puzzles we are interested in involve anything that looks like a singular term: some instead involve universal quantification over all objects of a

[9] In Chapters 5 and 6, we will attempt to further pin down the source of this kind of resistance to Hypertolerance claims.

certain sort—e.g. all pyramids, or all tables, or all tables made according to a certain design, or all tables made in a certain factory. For example, it seems plausible that *every* Egyptian pyramid could still have been a Egyptian pyramid while being somewhat shorter, whereas at least some of them could not have been Egyptian pyramids while being radically shorter; but this combination of judgements can be challenged in a way that parallels the corresponding combination of judgements about the Great Pyramid.[10]

So, in our most general schema for Tolerance Arguments, the role of the singular term *a* will be taken over by some predicate *K*—'table', 'pyramid', 'Egyptian pyramid', 'table in this room', 'table in front of us', etc. Instead of our old notions of tolerance and hypertolerance, we will work with definitions relativized to *K*:

> *x* is *tolerantly K* := *x* is *K*, and for any *F* and *G* such that *G* is close to *F* and *x* is *F*, it is possible for *x* to be *K* and *G*.
>
> *x* is *hypertolerantly K* := *x* is *K*, and for any *F* and *G* such that *G* is ancestrally close to *F* and *x* is *F*, it is possible for *x* to be *K* and *G*.

For example, when 'close' stands for the relation that holds between two properties when both are heights and their ratio is between 9/10 and 10/9, something is tolerantly a pyramid iff it is a pyramid which could have still been a pyramid while having any height within 10 per cent of its actual height.[11]

Unless the closeness relation holds only between properties that entail being *K*, being tolerantly *K* is stronger than being *K* and tolerant. For example, one could consistently hold that a certain Egyptian pyramid could have had any height close to its actual height, although some of these possible heights are not such that it could have been an *Egyptian pyramid* while having those heights. Nevertheless, the intuitive force of the premise that the Great Pyramid is tolerant extends to the premise that it is tolerantly an Egyptian pyramid. And since the appeal of this thought doesn't depend on the features that distinguish the Great Pyramid from

[10] Arguments that stay at the general level by using such quantifiers also bypass certain potential ambiguities in English singular terms that threaten to obscure the metaphysically interesting reading of Tolerance premises stated using such terms. As we discussed in the Introduction, *non-rigid* readings are available for many expressions standardly classified as singular terms in English—certainly for definite descriptions ('Benjamin Franklin could easily not have been the first Postmaster General of the United States'), but also, arguably, for demonstratives and deictic pronouns ('[That speaker/she] was very nearly someone much more famous' could mean something like 'The second keynote speaker of the conference was very nearly someone much more famous'). Kripke (1972) famously argued that *proper names* do not give rise to such ambiguities, but this has been contested (see Elbourne 2005: ch. 6). While we trust in our ability to zero in on the intended resolution of such ambiguities, it is also helpful to be able to fall back on quantified formulations for which no similar concerns of ambiguity arise.

[11] For many of the predicates that we will be plugging in to the role of *K*, there is an an ambiguity between an eternal and a non-eternal reading: e.g. 'table' could mean 'thing that is, was, or will be a table' or 'thing that is currently a table', just as 'king' could mean 'person that is, was, or will be a king' or 'current king'. We get interesting puzzles on both ways of resolving such ambiguities; the choice of a resolution will rarely matter for our purposes.

other Egyptian pyramids, it extends to the premise that *every* Egyptian pyramid is tolerantly an Egyptian pyramid. But if we accept this, we will be confronted with a hard-to-resist argument that every Egyptian pyramid is *hypertolerantly* an Egyptian pyramid, which implies the prima facie implausible claim that any Egyptian pyramid could still have been an Egyptian pyramid while being thimble-sized.

The form of this more general version of the argument is given in Figure 2.3. This can be seen to be valid using a a slight variant of the derivation from the previous section. As before, we have three intermediate lemmas:

Necessitated Tolerance It is necessary that every *K* object is tolerantly *K*.

$$\Box\forall x\forall F\forall G(G \sim F \rightarrow (Kx \wedge Fx) \rightarrow \Diamond(Kx \wedge Gx))$$

Stepwise Necessitated Tolerance For any *F* and *G* such that *G* is close to *F*, and any object *x*: it is necessary that if *x* is *F* and *K*, then it is possible for *x* to be *G* and *K*.

$$\forall x\forall F\forall G(G \sim F \rightarrow \Box((Kx \wedge Fx) \rightarrow \Diamond(Kx \wedge Gx)))$$

Possibility Transfer For any *F* and *G* such that *G* is close to *F*, and any object *x*: if it is possible for *x* to be *F* and *K*, then it is possible for *x* to be *G* and *K*.

$$\forall x\forall F\forall G(G \sim F \rightarrow \Box((Kx \wedge Fx) \rightarrow \Diamond(Kx \wedge Gx)))$$

The derivations of Necessitated Tolerance from Tolerance and Non-contingency, of Stepwise Necessitated Tolerance Necessitated Tolerance and Persistent Closeness, of Possibility Transfer from Stepwise Necessitated Tolerance and Iteration,

Tolerance Every *K* object is tolerantly *K*.

$$\forall\, x\forall\, F\forall\, G(G \sim F \rightarrow (Kx \wedge Fx) \rightarrow \Diamond(Kx \wedge Gx))$$

Non-contingency If every *K* object is tolerantly, then *K* necessarily every *K* object is tolerantly *K*.

$$\forall\, x\forall\, F\forall\, G(G \sim F \rightarrow (Kx \wedge Fx) \rightarrow \Diamond(Kx \wedge\ Ga)) \rightarrow$$
$$\Box\forall\, x\forall\, F\forall\, G(G \sim F \rightarrow (Kx \wedge Fx) \rightarrow \Diamond(Kx \wedge\ Gx))$$

Iteration Whatever is possibly possible is possible.

$$\forall\, p(\Diamond p \rightarrow \Diamond\Diamond p)$$

Persistent Closeness When properties are close, they are necessarily close.

$$\forall\, F\forall\, G(G \sim F \rightarrow \Box G \sim F)$$

Hypertolerance Every *K* object is hypertolerantly *K*.

$$\forall\, x\forall\, F\forall\, G(G \sim^{*} F \rightarrow (Kx \wedge Fx) \rightarrow \Diamond(Kx \wedge\ Gx))$$

Fig. 2.3 Tolerance Argument (quantified version).

and of Hypertolerance from Tolerance and Possibility Transfer are essentially the same as in the *de re* case.[12]

Our new schema for constructing Tolerance Arguments is a strict generalization of our initial *de re* version: we can get something whose premises and conclusion are logically equivalent to those of that version by choosing K to be *identical to a*, for a singular term *a*.

In many of the quantified Tolerance premises we will be considering, K is some predicate with a large number of instances, like 'table' or 'IKEA Melltorp table' (as opposed to 'table in this room', 'pyramid of Giza', etc.). Such premises do have a downside, in that they are a few steps further removed from our everyday practices. Whereas we often make claims that imply or presuppose Tolerance claims about specific tables, the question whether *every* table can survive changes of a certain kind is less likely to come up outside the context of philosophy. And it would be naïve to assume that the only basis we could have for our disposition to take it for granted that specific tables are tolerantly tables would be a commitment to a universal generalization to the effect that every table, or even every table of a certain design, is tolerantly a table. Habits of singular belief-formation are often not underwritten in any such way, and would often be more epistemologically problematic if they were.[13] On the other hand, there is something preposterous about the suggestion that while *most* IKEA Melltorp tables are such that they could still have been tables while having any one of their legs replaced by any other matching leg, a few of those tables are not like that. Even if the presence of a single intolerant Melltorp table lurking in the IKEA showroom in Burbank wouldn't disrupt our knowledge of the tolerance of a table in Stockholm, it is hard to take seriously the thought that Melltorp tables vary in this modal respect despite their uniformity in design. So, we think that many quantified Tolerance premises are on sufficiently firm ground that it would be quite shocking if the best way of escaping the associated Tolerance Puzzles required denying them.

[12] The one point where there is added potential for controversy is in the derivation of Stepwise Necessitated Tolerance. This works by applying CBF to Necessitated Tolerance to get $\forall x \forall F \forall G \Box(G \sim F \rightarrow (Kx \wedge Fx) \rightarrow \Diamond(Kx \wedge Gx))$, which implies Stepwise Necessitated Tolerance given Persistent Closeness. As discussed in note 3, contingentists will worry about the appeal to CBF here. To bridge the gap, it is no longer enough to strengthen Persistent Closeness to Persistent Closeness+ (as explained in that footnote); they will also need to add an extra premise such as $\forall x \Box(Kx \rightarrow \exists y(y = x))$, or more minimally $\forall F \forall G \forall x(G \sim F \rightarrow \Box((Kx \wedge Fx) \rightarrow \exists y(y = x)))$. This does not open up an interesting new resistance strategy, since in the puzzling cases, the remaining premises remain just as plausible or implausible if we explicitly add $\exists y(y = x)$ as an extra conjunct alongside Kx.

[13] For example, our habit of taking it for granted that apparently red things are red is certainly not underwritten by the universal belief that all apparently red things are red. As Williamson (2007: 146) remarks: 'The trouble with replacing a pattern of inference by a universal generalization is that it has us rely on all instances of the pattern simultaneously, by relying on the generalization … Epistemologically, folk "theories" seem to function more like patterns of inference than like general premises.'

2.3 Varying the Modality

When philosophers use words like 'necessary' and 'possible' without qualification, they normally have in mind the very demanding status of *metaphysical* necessity and the correspondingly undemanding status of *metaphysical* possibility. In our examples up to now, we have had these interpretations in mind. But the abstract Tolerance Argument remains valid under a wide range of interpretations of 'necessary' and 'possible', whether or not they are available meanings for those words in English. All we need require of the meanings assigned to 'necessary' and 'possible' in the argument is that they conform to the minimal basic modal logic presented in §1.4.[14] So, for example, we could interpret 'necessary' and 'possibly' as expressing *nomic* or *deontic* necessity and possibility. Or we could interpret 'necessary' as expressing the property of having an objective chance of 1 (at a certain time), and 'possible' as expressing the property of having an objective chance greater than 0 (at the same time). Or we could choose some sentence *P*, and replace 'it is necessary that . . . ' with 'it is metaphysically necessary that if *P* then . . . ' and 'it is possible that . . . ' with 'it is metaphysically possible that *P* and . . . ': no matter what *P* we choose, these operators will obey the basic modal logic on the assumption that metaphysical necessity and possibility do.[15]

All of these modalities are standardly thought to be *narrower* than metaphysical modality: everything metaphysically necessary is nomically and deontically necessary, has chance 1 at every time, and is necessitated by any given *P*, but the reverse is not true. Substituting narrower modalities for metaphysical modality thus amounts to a strengthening of both Tolerance and Hypertolerance. This will often sharpen the puzzle, since the ordinary judgements we have been appealing to in motivating Tolerance typically already concern some much narrower modalities, whereas even if we were prepared to bite the bullet and accept Hypertolerance on the metaphysical interpretation, we might well balk at the suggestion that the relevant metaphysical possibilities are nomically or deontically possible, or had at some point a positive objective chance of being actualized, or are metaphysically compossible with our chosen *P*.[16]

[14] In fact, we need much less than this: given that Iteration is one of the premises of the argument, we can harmlessly weaken $\Box^*$N, $\Box^*$K, and $\Box^*$Duality by replacing their initial $\Box^*$ with a $\Box$.

[15] Note that the Tolerance argument does not employ the T axiom, according to which anything necessarily true is true (and so anything true is possibly true). So the existence of falsehoods with an objective chance of 1 (e.g. that the dart does not hit a certain particular point) does not undermine the validity of chance-based Tolerance Arguments; likewise, the falsehood of *P* does not undermine the validity of Tolerance Arguments where the modalities are conditional on *P*.

[16] While we will be focusing on non-epistemic interpretations of the $\Box$, it is also interesting to consider interpretations of Tolerance Arguments on which $\Box$ stands for some *epistemic* expression like 'We know that . . . ' or 'We are in a position to know that . . . ' or 'It is certain that . . . '. It is far from obvious that such expressions conform to our basic modal logic; but even if they don't, one might think that one can idealize or regiment the epistemic operators in such a way that the argument goes through. In fact, if we interpret $\Box$ as 'one is in a position to know that', one can see Williamson's influential

Another reinterpretation that does not disrupt the logic of the argument is to read 'possibly' as 'sometimes' (or alternatively, 'it will be the case that' or 'it has been the case that') and 'necessarily' as 'always' (or 'it will always be the case that' or 'it has always been the case that').[17] For the temporal versions of Tolerance and Non-contingency to have any plausibility, we will need to suppose, at least in the spirit of thought-experiment, that there actually is some process of appropriately gradual change, in which properties that stand in the chosen closeness relation are instantiated at nearby times. But under the supposition of the right kind of history, temporal instantiations of the paradox can be quite gripping. A famous one from antiquity involves the Ship of Theseus, whose planks are replaced one by one until all the original planks have been discarded. This can be regimented as a temporal Tolerance Argument by defining '*G* is close to *F*' as follows:

> There are collections of planks *C* and *D* each of which composes a ship at some time and which differ by at most one plank, such that *F* is *being composed of C* and *G* is *being composed of D*.[18]

Necessitated Tolerance then says that whenever the Ship is composed of some planks (which at some time compose a ship), then for any other sometimes-ship-composing collection of planks differing by at most one, the Ship is *sometimes* composed by that collection. For this to be false, it would need to be the case that at least one of the plank-replacement operations did not have the effect of changing the composition of the Ship by one plank. This is compatible with various alternative hypotheses about what happened during those plank-replacements: the

"anti-luminosity argument" (Williamson 1994: ch. 8, 2000: ch. 4) as a special case of a Tolerance Argument, run as a *modus tollens* against Iteration ("the KK principle"). Thanks to Richard Roth for drawing our attention to this.

[17] In treating the temporal arguments as structurally parallel to modal Tolerance Arguments, we are adopting an "A-theoretic" position according to which there are true propositions that have not always been or will not always be true, so that the temporal operators really function as operators. According to the competing "B-theoretic" tradition, genuine *propositions* have their truth values eternally (Frege 1918), but words like 'sometimes' and 'always' make a difference because they function as, or analogously to, quantifiers. (See Dorr unpublished *a* (ch. 1) for more on this distinction.) B-theorists will regard our assimilation of the temporal puzzles to the modal puzzles as misleading. But they should still recognize the interest of the former, which by their lights would less misleadingly be treated on the model of certain spatial puzzles, such as the following (set in a world of gradual linguistic variation): Every language spoken in our village is spoken in all the neighbouring villages. But if every language spoken in our village is spoken in all the neighbouring villages, then every village is such that every language spoken in it is spoken in all its neighbouring villages. So, every language spoken in our village is spoken in every village reachable from it by a chain of neighbouring villages. This argument does not quite fit our schema for Tolerance Arguments, since it does not involve a sentential operator. Nevertheless, it is puzzling in a similar way, and our own approach to Tolerance Puzzles (Chapter 11) extends naturally to these spatial puzzles. By contrast, those who favour Iteration-denial as a response to Tolerance Puzzles will need a differential treatment. The worry that Iteration-denial provides an insufficiently general treatment of our puzzles is a theme that we will return to in §7.1.

[18] We ignore parts of the Ship other than the planks, such as its mast.

Ship might have been destroyed, so that after the operation it was not composed by any collection of planks; or it might have been scattered, so that it was no longer ship-shaped; or it might have shrunk, taking on the shape of a ship with a plank-shaped hole.

In Tolerance Arguments constructed with 'always' and 'sometimes', Iteration says that if it is sometimes the case that it is sometimes the case that *P*, then it is sometimes the case that *P*. Whereas Iteration for metaphysical possibility is a rather theoretical claim that some philosophers have resisted, this temporal Iteration claim seems especially secure: it is very hard to make sense of the idea that something that has never happened was once such that it had happened, or that something that is never going to happen is such that it will later be such that it will happen.[19]

Hypertolerance in this instantiation is equivalent to the claim that the Ship is at some time composed by all of the relevant collections of planks that sometimes compose *a* ship, including those collections that do not overlap the initial set of planks at all. This particular Hypertolerance claim is not terribly counterintuitive, and many of those who have considered the Ship of Theseus have embraced it.[20] After all, ocean waves and storms seem to routinely survive complete replacement of the water and air molecules that make them up, and likewise certain organisms seem able to survive complete replacement of the cells that make them up, so it is not so outlandish to suppose that ships can analogously survive complete replacement of the planks that make them up.[21] In other cases, we are perfectly well able to take in our stride temporal Hypertolerance claims that involve radical changes in structure and organization rather than changes in matter. For example, we sometimes seem to talk as if the town of Pompeii is still a place that tourists can visit despite the radical change it underwent in the year 79 CE, and we will happily point to the thing that comes out of a compactor and say 'That's my car.'[22, 23] And there is no obvious inconsistency in allowing that many objects are capable

[19] We will say a bit more about temporal Iteration in §7.1.

[20] Examples include Wiggins (1967: 37) (but see Wiggins 1980: ch. 3 for second thoughts), Smart (1972, 1973), Lowe (1983), and Chandler (1975: 365).

[21] There are, however, some interesting further arguments that can be given against the claim that the Ship survives the replacement of all its planks, including a famous one by Hobbes (1655) that we will return to in §5.2.

[22] Note that the true reading of 'That's my car' may involve a use of 'car' analogous to the use of 'U.S. President' in 'There were three U.S. Presidents at the funeral.'

[23] Of course, we don't always talk like this. 'Pompeii no longer existed after the eruption of Vesuvius' sounds good too; and it would be natural enough to say, 'That car no longer exists' when pointing to a photograph of a car whose compacted remains are still around in some scrapyard. The positive view we will presenting in Chapter 11 allows for a high enough level of context-sensitivity that there is no need to read error into any of the apparently competing natural descriptions of such cases.

of undergoing both complete changes of matter and radical changes of structure and organization.[24]

Still, there are some temporal Tolerance Arguments where Hypertolerance is harder to take in one's stride. For example, we could run a temporal analogue of our modal puzzle about the Great Pyramid under the supposition that some enormous grinding machine was applied to the base of the Great Pyramid, and left running until all that remained was a thimble-sized remnant. Embracing Hypertolerance in this case will amount to claiming that the Great Pyramid would still be around, but thimble-sized, at the end of such a process. This may be less comfortable than the conclusion that the Ship of Theseus is still around after any number of plank-replacements. Likewise, we can run a temporal analogue of our modal puzzle about the painting by imagining a history where a long-lived artist with evolving taste dabs and scrapes at an oil painting a little bit each day: at the beginning we have a landscape in the style of Constable's *Hay Wain*, at the end a surrealist effort in the style of Dalí's *Soft Construction with Boiled Beans*. Embracing Hypertolerance in this case amounts to claiming that the original painting is still around at the end of the process. This conclusion is more worrisome than the conclusion that ships can survive total replacement of their planks: in ordinary life one would not hesitate to describe the story we have told as one in which the original painting was destroyed, and would no longer count it as one of the extant paintings by that artist. (A real-life case involves an amateur restorer who dabbed and scraped at a fresco in a church in Borja, Spain, with the grotesque results visible in Figure 2.4. The judgements about this particular case seem to go both ways: the *Telegraph*'s story on the case (Govan 2012) is headlined "19th century fresco destroyed by rogue DIY pensioner", whereas according to Wikipedia (2020), "The intervention transformed the painting and made it look similar to a monkey".)[25]

That completes our presentation and explanation of our general schema for Tolerance Puzzles. In the remainder of this chapter, we will consider the specific family of puzzles having to do with variation in objects' original material composition that has dominated the literature, thanks in good measure to the seminal work of Chandler (1976) and Salmón (Salmon 1979, Salmoin 2006). Our aim will be,

[24] It is sometimes assumed, following Hobbes (1655), that the view that the Ship of Theseus can survive total replacement of its matter would have to go hand in hand with the view that "continuity of form" is the crucial factor in survival (for ships), which suggests that the Ship of Theseus couldn't also survive a run through the ship-compactor (see Simons 1987: 200). But one can perfectly well imagine a model where continuity of ship-form trumps continuity of matter when there is continuity of ship-form, but continuity of matter takes over as the decisive factor after a sufficiently radical discontinuity of form. Smart (1973) suggests a model along these lines, although on his picture the object ceases to exist once it becomes 'causally impossible' to reassemble the parts according to the relevant form.

[25] Williamson (1990: ch. 7) discusses the evolution of languages over time as another case where temporal Tolerance thoughts are natural but temporal Hypertolerance seems crazy. We could also consider a case where chess will slowly morph into Twister over thousands of years.

Fig. 2.4 The Borja *Ecce Homo*.

first, to show how they can be assimilated to our schema, and second, to sharpen them so as to bypass some shallow objections faced by various initially tempting ways of stating the premises. Spelling out the details takes some work, but we believe the exercise will be instructive.

2.4 Originating Matter

In what remains the most detailed treatment of a Tolerance Puzzle in the literature, Salmón (Salmon 1989*a*) develops an argument against Iteration whose central part can be assimilated to the structure of a Tolerance Argument. The following thesis (whose motivation we will examine below) plays a central role:

Salmón Tolerance If a [. . .] table x is the only table originally formed from a hunk of matter y according to a certain plan P, and y' is any (possibly scattered) hunk of matter that sufficiently substantially overlaps y and has exactly the same mass, volume, and chemical composition as y, then x is such that it might have been the only table originally formed according to the same plan P from y' instead of from y. (Salmon 1989*a*: 77)

We have taken a small liberty here by deleting the word 'wooden', which occurs where we have ellipsis dots. If we put it back, the premise becomes weaker, but we see no way of using it to raise a puzzle or argue against Iteration without relying on some distracting and debatable supplementary premise having to do

with woodenness.[26] Since Salmón gives no positive reason for thinking that the restriction to wooden tables is needed, it seems wisest to just pass over it.

With this modification made, Salmón Tolerance is equivalent to our Tolerance premise where the modality is metaphysical; *K* is 'table'; and 'close' is defined as follows:

> *G* is *Salmón-close* to *F* ≔ for some plan *P* and some hunks of matter *y* and *y′* that sufficiently substantially overlap and have exactly the same mass, volume, and chemical composition: *F* is *being the only table originally formed from y according to P*, and *G* is *being the only table originally formed from y′ according to P*.

The corresponding Non-hypertolerance claim—call it Salmón Non-hypertolerance—then says that there is a table *x* that is the only originally formed from some hunk of matter *y* according to some plan *P*, such that for some other hunk of matter *y′*, *x* could not have been the only table formed from *y′* according to *P*, although *y′* can be connected to *y* by a chain of hunks of matter, all with exactly the same mass, volume, and chemical composition, in which adjacent elements sufficiently substantially overlap. (Salmón has an interesting argument for this Non-hypertolerance claim, which we will discuss in Chapter 5.) The denial of Iteration follows from Salmón Non-hypertolerance, Salmón Tolerance, the Non-contingency of Salmón Tolerance, and the appropriate Persistent Closeness premise, which is equivalent to the claim that when two hunks of matter have the same mass, volume, and chemical composition and substantially overlap, it is metaphysically necessary that they do so.[27]

[26] The supplementary premise will need to rule out odd possibilities like the following: for any wooden table *x* originally formed from a hunk *y*, it is impossible for *x* to be a *wooden* table without being formed from *y*, though *x* could have been a *non-wooden* table formed from any hunk sufficiently substantially overlapping *y* (and having the same mass, etc.). Here is one premise that would do the job: necessarily, if a wooden table could be made of a certain hunk of matter *y*, then necessarily any table made of a hunk of matter with the same mass, volume, and chemical composition as *y* is wooden. However, this seems quite dubious, since woodenness plausibly requires having a certain kind of biological origin.

A different way of strengthening salmon's principle which also gets around the need for such supplementary premises would be to replace '... might have been the only table ...' with '... might have been *wooden*, and the only table ...'.

[27] The "Four Worlds Paradox" that Salmón generates based on Necessitated Salmón Tolerance and Salmón Non-hypertolerance is also somewhat different from the corresponding Tolerance Puzzle. As presented in Salmon 1989*a* (building on Salmon 1979: 722–3 and Salmon 1981: app. I), the paradox turns on the observation that the following six sentences are inconsistent given Iteration and the necessity of distinctness:

1. $M(a, h)$
2. $\Diamond \exists x M(x, h')$
3. $M(a, h) \rightarrow \neg \Diamond M(a, h')$
4. $\Box \forall z (M(z, h') \rightarrow \Diamond M(z, h''))$
5. $M(a, h) \rightarrow \Diamond M(a, h'')$
6. $\Diamond M(a, h'') \rightarrow \Box \forall x (M(x, h'') \rightarrow x = a)$

The closeness relation that generates Salmón's argument is very similar to the "origin-closeness" relation that featured as one of our examples at the beginning of this chapter:

> *F* is *origin-close* to *G* ≔ for some finite collections *C* and *D* of atoms such that *C* and *D* contain equally many instances of each atomic number, and at least 90 per cent of the members of *C* are also in *D*, *F* is *being originally composed by C* and *G* is *being originally composed by D*.

The differences are as follows: (i) We have been using 'collection of atoms' where Salmón, following Kripke (1972: 114, n. 56), has 'hunk of matter'. We mildly prefer doing it our way, since hunks of matter might be suspected of being an unwelcome holdover from some antiquated (Newtonian? Aristotelian?) physical theory. (ii) We did not require sameness of mass or volume.[28] (iii) For the sake of

(*Sketch of derivation:* get $\Diamond\exists x(M(x,h') \wedge x \neq a)$ from (1–3); infer $\Diamond\exists x(\Diamond M(x,h'') \wedge \Box x \neq a)$ from this, (4), and ND; infer $\Diamond\exists x\Diamond(M(x,h'') \wedge x \neq a)$ by K, hence $\Diamond\Diamond\exists x((x,h'') \wedge x \neq a)$ by CBF, and thus $\Diamond\exists x(M(x,h'') \wedge x \neq a)$ by Iteration. This contradicts $\Box\forall x(M(x,h'') \rightarrow x = a)$, which follows by Modus Ponens from (1), (5), and (6).) '$M(x,y)$' is to be read as 'x is the only table made from y according to P', for some specific P. a is to be some table that is the only table made according to P from a hunk h. h'' and h' are to be two hunks such that a could have been the only table made from h'', but could not have been the only table made from h', even though h'' has the same mass, etc., and sufficiently overlaps h'. (Salmón suggests that h'' might be the result of replacing one of the legs in h, and h' the result of replacing two legs; but as he recognizes, nothing turns on this: given Salmón Non-hypertolerance, there must be some way of choosing P, a, h, h'', and h' in such a way that h' has the same mass, etc., and sufficiently overlaps h'', and premises (1), (3), and (5) are true.)

Salmón argues for (4) from the necessary truth of Salmón Tolerance. He does not explicitly state a Persistent Closeness premise, but something like it seems to be needed to justify this inference; otherwise, one might think that the only worlds where a table is made from h' are worlds where it is false that h' and h'' have the same mass, etc., or false that they sufficiently overlap.

So far, then, we have the same ingredients as a Tolerance Puzzle. But Salmón's paradox is unlike the corresponding Tolerance Puzzle in that it also relies on (6). Salmón derives (6) from a "sufficiency of origins" premise which we will discuss in Chapter 5; as we will see, it is quite controversial. But in the present context, (6) is a red herring, since Salmón's set of sentences remains inconsistent in the same logic if we drop (2) and (6) while replacing (4) with (4′):

$$4'. \qquad \Box\forall z(M(z,h'') \rightarrow \Diamond M(z,h'))$$

The derivation of a contradiction is now simpler: given (1), (3), and (5), we have $\neg\Diamond M(a,h')$ and $\Diamond M(a,h'')$, but given (4′), the latter implies $\Diamond\Diamond M(a,h')$, and hence $\Diamond M(a,h')$ by Iteration. We no longer need ND. But the only motivation that Salmón gives for (4) applies in exactly the same way to (4′).

Thus, once we have Salmón Tolerance and Salmón Non-hypertolerance, the sufficiency-of-origins premise that Salmón uses to motivate (6) is not needed to generate the puzzle, or to motivate the rejection of Iteration. However, the premise is potentially relevant in a more indirect way, since Salmón also uses it to argue for Salmón Non-hypertolerance, as we will discuss in Chapter 5.

[28] It is in fact far from obvious that there are any table-sized collections of atoms (or hunks of matter) that *exactly* match with regard to their mass or volume: if there are only countably many such collections but mass and volume vary continuously, perfect match would seem to be vanishingly unlikely. Note that there is considerable vagueness as regards attributions of volume to collections of atoms, since it is unclear how to delineate an atom's boundary within the cloud of electrons that surround it, and unclear what to say about pairs of atoms whose clouds overlap.

concreteness, we specified 90 per cent overlap where Salmón says 'sufficient'. (iv) The properties related by our closeness relation do not require being made according to to any particular "plan", which seems an advantage given the potential for unclarity as regards what counts as a plan.[29] (v) Salmón has 'is the only table originally formed from' rather than just 'is originally formed from', ruling out cases where the same hunk is formed into multiple tables over the course of its career. (iv) and (v) make Salmón's version of Non-hypertolerance a bit weaker (and his version of Tolerance correspondingly stronger) than it would be if they didn't mention plans and didn't say 'only'. The weak version of Non-hypertolerance is easier to argue for on the basis of "sufficiency of origin" principles which we will be considering in Chapter 5. But it would be bizarre to suppose that the weak version is true and the strong version is false: that would mean that while any table could be originally composed by any hunk of matter (or collection of atoms) matching the one that in fact originally composed it, there are some cases where this can only be achieved by varying its plan, or by seeing to it that some other table also gets made from the same hunk or collection.

How gripping are the Tolerance Arguments based on either Salmón-closeness or our origin-closeness? Not very. Their most obvious flaw is that Persistent Closeness is quite implausible, since it seems quite likely that there are collections of atoms *C* and *D* for which it is true, but not metaphysically necessary, that *C* and *D* 90 per cent overlap and chemically match.[30] For one thing, the atomic number of an atom might be contingent—for example, a uranium atom can plausibly become a thorium atom by emitting an alpha particle.[31] For another thing, the *concrete existence* of a given atom (or hunk) is surely a contingent matter. This also makes for contingency in the closeness facts, since *having the same atomic number* entails *both concretely existing* (as does *having the same mass, etc.*). But in any case where it is only contingently true that *C* and *D* 90 per cent overlap and chemically match, it will only be contingently true that *being the only table originally formed from C* is origin-close to *being the only table originally formed from D*, so these two properties will be a counterexample to Persistent Closeness

[29] Salmón (Salmon 1981: 273) glosses 'plan' as 'form, structure, design, configuration'. Note that if "plans" were understood in such a way as to sometimes include *de re* reference to specific materials—e.g. 'make a square tabletop from those planks and attach some three-foot-long legs'—then Salmón Tolerance would become a great deal harder to justify.

[30] In Salmón's Four Worlds Paradox (see note 27), the problems for Persistent Closeness show up in the inference to (4) from Necessitated Salmón Tolerance and the claim that h' has the same mass, etc., and sufficiently overlaps h''. If we thought that there were some cases where one hunk only contingently has the same mass, etc., and sufficiently overlaps another, then even if we were convinced of Necessitated Salmón Tolerance, it is unclear what motivation there would be for thinking that we could choose a, h, h'', and h' in such a way as to make (4) (or (4′)) true along with (1), (3), and (5).

[31] The contingency of the *volume* of hunks of matter, which Salmón includes, is even more obvious given its dependence on conditions such as temperature.

(and similarly for Salmón-closeness). And without Persistent Closeness, we don't yet have any puzzle.[32]

But pointing out the implausibility of Persistent Closeness does not get to the heart of the puzzle: there are ways of modifying the argument which make Persistent Closeness unproblematic without making much of a difference to the plausibility of the other premises or the conclusion. The most straightforward option is simply to replace all talk of metaphysical possibility and necessity throughout the argument with "atomic possibility" and "atomic necessity", defined as follows:

> *p* is *atomically possible* ≔ *p* is metaphysically compossible with every collection of true propositions of the form *a is an atom with atomic number n* or *a and b are two distinct atoms*.
>
> *p* is *atomically necessary* ≔ *p* is metaphysically necessitated by some collection of true propositions of the form *a is an atom with atomic number n* or *a and b are two distinct atoms*.[33]

We include the distinctness propositions in the collections so that, for example, we don't have to worry about a scenario where, e.g., two 90 per cent-overlapping collections overlap by less than 90 per cent at some possible worlds at which certain pairs of distinct atoms belonging to both collections are identical. If metaphysical necessity obeys the Necessity of Distinctness, then dropping the distinctness propositions would of course make no difference.

Replacing metaphysical with atomic modality makes Persistent Closeness automatic, since whenever two finite collections of atoms *C* and *D* chemically match and overlap by at least 90 per cent, this fact is metaphysically necessitated by the

[32] The following model shows how the failure of Persistent Closeness lets us consistently combine Tolerance, Non-contingency, Iteration, and Non-hypertolerance in this instantiation. There are twenty-one metaphysically possible worlds: w_0 (which is the actual world) and $w_{i,j}$ for all $1 \leq i \leq 10$ and $j \in \{0, 1\}$. Each is accessible from itself and every other. At w_0, there are twelve gold atoms $a_1 - a_{12}$. There is also an object x that is a table at every world; there are no other possible tables. At w_0, x is originally formed from a_1–a_{10}. At $w_{i,j}$, x is instead originally formed from all of a_1–a_{10} except for a_i, together with with a_{11+j}. Each of $a_1 - a_{10}$ is still a gold atom at $w_{i,j}$, as is a_{11+j}; however, the remaining possible atom (a_{12-j}) is not—perhaps because it has some different atomic number, or perhaps because it does not concretely exist at all. So, Tolerance is true at w_0, since for each of the twenty-one collections of atoms matching $a_1 - a_{10}$ and including at least nine of them, there is a possible world where x is composed by that collection. And Tolerance is also true at $w_{i,j}$, since at $w_{i,j}$ there are just eleven collections that match $a_1 - a_{i-1}, a_{i+1} - a_{10}, a_{11+j}$ and include at least nine of them, and all of these compose x at some world. So Non-contingency is true at every world. Iteration is also true at every world, since the accessibility relation is universal. But Hypertolerance is false at w_0. *Being originally formed from* $a_3 - a_{12}$ is a property that x could not have had, although at w_0 it is close to *being originally formed from* $a_2 - a_{11}$, which is in turn close to *being formed from* $a_1 - a_{10}$, which x instantiates.

[33] For the version with Salmón-closeness, we could instead use "hunk-possibility" and "hunk-necessity" where '*p* is hunk-possible' means '*p* is metaphysically compossible with every collection of all true propositions of the form *y is a hunk of matter with mass m, volume v, and chemical composition C* and *y and y′ are hunks of matter that sufficiently substantially overlap*'.

collection of all true propositions of the form *a has atomic number n* or *a and b are distinct atoms*. It makes Tolerance a bit stronger: unlike the old version of Tolerance, the new version is falsified if there is a table *x* and a collection *C* overlapping *x*'s originating collection by at least 90 per cent, such that *x* could have been originally formed from *C*, but only in a world where some of the actual atoms aren't around anymore. But there is no interesting picture on which the old version of Tolerance is true while the new version is false: why should some possible modest changes as regards which atoms originate which tables require changes in what atoms there are or what their atomic numbers are? We are putting off discussing motivations for Non-contingency until the next chapter, but the motivations we have found apply in the same way whether we use atomic or metaphysical necessity. The new version of Non-hypertolerance is of course weaker than the old one (since atomic possibility is stronger than metaphysical possibility), and thus at least as plausible as it. And finally, the switch doesn't diminish the case for Iteration, since if Iteration holds for metaphysical necessity, it automatically also holds for atomic possibility.[34, 35]

Even when the worries about Persistent Closeness are taken care of by a narrowing of the operative modality along these lines, the originating-matter Tolerance

[34] Understand *being an atom with atomic number n* as *being an atom that contains exactly n protons*, where 'exactly *n*' is spelled out in terms of identity in the usual way. Then, the property of being a collection of propositions each of which is of the form *a is an atom with atomic number n* or *a and b are two distinct atoms* is a necessary property of any collection of propositions that has it (even if the property of being a collection of *true* propositions of these forms is not). Suppose *p* is necessitated by a collection of propositions *C* with this property. Then by Iteration for metaphysical necessity, it is necessary that *C* necessitates *p*. But since it is also necessary that *C* is a collection with the relevant property, it follows that it is necessary that if every member of *C* is true, then there is a collection of truths with the relevant property that necessitates *p*. Hence *p* is atomically necessarily atomically necessary.

[35] There are several other ways of modifying the argument to shore up Persistent Closeness. One is to insert an *actuality* operator into the definition of closeness, so that properties are close in the new sense iff they are *actually* origin-close. Persistent Closeness will then follow from the principle that whatever is actually the case is necessarily actually the case. The new version of Tolerance will follow from the old version together with the principle that whatever is the case is actually the case, and the old version of Hypertolerance will follow from the new one together with the principle that whatever is actually the case is the case. The cost of this move is that Non-contingency becomes harder to motivate. We are putting off discussing the motivations for Non-contingency until Chapter 3; but even without considering them, it is clear that when actuality operators are floating around, one has to be even more cautious than usual about making offhand assumptions to the effect that the considerations that support certain propositions also support their necessary truth.

Another option is to insert a *metaphysical possibility* operator into the definition of closeness, so that properties are close in the new sense iff they are *possibly* origin-close. Persistent Closeness will then be derivable from the widely accepted 5 axiom that whatever is possible is necessarily possible, which is widely accepted for metaphysical modality. However, this axiom is controversial. Iteration-deniers like Salmón have good reason to reject it (since its necessary truth entails Iteration when combined with that of the T axiom according to which everything true is possible). And more problematically for the current dialectic, there are some considerations which might lead even friends of Iteration to reject it, as we will be discussing in §4.1 and §8.1. Moreover, the possibility-based definition makes Tolerance stronger and subject to new objections: it would be false if, for example, there is a table *x* originally composed of a collection *C* of a trillion iron atoms such that *C* could instead have been a collection of a trillion xenon atoms, and a different collection *D* which is *necessarily* a collection of a trillion xenon atoms and could not have originally composed a table.

Arguments face another obvious worry, albeit one that again does not seem to get to the heart of the puzzle: namely, their Tolerance premises are rather strong, and go well beyond anything that can be extracted from ordinary judgements. Consider a table, Woody, originally composed by a collection of atoms *C*, and a second collection *C′*, comprising 95 per cent of *C* together with some more atoms, chemically matching the remaining 5 per cent of *C*, which were recently created by nuclear fusion during a supernova in a very distant galaxy. It is not obviously *physically* possible for *C′* to have originated *any* table at all, let alone Woody: perhaps gathering *C′* into an appropriately table-shaped arrangement would have required something physically impossible, like superluminal transportation.[36] And even though the claim that it would be *metaphysically* (or atomically) possible for *C′* to originate a table is not quite as dubious as the claim that it is *physically* possible, it is also far from obvious. Even on the assumption that the laws of nature are contingent (pace Shoemaker 1998), it would not be absurd to think that none of the actual atoms could have been an atom unless the actual laws of nature were true. Highly scattered hunks raise parallel concerns for Salmón's tolerance premise Salmón Tolerance.

Salmón does not simply appeal to Salmón Tolerance as a premise: he provides the materials for a deductive argument for it, which is worth a look in the light of the above worries. The argument has two premises (which he states without argument: Salmon 1989*a*: 75):

A If a [. . .] table *x* is the only table originally formed from a hunk . . . of matter *y* according to a certain plan . . . *P*, then *x* is such that it might have been the only table formed according to the same plan *P* from a distinct but overlapping hunk of matter *y′* having exactly the same mass, volume, and chemical composition as *y*.

B If a [. . .] table *x* originally formed from a hunk of matter *y* is such that it might have been [the only table] originally formed from a hunk of matter *y′* according to a certain plan *P*, then for any hunk of matter *y″* having . . . [at least as much] matter in common with *y* that *y′* has, and having exactly the same mass, volume, and chemical composition as *y′*, *x* is also such that it might have been [the only table] originally formed from *y′* according to the same plan *P*.[37]

[36] One might think it obviously is physically possible on account of the physical possibility of the initial conditions being such as to guarantee that *C′* is never very widely scattered. But it is far from clear that radically different initial conditions could generate *C′* at all, since some form of origin-essentialism might apply to the atoms in *C′*. Perhaps, for example, some members of *C′* are such that necessarily, if they concretely exist at all, they originate in a particular supernova.

[37] The original versions of (A) and (B) have 'wooden' in place of our first ellipses: we deleted these for the same reason as with Salmón Tolerance. We also added the words 'the only table' at two points to (B), and replaced Salmón's 'exactly the same matter' with 'at least as much matter', in order

Together, (A) and (B) imply Salmón Tolerance; or to be more careful, they imply that there is *some* percentage of overlap *n* less than 100 per cent such that Salmón Tolerance is true when we interpret 'sufficiently substantially overlap' as 'overlaps by at least *n*%'. And (A) certainly seems plausible. But our worries about super-scattered hunks apply just as much to (B) as to Salmón Tolerance, so this argument does not provide a response to those worries.

Indeed, even if we ignore scattered hunks, (B) is on reflection an extremely bold claim. Assuming Salmón Non-hypertolerance, there is a table *x* originally formed from a hunk *y* in accordance with a plan *P* such that some, but not all, of the hunks sharing *y*'s mass, volume, and chemical composition are such that *x* could have been the only table originally formed from them in accordance with *P*. (B) asserts that the *only* factor relevant to whether a given hunk *y′* belongs to one or other of these subclasses is the fraction of overlap between *y′* and *y*. But on what grounds could we be confident that the line is drawn as neatly as this? Might not certain subtler differences between hunks which agree as regards their degree of overlap with *y* be relevant—for example, how easy it would have been for them to have been selected at the time when *y* was in fact selected to be formed into a table?[38]

The worry about Tolerance raised by scattered collections feels like it should be avoidable by judiciously strengthening the closeness relation, so that *being originally composed of C′* will no longer count as "close" to *being originally composed of*

to make the argument valid. The 'exactly as much' version is compatible with there being a table *x* formed from a hunk *y* such that among all the other hunks with exactly the same mass, volume, and chemical composition as *y*, *x* could have been formed according to the same plan from all those with exactly 98 per cent of their matter in common with *y*, but not from those with between 99 per cent and 100 per cent of their matter in common with *y*, and thus fails to imply Salmón Tolerance in conjunction with (A).

[38] Salmón's (B) exemplifies a certain general schema we could call "Uniform Tolerance". To instantiate this schema we need to isolate a ternary relation of *comparative* closeness on the properties in some family. The claim (see Yablo forthcoming) is the "zone of tolerance" of every *K* object is a sphere centred on a property it instantiates, with all of the properties it couldn't instantiate while being *K* further away than any of the properties it could instantiate while being *K*. We can break this up into two parts:

Spherical Zone For every *K* object *x*, there is a property *F* such that every property in the family that *x* could have while being *K* is closer to *F* than any property in the family that *x* couldn't have while being *K*.

Sweet Spot Every *K* object *x* has some property *F* such that for any properties *G* and *H* in the family: if *x* couldn't have *G* while being *K*, there is a property *J* that *x* couldn't have while being *K* that is at least as close to *H* as *G* is to *F*.

Sweet Spot says, intuitively, that every *K* object as far as can be from the edge of its zone of tolerance. Both parts are potentially quite problematic. Accepting Spherical Zone would require a hard-to-justify confidence that the selected comparative closeness relation exhausts everything relevant to the modal behaviour of the relevant objects, as discussed above with regard to Salmón's (B). For a potential counterexample to Sweet Spot, consider a table *x* which was intended to be made of a certain hunk *y* but where the original artificer died, leaving it unfinished, and some other artificer finished it hastily using some lower-grade material. One might think that some hunks that *x* couldn't have been made from overlap *x*'s actually originating hunk more than any such hunk overlaps *y*.

Fortunately, there are many Tolerance Arguments where Tolerance is plausible quite independently of any Uniform Tolerance premise.

C when C' is unduly scattered and C is not. Flat-footedly, we could add a clause to the definition of closeness requiring that both of the relevant collections (or hunks) be in some sense "not too scattered". However, since scatteredness is a contingent matter, this will reintroduce a problem for Persistent Closeness.[39] More promisingly, we could require both collections to be such that they *could* have composed tables. So long as atomic possibility obeys the 5 axiom (see above), this condition is necessarily satisfied whenever it is satisfied, and hence will not introduce any new problems for Persistent Closeness. However, it is not clear that it suffices to dispel all the worries that scattered collections raise for Tolerance. For one might hold that it is metaphysically possible for some intergalactic collection C' to have originally composed a table after undergoing a miraculous superluminal transportation, while still denying that it is metaphysically possible for C' to have originally composed *Woody*, on the grounds that it is essential to Woody to be created, if at all, by some unremarkable, non-miraculous process (such as we find in the actual world).[40]

Our best shot at avoiding this worry is to strengthen the definition of closeness further, requiring both relevant collections to be such that it is not only atomically possible for them to be formed into tables, but atomically possible for them to be so formed without anything miraculous or otherwise remarkable happening to them during the process.[41] It's perhaps a little disappointing, but on reflection not surprising, that the relevant restriction is specified in such a vague manner. As far as we can see, the vagueness in 'miraculous' and 'remarkable' doesn't introduce any new problems, and leaves Tolerance impervious to worries based on scattered or otherwise "special" collections of atoms. However, to keep things simple, we will generally ignore this restriction from now on.[42]

[39] One might attempt to address this by further narrowing the operative modality from atomic possibility to one that also requires holding fixed the facts about which collections are not too scattered. But this is now a rather demanding notion of possibility, since it is hard to make any interesting changes to the distribution of matter without *some* collections thereby becoming somewhat more scattered. So no matter how we spell out "too scattered", there is a risk that possibility in the new sense will be so demanding that the new Tolerance will be implausible for a completely different reason.

[40] This idea is quite in keeping with the tradition of essentialist thinking inspired by Kripke (1972: n. 57), which is enthusiastic about the idea that certain rather specific facts about a particular object's originating process were necessary for that very object to be created.

[41] To control for the fact that some collections may be such that they could have unremarkably composed a table a thousand years ago, but not more recently, we had better make the argument be about properties which specify not only an originating collection of molecules but also a time of manufacture. Then we can, for example, incorporate remarkableness into a closeness relation among such properties as follows: *being formed from C at t* is close to *being formed from C′ at t′* iff $t = t'$, C and C' 90 per cent overlap and chemically match at t, and both C and C' could (atomically) have been formed into tables at t without anything remarkable happening.

[42] A different strategy for circumventing the worries would be to introduce some *de re* restriction to a particular collection of collections of atoms, e.g. all the collections of atoms *actually* within a certain workshop during a certain period, comprising reasonably contiguous portions of wood, etc. However, as discussed noted in note 35, the use of the actuality operator in specifying the Tolerance premise tends to make Non-contingency harder to justify. This point also applies to names introduced with the help of reference-fixing descriptions.

When fixed up along these lines, the Tolerance Arguments involving originating matter make for some quite compelling puzzles. But like all Tolerance Puzzles, they involve a Non-contingency premise, whose attraction is less immediate. In the next chapter, we will discuss how Non-contingency might best be motivated.

3
Motivating Non-Contingency

In the Introduction, we showed how ordinary judgements provide strong support for the *truth* of many Tolerance premises. But such judgements do little to support their *necessary* truth. We thus still stand in need of a motivation for the Non-contingency premises, without which the Tolerance argument gets nowhere.

Some philosophers have suggested that Tolerance Puzzles should be seen as a special case of Sorites Paradoxes, since they depend on the thought that small enough differences can't matter in a certain respect. If that *was* the thought, it would support Necessitated Tolerance just as much as Tolerance itself, so there would be no interesting prospect of holding that Tolerance is true but not necessarily true (i.e. of denying Non-contingency). But as we will discuss in §3.1, Tolerance Puzzles are *not* properly assimilated to the Sorites, since there are interesting motivations for Tolerance premises that have nothing to do with the deeply problematic "small differences can't matter" idea. Non-contingency thus needs a separate motivation. In §3.2 we will discuss some initially tempting motivations that are hard to develop in a satisfying way; §3.3 introduces what we think is the most promising motivation.

3.1 Tolerance Arguments and the Sorites

In §2.2 we saw how the ancient paradox of the Ship of Theseus can be regimented as a Tolerance Argument. Another celebrated ancient paradox, the Sorites, can also be more or less fitted into our mould. Use metaphysical modality; take K to be 'heap'; and define 'G is close to F' as 'for some positive integer n, F is *being composed of* $n + 1$ *grains of sand* and G is *being composed of* n *grains of sand*'. Tolerance and Hypertolerance in this "Heap Argument" are equivalent to the following:

Heap Tolerance For any positive integer n: every heap composed of $n + 1$ grains of sand could have been a heap composed of n grains of sand.

Heap Hypertolerance For any positive integers $n < m$: every heap composed of m grains of sand could have been a heap composed of n grains of sand.

The Bounds of Possibility: Puzzles of Modal Variation. Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri, Oxford University Press. © Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri 2021.
DOI: 10.1093/oso/9780192846655.003.0004

Since Tolerance Arguments are valid, Heap Hypertolerance follows from the necessitation of Heap Tolerance given Iteration and Persistent Closeness (which is uncontroversial in this case). But Heap Hypertolerance is completely absurd: there is in fact a heap composed of some finite number of grains of sand, but it could not have been a heap composed of just one grain of sand, since there could not be a heap composed of just one grain of sand.

We can also consider a temporal variant of this argument, which we would run under the supposition that some heap will be subjected to a process of gradual grain removal. In this version, Necessitated Tolerance says (in effect) that if at some point in the process there is a heap composed of the remaining grains of sand, then that heap will still be a heap after the next grain-removal, while Hypertolerance implies that every heap initially composed by some finite number of grains will eventually be a heap composed of just one grain.

Other familiar versions of the Sorites can be developed in a similar way: for instance, take *K* to be 'person who is not bald', define '*G* is close to *F*' as 'for some *n*, *G* is *having exactly n hairs* and *F* is *having exactly n* + 1 *hairs*', and either use metaphysical modality or use a tense operator and run the argument under the supposition that some initially hairy person will undergo a gradual total hair-removal (or that some initially bald person underwent a gradual thorough hair restoration). There is of course no temptation to think that the original person is no longer there at all at the end of this process. But for Hypertolerance to be false, it suffices that the original person is bald at the end of the process.[1]

The Sorites, in its many versions, fully deserves its position as a central puzzle of philosophy. For some reason that is not so easy to identify and articulate, the thought 'How could removing one insignificant little grain of sand possibly make the difference between a heap and a non-heap?' has a lot of pull on us. But the dominant consensus, which we share, is that thoughts like this are bad thoughts, ultimately no better than the thought that many small differences cannot add up to a large difference.[2] Given this, the central philosophical challenge posed by the Sorites is to diagnose why we were tempted by the thoughts in a way that lets us develop a stable theoretical perspective on so-called borderline cases. This is an important task. But it is not the task of this book. We will be proceeding on the assumption that the 'How could such a small change matter?' intuition that drives Sorites Paradoxes must be rejected outright. We just have to get used to the idea

[1] Those who are tempted by reactions like 'only those with zero hairs are literally bald, so what is the puzzle?' should replace 'not bald' with something like 'very hairy' or 'nowhere near bald'. Similarly, if you are tempted by thoughts like 'four grains but not three could make a heap, since four are required for one to balance stably on the rest' (Hart 1992), replace 'heap' with 'large heap', following Williamson (1994: 213).

[2] Proponents of classical logic (see §1.1 for a list of some members of this large and varied group) of course think that one grain of sand *can* make the difference between a heap and a non-heap. But opponents of classical logic, even if they refuse to assert that one grain can make a difference, tend to stop short of asserting that no grain can.

that even the tiniest changes do, on occasion, make a difference of the relevant sort: from a situation in which a heap is present to one in which no heap is present, or from a situation in which a person is not bald to one in which they are bald.[3]

The Heap Argument discussed above isn't exactly the same as the usual Sorites Argument familiar from the ancient and modern literature. It concerns *de re* modality: the absurd conclusion is that each *particular* heap could still have been a heap while consisting of just one grain of sand. The usual Sorites Argument has the weaker (but still absurd) conclusion that there could have been *a* heap consisting of one grain of sand. If we wanted to eliminate *de re* modality from the argument, we could do so by replacing Heap Tolerance and Heap Hypertolerance with the following:

Unspecific Heap Tolerance For any positive integer n: if there is a heap consisting of $n + 1$ grains, then there could be a heap consisting of n grains.

Unspecific Heap Hypertolerance For any positive integers $m > n$: if there is a heap composed of m grains of sand, there could have been a heap composed of n grains of sand.

Unspecific Heap Hypertolerance is absurd, and follows from the necessitation of Unspecific Heap Tolerance given Iteration. But the appeal to Iteration in this argument (which does not fit our schema for Tolerance Arguments) looks like a red herring. Its role is to enable us to derive the following lemma from the necessitation of Unspecific Heap Tolerance:

Unspecific Heap Possibility Transfer For any positive integer n: if it is possible for there to be a heap composed of $n + 1$ grains of sand, it is possible for there to be a heap composed of n grains of sand.

But the 'How could a tiny difference matter?' thought that drives Unspecific Heap Tolerance can just as well be used to motivate Unspecific Heap Possibility Transfer directly. Nothing dialectically interesting is added by the phase of the argument that depends on Iteration. If you have thought about the Sorites sufficiently to internalize the case that Unspecific Heap Possibility Transfer is false—i.e. that there is some number n such that it is possible for there to be a heap composed of n grains of sand, but impossible for there to be a heap composed of fewer than n grains of sand—then you should not be at all impressed by an argument that depends on Unspecific Heap Tolerance. If there is a minimum *possible* size for heaps, it seems quite likely that somewhere in the vast expanse of the Sahara desert, there already *is* a heap of that minimum size. And even if for some reason you were

[3] See §1.1 for more on this.

confident that the actually instantiated array of heap-like configurations was gappy in such a way that there weren't any minimum-sized heaps, it is hard to imagine what could convince you that this was necessarily true.

These remarks also apply to the original *de re* Heap Argument. Iteration in that argument also looks like a red herring. Its job is to get us from the necessitation of Heap Tolerance to the following lemma, which implies the absurd Heap Hypertolerance:

Heap Possibility Transfer For any positive integer n and heap h: if it is possible for h to be a heap composed of $n + 1$ grains of sand, it is possible for h to be a heap composed of n grains of sand.

But again, insofar as the 'How could a small change make the difference?' intuition is what drives the appeal of Heap Tolerance, one might as well appeal to the intuition to directly support Heap Possibility Transfer. Nothing dialectically interesting is added by the argument that starts with Heap Tolerance and derives Heap Possibility Transfer using Non-contingency and Iteration. Moreover, if you have thought enough about the Sorites to have internalized the case for thinking that there is a minimum number of grains that could compose a heap (i.e. that Unspecific Heap Possibility Transfer is false), you should not be at all impressed by either Heap Possibility Transfer or Heap Tolerance. Any minimal-sized heaps that might be lurking in the Sahara will be a counterexample to Heap Tolerance as well as Unspecific Heap Tolerance; if there somehow happen not to be any in the actual world, the problem instead lies with Non-contingency.

Fortunately for the interest of this book, there are many Tolerance Arguments in which Tolerance has an appeal that does not rest merely on the disastrous 'How could small differences matter?' intuition. One example is the Bucket Argument discussed in §2.1. Its closeness relation is just the relation of following in a certain series of five properties, and the steps along that series are not at all small: for example, one of them involves doubling the size of a bucket by adding an extra chunk of plastic to its original composition. The judgement that a thin bucket could have been fattened in this way does not rely on the thoughts that give the Sorites its grip, such as the thought that it would be mysterious how our words could possibly have ended up with meanings which drew a line in this place rather than some other place. The most compelling motivation comes down, rather, to a range of modal judgements of a sort that often come up in everyday life, such as the judgements we might express by saying, 'There wouldn't be paint all over the floor if the walls of that bucket had been thicker.' True, if someone was inclined to resist the Bucket Argument at the "fattening" step, we might want to push back against them by considering a longer series of properties separated by smaller steps—e.g. we could fatten up gradually, adding only 10 per cent each time, and then thin down gradually in the same way; or we could substitute an entirely new piece of

plastic a little at a time while leaving the size of the bucket the same at each step. While the resulting multi-step arguments are perhaps a bit harder to resist than the original five-step Bucket Argument, the impression that they are *much* more forceful seems due primarily to the availability of a bad, Soritical motivation for Tolerance on top of the good motivation from ordinary modal judgements.[4]

This also applies to other gradual Tolerance Arguments. The interesting motivation for the judgement that Woody could still have been a table while being composed of any slightly different collection of atoms (of appropriate sorts) has nothing to do with the idea that it would be mysterious how something so slight could be relevant. Indeed, the best-motivated judgements in this vein involve larger differences of a sort more likely to be relevant to the purposes of everyday life, such as the judgement that Woody could have had a different leg—for example, if we observed that one of the legs was shorter than the rest, we would naturally conclude that it would have been more stable if the carpenters had instead picked a different leg when they were making it. (We can even quite readily get into a mood where we are happy to say that the table would have been a more or less useful or beautiful table if it had had *four* different legs.) The judgements about tiny differences involving just a few atoms are best motivated by extrapolation from judgements about these larger, more practically relevant differences.

The dangers of slipping into Soritical thinking are greater when it comes to quantified Tolerance Arguments, especially when the restricting predicate *K* is something broad like 'table' as opposed to something narrow like 'table made in this factory this year'. Given that Sorites premises are false, we have good reason to think that some tables are so low or irregular or flimsy that they are only *barely* tables, which could have failed to be tables at all while being only slightly different in shape or size or composition. Thus we need to be careful to steer clear of Tolerance premises about all tables which would seem plausible if we were thinking only of paradigmatic tables, e.g. that any table could still have been a table while having any somewhat different ratio of height to length, ratio of longest leg length to shortest leg length, etc. Other kinds of exotically non-paradigmatic tables will put pressure on certain Tolerance premises involving changes in original composition without changes in shape, size, or materials. For example, perhaps someone somewhere has made a one-legged table whose surface is so incredibly thin that 99 per cent of its mass is in its one leg, in which case the Tolerance premise 'Every table could still be a table while having any of its legs replaced by any other leg of the same size and design' is far more tendentious than it might have seemed if we were only thinking about paradigm cases. Likewise, nanoscale tables made of

[4] The relevant ordinary judgements may in addition be a bit more forceful in the more gradual versions of the argument. But working with those versions carries the dialectical risk that people will think that the entire motivation for Tolerance is undermined once the Soritical motivations have been set aside.

just a few atoms will put pressure on Tolerance premises like 'Every table could still be a table made of any collection of matching atoms differing only by ten atoms.' Still, as discussed in §2.4, some Tolerance premises where *K* is 'table' remain *fairly* plausible even when we remind ourselves of the range of marginal and exotic tables there might be. And many more become plausible when *K* is something narrower along the lines of 'IKEA Melltorp table'.

Forbes (1984) points out that in the modal logic S5, the formula $\Box(AP \rightarrow \Diamond BQ)$ is equivalent to $\Diamond AP \rightarrow \Diamond BQ$, so that Stepwise Necessitated Tolerance not only implies, but is equivalent to, Possibility Transfer. He takes this observation to support his thesis that the "Sorites formulation" of the paradox—in our nomenclature, the argument from Tolerance and Possibility Transfer to Hypertolerance—is a better formulation of the paradox than the version that goes by way of Stepwise Necessitated Tolerance and Iteration:

> The intuitive thought which gets the Paradox going is that a sufficiently small degree of change in the original constitution of an object preserves possibility for that object…. So one who assents to the premisses of Chisholm's Paradox in its first version should also assent to its premisses in the Sorites version, and a solution which works for one version should be accepted only if it works equally well for the other. (Forbes 1984: 173; see also Forbes 1985: 163–8)

Forbes argues on this basis that his own account—which invalidates Modus Ponens, thus enabling him to treat Sorites premises like Possibility Transfer as "inviolable" (Forbes 1985: 168)—is superior to Chandler's and Salmón's Iteration-denying solution to Tolerance Puzzles.

Since our background logic includes classical propositional logic, we will not be discussing Forbes's view in any detail. He is doubtless right that if one takes the Sorites Paradox to independently justify a complete overturning of logical orthodoxy—on Forbes's view, even that old classic 'All men are mortal; Socrates is a man; therefore Socrates is mortal' is invalid, thanks to the vagueness of 'man' and 'mortal'!—there isn't much point in bothering with solutions to Tolerance Puzzles that require denying any of the premises, or that require accepting Hypertolerance. If we can without inconsistency accept (or at least maintain a high level of confidence in) all five of the prima facie plausible, but classically inconsistent, claims that comprise a Tolerance Puzzle, why wouldn't we? But like Salmón (Salmon 1993: n. 26), we firmly reject Forbes's claim that the "intuitive thought which gets the Paradox going" is just the Sorites-like one. And so we think that proponents of classical logic—and indeed, anyone who accepts the validity of the Sorites Argument, which also includes many nonclassical theorists—should regard all of the possible ways out of Tolerance Puzzles as options to be carefully weighed.

Even though Tolerance Puzzles aren't just warmed-over Sorites Paradoxes, it is still plausible that there is vagueness in the vocabulary used in setting them up.

In particular, if the right solution to some Tolerance Paradox involves accepting Tolerance and rejecting Hypertolerance, we should expect that there will often be a lot of vagueness as regards the question 'How many steps do we have to go, starting with a property instantiated by the relevant object, before we get to a property that it could not instantiate?' For example, if we think that the Great Pyramid could have been 90 per cent of its actual size but couldn't have been 1 per cent of its actual size, we should surely expect 'What is the greatest *n* such that the Great Pyramid couldn't have been *n* per cent of its actual size?' to suffer from vagueness analogous to that of 'What is the greatest *n* such that there couldn't have been a heap made of *n* grains of sand?'[5] As usual, such vagueness will plausibly go along with a kind of context-sensitivity, where we are free to talk in different ways that place different, incompatible demands on the answers to the question.[6] It will seem wrongheaded to attempt to organize a serious philosophical debate about the answer to the question. And while we may be puzzled by the apparent arbitrariness involved in the claim that *there is* such a greatest *n*, we should no more take this puzzlement as a reason to reject the conjunction of Tolerance and Non-hypertolerance than we should take the corresponding puzzlement about heaps as a reason to reject the conjunction of the claim that there could have been a heap made of a million grains with the claim that there couldn't have been a heap made of one grain. The interesting and distinctive challenge to the conjunction of Tolerance and Non-hypertolerance is not based on any such aversion to arbitrariness, but rather on the combination of Non-contingency, Persistent Closeness, and Iteration.

3.2 Non-Soritical Motivations for Non-Contingency

As we have seen, many Tolerance premises can be motivated by appeal to ordinary modal judgements which don't depend on any specifically philosophical modes of thinking, and in particular don't depend on the 'How could small differences matter?' thoughts that drive the Sorites. But the situation with Non-contingency is more delicate. Judgements involving iterated modality are rather peripheral to our everyday modal thought, so it is far from clear that denying Non-contingency would lead to any very drastic conflict with pre-philosophical modal opinion.

True, if we look hard enough, we will find some scattered examples in ordinary life of plausible-looking iterated modal claims that suggest that the truth of Tolerance is at least somewhat modally robust. For example:

[5] It's an interesting question how such vagueness at the level of sentences should be explained in terms of vagueness at the level of words—we will discuss this in Chapter 11—but the fact that the answer to this isn't obvious does nothing to undermine the reasons to expect there to be vagueness.

[6] For the idea that vagueness goes hand in hand with such context-sensitivity, see Lewis 1973, 1979, Raffman 1994, Soames 1999, Graff 2000, and Dorr 2003.

> If we hadn't replaced the blunt blade on that saw, then every table we made using it would have been in danger of being too crudely finished to be saleable.
>
> If we had made the top of that walnut table out of mahogany, then we would have soon realized that it would have looked a lot better if we had made its legs out of mahogany as well.
>
> If we had made a thick purple bucket by combining the red and blue pieces of plastic, we would have recognized it to be more attractive and sturdier than it would have been if we had made it out of just the red plastic or just the blue plastic.
>
> If I had made this a spork rather than a fork, I would have regretted not having taken the opportunity to just make it a spoon.

But some of these judgements are ones that even ordinary speakers might balk at unless you catch them in the right mood. And in any case, it is not obvious how one can get from these kinds of judgements to anything strong enough to play the role of a Non-contingency premise. Moreover, even if the strategy of denying Non-contingency did require some kind of localized error theory, or a localized appeal to some form of nonliteral discourse, with regard to certain iterated modal judgements such as those above, this is not remotely as radical as a strategy of appealing to error theory or nonliteral speech to dismiss the ordinary judgements that support Tolerance.

We will have more to say in Chapter 11 about the possibilities of accommodating the relevant aspects of our ordinary practice without embracing Non-contingency. But in any case, if one wants a case for Non-contingency strong enough to have any chance of supporting such far-reaching metaphysical conclusions as accepting Hypertolerance or denying Iteration, it is fair to expect some more theoretical argument in addition to whatever modest support can be extracted from ordinary claims like those above. And it is natural to worry that the only available arguments will just send us right back to the kind of Soritical thinking we are striving to avoid.

The literature on our puzzles doesn't provide much help here. Chandler (1976: 106) just breezily asserts that 'surely', if a certain bicycle had been made with a different spoke it would still have been capable of having been made with a different handle-grip. Salmón (Salmon 1986*a*: 77) likewise takes care of Non-contingency in a single sentence: 'Paradox arises when it is noted that none of these modal principles is the sort of proposition that merely happens to be true as a matter of fact.' Salmón seems to be gesturing towards an argument from a premise to the effect that every true proposition in a certain category is necessarily true, together with the premise that his Tolerance principle belongs to that category. But he provides no hint as regards what the relevant category might be.

One suggestion worth considering is that the relevant category is or involves *knowability a priori*. Most simply, one might argue that since Tolerance is knowable a priori, and everything that can be known a priori is necessarily true, Tolerance is necessary true. But the second premise of this argument is highly controversial, especially in view of Kripke's (1972) examples of propositions which are, putatively, knowable a priori despite being contingently true (e.g. that the Standard Metre is a metre long if it exists).[7] However, even among those who accept Kripke's counterexamples, many have thought that they depend on some special feature, e.g. the presence of names that were introduced by 'reference-fixing descriptions' rather than in the usual way. So there is still the prospect of a more complex argument that derives the necessity of Tolerance from its a priori knowability together with some premises of the form 'Any proposition that is knowable a priori and lacks special feature *F* is metaphysically necessary' and 'Tolerance lacks special feature *F*'. A survey of the candidates that might be thought to play the role of feature *F* would take us too far afield, but we are generally pessimistic about this argumentative strategy, in part because we are doubtful that there is any epistemological joint in the vicinity of the traditional a priori/a posteriori distinction.[8]

Even setting this general pessimism aside, it is quite doubtful whether the Tolerance premises we have been talking about are really a priori. For example, the origin-theoretic Tolerance premises discussed in §2.4 involve a restriction to collections of atoms which match as regards how many atoms they contain having each atomic number. The reasons for worrying about the stronger versions of Tolerance that drop this restriction are empirical ones. And we could make further discoveries about the importance of properties which distinguish atoms with the same atomic number—e.g. that some carbon atoms have a special feature that precludes them from ever becoming part of a piece of wood—which would prompt similar worries about the current version of Tolerance. These considerations carry over to Tolerance premises that utilize less scientific ideology. For example, we could also in principle discover that bicycles are intelligent organisms whose wheels serve a similar function to brains, in which case we would have a very different attitude to the premise that any bicycle could have had different wheels.

Proponents of the argument from a priority might respond to this by saying that what is really a priori is not the Tolerance premise itself, but some complicated conditional with Tolerance as its consequent and all the relevant background beliefs packed into the antecedent. But without having a worked-out example before us, it is extremely hard to say whether any given choice of antecedent would

[7] For other examples of putative contingent a priori knowledge with a different flavour see Hawthorne 2002.

[8] See Hawthorne 2007 and Williamson 2013*a*.

lead to something plausibly a priori while still being such that the necessity of the conditional could do the same argumentative work as Necessitated Tolerance.[9]

A different, and even harder to pin down, source of pressure towards Non-contingency involves some idea to the effect that the falsity of Non-contingency would mean that *the actual world was special* in a certain respect, and that it is somehow repugnant to think that the actual world is in this respect special. For example, Forbes (1985: 160) justifies an initial 'Necessarily' in his version of Necessitated Tolerance by the remark 'we do not believe that there is some special property of actual artefacts or the actual world which makes this so'. However, it is unclear what he means by 'special', and he does not say what would be bad about the relevant belief. Kment (2018: 559–60), more helpfully, suggests cashing out the relevant form of specialness as a kind of *orderliness*. The suggestion (not one he ultimately endorses) is that if Tolerance is true but could easily have been false, then that means that the actual world, though itself metaphysically orderly, is 'surrounded in modal space by metaphysically disordered worlds' that contain a mixture of tolerant and intolerant tables. But turning this into a fully worked-out argument for Non-contingency is a tall order. After all, there are many respects in which it is perfectly orthodox to think that actuality is orderly in a way that not all possible worlds are orderly—for example, it is subject to simple and exceptionless laws of nature, whereas some possible worlds are much more chaotic.

That said, this talk of specialness and orderliness does succeed in bringing out a prima facie strangeness to such thoughts as that actually all tables are tolerantly tables although there could easily have been tables that were not tolerantly tables. We should hope that any view that denies Non-contingency will have something to say to ameliorate this initial impression.

So far, the case for Non-contingency may be looking somewhat slender, depending either on some rather scattered and peripheral ordinary judgements involving iterated modality, on some highly tendentious doctrines involving the vexed concept of a priori knowledge, or on some hard to precisify sensibility about the actual world not being special.[10] If one is not impressed by these considerations, one might start to suspect that there is no real puzzle here at all beyond the familiar general puzzles posed by Sorites Arguments. But there is, we

[9] That said, the Non-contingency-denying view we will be developing in Chapter 11 will suggest that Tolerance premises have much more in common with the standard Kripkean examples of the contingent a priori than one might have supposed. So even if we bracket doubts about whether they are in fact a priori, and take there to be a special category within which a priority guarantees necessity, the claim that Tolerance premises belong to this category will be quite dubious.

[10] Another style of argument, which McKay (1999) thinks fully explains the attraction of Non-contingency, derives it from the combination of some universally quantified possibility claim such as 'Every hunk that could have been made into a table could have been made into a table that was tolerantly a table' with a "sufficiency" principles along the lines of 'Necessarily, if a table *x* is made of hunk of matter *y*, it is necessary that any table made of *y* is identical to *x*.' §5.2 will further discuss such sufficiency principles; like McKay, we are somewhat sceptical. The other premise also seems far from obvious.

think, a more compelling case to be made for Non-contingency, which has nothing to do with the Sorites, and also does not require bringing in difficult concepts like a priority and specialness. In the next section we will attempt to spell out this argument.

3.3 The Security Argument

The motivation for Non-contingency that strikes us as most compelling turns on the thought that when we endorse Tolerance and the everyday modal judgements that motivate it, it is not just a matter of *luck* that we aren't making a mistake. Consider the Table Argument (the Tolerance Argument where *K* is 'table', the modality is atomic possibility, and the closeness relation is origin-closeness). If Tolerance is only contingently true, then there are atomically possible worlds where some tables are not tolerantly tables. But if there are such worlds, then it seems plausible that some of them are very much like the actual world in lots of other ways—in particular, at some of them, we lead roughly similar lives, and hold similar beliefs on general metaphysical questions to those we in fact hold, as well as holding similar tolerance-friendly beliefs concerning the particular tables we encounter. Why not? If it is possible to make a table that isn't tolerantly a table, it doesn't take much. For example, suppose that each of Woody's legs contains 10 per cent of its atoms, so that if Woody is tolerantly a table, it must be able to swap out all the matter of any one of its legs. Suppose further that Woody could in fact have still been a table with any *three* legs different, although it could not have been a table with four legs different. Then all one would have had to do to make Woody be intolerantly a table is make Woody as a table with three of its legs different. And this seems like it should be a pretty easy thing to do, given the assumption that it is possible. Moreover, it is hard to see how doing this could require making any changes to *us* and our Tolerance-favouring patterns of modal judgement.

Thus if one takes Tolerance to be a contingent truth, one seems forced to think that one is just *lucky* not to be mistaken about it. Indeed, if the truth of Tolerance is contingent, it seems to be a rather striking metaphysical coincidence: on each and every occasion where a table was constructed, the materials selected happened not to be composed of atoms that would have made the table in question not be tolerantly a table, although it was in each case possible that this would happen. But Tolerance certainly doesn't *feel* like a risky bet in the way that this picture suggests.

We can boil this line of thought down to the three-premise argument displayed in Figure 3.1, regimenting talk of luck and risk in the language of 'easy possibility'. The argument is obviously valid, treating all the conditionals as material.

Like the Tolerance Argument which it supplements, the Security Argument is a schema. To instantiate the schema, we will need all the ingredients required for the Tolerance Argument (a choice of closeness relation, a modality to interpret 'could',

Independence If Tolerance could easily have been false, we could easily have falsely believed it.

Security If Tolerance is true, we couldn't easily have falsely believed it.

Easiness If Tolerance could have been false, it could easily have been false.

Non-contingency If Tolerance is true, it couldn't have been false.

Fig. 3.1 Security Argument.

and a singular term or predicate *K*), and in addition an interpretation for 'could easily'—which need not be the most salient or natural one—and an interpretation for 'we'—which can, but need not, include the authors of this book. And just as some Tolerance Arguments present no puzzle at all, some Security Arguments generate no interesting challenge. For example, if we made the mistake of fixating on a table made by a slow-moving carpenter from all the wood available for miles around, and interpreted 'easy' in such a way that the carpenter couldn't easily have used different wood, then there would be no plausible route from the mere possibility of Tolerance being false to the easy possibility of its being false, or of our being wrong about it. But for the Security Argument to be interesting, it is enough that there are some instantiations where there is a prima facie case for each of its premises, but where the relevant Tolerance Puzzle motivates the negation of its conclusion. And as it turns out, there are many such cases. To see this, we will go through the premises of the Security Argument one by one.

Let's begin with Independence. For many choices of closeness relation, Independence is not on reflection plausible, since the instantiation of different properties in the family would tend to make for relevant differences in our tolerance beliefs. For example, suppose *K* is 'bicycle belonging to Cian', and our family of properties comprises different "degrees of sphericity" ranging from perfect sphericity to certain quite undemanding geometric properties instantiated even by ordinary bicycles.[11] Tolerance in this instantiation says that each of Cian's bicycles could still have been a bicycle while being, say, 1 per cent more spherical. Since perfect sphericity is ancestrally close to every property in the family, Hypertolerance in this argument has the implausible consequence that each of Cian's bicycles could still be a bicycle while being perfectly spherical. But in possible worlds where one of Cian's bicycles fails to be tolerantly a bicycle, in the sense that it is a bicycle but could not be a bicycle while being even 1 per cent more spherical than it is, we would likely know about its odd shape, and

[11] An object's degree of sphericity can be identified with the maximum, for all spatial regions entirely containing it, of the fraction of the region's volume it occupies.

consequently be much less confident in the Tolerance premise than we actually are. Thus Independence has little purchase in this instantiation. Of course, we might imagine that Cian is so lax in monitoring his bicycle collection that it could easily happen that one of his bicycles was freakishly distorted without him noticing. But in that case, Security has little appeal. If we are people whose bicycles could easily get freakishly distorted without their noticing, then we shouldn't have been very confident in Tolerance in the first place. And if we were confident, and we managed to avoid error, then we *were* just lucky to do so.[12] Similar remarks apply to the Heap Argument from §3.1.

There are other kinds of Tolerance Arguments that don't feel particularly akin to the Sorites, but where it is similarly hard to make a compelling case for Independence. For example, consider the Tolerance Argument about chess that we gestured at in the Introduction, where Hypertolerance implies that chess could have been played according to the rules of Twister. To resist this, while acknowledging that chess is tolerant with respect to slight variations in its rules, we need to think that there are some fairly nearby worlds where chess is played according to some rules such that it couldn't have been played according to certain slight variants of those rules. Presumably at some such worlds the word 'chess' is still in use, and there are philosophers like us who make the speech 'Chess is tolerant with respect to slight variations in its rules.' But to conclude from this that the philosophers in question are making a mistake, we would need to assume that their word 'chess' refers to *chess*; and such homophonic translation seems rather naïve in this case. It is natural to think that, roughly speaking, our meaning of 'chess' has the actual rule-set as a paradigm, while their meaning for 'chess' instead has their variant rule-set as a paradigm. Their Tolerance speeches are true, because their word 'chess', unlike ours, applies to games played according to all close variants of their paradigm rule-set.

But Independence is much more compelling in many other cases. For example, insofar as we are convinced that the Tolerance premise that every table is tolerantly a table with respect to originating matter (or shape) could easily have been false, we will think that the worlds where it is false can differ from the actual world just by, e.g., some moderate shift in the positioning of one saw in one workshop. That doesn't *seem* like the sort of thing that would stop us from meaning *table* by 'table', or from meaning what we actually mean by any of the other words we used in formulating the relevant Tolerance premise. By contrast with the chess case, the factors that supposedly make the difference between the truth and falsity of Tolerance in the table case are not factors to which even expert users of the word

[12] Independence would also plausibly be true in this instantiation if we were the kind of people who routinely believe the Sorites Premises that the Sorites Argument shows to be false. But that belief-forming disposition clearly is one that risks error, so Security will not plausibly be true if the referent of 'we' is some group of people in the grip of bad Sorites-generating dispositions.

'table' are sensitive, or in a position to use as a basis for constructing a paradigm that shifts from world to world.[13]

Let's turn next to Security—the idea that we are not just *lucky* to be right about Tolerance. One way to motivate this premise is by way of considerations about knowledge. We seem to *know* (and not merely truly believe) that Tolerance is true. And the status of our belief as knowledge would seem to be threatened if we could easily have been wrong, thanks to a different selection of parts by some table-maker.

Admittedly, the general principle that one can't know that *p* if one could easily have been wrong about *p* is quite tendentious. Suppose you come to believe that Tim is standing by seeing him standing. Even if you could easily have come to believe that Tim was standing when he wasn't, by trusting a liar, that would not show that your actual vision-based belief fell short of knowledge. What seems germane here is the fact that the method whereby one could easily have acquired the false belief is very different from the method whereby one in fact acquires the true belief. But the table-maker case doesn't seem anything like this: our method of coming to believe Tolerance involves nothing like a careful survey of the part-choices made by all the world's table-makers, and indeed some distant table-maker's choosing different parts would intuitively make no difference at all to the method whereby we acquired our belief in Tolerance. More generally, while 'could easily' is a very flexible expression, and we certainly wouldn't want to suggest some context-invariant connection between it and 'know', the possibilities of error that seem to be presented by the table case seem exactly the kinds of possibilities of error that do threaten actual knowledge. You wouldn't know you had avoided choosing poisonous grapes if you grabbed the grapes at random from the bunch and a modest repositioning of your hand would have been enough to snare some poisoned ones. Similarly, if a modest repositioning of some distant carpenter's hand would have made for the falsity of our Tolerance beliefs, it seems ridiculous to suggest that those beliefs nevertheless in fact constitute knowledge.

Security can also be motivated in ways that don't turn at all on judgements about the epistemological status of our actual Tolerance beliefs. There is just something quite bizarre about the view that the modes of thought rooted in our ordinary modal practices that lead philosophers like us to endorse Tolerance do not in fact lead us into error, but could easily have done so. Can we really make our peace with such thoughts as 'Nathan Salmón is right about the first important premise in his argument against Iteration, but if the saw in the IKEA factory had been repositioned a few feet to the right yesterday morning, he would have been wrong about that premise'?[14]

[13] We'll be subjecting this line of reasoning to a great deal of critical scrutiny in Chapter 11.

[14] One recently influential current of thought in several domains attempts to make the prospect of widespread nearby false belief more palatable by appealing to the distinctive ideology of "relative truth"

Finally, we turn to Easiness. Its plausibility in any particular instance is of course dependent on the operative interpretations of 'could' and 'could easily'.[15] And there

as developed by, e.g., MacFarlane (2015), Egan, Hawthorne, and Weatherson (2005), Kölbel (2002), and many others. Here, for example, is the sort of thing that Relativists have wanted to say about "predicates of personal taste" such as 'tasty': 'Marmite-lovers (like us) and Marmite-haters use "Marmite is tasty" to express the same proposition. They think and say that Marmite isn't tasty when it is, and thus believe and assert something that is not the case. But since the proposition that Marmite is tasty is true *relative to them*, they still achieve the epistemically more important goal of believing and asserting something that is true relative to them. Because of this, the existence of these people who form false beliefs using similar methods to those we use to form our true beliefs shouldn't get us Marmite-lovers worried that when we think and say that Marmite is tasty, we might be believing and asserting something that is not the case. Nor should we be concerned by the fact that we ourselves would have had false (but true relative to us) beliefs if our parents hadn't exposed us to Marmite at a tender age.'

As one of us has explained at some length (Cappelen and Hawthorne 2009), we don't really understand what it could mean for a proposition to be true or false relative to a person: at least, we don't understand what it could mean in such a way that the absence of nearby possibilities where one believes things that are false relative to oneself would do anything to soothe the worries induced by the existence of nearby possibilities where one believes things that are false (i.e. that are not the case). But those who do understand this ideology and find it epistemically calming might consider deploying it to defuse the case for Security. 'If Woody had been made of somewhat different matter in such a way that Woody wasn't tolerant, we would have held the false belief that Woody was tolerant. But despite its falsehood, the proposition that Woody is tolerant would still have been *true relative to us*; so in believing it, we would have achieved the epistemically important goal that some philosophers mistakenly suppose can only be achieved by believing truths. Since we enjoy that kind of success in all the relevant nearby possibilities, the existence of abundant nearby possibilities where we believe and assert things that aren't the case shouldn't get us worried that when we utter "Woody isn't tolerant" in the actual world, we might in fact be asserting something that isn't the case.'

Murray and Wilson (2012) develop an approach to Tolerance Puzzles that seems to be in this Relativist spirit. They suggest that for "non-basic" propositions, like the proposition that Woody is tolerant, there are two dimensions of world-relativity: a proposition can be true at a world "from the perspective of" another world. The fact that Woody could easily have been intolerant only commits us to the existence of a nearby world w such that the proposition that Woody is tolerant is false at *w from the perspective of the actual world*. But the proposition is still true at w from the perspective of w. They don't explicitly discuss how this claim might help with Tolerance Puzzles like ours (which are expressed using modal operators); but they might have in view a response to Independence along the lines of the foregoing Relativist speech, interpreting 'p is true relative to x' as 'p is true from the perspective of whichever world is actualized'. However, we have some hesitation in attributing this line to Murray and Wilson: the later Hellie, Murray, and Wilson 2021 seems like a more straightforwardly *contextualist* proposal much closer to our own (see Chapter 11 and §13.1), on which 'Woody is tolerant' does not *express the same proposition* across the relevant worlds, but merely has the same "character" in the sense of Kaplan (1989).

[15] Easiness will become trivial if we interpret 'could easily' as equivalent to 'could' (whose interpretation is fixed by the particular Tolerance Argument we are concerned with). If 'could' is something broad, like metaphysical modality, this will make the argument not very interesting, since unless one were antecedently convinced that Tolerance isn't a metaphysically contingent truth, it is hard to see why one would accept Security on this interpretation. But when the relevant sense of 'could' is narrower, there may be an interesting prima facie case for Security and Independence even when the word 'easily' is deleted. For example, when the Tolerance Argument we are interested in uses a tense operator, the Security Argument may have force when 'could easily' is replaced with that same tense operator. Consider the *de re* Tolerance Argument where the modality is the future tense, the singular term is 'English', the closeness relation is the one F bears to G iff there are two adjacent years y_1 and y_2 such that F is the property of being a spoken language and such that it is y_1, and G is the property of being a spoken language and such that it is y_2. Tolerance is, then, equivalent to the claim that English will still be spoken next year, and Hypertolerance entails that English will still be spoken in a million years. Assuming that the course of linguistic evolution will remain slow and steady for the next million years, Security is appealing: given a course of history like that, it seems very strange to suppose that at some point, English-speakers will be making a mistake in assuming that English will be spoken a year later.

are plenty of interpretations on which the denial of Easiness can on reflection be made completely plausible.[16] But for many of our stereotypical Tolerance premises, such as those involving the originating matter of tables, we can readily flesh out the background facts in ways that make the denial of Easiness seem very strange. When an appropriate range of materials is ready to hand, and some collections of these materials are such that Woody could have been made of them but would in that case have been intolerant, it is hard to see how one could sensibly deny that Woody could easily have been intolerant.

Such a denial is not completely out of the question. For example, one could suppose that while it is impossible for Woody to be made with four legs different and possible for Woody to be made with three legs different (and thus intolerant), Woody could not easily have been made with three legs different. Although the carpenter could easily have picked three different legs during the process that actually resulted in Woody, what would have resulted if she had done so would not have been Woody but some other table, which would have been tolerant. On this vision, Tolerance enjoys a kind of *counterfactual* robustness despite the fact that there are metaphysically possible worlds where it fails: it would have been true no matter which collections of parts were chosen to make tables, and could not easily have been false.[17] One could take this robustness to be brute; alternatively, one one could hold that despite being metaphysically contingent, Tolerance is metaphysically necessitated by certain uncontroversially robust facts, such as facts about the laws of nature or about the distant past. We don't have much sympathy for this picture. Moreover, it naturally invites a revenge argument: even if some oddball view succeeds in undermining Non-contingency in a Tolerance Argument based on some broad modality like metaphysical or atomic possibility, the natural response is to just set those Tolerance Arguments aside and focus instead on arguments where the modality is something more restrictive. As we discussed in Chapter 2, the ordinary modal judgements that motivate Tolerance premises involving metaphysical possibility carry over to many much narrower modalities, including nomic possibility.

And Independence also has some force: if there's a last year when English is being spoken, then prima facie, people who then utter whatever sentence is then analogous to 'The language we speak will still be spoken next year' will be falling into error.

[16] In some cases Easiness isn't tempting to begin with; in others, any initial temptation can be overcome by appropriate imaginative exercises. For example, suppose we count *F* as close to *G* iff for some chemically matching collections of atoms *C* and *D* such that there are at most 100 atoms in *C* but not *D* and at most 100 atoms in *D* but not *C*, *F* is being originally composed of *C*, and *G* is being originally composed of *D*. It seems plausible that every postage stamp is tolerantly a postage stamp (in the sense defined by this notion of closeness). But it seems quite tendentious to suppose that this is necessarily true, even on the assumption that it couldn't easily have been false. For example, it is plausibly physically and sociologically possible for a fad of sending nanoscale physical letters to evolve, as part of which the postal services would go to the trouble of officially issuing stamps made of only 100 atoms in total, installing high-powered electron microscopes capable of reading such stamps in certain special post offices.

[17] Thanks to Ofra Magidor for prompting us to think about this kind of view.

The Security Argument as we have stated it depends on the reader's interpreting 'could easily' in the right spirit. The other words we have been throwing around, such as "lucky", "danger", "fragile", and "robust", provide some guidance, but are still rather impressionistic. One helpful way to sharpen things up is to replace claims about what could easily be the case with claims about *objective chance*. We could just replace 'it could easily be that *P*' in the Security Argument with 'there is a nonzero chance at *t* that *P*' for some appropriately chosen *P*. But the argument is more forceful if we instead focus on thresholds considerably higher than zero. In Figure 3.2 we present the argument using 'substantial' and 'very substantial' as labels for two relevant thresholds, one higher than the other. We can experiment with many different interpretations for these labels: for many Tolerance premises there are available interpretations that make each premise prima facie compelling.

The arguments we earlier considered for Independence, Independence, and Easiness can be adapted (and to some extent tightened up) in the chance-theoretic setting. Chance Independence can be motivated by some combination of the thoughts that (a) the chance at *t* of our believing Tolerance is sufficiently high, and (b) the question whether we believe Tolerance is (approximately) probabilistically independent of the truth of Tolerance. (b) can in turn be motivated by the claim that the factors responsible for our believing Tolerance are not sufficiently causally sensitive to the factors that determine whether Tolerance is true or false (e.g. which collections of atoms are used to make tables). If the chance of our believing Tolerance is high enough, (b) will not be needed: for example, if the chance of our believing Tolerance is at least 95 per cent and the chance of Tolerance being false is at least 30 per cent, the chance of our falsely believing Tolerance must be at least 25 per cent. Indeed, if the chance of our believing Tolerance was 1, then there would be no drop off at all between the chance of Tolerance being false and the chance of our falsely believing it. This would arguably be true, for example, if the Tolerance premise in question is quantified and eternal (like 'Every table is

Chance Independence If the chance at *t* of Tolerance being false is very substantial, the chance at *t* of our falsely believing Tolerance is substantial.

Chance Security If Tolerance is true, then at *t* the chance of our falsely believing it is not substantial.

Chance Easiness If Tolerance could have been false, the chance at *t* of its being false is very substantial.

Non-contingency If Tolerance is true, it couldn't have been false.

Fig. 3.2 Chance Security Argument.

tolerantly a table', when this is understood as 'Everything that is ever a table is tolerantly sometimes a table') and *t* is a time after we write this book. In that case, one could derive that the chance of our believing Tolerance is 1 from some general premise to the effect that for any *t*, facts entirely about times earlier than *t* have chance 1 at *t*, and facts about what people believe at a certain time—at least when the belief in question isn't about *de re* matters—count as being "entirely about" the period up to and including that time.[18] But even if our belief in the Tolerance premise isn't formed until after *t*, so long as our general philosophical sensibilities are in place by *t* and our chance of dying or undergoing a major philosophical conversion is low, it seems that in many cases the chance of our believing Tolerance will be high enough to preclude a radical drop-off.[19]

The case for Chance Security will be stronger the higher we set the threshold for 'substantial'. One way of arguing for it is by way of the claim that if Tolerance is true we know it, together with a general principle according to which a belief cannot constitute knowledge when there is a substantial chance of a relevant error. Such principles are attractive (see Dorr, Goodman, and Hawthorne 2014), although not completely uncontroversial. For example, if you think that you can know each of a thousand probabilistically independent propositions about the future each with a chance in the region of 99 per cent, and you think that deduction from multiple known premises preserves knowledge, then you will have to allow that there can be knowledge of the conjunction of the propositions despite its tiny chance of being true.[20] But one does not need to invoke knowledge to motivate Chance Security; one can also do quite a lot with some notion like reasonableness. There is something intuitively pathological about the theoretical stance of someone who is highly confident that some proposition is true, but also highly confident that there is or was a high chance of that proposition being false while they remain highly confident in it, without any relevant difference in the "method" responsible for their having that high confidence. Such a character seems too much like someone that is confident that a certain horse is going to win while also confident that the objective chance of its winning is low. But the position of someone who is confident

[18] For more on the former premise see the discussion of History Fixity in §9.1; for more on the latter, see the discussion of temporal externalism in §11.5.

[19] There is a complication when the Tolerance premise in question is a *de re* one like the proposition that Woody is tolerant. If Woody isn't created at all then this proposition is vacuously true (since no collection of atoms composes Woody); so if the chance of Woody being created at all is low, the chance of our falsely believing that *Woody* is tolerant cannot be high. Still, if thanks to our already-formed philosophical sensibilities there is a high chance that we will believe of whichever table is created that it is tolerant, and this is approximately probabilistically independent of which table is made, the chance of our falsely believing Woody to be tolerant will still be close to the chance of Woody not being tolerant.

[20] Hawthorne and Lasonen-Aarnio (2009) point out the further difficulties for principles linking knowledge and objective chance driven by the fact that knowledge is "guise-dependent" and chance isn't. Goodman and Salow (2018) (building on Dorr, Goodman, and Hawthorne 2014) show how a commitment to the KK principle—with which they express some sympathy—rules out some otherwise attractive principles linking chance to knowledge, at least given further plausible premises about the retention of knowledge over time in situations where no new evidence is acquired.

that Chance Security is false—i.e. that Tolerance is true while the chance of their falsely believing it is substantial—seems to display exactly this pathology.

Finally, even setting all epistemological ideology to one side, there is something quite unsettling in the idea that there are the kinds of thoughts we expressed in the Introduction while explaining our motivations for believing Tolerance premises were in fact true but had a high chance of being false. It just seems bizarre to take the distribution of truth and error over philosophers working on this topic to be hostage to the outcomes of chancy processes such as those in carpenters' workshops. Indeed, it even seems bizarre to think that there is *any* nonzero chance of philosophers using the same belief-forming method that actually leads to true Tolerance beliefs in almost exactly similar physical circumstances and coming up short. We can live with the idea that there is some astronomically tiny chance that the world will come to any end in five minutes, so that falsehood is rife among ordinary future-directed beliefs. By contrast, it is much stranger from a metaphysical point of view to suppose that there is even an astronomically small chance that we falsely believe some table to be tolerant despite the physical facts being much as they actually are.

To bring out the motivation for Chance Easiness, we can break it up into two steps: (i) if Tolerance could have been false, it has a nonzero chance of being false; (ii) if Tolerance has a nonzero chance of being false, it has a very substantial chance of being false. If the Tolerance claim we are interested in involves some broad modality like metaphysical or atomic possibility, (i) will be rejected by proponents of oddball views like those we mentioned when discussing the original Easiness premise: if one liked the idea that Tolerance was 'counterfactually robust' despite being metaphysically contingent, one would probably also like the idea that its chance of being false is 0. But as we have said, we don't think such views have much going for them. And in any case, we can set them aside just by focusing on Tolerance Arguments involving a narrower modality. In the present context, the obvious such modality to focus on is that of having nonzero chance: on this interpretation, 'could have' just means 'has positive chance', so (i) becomes a triviality. Such chance-theoretic Tolerance Arguments are also especially good to focus on for other reasons (for example, because there are distinctive motivations for accepting Iteration in them); Chapter 9 will be entirely devoted to them.

The motivation for (ii) is especially strong when we are dealing with quantified Tolerance Arguments, but it is also easy to bring out for many *de re* Tolerance Arguments. Here is a simple model where the chance-calculations are easy to make precise. A square paper plate measuring 10 by 10 inches is to be punched out of a rectangular strip of cardboard measuring 10 by 25 inches. At a certain time t, after the process was set in motion but before the positioning of the punch was settled, there was chance 1 that some plate would be punched, and any two equally long segments of the 15-inch central portion of the strip (excluding the first and last 5 inches) had an equal chance of having a plate centred on them. In the actual

world, the punch was centred 7 inches from the beginning of the strip, and a particular plate, Plato, was made (with a two-inch strip left over on one side and a 13-inch strip left over on the the other side). Suppose moreover that which plate is made depends entirely on where the punch that created it was centred, in the sense that the central portion of the strip can be partitioned into two regions such that the chance of Plato being made conditional on the punch being centred in the first region is 1 and the chance of Plato being made conditional on the punch being centred in the second region is 0. For concreteness, we might imagine that any punch centred within 11 inches of the beginning of the strip gets you Plato, and any punch centred further than 11 inches from the beginning of the strip gets you a plate that isn't Plato. On that assumption, if the punch had been centred anywhere between 10 and 11 inches of the beginning of the strip, Plato would have been produced but would have been intolerant in the origin-theoretic sense: there would have been a chemically matching collection of atoms—namely, those between 11 and 21 inches from the beginning of the strip—which could not have originally composed Plato despite the fact that they overlapped the atoms which in fact originally composed Plato by at least 90 per cent. So, the chance of Plato being intolerant is at least 1/15, and the chance of some intolerant plate being produced is even higher.[21]

When we turn to quantified Tolerance claims such as the claim that every table is tolerantly a table, or even that every table that will be produced in a certain factory in the next few years will be tolerant, the view that they are true and such that the chance of their being false is low but nonzero becomes even harder to maintain. Barring some mysterious, physically unrealistic correlation between the many chancy processes that go into the selection of materials for any given table-construction project, most pairs of tables should be such that processes which might have led to one or the other being intolerantly a table are approximately probabilistically independent. So if the chance of any particular table-making process issuing in a table that wasn't tolerantly a table was small but not tiny, then

[21] This doesn't depend on the simplifying assumption that all and only cuts centred at most 11 inches from the beginning of the strip would generate Plato, or even on the assumption that the set of cut-positions that generate Plato is topologically connected. Let n be the least upper bound of the set of cut-positions that get you Plato. Since Plato can tolerate replacement of at least 10 per cent of its originating atoms, $n \geq 7.5$. Since Plato could not be created from entirely different atoms, $n < 17$. If the cut had been made anywhere x inches from the beginning of the strip, where $n-1 < x < n+1$, whichever plate would have been created would have been intolerant. If it was Plato, then it would have been intolerant since Plato wouldn't have been made if the cut had been made $x + 1$ inches from the beginning. And if it was not Plato, it would have also been intolerant, since there is a possible Plato-generating cut-position in the interval $(x - 1, n]$ and hence in the interval $(x - 1, x + 1)$. The objective chance of the cut being made in this 2-inch open interval is 2/15.

Note that in this model there is also a substantial chance of an intolerant plate being produced that is not Plato. Assuming that there is a high chance that we will believe that whichever plate is produced is tolerant, this will mean that there is also a substantial chance of our having false tolerance beliefs about a plate other than Plato. These other false beliefs will make the denial of Chance Security even harder to live with.

the chance that some table-making process or other would issue in such a table would still be close to 1. To keep the latter chance low, the chance of each particular table being intolerantly a table would have to be *incredibly* small. It is completely mysterious why that should be so.

The chance-based sharpening of the Security Argument confirms our suspicion that there is a strong prima facie case for Non-contingency based on the thought that our Tolerance judgements are secure. The Security Argument for Non-contingency has turned out to be more promising than the motivations that we looked at in §3.2 involving concepts like 'a priori knowledge', 'specialness', and the like. And it marks a clear distinction between Tolerance Puzzles and tempting but certainly bad Sorites paradoxes.

That said, the Security Argument is not watertight. In Chapter 11 we will be presenting some ideas which provide a principled way of resisting the crucial Independence premises. Still, the force of the argument is certainly strong enough to motivate a careful look at the other possible responses to Tolerance Puzzles, namely accepting Hypertolerance and rejecting Iteration. That will be the task of Chapters 5–9. But first, Chapter 4 will introduce a new family of puzzles that, as well as being interesting in its own right, is connected to Tolerance Puzzles in several instructive ways.

4
Coincidence Puzzles

Are distinct material objects ever in exactly the same place? On first encountering the question there is some temptation to think not. But there are well-known arguments that such "coincidence" is in fact quite routine. Every clay statue coincides with a lump of clay; but no ordinary clay statue *is* a lump of clay, since any lump of clay has properties which no ordinary statue has. The distinguishing properties include modal properties like the property of possibly being ball-shaped, and in most cases also temporal ones like the property of having once been ball-shaped.[1]

These arguments for pervasive coincidence are sometimes presented as paradoxes, on the grounds that while the premises seem true, the conclusion seems false.[2] But it seems to us that whatever mild prima facie appeal the 'no coincidence' thesis might have, it is so completely swamped by the arguments against it as to make words like 'paradox' quite inappropriate. Whatever explains the initial temptation to think that distinct objects never coincide, the disposition does not seem to have any deep roots in ordinary practice.[3]

[1] See Wiggins 1968, Doepke 1982, and Thomson 1998. Fine (2003) points out that the case can also be made using only properties like *being well-made* whose expression involves neither modal nor temporal vocabulary.

[2] For this framing, see Wasserman 2018. Locke (1689: ch. 27) avows that it is 'impossible for two things of the same kind to be or exist in the very same instant, in the very same place', on the grounds that to deny this 'takes away the distinction of identity and diversity . . . and renders it ridiculous'. (Since he recognizes just three kinds—God, finite intelligences, and bodies—it is clear in context that statues and lumps would both be of the kind 'bodies'.) Descartes (1644: 2.10–15) uses anti-coincidence for denying even the distinction between a body and the space it occupies. Van Inwagen (1981: 81)—deploying one of his signature rhetorical moves—claims that if someone posited coincidence, 'I should not understand him and I suspect that no-one else would either.' Some authors limit their abhorrence to certain cases of *permanent* (mereological) coincidence—which Lewis (1986*a*: 252) calls 'absurd on its face', and Noonan (1988) diagnoses as 'a bad case of double vision'—though this reaction is surely channelling some Relativity-influenced assimilation of time and space, rather than any ordinary sensibility.

[3] We are here referring only to the idea that the coincidence thesis is implausible or counterintuitive in its own right, not high-handedly dismissing any philosophically substantial *arguments* that opponents of coincidence might come up with. The most influential such argument (see Burke 1992; Olson 2001; Bennett 2004) involves theses about *explanation*. The idea is that when objects differ in the ways in which believers in coincidence think coincident objects differ—e.g. in whether they are statues or lumps—there must be some explanation of the difference, where the similarities forced by coincidence are supposed to make such an explanation impossible or challenging. For what it's worth, we are unmoved: from our point of view, 'How can these objects differ in their modal properties given that they are coincident?' is no more gripping than 'How can these *properties* differ in their modal properties given that they are coextensive?' Chapters 7 and 11 will discuss some ways of thinking of non-fundamental objects that make this analogy particularly forceful.

The Bounds of Possibility: Puzzles of Modal Variation. Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri, Oxford University Press. © Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri 2021.
DOI: 10.1093/oso/9780192846655.003.0005

By contrast, our ordinary practices do seem to take a quite firm stand against the claim that it is routine for two distinct statues, or lumps of clay, or tables, or buckets, or boats… to be in exactly the same place, or even in substantially overlapping places. For we are often quite confident that there are only two tables in a certain room, only one statue on a certain plinth, only seven ships in a certain harbour, etc. In forming such everyday counting judgements, we apparently take it for granted that the kinds of objects in question don't have coinciding instances in the case at hand. True, there may be some esoteric cases where we would be open to positing coincident instances of certain familiar kinds—for example, Fine (2000) argues that there could be two coincident *letters* if they were written on opposite sides of a single sheet of paper, and Johnston (2006) describes what is arguably a case of coincident *road signs*.[4] But those arguments do not generalize in any obvious way to other kinds like tables, buckets, and boats.[5] And in any case, they do nothing to suggest that such coincidence is anything but exceptional. And while philosophers occasionally proceed as if rampant coincidence of *objects* would bring with it rampant coincidence within familiar kinds, it is obscure how such an inference could be justified.[6]

This chapter will introduce a family of puzzles that are much more challenging than standard so-called "paradoxes of coincidence", since they threaten not just the ill-considered judgement that objects never coincide at all, but the far more entrenched judgements to the effect that tables, buckets, and so on do not coincide in ordinary circumstances. Such "Coincidence Puzzles" have come up from time to time in discussions of modal variation. It will be instructive to consider them alongside the Tolerance Puzzles that are the primary topic of this book: some of the strategies that might help with Tolerance Puzzles extend naturally to Coincidence Puzzles, and vice versa, whereas other strategies for dealing with Tolerance Puzzles, including Iteration-denial and Hypertolerance-acceptance, do not suggest any particular approach to Coincidence Puzzles, and may look less plausible when paired with certain treatments of Coincidence Puzzles.

[4] Wiggins (1968), by contrast, takes a harder line. He thinks that there is a crucially important notion of "kind" on which coincident objects cannot belong to the same kind, and we suspect that he would take expressions like 'letter', 'road sign', and 'sculpture' to express kinds. When we use the word 'kind', we are not trying to channel the distinctive doctrines that philosophers like Wiggins have associated with words like 'kind' and 'sortal'.

[5] That is not to say one can't contrive somewhat plausible examples of coincident boats. We can imagine a cleverly designed folding boat that can be either partially unfolded to form a small canoe, or fully unfolded to form a large lifeboat.

[6] For example, Rea (1995: 535) writes that 'To reject the Identity Assumption [i.e. the thesis that distinct objects cannot coincide] is to admit that there might have been two (or more) co-located ships.' Kment (2018: 660) likewise seems to assume that if a certain principle guaranteeing that every object coincides with a great many other objects is true, then every *table* coincides with many other tables; Leslie (2011) mostly seems to assume the same thing about *axes*, though a couple of parenthetical remarks suggest she is not fully committed to this (see note 26 below).

4.1 Coincidence Puzzles

Let's return to the case of Flimsy, the plastic bucket we met in §2.1. Suppose Flimsy was made by a machine that pours molten plastic into a bucket-shaped mould, spins the mould so that the plastic coats its interior, waits for it to harden, and then ejects the resulting bucket. The machine can make thinner- or thicker-walled buckets depending on how much plastic is put in. It seems plain that Flimsy could have been thicker—indeed twice as thick—by the addition of more plastic. (Consider counterfactuals like 'That bucket wouldn't have collapsed if it had been made twice as thick.') Moreover, it seems quite implausible to suppose that among the portions of plastic of the right size and type (at least among those in the bucket factory), only some are such that if they had been added to the machine along with the portion of plastic that in fact went into Flimsy, Flimsy would have then been a thick-walled bucket. ('If we had added any of those portions of plastic this bucket would have been more useful and robust, though if we had instead added sand, it would have been so fragile as to be completely useless.') And the same goes for all of the other flimsy buckets made by the machine. But now a puzzle looms. Consider another of the buckets—call it Frail. Suppose Frail was made next after Flimsy, out of a portion of plastic *B*, while Flimsy was made from *A*. Given what we have said, if *B* had been added to the mould along with *A*, Flimsy would have been a thick-walled bucket made of *A* and *B*. By parity of reasoning, if *A* had been added to the mould along with *B*, Frail would have been a thick-walled bucket made of *A* and *B*. But if we assume the necessity of distinctness, it follows from what we have said that if *A* and *B* had been melted together, Flimsy and Frail would have been two distinct buckets both made of the exact same portion of plastic, and thus in exactly the same place. And that is quite odd, because we seem to assume that buckets do not ordinarily coincide with other buckets distinct from them, as for example when we say things like 'There is only one bucket in the closet', never even considering the possibility that we bought several coincident buckets for the price of one.

We can lay out the paradox as an argument from individually plausible premises to an absurd conclusion:

Resiliency 1. If *A* and *B* had been combined, Flimsy would be a bucket made of *A* and *B*.

2. If *A* and *B* had been combined, Frail would be a bucket made of *A* and *B*.

Non-coincidence If *A* and *B* had been combined, there would be at most one bucket made of *A* and *B*.

Distinctness If *A* and *B* had been combined, Flimsy would not be identical to Frail.

Vacuity If *A* and *B* had been combined, Flimsy would be both identical to and not identical to Frail.

The argument is valid given a very basic logic for counterfactuals: all we need is that the result of embedding the premises and conclusion of a classically valid argument under the same counterfactual antecedent is truth-preserving. Indeed, since everything follows from a contradiction, the argument would still be valid in this logic if the consequent of Vacuity were replaced by any other sentence.

We have already tried to bring out why the Resiliency premises and Non-coincidence are plausible. Distinctness can be motivated by appeal to the following widely accepted general principle about counterfactuals:

Counterfactual ND If $x \neq y$, then if P it would be that $x \neq y$.

Weaker principles would also suffice, as we will discuss later. The implausibility of Vacuity, meanwhile, should be evident. If an argument against Vacuity is required, one could appeal to the premise that it is metaphysically possible for A and B to be combined, together with the principle that counterfactuals transmit metaphysical possibility:

Possibility Preservation If it is metaphysically possible that P, and if P it would be that Q, then it is metaphysically possible that Q.[7]

For by the basic modal logic, it is not metaphysically possible that Flimsy is both identical to and not identical to Frail; so given Possibility Preservation, Vacuity could not be true unless its antecedent were metaphysically impossible.

Possibility Preservation is controversial: some theorists will deny that its instances are true in all contexts, namely those who hold that counterfactuals are context-sensitive in such a way that in a typical context of use, various true, metaphysically contingent propositions are "held fixed" (in the sense that any counterfactual whose consequent expresses one of these propositions is automatically true in the context).[8] But these theorists also hold that uttering a counterfactual with a metaphysically possible antecedent has a strong tendency to shift the context to one in which the negation of that antecedent is *not* "held fixed" in that sense.[9] If so, then even if Vacuity is true in some contexts, it will not be true in the contexts it naturally evokes, or in the contexts naturally evoked by any of the premises of the Coincidence Argument.

[7] See Williamson 2007:156. Other proponents of this principle include Stalnaker (1968), Lewis (1973), and Kment (2014). Nolan (1997), by contrast, suggests that Possibility Preservation may fail in some contexts (those where his 'Strangeness of Impossibility Condition' is not adhered to). However Nolan would also deny that counterfactuals conform in those contexts to our basic modal logic; so if one wanted to appeal to his views as a solution to Coincidence Puzzles, it would more plausibly involve accepting all four premises while still denying Vacuity.

[8] See von Fintel 2001 and Williamson 2020.

[9] This tendency might be explained in terms of the idea that 'If P it would be that Q' *presupposes* the falsity of 'If P it would be that not-Q': see von Fintel 2001.

Resiliency	1. $\forall x(Gx \rightarrow \Box Fx)$ 2. $\forall y(Hy \rightarrow \Box Fy)$
Non-Coincidence	$\Box\forall x\forall y((Fx \wedge Fy) \rightarrow x = y)$.
Distinctness	$\exists x\exists y(Gx \wedge Hy \wedge \Box x \neq y)$
Vacuity	$\exists x\exists y\Box(x = y \wedge x \neq y)$

Fig. 4.1 Coincidence Argument.

The above Coincidence Argument can be seen as an instantiation of a general schema, which is given in Figure 4.1. For the sake of generality, we have formalized it in a way that doesn't require us to introduce proper names like 'Flimsy' and 'Frail', but also lets us pick out the relevant objects using predicates (e.g. 'bucket created first' and 'bucket created second'); this also screens off diagnoses that turn on special behaviour in proper names. We recover the above instantiation by setting $\Box$ as 'If A and B had been combined it would have been that . . . '; G as 'identical to Flimsy'; H as 'identical to Frail'; and F as 'bucket originally made of A and B'. But the argument is valid for any predicates F, G, H, and operator $\Box$, so long as $\Box$ obeys our basic modal logic.[10]

Here is another puzzling instance of the schema. Suppose we have a Y-shaped network of roads: one road goes from village A to village B and then on to village C, and another goes from A to B (overlapping in that segment with the first road) and then on to village D. If we focus on the road from A to C, it is tempting to think that if no roads to C or D had been constructed, this road would have been shorter, running just from A to B. Similarly, if we focus on the road from A to D, it is tempting to think that if no roads to C or D had been constructed, *it* would have just run from A to B. It seems plausible that if no roads to C or D had been constructed, there would only have been one road running from A to B. And finally, since the roads from A to C and from A to D are not in fact identical, if we endorse Counterfactual ND we have to think that they would still not have been identical if no roads to C or D had been constructed. But these four claims jointly imply the absurd conclusion that there are two things that would have been both identical and distinct if no roads to C or D had been constructed.

$\Box$ need not be defined in terms of a counterfactual conditional. A different way of filling in the schema that is also worth considering takes $\Box$ to be 'It is metaphysically necessary that if P then . . . ', for some appropriately chosen P. For example, we could replace all the counterfactuals in the argument about the

[10] Indeed, we don't need very much of the basic modal logic: all we need are the weakenings of $\Box$*N and $\Box$*K that drop the initial $\Box$*.

buckets with strict conditionals of the form 'Necessarily if *A* and *B* are combined, then . . . ', or perhaps 'Necessarily if *A* and *B* are combined at such-and-such time, in a machine of such-and-such sort, operated by so-and-so, then . . . '. However, the Resiliency premises in these strict-conditional variants of the argument do not enjoy nearly as much pre-theoretical support as the original counterfactual versions. Claims of metaphysical necessity are, in general, less central to our everyday practices and more in need of argument than counterfactuals. Moreover, once one has noticed the apparent counterexamples to the metaphysical necessity claims involving simple choices of *P*, it is hard to muster much confidence that more complex choices of *P* will avoid all counterexamples.[11]

Another family of instantiations takes □ to be 'The objective chance at *t* of . . . conditional on *P* is 1', for an appropriately chosen *t* and *P*.[12] In the case of Flimsy and Frail, *P* could just be '*A* and *B* are combined', and *t* could be some time shortly before the buckets were actually made, when there was still a nonzero chance that *A* and *B* would be combined rather than being made into separate buckets as they actually were. These Resiliency premises are somewhat less tendentious than the corresponding metaphysical necessity claims, though they are still not *that* hard to deny, when we bear in mind the vast array of bizarre outcomes that turn out to have nonzero chances according to our best physical theories. The prima facie case for them becomes a lot stronger if we weaken 'conditional chance 1' to 'high conditional chance', say 0.9 or more. Of course, *there is a high conditional chance at t that . . .* does *not* obey the basic modal logic: two propositions can have a high chance conditional on some third proposition although their conjunction does not. But we can still run a modified version of the argument. According to standard probability theory, when each premise of some valid four-premise argument has chance at least 0.9 (conditional on some given proposition), its conclusion has chance at least 0.6 (conditional on that same proposition). Given this, we can construct a valid variant of the Coincidence Argument where each premise is of the form 'The chance at *t*, conditional on *A* and *B* being combined, that . . . is at least 0.9' and the conclusion is of 'The chance at *t*, conditional on *A* and *B* being combined, that Flimsy is both identical to and not identical to Frail is at least 0.6'. But the conclusion of this argument conflicts with the principle that for any *p* the chance that $p \wedge \neg p$ is 0, which follows immediately from basic rules of probability theory.

Finally, there are also interesting Coincidence Arguments where □ is a temporal operator of the form 'At such and such past or future time it was/will be the case that . . . '. For example: suppose that at present a small forest, Sherwood Forest, is part of a large forest, the Boreal Forest.[13] Suppose that the history is as follows:

[11] Chapter 5 will consider such 'sufficiency of origin' principles and the counterexamples they face in more detail.

[12] The conditional chance of *q* on *p* can be equated with the result of dividing the chance of $p \wedge q$ by that of *p*; we need not concern ourselves with what happens when the chance of *p* is 0.

[13] Forests can have other forests as parts—e.g. the Abernethy Forest was once part of the Caledonian Forest.

a thousand years ago, only the area inside the current boundaries of Sherwood Forest was forested; since then, the trees have gradually spread outward to cover the current area of the Boreal Forest. Given this history, each premise of the following Coincidence Argument seems prima facie compelling:

Forest Resiliency 1. A thousand years ago, Sherwood Forest was a forest occupying exactly the area it currently occupies.
2. A thousand years ago, the Boreal Forest was a forest occupying exactly the area currently occupied by Sherwood Forest.

Forest Non-coincidence A thousand years ago, there was at most one forest exactly occupying the area currently occupied by Sherwood Forest.

Forest Distinctness A thousand years ago, Sherwood Forest was not identical to the Boreal Forest.

Forest Vacuity There are two things that were both identical and not identical a thousand years ago.

The Ship of Theseus case, as elaborated by Hobbes (1655: ch. 11), can also be regimented as a temporal Coincidence Argument. Hobbes asks us to consider what would have happened 'if some Man had kept the Old Planks as they were taken out, and by putting them afterwards together in the same order, had again made a Ship of them'. At the end of Hobbes's thought-experiment, we have two ships: one in the harbour, connected to Theseus by a history of gradual plank-replacements, and another, let's say, in a collector's museum, assembled from planks that were once assembled shipwise and carrying Theseus. With this history in view, we can set up a Coincidence Argument as follows:

Ship Resiliency 1. Every ship in the harbour was a ship captained by Theseus when he was king.
2. Every ship in the museum was a ship captained by Theseus when he was king.

Ship Non-coincidence When Theseus was king, he only captained one ship.

Ship Distinctness There are are two ships, one in the harbour and one in the museum, which have never been identical.

Ship Vacuity There are two things such that when Theseus was king, they were both identical and not identical.

The conclusion is absurd, but each premise seems plausible in isolation.[14] The literature on personal identity (Parfit 1971; Lewis 1976) and material

[14] Hughes (1997), who denies Ship Non-coincidence, describes an ingenious temporally inverted variant of the example where the other premises are arguably a bit harder to deny. Chandler (1975)

constitution (van Inwagen 1981; Burke 1994) contains many other examples that can be regimented as temporal Coincidence Puzzles, although the plausibility of the premises in these different instantiations varies widely.[15]

Whereas the Tolerance premises that drive Tolerance Arguments are just about *possibility* of things being different from how they actually are in certain respects, the Resiliency premises that drive Coincidence Puzzles are stronger than mere possibility claims.[16] One might wonder whether one could generate a puzzle just from the possibility claims that each of Flimsy and Frail is such that it *could* have been a bucket made of *A* and *B*. But it is hard to see how. Even if it is impossible for distinct buckets to coincide, and impossible for buckets that are in fact distinct to be identical, one might perfectly well think that among the possible worlds where *A* and *B* are combined, there are some where the fat bucket they compose is Flimsy and some where it is Frail (and perhaps others where it is neither). There is admittedly something rather weird about this view, over and above its rejection of the Resiliency premises: one wants to ask what *makes the difference* between the

responds to the temporal puzzle by denying Ship Resiliency 2 (with the qualifying remark "it is entirely reasonable to decide to talk in the following way"), but presents an interesting follow-up modal puzzle, which can also be regimented as a Coincidence Argument. In the follow-up puzzle, we consider what would have happened had there been no replacement of planks, so that at the end the original planks were assembled ship-wise in the museum and there were no relevant planks in the harbour:

Modal Ship Resiliency
1. Every ship in the harbour would have been in the museum if there had been no plank-replacement.
2. Every ship in the museum would (still) have been in the museum if there had been no plank-replacement.

Modal Ship Non-coincidence There would have only been one ship in the museum if there had been no plank-replacement.

Modal Ship Distinctness There are are two ships, one in the harbour and one in the museum, which would not be identical if there had been no plank-replacement.

Modal Ship Vacuity There are two things which would have been both identical and not identical if there had been no plank-replacement.

Chandler sets up his puzzle using proper names, and bills it as a counterexample to the theory of rigid designation; but as the above quantified regimentation makes clear, claims about the semantics of proper names will not get to the heart of the puzzle. Salmón (Salmon 1979, 1981) denies Modal Ship Resiliency 2, suggesting that in taking it for granted, Chandler is guilty of "unconscious and uncritical reliance" on a certain sufficiency-of-origins principle (which we will discuss in Chapter 5 under the name "Plan Sufficiency"). But this seems an unlikely diagnosis to us, since that principle would also imply Ship Resiliency 2 in the temporal case, which Chandler rejects. We think Modal Ship Resiliency 2 is pretty compelling on its face—'This ship here in the museum would never have been made if they hadn't replaced the planks with new ones' is a really strange speech! We are sceptical that any judgements in this area are driven by tacit reliance on sufficiency principles.

[15] In some of the instances, such as those discussed by van Inwagen (1981) and Burke (1994), the only motivation offered for the Non-coincidence premise comes from the thesis that no two distinct *material objects* ever coincide at a time, a thesis which we take to be refuted by the fact that statues sometimes coincide with lumps of clay.

[16] Of course, there is no general implication from 'If *P* it would be that *Q*' to 'Possibly *Q*': given the very plausible principle that 'If *P* it would be that *P*' is true for any *P*, 'If $(P \wedge \neg P)$ it would be that $(P \wedge \neg P)$ is true', although 'Possibly $(P \wedge \neg P)$' is false. However, given Possibility Preservation and the uncontroversial metaphysical possibility of *A* and *B* being combined, the truth of the Resiliency premises guarantees the metaphysical possibility of their consequents.

possible situations where the bucket made of *A* and *B* is Flimsy and the ones where it is Frail, and the proponent of the view doesn't seem to have a satisfying answer. In §2.1 we noticed that a similar 'what makes the difference?' intuition seemed to lie behind the intuitive repugnance of certain Hypertolerance claims, such as the claim that Flimsy could have been made of *B* while a different bucket (e.g. Frail) was made of *A*. Chapters 5 and 6 will consider how such reactions might be put on a firmer theoretical footing. But for now, it suffices to note that the Resiliency premises have plenty of intuitive plausibgility in their own right, quite apart from any argument based on the corresponding possibility claims.

4.2 Denying Distinctness

Let's now survey the range of possible ways out of our sample Coincidence Puzzle. The most radical option of all would be to accept Vacuity, maintaining that all propositions whatsoever would have been true if *A* and *B* were combined. But this involves a level of error theory far beyond what we can stomach, and has no obvious advantages; let's set it aside. The next most radical option is to deny Distinctness while endorsing Resiliency and Non-coincidence, thus accepting their logical consequence, Identity:

Identity If *A* and *B* had been combined, Flimsy would have been identical to Frail.

We have already noted that this option requires rejecting Counterfactual ND. It also plausibly requires giving up the Necessity of Distinctness for metaphysical necessity:

ND $\forall x \forall y (x \neq y \rightarrow \Box x \neq y)$

This inference can be justified by appeal to Possibility Preservation: since the antecedent of Identity is uncontroversially metaphysically possible, Possibility Preservation implies that Identity can be true only if its consequent is metaphysically possible as well (something ruled out by ND). It could also be justified by appeal to the principle that counterfactuals with metaphysically necessary consequents are invariably true (dubbed 'Necessity' in Williamson 2007: 156), or to the less controversial principle that counterfactuals with metaphysically possible antecedents and metaphysically necessary consequents are invariably true.[17]

[17] Stalnaker (1968), Lewis (1973), and Williamson (2007) accept the stronger principle. Kment (2014) accepts the weaker but not the stronger. Nolan (1997) suggests that even the weaker principle fails in some contexts.

ND is not part of our basic modal logic, by contrast with the Necessity of Identity, which (as we noted in §1.4) follows immediately by Leibniz's Law from the thesis that everything is necessarily self-identical:

NI $\forall x \forall y(x = y \rightarrow \Box x = y)$

Interpreting $\Box$ as 'If P it would be that . . . ', this becomes

Counterfactual NI If $x = y$, then if P it would be that $x = y$

which likewise follows by Leibniz's Law from the thesis that everything would have been self-identical if P.

Someone might think that there is no interest in a package that resolves our Coincidence Puzzles by appealing to failures of ND but retains NI, based on putatively analogous puzzles for which the analogous solution would seem to require rejection of NI.[18] Suppose that *in fact*, a thick-walled bucket was made by combining two pieces of plastic, A and B. Let's introduce the names 'A-bucket' and 'B-bucket' by saying 'Let A-bucket be the thing that would have been a bucket made of A if A and B had been made into separate buckets, and let B-bucket be the thing that would have been a bucket made of B if A and B had been made into separate buckets.' Then we could argue as follows: A-bucket is made of A and B; B-bucket is made of A and B; only one bucket is made of A and B; therefore A-bucket is B-bucket, so by Counterfactual NI, A-bucket would have been B-bucket if A and B had been made into separate buckets. But if A and B had been made into separate buckets, A-bucket would have been made of A and not B, while B-bucket would have been made of B and not A, so A-bucket would not have been B-bucket. So, a contradiction would have been true if A and B had been made into separate buckets.

In response: it seems to us that the first two premises of this argument are considerably less immediately compelling than the Resiliency premises of our original Coincidence Puzzle. Names introduced by counterfactual descriptive reference-fixing are an esoteric phenomenon, and judgements expressed using such names warrant a level of caution well beyond what is required for judgements involving names introduced in the ordinary way. Moreover, names for buckets weren't really crucial to the original Coincidence Puzzle, since we could unproblematically replace them with quantified premises (as we did when stating the formalized version of the argument). But when we try to construct an analogous quantified

[18] Chandler (1975) focuses on schematic rather than quantified versions of ND and NI, and suggests that the same example can be used to undermine both depending on which world one takes as actual. We have tried to be more explicit than he is about how the two names are supposed to be introduced when the actual world is the one-object world.

argument against Counterfactual NI, we will end up with something like the following: 'There is a bucket made of A and B that would have been made of A if A and B had been used to make separate buckets; there is a bucket made of A and B that would have been composed of B if A and B had been used to make separate buckets; but there is only one bucket made of A and B; so there are identical buckets that would have been distinct if A and B had been used to make separate buckets.' But Unlike our quantified Resiliency premises, the first two premises of this argument don't even seem initially gripping.[19]

The ND-rejecting strategy thus cannot be easily dismissed on the grounds that it would need to be packaged with NI-rejection in order to provide a sufficiently general treatment of the full range of relevant Puzzles. So let's turn to a consideration of the arguments that might be given for ND.

First: it is often suggested that there is some way of arguing from the *transitivity* of identity to ND (as well as NI). In a much-discussed footnote we will return to in the next chapter, Kripke says as much, offering the following argument for ND:

> Suppose $X \neq Y$; if X and Y were both identical to some object Z in another possible world, then $X = Z$, $Y = Z$, hence $X = Y$. (Kripke 1972: n. 56)

This argument seems fallacious: clearly the transitivity of identity cannot be enough to derive ND, since there are many transitive relations—e.g. *being the same height as*—that can fail to hold between two objects although they possibly hold between those objects. If we replace identity with *being the same height* in the above argument, we can see what goes wrong. The most that follows from the supposition that it is possible for X and Y to be the same height is that there is some object Z and world w such that *at* w X and Z are the same height and X and Y are the same height. But, uncontroversially, being the same height as something at a world does not imply being the same height as it. (Being the same height as someone *in Australia* does indeed imply being the same height as that person; but *pace* (Lewis 1986*a*: §1.1), 'at w' does not work like 'in Australia'; its logic is more like that of 'according to such-and-such story'.) Likewise, the most that follows from the supposition that possibly $X = Y$ is that there is some object Z and possible world w such that *at* w $X = Z$ and $Y = Z$. (X and Y themselves are both such objects Z.) But according to deniers of ND, this can be true without there being any Z such that $X = Z$ and $Y = Z$.

Second, as Prior (1963) showed, and as Kripke notes immediately after the sentence quoted above, there is an argument for ND from the B axiom of propositional modal logic, which says that anything that is possibly necessarily the case is the

[19] Of course, you could try to talk yourself into them, using arguments that deploy names introduced by counterfactual-involving reference-fixing descriptions. But as we have emphasized, such arguments should be treated with extreme caution.

case ($\forall p(\Diamond\Box p \rightarrow p)$). (We already mentioned this argument in §1.4.) Given the validity of Leibniz's Law, we have that $\Box(x = y \rightarrow (\Box x = x \rightarrow \Box x = y))$; given the necessary reflexivity of identity, this is equivalent to $\Box(x = y \rightarrow \Box(x = y))$, hence $\Diamond x = y \rightarrow \Diamond\Box(x = y)$. But by the B axiom, $\Diamond\Box(x = y) \rightarrow x = y$, so we get $\Diamond x = y \rightarrow x = y$, or equivalently, $x \neq y \rightarrow \Box x \neq y$.[20] However, while the B axiom is orthodox, its status as orthodoxy does not seem to have been earned by anyone's having given any good arguments for it.[21] So it is unclear whether there would be any great theoretical cost if one were to give it up as part of a strategy for resolving Coincidence Puzzles.[22]

The situation in tense logic is different: the versions of ND for the future tense operators (G and F) and the past tense operators (H and P) can be derived from certain much more central and hard-to-deny principles of tense logic, namely the "mixing" axioms, according to which whatever is the case will always have been the case ($p \rightarrow \mathsf{GP}p$) and has always been going to be the case ($p \rightarrow \mathsf{HF}p$). In the future direction: $x = y \rightarrow (\mathsf{H}x = x \rightarrow \mathsf{H}x = y)$ is valid by Leibniz's Law, so $x = y \rightarrow \mathsf{H}x = y$ is valid by necessitation for H (the 'always has been' operator) H, so $\mathsf{G}(x = y \rightarrow \mathsf{H}x = y)$ is valid by necessitation for G (the 'always will be' operator), so $\mathsf{F}(x = y) \rightarrow \mathsf{FH}x = y$ by the temporal analogue of $\Box(p \rightarrow q) \rightarrow (\Diamond p \rightarrow \Diamond q)$, and thus $\mathsf{F}(x = y) \rightarrow x = y$, or equivalently $x \neq y \rightarrow \mathsf{G}x \neq y$. The mixing axioms play a very central role in much ordinary reasoning about time: e.g. from 'John lives in LA, used to live in Oxford, and will live in Melbourne' we will infer 'There will be a former inhabitant of LA living in Melbourne' and 'There has been a future inhabitant of LA living in Oxford'. Given this, the option of denying Distinctness is particularly implausible for temporal Coincidence Puzzles. It is thus a cost of the Distinctness-denying approach to the modal puzzles that it will either have to embrace this implausible view (and so deny the compelling mixing axioms), or provide surprisingly different solutions to the temporal and modal puzzles. But this point is not decisive against the Distinctness-denying approach to the modal Coincidence Puzzles, since the case for the remaining premises seems stronger in the most forceful modal puzzles than in any of the temporal instantiations we have managed to contrive (such as the forest puzzle).

[20] Given the doctrine of Intensionalism (see §1.4), the B axiom not only implies but is *equivalent* to (the type-$\langle\rangle$ instance of) ND. For suppose p; then since $\neg(p \leftrightarrow (p \wedge \neg p))$, $p \neq (p \wedge \neg p)$ by LL, so by ND, $\Box(p \neq (p \wedge \neg p))$, so by then necessitation of Intensionalism, $\Box\neg\Box(p \leftrightarrow (p \wedge \neg p))$; this implies $\Box\neg\Box\neg p$ by substitution of logical equivalents, and hence $\Box\Diamond p$ by $\Box$Duality.

[21] For example, Plantinga (1974: 54), while working towards an argument for the existence of God in which the 5 axiom ($\forall p(\Diamond\Box p \rightarrow \Box p)$) plays a crucial role, contents himself with setting aside some bad objections and then remarking 'I think we can see that' it is true. Note that the 5 axiom implies the B axiom given the uncontroversial T axiom ($\forall p(\Box p \rightarrow p)$).

[22] Salmón (Salmon 1989*b*) says that the B axiom 'may well be necessarily true', though it 'does not seem logically true'. As noted in §1.1, we prefer to avoid the invocation of categories like that of logical truth. Anyway, for the purposes of arguing for the *truth* of Distinctness, one only needs the *truth* of B. §8.2 will introduce a strategy for arguing for B which we take quite seriously, but which goes by way of ND and is thus not dialectically useful in arguing for ND.

Third, quite independent of these rather theoretical arguments, there is a pretty straightforward case for Counterfactual ND based on ordinary judgements involving counting. Consider such speeches as the following:

> Peter has three loaves of bread that he would have brought if there had been a party, and Matthew has two loaves of bread that that he would have brought if there had been a party, therefore if there had been a party, then at least five loaves of bread would have been brought.

This seems like an impeccable piece of reasoning: we do not seem to need to control for the possibility that two of the relevant loaves would have been identical if there had been a party.

Someone might respond that such possibilities are real but too far-fetched to be worth bothering with in ordinary reasoning, so that although Counterfactual ND is not true in full generality, it is reliable when it comes to the kinds of counterfactual antecedents that come up in ordinary life. (By analogy, many philosophers have noted that our ordinary reasoning treats laws of nature as counterfactually robust, but few nowadays would endorse the claim that when p is a law of nature, 'If q it would be that p' is true for absolutely any q.) But even if one resists the argument for Counterfactual ND by playing this card, it doesn't seem adequate as a defence of the Distinctness-denying approach to Coincidence Puzzles. For according to that approach, the kind of reasoning illustrated by the above speech will be unreliable even when we are dealing with quite ordinary counterfactual suppositions. Perhaps, for example, being invited to the party would have caused Matthew to throw all of his dough into one big pan, with the result that the two loaves he actually has would have been identical, just like Flimsy and Frail. Thus the view that Counterfactual ND fails in such everyday cases as that of our buckets seems to lead to a rather offputting sort of error theory with regard to our ordinary counterfactual reasoning.

Fourth, Williamson (1996) has given an important argument for ND, which can also be run directly in support of Counterfactual ND, from considerations relating to the logic of 'actually' (or so-called 'modal anaphora'). Here is the intuitive idea, applied to our present example. Let's imagine that Flimsy is in fact made of dark-blue plastic, and Frail is made of light-blue plastic, while if A and B had been combined, any resulting buckets would have been of an intermediate shade. Suppose for *reductio* that if A and B had been combined, Frail and Flimsy would both have been buckets composed of A and B and would have been identical. Given that Frail is in fact light blue and would have been medium blue if A and B had been combined, Frail would have been darker than it actually is if A and B had been combined. By the necessitation (or "counterfactualization") of Leibniz's Law, it follows from this, together with the hypothesis that if A and B had been combined, Flimsy and Frail would have been identical, that if A and B had been

combined, *Flimsy* would have been darker than it actually is. But this conclusion seems absurd! Clearly, if *A* and *B* had been combined, Flimsy would have been *lighter* than it actually is. And it seems incoherent to suppose that it would have been simultaneously lighter than it actually is and darker than it actually is.

Williamson's actual argument is developed in a formal language with an operator @ (for 'actually') alongside the $\Box$ of metaphysical necessity, as in Crossley and Humberstone (1977). Here is a condensed version:

1.	$\forall x \forall y \Box(x = y \rightarrow @x = x \rightarrow @x = y)$	Premise
2.	$\forall x \Box @x = x$	Premise
3.	$\forall x \forall y \Box(\neg @x = y \rightarrow x \neq y)$	1, 2
4.	$\forall x \forall y \Box(@x \neq y \rightarrow \neg @x = y)$	Premise
5.	$\forall x \forall y (x \neq y \rightarrow \Box @x \neq y)$	Premise
6.	$\forall x \forall y (x \neq y \rightarrow \Box x \neq y)$	3, 4, 5

Line 1 is the universal generalization of the necessitation of an instance of $\mathrm{LL}_{\mathrm{sub}}$. Lines 2 and 5 can both be derived from the general principle $\forall p(p \rightarrow \Box @p)$ ('Whatever is the case is necessarily actually the case'). Line 4 can be derived from the principle $\forall p \Box(@\neg p \rightarrow \neg @p)$ ('Nothing could be both actually the case and actually not the case'); or alternatively, from $\forall p(\Diamond @p \rightarrow p)$ ('Anything that could be actually the case is the case'). Both the inferences from 1 and 2 to 3 and from 3–5 to 6 use only the basic modal logic H_{K} (for $\Box$). Exactly the same argument could be given interpreting $\Box$ is 'If *P* it would be the case that . . . ' for any given antecedent *P*.

The argument is not irresistible: our ability to evaluate sentences in the formal language containing @ is not as firm as one might wish, since the relation between that language and the natural language phenomenon it is supposed to model is rather loose. In English, the use of the word 'actually' doesn't seem to be required to get the relevant thoughts across. Moreover, some of the central theorems of the standard logic of @ don't seem to correspond to any possible interpretation of the corresponding natural language sentences (Yalcin 2015): for example, there is no reading of 'I could actually have run you over' on which it entails 'I ran you over'. However, our informal argument using 'darker than it actually is' suggests that Williamson's argument is not just an artefact of an unfortunate choice of formalism: the force of the informal argument is readily appreciated already in English. And this is true despite the fact that the word 'actually' is entirely redundant in stating the argument: for example, 'If *A* and *B* had been combined, Frail would have been darker than it is' seems unambiguously equivalent to 'If *A* and *B* had been combined, Frail would have been darker than it actually is.'[23]

[23] If you want to deny ND, which step of Williamson's argument should you resist? Here we can profitably deploy the higher-order regimentation of 'world'-talk which we introduced in §1.6, on which

In sum, there are weighty objections to the way out of the puzzle based on denying Distinctness, although we will not pretend that they are entirely decisive.[24]

4.3 Denying Non-Coincidence

Let's turn next to the strategy that accepts Distinctness and the two Resiliency premises while rejecting Non-coincidence. Here we can usefully distinguish two sub-possibilities. The first option is to say that already at the actual world, some ordinary buckets coincide with other buckets. The second option is to say that *in fact* no two ordinary buckets coincide, but this would have been false if *A* and *B* had been combined.[25] Let's focus to begin with on the first option.

Some philosophers have in fact been willing to embrace a view on which, for many familiar kinds *K*, it is quite typical for *K*s to be coincident with, or at

worlds are identified with certain collections (rigid properties) of propositions. This makes it natural to interpret the constant @ as expressing a rigid property of propositions which is coextensive with truth. In that case, the place to resist Williamson's argument is line 4. For suppose $x \neq y$ but $\Diamond x = y$. Then we also have a failure of ND in type $\langle\langle\rangle\rangle$: $(x = y) \neq (x = x)$, but $\Diamond((x = y) = (x = x))$. Since @ is coextensive with truth, we have $@x = x \wedge @x \neq y$, hence $\Box(@x = x \wedge @x \neq y)$ by Persistence, and hence $\Box(x = y \rightarrow (@x = y \wedge @x \neq y))$ by the necessitation of LL; given $\Diamond x = y$, this implies that $\Diamond(@x = y \wedge @x \neq y)$. Although the collection @ in fact doesn't contain any pair of a proposition and its negation, the failure of ND means that this is not necessarily the case, since there are pairs of truths one of which could be identical to the negation of the other.

[24] Giving up Distinctness might seem like the most natural option for those who favour treating Tolerance Puzzles as refutations of Iteration, on the principle that if one is in the business of solving puzzles about the variability of material objects by giving up widely accepted principles of modal logic, one might as well be hanged for a sheep as for a lamb. Indeed, Lewis (1986*a*), whose Iteration-denying approach to Tolerance Puzzles will be considered in Chapter 10, is also an opponent of ND. On the other hand, Salmón, the most vocal opponent of Iteration, is firmly committed to ND. He suggests that the ND-implying principle that "objects identical *within* a possible world are identical in the absolute sense of being numerically one and the same object" (Salmon 1979: 719) should be accepted as "a presupposition of the theory of direct reference". His thinking seems to be that if $a \neq b$, although at a certain possible world *w*, $a = b$, then the name '*b*' would fail to be *rigid* since it refers to *a* with respect to *w* but not respect to the actual world. In response, it seems to us that even taking on board the ideology of reference with respect to a world (which is not our preferred way of theorizing about direct reference), there is no good reason to accept the inference from '*N* refers to *a* with respect to *w*' and 'at *w*, $a = b$' to '*N* refers to *b* with respect to *w*'. The temptation to make this inference seems to conflate the question what a name refers to *with respect to* a world with the question what it refers to *at* that world. The three-place relation '*N* refers to *x* with respect to *w*' is supposed to capture something about the *actual* semantics of *N*: it is quite different from 'at *w*, *N* refers to *x*'. (Moreover, even if one wanted to endorse the inference, it would not be an important departure from the theory of direct reference to define a rigid name as one such that for some *x*, it refers with respect to any world *w* to all and only those objects that are identical to *x* at *w*.)

In a passage from Salmon 1979 that does not reappear in the corresponding portion of Salmon 1981, Salmón also suggests that it is "quite easy to show" that the above ND-implying principle is equivalent to the principle that "it makes sense to identify, in an *absolute* sense, individuals in different possible worlds". But the argument he sketches (Salmon 1979: n. 19) turns on the assumption that for an individual *y* and world W_2 there is such a thing as "*the* individual that *y* is intra-world identical with in W_2" (emphasis ours), which ND-deniers will reject.

[25] In the temporal cases, the contrast is between present coincidence and coincidence at other times.

least *almost* coincident with, many other *K*s. For example, Leslie (2011) defends an approach to Tolerance Puzzles on which every *axe* is in the same place as, and shares the same parts as, many other axes, which differ from it in their modal properties.[26] A precedent for this is Lewis (1993), who develops (without unequivocally endorsing) a view on which wherever there is one cat, there are enormously many other cats having almost the same spatial boundaries as the first cat, differing only in that their boundaries include or exclude certain hairs.

Leslie and Lewis both consider the obvious worry that their view leads to flagrant error in ordinary counting practices, and suggest that this worry can be addressed by rejecting the orthodox conception of how counting-words relate to identity, in favour of a view that ties counting-words instead to some less demanding relation such as coincidence or overlap. For example, maybe 'There is only one *F*' is equivalent not to 'There is an *F* such that every *F* is identical to it' but to 'There is an *F* such that every *F* coincides with it', or perhaps 'There is an *F* such that every *F* almost entirely overlaps it', or maybe even just 'There is an *F* such that every *F* overlaps it'. But it seems to us that the case for the orthodox counting-identity link is far stronger than Leslie and Lewis realize. For the departures from orthodoxy also risk denying the validity of certain paradigmatically valid-looking argument-schemas involving quantifiers and counting-words, such as the schema 'Some *F* is *G*; some *F* is not *G*; therefore there is more than one *F*.' This form of argument seems about as good as it gets, and not one that we need to test by making a thorough survey of the instantiations we get by plugging in particular values of *F* and *G*. But on the face of it, Leslie's and Lewis's proposals will require flatly denying its validity, since if we, say, set *F* to be 'table in this tiny office' and *G* to express some property that according to them distinguishes the objects that we count as one—e.g. *containing this wood chip* for Lewis, or *possibly not containing this wood chip* for Leslie—the premises will be true and the conclusion false.[27]

[26] Two parenthetical interjections in Leslie's paper back off from the 'many axes' thought: on p. 290 she writes that 'there is multiple co-location of axes (or axe-like entities) in the worlds under discussion'; similarly, on p. 281 she says that if her view is true, 'it follows that many ships (or ship-like entities) were entirely coincident with our ship'. But we do not see how to integrate the rest of Leslie's paper with the neutrality as regards whether the relevant entities are axes (or ships, etc.) suggested by these remarks. We will have more to say about Leslie's approach to Tolerance Puzzles in §11.5 below.

[27] One might, as a rather desperate expedient, attempt to restore the validity of this inference by appealing to a devious semantics for 'some' on which 'some *F* is *G*' is true if and only if there is an *F* such that everything coincident with it is *G*. But this disrupts further inferences such as 'Something is *F*; therefore either some *F* is *G* or some *F* is non-*G*', and moreover will make even speeches like 'Some table is tolerant' turn out false.

It is worth noting that the argument we used against the 'counting is not by identity' also works against those (e.g. Caie, Goodman, and Lederman 2020) who *do* think that counting is by identity, but reject Leibniz's Law, holding that in some cases coinciding objects are identical despite not having all the same properties. As discussed in §1.3, that view requires rejecting not only 'Some *F* is *G*; some *F* is not *G*; therefore there is more than one *F*', but also 'Some *F* is *G*; some *F* is not *G*; therefore some *F* is not identical to every *F*'.

Lewis (1976: 63) gives an example that convinced many metaphysicians that we often 'count not by identity but by a weaker relation'. He considers two roads which share a segment, and observes that if one got from *A* to *B* just by crossing this segment, 'I only crossed one road to get here' would ordinarily express a truth. He proposes to explain this by appeal to a semantics for 'one' that makes the remark equivalent to something like 'There is a road *x* that I crossed to get here, and for any road that I crossed to get here, *x* and *y* share a segment that I crossed to get here.' This also involves giving up the argument-form we like: e.g. take *G* to be 'road I crossed to get here' and *F* to be 'identical to the Chester Arthur Thruway'. But there are other workable ways to explain Lewis's datum without giving up on the counting-identity link. One is to appeal to context-sensitivity in the word 'road': perhaps, in the context of the speech, it expresses a property that only has one instance crossed by the speaker.[28] Another is to appeal to the well-known phenomenon of contextual quantifier domain restriction: perhaps, while *unrestrictedly speaking* there are several roads which the speaker crossed, only one of them is in the domain of the quantifier 'one' (maybe it is vague which one).

While contextual quantifier domain restriction can be invoked to block Lewis's independent case against the standard counting-identity connection, it also provides an alternative strategy for blocking the counting-based objections to pervasive coincidence within ordinary kinds. Just as we might say, 'Every graduate student has to write a qualifying paper' and mean only that every graduate student *at NYU* has to write a qualifying paper, perhaps when we say things like 'There is only one table in the room' or 'Every table in my office is tolerantly a table', we mean something true of the form 'There is only one *F* table in the room' or 'Every *F* table in my office is tolerantly a table.'[29] The central worry we have about this appeal to domain restriction is that it makes unrestricted quantification look like an achievement of philosophers that is radically discontinuous with the uses of quantification in ordinary thought and talk. As we see it, even though contextually restricted quantification is common, so is unrestricted quantification. While a typical utterance of 'There are no wine-glasses' would involve some contextually provided restriction, a typical utterance of 'There are no wine-glasses in my dishwasher' would not: if there is any object in the whole of reality that is a wine-glass in the speaker's dishwasher, the utterance will say something false. Ordinary people aren't blind to this contrast. Having said 'Every graduate student has to write a qualifying paper', they will readily acknowledge 'Of course, I don't mean every graduate student in the whole of reality' if pressed. But after a speech

[28] This would be natural for those like ourselves (see Chapter 11) whose solution to Tolerance Puzzles turns on context-sensitivity in 'table'.

[29] One might object that there is no non-arbitrary way for the context to single out a particular domain-restricting property *F* that applies to only one of the many tables in the office; but it is hard to make this objection stick, since we can make our peace with this arbitrariness in whatever way we make our peace with the arbitrariness afflicting canonical examples of vagueness.

like 'There are no wine-glasses in my dishwasher' or 'No donkeys can talk', they will vehemently insist that there really is absolutely nothing whatsoever that is both a wine-glass and in their dishwasher or that is both a donkey and able to talk.[30]

But for the appeal to domain restriction to block the counting objection, we would have to posit domain restriction in a huge variety of cases where all of these diagnostics point towards an unrestricted interpretation. If we follow up an utterance of 'This is the only table in my office' with 'Do you really mean that in the whole of reality there is nothing that is a table in your office other than this?', an ordinary speaker will vehemently answer 'Yes'. Of course, we can then engage them in some philosophy; perhaps if we give them the right reading list, they might end up with a 'many tables' view like Lewis's or Leslie's, and start backing off from that insistence. But their disposition to succumb to such views under dialectical duress is no evidence that their original quantifiers were restricted: by analogy, it would be poor form to take someone's disposition to be swayed by reading Lewis's *On The Plurality of Worlds* as evidence that their original utterance of 'No donkeys can talk' already involved some domain restriction.[31]

Thus, the obvious worries for views on which coincidence of tables, buckets, cats, etc. is pervasive in the actual world seem robust. Could deniers of Non-coincidence do better by instead going for the alternative picture, on which the actual world is special in being free of bucket-coincidence?[32] Such special pleading for actuality is prima facie weird. In §3.3, we suggested that the most compelling argument in the vicinity of this feeling of weirdness is one about the security of our judgements: it is uncomfortable to suppose that while we are are in fact right when we assert or presuppose 'No two buckets coincide', we could easily have been in error, simply because one machine operator decided to make thick-walled rather than flimsy buckets on one occasion. Moreover, quite apart from such reflections about security from error, the idea that coincidence of buckets would

[30] See Williamson 2003: 419–23.

[31] It might not be so bad to have to appeal to domain restriction in a more limited array of cases. In particular, for kinds like *forest*, *road*, and *pair of socks*, where it is already plausible that many instances of the kind have other instances of the same kind among their parts, or overlap them, there is good independent reason to think that the quantifier domains in play when we are making ordinary counting speeches are almost always limited so as to exclude any overlapping pairs of members of the kind. (One would not normally say, 'I brought six pairs of socks' if one only brought four matching socks.) Given this, it would not be so strange to think that distinct members of these kinds often perfectly coincide. As noted below, the most initially gripping *temporal* Tolerance Puzzles, like the one about the forests, involve kinds of this sort.

[32] One might think that this alternative picture should be rejected as insufficiently general on the grounds that it will not allow a 'coincident buckets' solution to the variant puzzle discussed in the previous section, in which we suppose that at the actual world there is a fat bucket, and use names 'Flimsy' and 'Frail' introduced via counterfactual descriptive reference-fixing. But as we said, the esoteric analogues to Resiliency in that puzzle strike us as far less immediately compelling than the Resiliency premises of our main argument. So we do not regard the need for a differential treatment of the two puzzles as a significant cost.

have occurred under certain everyday counterfactual circumstances raises obvious worries about ordinary speeches where counting judgements are embedded in counterfactuals or other modal environments: 'If we had combined those two portions of plastic, we would only have made 499 buckets today'; 'Because it only has one mould, this machine can only make one bucket at a time'; 'Since I was well aware how small my closet was, there was little chance that I'd end up storing two buckets in it.' Here again one might try invoking either domain restriction or the claim that we don't count by identity, but the objections to these manoeuvres we discussed above will apply equally when they are applied in modal contexts.

So far in this section we have been focusing on the modal Coincidence Puzzles. In the case of the temporal Coincidence Puzzles, there is a Non-coincidence-denying approach that neither requires claiming that coincidence (among entities of the relevant kind *K*) is *pervasive*, nor requires any prima facie worrisome special pleading on behalf of the actual world or the present time. Namely, one could claim that *K*s whose future or past history involves the patterns that generate Coincidence Puzzles coincide with other *K*s, while denying that this is true of *K*s that don't have that sort of history. For example, Lewis (1976) wants to say that people who are going to enter fission machines or have emerged from fusion machines coincide (temporarily) with people distinct from themselves, whereas run-of-the-mill people do not. One could say something similar about our forest example, where the Resiliency premises are much harder to dismiss as the result of a mistaken extension to exotic science-fictional possibilities of heuristics that work well enough under ordinary circumstances.[33]

However, views of this sort have the weird property that they make the question how many *K*s there are in a given place at a given time modally depend on how things are going to go at later times. For *K*s like 'person', 'bucket', 'forest' (though of course not for 'future farmer' and the like), this is very hard to get used to: if there were two people in the room yesterday, it doesn't seem like there could be a positive objective chance today that there was only one person in the room yesterday. Moreover, accepting it leads directly to some particularly biting *modal* Coincidence Puzzles. Suppose that Calvin is going to build and enter a fission machine. On Lewis's view Calvin now coincides with a distinct person, Calvin′. What would Calvin and Calvin' be like if no fission machine were built? There is a lot of pressure to think that they would still coincide: after all, it seems they might speak truly if they said, 'If I gave up on building the fission machine, I would go to the treehouse instead.' But on the view in question, no fission means no

[33] Our impression is that the Resiliency premises in temporal Coincidence Arguments are much more compelling when the relevant kind is one for which it is uncontroversial that some instances have other instances as parts. By contrast with forests, there is little temptation to think that any *bucket* would survive as a bigger bucket if we melted it down and combined its matter with that of another bucket. Conveniently, the appeal to contextual quantifier domain restriction to defuse the counting objection is also considerably more plausible in the case of kinds like *forest*, as discussed in note 31 above.

coincidence of distinct people, so we are forced to the radical view that gives up ND and maintains that the actually distinct people would in that case be identical.[34]

Denying Non-coincidence thus looks like a costly way out of our Coincidence Puzzle. But of course all the ways out have some costs, and some may regard it as the least costly. Those that take this route should note that it may have important implications for Tolerance Puzzles. If one goes for a view where coincidence of ordinary buckets is common in the actual world, one will need to rethink one's initial attraction to quantified Tolerance premises along the lines of 'Every table is tolerantly a table'. On a view where every table is one of many distinct tables which are all in the same place and made of the same matter, but differ in their modal properties (e.g. in that one would have been made of this half of the matter and one would have been made of the other half if the two halves had been made into distinct tables), the claim that *every* table is tolerantly a table looks far riskier than it might have initially seemed. It constrains the range of variation among the modal profiles instantiated by coincident tables in a way that would cry out for some kind of explanation, and it is hard to think of a principled story that would provide such an explanation. Meanwhile, if one goes for the view where coincidence of ordinary buckets is absent from the actual world but occurs in close counterfactual possibilities, one will need to provide some therapy to help overcome the sense that it gives a bizarre privilege to the actual world, and one will need somehow to resist the argument that it makes our avoidance of error (as regards whether buckets coincide) a matter of luck in an unacceptable way. But anything one might say that would help with this would tend to undermine the security-based motivation for Non-contingency in Tolerance Arguments.

4.4 Denying Resiliency

Finally, we come to strategies that reject the conjunction of the two Resiliency premises. Here we can usefully distinguish two different ways in which such a strategy might be developed. On the *contextualist* approach, both Resiliency premises are context-sensitive, and each is true in some interpretations, perhaps including the interpretations that would be most natural out of the blue. But they are not *both* true in any single context. By contrast, on the *hardline* approach, one or both of the Resiliency premises is context-insensitively false.

Let's begin with the hardline approach. There are many ways in which this might be developed. On the most straightforward version, *both* Resiliency premises are context-insensitively false. Perhaps it is even true in full generality that nothing—or at least no artefact—could have been originally made of twice as much material as it was actually made with. Or perhaps the buckets could have been made twice

[34] Lewis himself is no friend of ND: Chapter 10 will discuss the counterpart-theoretic setting for his rejection of it.

as big in some ways, but not with any of the particular pieces of plastic that are actually used to originate a single bucket. Or perhaps both buckets are such that they *could* have been made of *A* and *B*, although neither is such that it *would* have been made of *A* and *B* if *A* and *B* had been combined.

Alternatively, one could maintain that exactly one of the Resiliency premises is context-insensitively false, the other being true in at least some contexts. Perhaps our description of the case settles which of the the two it is: perhaps, for example, there is some general principle about the significance of temporal order according to which, because Frail was made after Flimsy, Flimsy takes priority over Frail in any situation in which just one of them is made. Or perhaps the question which of the Resiliency premises is context-insensitively false turns on some aspects of the setup that our description of the case didn't settle, although it could have been settled by a more detailed description of the case—e.g. on the question whether, if *A* and *B* had been combined, the resulting bucket would have been made closer to the time Flimsy was actually made or to the time when Frail was actually made. Or perhaps there is just no way for us to find out which of the counterfactuals is true, although we can know that they are not both true, just as some think there is no way for us to find out whether a coin would have landed heads if it were tossed one time more than it was actually tossed.

Our basic worry about all these versions of the hardline approach is that they are unduly error-theoretic about ordinary judgements. Talk about counterfactual possibilities where things are much bigger or smaller than they actually are is fairly common, and we have no sense that in engaging in such talk we are going out on a limb or speaking nonliterally. Consider, for example, the various counterfactuals a baker might assert about how various baked goods would have been different if more or less of various ingredients had been used: 'If I had been more careful, that doughnut would have been the same size as all the others instead of half their size'; 'Instead of saving the batter for the second cake [/bothering to make the first cake], I should have used that batter to make this one twice as thick.' And note too that there are many cases in which it seems quite obvious that an object of some ordinary kind could have gained or lost a lot of matter *after* being made: for example, we could add a second layer to a tabletop to make the table stronger, or extend a bungalow by adding a second floor. But once this is granted, it seems odd to think that the relevant object wouldn't have been around, or wouldn't have been around as a large-sized thing of the relevant kind, if the extensions had been integrated into the original creation process: 'We would have been a lot happier with this table if we had added the extra layer to the top right from the beginning.'

This leaves us with the contextualist approach. Indeed, some may have been feeling all along that such an approach was so obviously correct as to make 'Coincidence *Puzzle*' something of a misnomer. After all, counterfactuals are famously highly context-sensitive: 'If Caesar had fought in Korea he would have used nuclear weapons' and 'If Caesar had fought in Korea he would have used

catapults' each seems true, but clearly they are not both true in the same context.[35] But as we observed above, Coincidence Puzzles can also be used for claims of high objective chance in place of counterfactuals. And it is by no means obvious that attributions of objective chance are relevantly context-sensitive. In order to think they were, one would have to either hold that the expression 'objective chance' is context-sensitive in an unexpected and far-reaching way, or (perhaps more promisingly) that there is relevant context-sensitivity coming from other material in the sentence—e.g. the names 'Flimsy' and 'Frail' and/or the noun 'bucket'.

In the temporal case, the analogue of taking 'objective chance' to be context-sensitive would be a view on which tense operators like 'A thousand years ago' are context-sensitive. This is challenging: it is hard to get a grip on what the different possible meanings for such operators might be. However, in the case of some temporal Coincidence Puzzles, appealing to context-sensitivity (or plain ambiguity) in the relevant *count nouns* is an attractive and popular option. For example, several authors have wanted to block the Ship Argument of §4.1 by positing (at least) two readings of 'ship'. The ship_1 in the harbour was once a ship_1 captained by Theseus; the ship_2 in the museum was once a ship_2 captained by Theseus. But the ship_1 in the museum was never captained by Theseus: it was created only when the planks were reassembled in the museum. And the ship_2 in the harbour was never captained by Theseus: it too was created much later, perhaps only at the last plank-replacement. So, each Resiliency premise has a natural reading on which it is true, but there is no single reading on which both are true. Non-coincidence and Distinctness are true on both readings.[36]

We find this line pretty plausible. However, it is not so obvious that it can plausibly be extended to more symmetric Coincidence Puzzles, like the Flimsy/Frail case. It is not so hard to get a rough feel for what the two relevant interpretations of 'ship' are supposed to be like, or to imagine a language in which they are expressed by separate words. But it is a lot harder to do anything similar with 'bucket'.

In Chapter 12, we will develop the contextualist approach further and see how this challenge might be answered. For now, the main observation we want to make is that if we appeal to fine-grained context-sensitivity in these kinds of cases, that will again have some tendency to undermine the Security Argument for Non-contingency in Tolerance Arguments. If the Resiliency premises in Coincidence Arguments are importantly context-sensitive, it is natural to think that the same is true of the Tolerance premises in Tolerance Arguments. But insofar as we posit context-sensitivity in Tolerance premises, we have to be very careful about the inference from the nearby falsity of the propositions we in fact express by those sentences to nearby error.

[35] These examples are from Lewis (1973: 66), who credits them to Quine.

[36] Simons (1987) accepts a view like this, where 'ship' is ambiguous between 'form-constant ship' and 'matter-constant ship'. Hughes (1997) points out that 'ship' has uses that don't fit Simons's characterizations of either of the two senses, and instead rejects Non-coincidence. Wiggins (1980: ch. 3) considers but rejects a similar view, which he dubs the 'naïve conceptualist view'.

5
Accepting Hypertolerance

As we saw in Chapter 3, there are Tolerance Arguments whose premises remain prima facie plausible even when one has rid oneself of the temptation to go in for Sorites reasoning. Given the force of these arguments, the option of embracing Hypertolerance—i.e. holding that the relevant objects could have any of the properties ancestrally close to properties they instantiate, in the relevant sense of "close"—deserves careful consideration, even in cases where it initially seems bizarre to suppose that the relevant objects could be so dramatically different from the way they in fact are.

Sometimes the intuitive outrage prompted by some Hypertolerance claim can be softened with a little metaphysical imagination. For example, perhaps the thought that the Great Pyramid couldn't have been thimble-sized could be countered by imagining a Lilliputian scenario where everything is a tiny fraction of its actual size. In other cases, the intuitive resistance to Hypertolerance will be robust even in the face of such imaginative adventures. But given that, as we pointed out in the Introduction, Non-hypertolerance is not particularly deeply rooted in ordinary practice, it is important to look for systematic theoretical arguments against Hypertolerance rather than merely relying on intuition. In this chapter we will consider some such arguments.

5.1 Chisholm

In the paper that kicked off contemporary discussion of our puzzles, Chisholm (1967) describes something like a Tolerance Argument for a conclusion he takes to be unacceptable, namely, that the world could be indiscernible from the way it actually is while Adam plays the role of Noah and Noah plays the role of Adam. Here is his sketch of a gradual route from the actual world 'W^1', where Adam lived for 930 years and Noah lived for 950 years, via a world 'W^2' where Adam lived for 931 years and Noah lived for 949 years, to a world where Adam and Noah have swapped roles in the supposedly unacceptable way:

> Once again, we will start by introducing alterations in Adam and Noah and then accommodate the rest of the world to what we have done. In W^3 Adam lives for 932 years and Noah for 948. Then moving from one possible world to another, but keeping our fingers, so to speak, on the same two entities, we

The Bounds of Possibility: Puzzles of Modal Variation. Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri, Oxford University Press. © Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri 2021.
DOI: 10.1093/oso/9780192846655.003.0006

> arrive at a world in which Noah lives for 930 years and Adam for 950…. Now let us continue on to still other possible worlds and allow them to exchange still other properties. We will imagine a possible world in which they have exchanged the first letters of their names, then one in which they have exchanged the second, then one in which they have exchanged the fourth, with the result that Adam in this new possible world will be called 'Noah' and Noah 'Adam'. Proceeding in this way, we arrive finally at a possible world W^n which would seem to be exactly like our present world W^1, except for the fact that the Adam of W^n may be traced back to the Noah of W^1 and the Noah of W^n may be traced back to the Adam of W^1. (Chisholm 1967: 3)

(In context it is clear that when Chisholm says 'The Adam of W^n' he means something like 'The person called "Adam" in W^n', and that 'may be traced back to' can be read as 'is identical to'.)

Chisholm is unwilling to embrace the conclusion of this argument: he asks, rhetorically, 'Could God possibly have had a sufficient reason for creating W^1 instead of W^n?'.

If we interpreted Chisholm's characterization of the sequence in such a way that by the end Noah has to share *all* of Adam's actual properties—even properties like *being distinct from Noah*—Chisholm would certainly be right that there is no possible world that plays the role of W^n. By the necessity of identity, *being distinct from Noah* is not a property Noah could have. But by the same token, on this interpretation there is little appeal to the Necessitated Tolerance thought that at each world in the sequence the next world is possible: at some point in the sequence the time will come for swapping properties like *being distinct from Noah*, and there isn't even a prima facie reason to think that the required Tolerance premise is true at a world just before this. In fact, however, Chisholm makes it clear that he does not intend these kinds of identity-properties to be included on the list of properties to be swapped.[1] It seems most reasonable to interpret him as intending a restriction to what the subsequent literature in metaphysics has come to call *qualitative* properties.

Intuitively, a qualitative property is one that doesn't "involve the identity" of any particular objects. *Being distinct from Noah*, *standing beside Noah*, and *being the same height as Noah* would all be paradigmatically *non-qualitative*, or 'haecceitistic', properties, whereas Chisholm's examples (*living for 930 years*, *having a name that begins with 'A'*, etc.) are plausibly qualitative. Chapter 14 will consider the theory of qualitativeness in more detail. For now, let's grant that the notion of

[1] 'The two Adams could be called "discernible" in that the one has the property of being Noah in the other world and the other does not, and similarly for the two Noahs. But in the sense of "indiscernible" that allows us to say that "Indiscernibles are identical" tells us more than "Identicals are identical", aren't the two Adams, the two Noahs, and the two worlds indiscernible?' (4)

qualitativeness is in good standing, and consider how Chisholm's argument fares when we interpret him as intending a restriction to qualitative properties.

The conclusion of the argument on this interpretation is that it is possible that Adam instantiates all the qualitative properties that Noah in fact instantiates and Noah instantiates all the qualitative properties that Adam in fact instantiates. What's supposed to be wrong with that? Chisholm invokes God's choice, and gestures towards the Principle of Sufficient Reason. It is not so clear what version of this principle Chisholm has in mind. But whatever it is, it is hard to see how it could rule out the possibility that Adam and Noah swap qualitative roles unless it entails the following principle:

Weak Anti-haecceitism Every truth is necessitated by a qualitative truth.[2]

Given certain basic principles about qualitativeness, Weak Anti-haecceitism is equivalent to the claim that every possible world that differs from the actual world in any way differs with regard to the truth value of some qualitative proposition.[3] This restatement makes the connection to the Principle of Sufficient Reason a bit clearer: if there were a non-actual possible world that agreed with the actual world on all qualitative propositions, it is hard to see how God could have had sufficient reason to prefer actuality over that world.[4]

(We called the principle 'Weak Anti-haecceitism' because it does not rule out the hypothesis that some pairs of possible worlds agree on all qualitative propositions but disagree on some non-qualitative proposition: it merely implies that in any such pair, both of the worlds in question are non-actual. There is a stronger principle which does rule out the existence of such pairs:

Anti-haecceitism Every possibly true proposition is necessitated by a possibly true qualitative proposition.

Anti-haecceitism is equivalent to the thesis that *every proposition is necessarily equivalent to some qualitative proposition*. For let p be any proposition; let C be

[2] The label 'haecceitism' is due to Kaplan (1975). The terminology in this area is muddied by the fact that Lewis (1986*a*) embraced the label 'Anti-haecceitism' for a different view, having to do with worlds, but consistent (according to his idiosyncratic theory of worlds) with the denial of Weak Anti-haecceitism. See Skow 2008 for a careful discussion of the situation.

[3] See §14.1 for a statement of the relevant principles.

[4] We don't actually need any principle like 'God only cares about qualitative truths' to derive Weak Anti-haecceitism from the PSR. It suffices to assume (a) that the proposition that every divine being does everything it has sufficient reason to do is a qualitative one, and (b) that if there is sufficient reason not to actualize a world, then it is necessary that there is sufficient reason not to actualize that world. For suppose w is a non-actual possible world. By the PSR there is sufficient reason for God not to actualize w. By (b) this is necessary, and hence true at every possible world including w itself. But then the proposition that every divine being does everything it has sufficient reason to do is a qualitative truth that is false at w.

the collection of all qualitative propositions which necessitate p; and let q be the proposition that some member of C is true. q is qualitative since every member of C is.[5] p necessitates q, since if $p \wedge \neg q$ were possible, by Anti-haecceitism, it would be necessitated by some possible, qualitative proposition m, and since this m would necessitate p, it would belong to C, and thus necessitate q as well as $\neg q$, contradicting the possibility of m. And q also necessitates p, since each member of C does.[6] This clarifies why Anti-haecceitism entails Weak Anti-haecceitism. In the other direction, we can show that given an S5 modal logic and the (plausible) principle that every qualitative proposition is necessarily qualitative, Anti-haecceitism is equivalent to the claim that Weak Anti-haecceitism is necessarily true. However, against the background of weaker modal logics, the necessitation of Weak Anti-haecceitism does not imply Anti-haecceitism.[7])

Even Weak Anti-haecceitism is quite a strong and surprising claim. It implies that for any qualitative profile q that is uniquely instantiated by some object x, it is impossible for any object other than x to instantiate q. The most obvious way for this to be true is for some thoroughgoing hyperessentialism to be true, according to which any given object can only be concrete if everything is exactly as it is in every respect (qualitative and non-qualitative). Such hyperessentialism would require giving up all interesting Tolerance premises. Granted, there is room for a view that endorses Weak Anti-haecceitism while still accepting a wide range of ordinary Tolerance judgements. However, there is arguably something a bit unprincipled about the idea that the essences of objects are demanding enough that no object could instantiate any qualitative profile that is in fact uniquely instantiated by some other object, while also being weak enough that any collection of atoms sufficiently overlapping a table's actual originating atoms could have been that table's originating atoms.

Most metaphysicians reject Weak Anti-haecceitism. It's a compelling thought regarding two atoms, Tony and Jacky, that Tony could have been alone in empty space and that Jacky could also have been alone in empty space, and that such possibilities, while distinct, need not be qualitatively discernible.[8] Similarly, if there are just three atoms a, b, c such that b is 2 nanometres from each of a and c while

[5] This follows from the theory of qualitativeness we will present in §14.1.

[6] This is an appeal to Inextensibility (see §1.5): given $\mathsf{R}(C)$ and $\forall m(Cm \rightarrow \Box(m \rightarrow p))$ we have $\Box\forall m(Cm \rightarrow (m \rightarrow p))$ and hence $\Box(\forall m(Cm \rightarrow m) \rightarrow p)$.

[7] The necessitation of Weak Anti-haecceitism is compatible with there being two possible worlds w_1 and w_2 that agree on every qualitative proposition but disagree on some non-qualitative proposition, provided that neither of these worlds is possible at the other, as can happen if the 5 axiom fails. If Iteration fails, the truth of Anti-haecceitism is compatible with there being a possible world w at which Weak Anti-haecceitism is false because of some impossible w' which is possible at w and agrees with w on all microphysical propositions. And if the 5 axiom fails (even if Iteration holds), the truth of Anti-haecceitism is compatible with there being a possible world w at which Weak Anti-haecceitism is false because at w there are *new* propositions that are true but not necessitated by any qualitative truth.

[8] The most influential arguments against anti-haecceitism, very much in this spirit, are due to Adams (1979).

a and *c* are 3 nanometres apart, it is natural to think that it could instead have been that *a* was 2 nanometres from *c* as well as *b* while *b* was 3 nanometres from *c*. So, if Weak Anti-haecceitism follows from some version of the Principle of Sufficient Reason, that looks more like bad news for Sufficient Reason than like good news for Weak Anti-haecceitism. If "Chisholm's Paradox" (as it is sometimes called) amounts to a conflict between Weak Anti-haecceitism, or the Principle of Sufficient Reason, and some other plausible premises, it isn't much of a paradox.

For whatever reason, Chisholm finds the conclusion that Adam and Noah could have swapped qualitative roles repugnant. Others may find it congenial. It is suggested by a certain radical but simple theory according to which for any object, and any possibly instantiated qualitative property, it is possible for that object to instantiate that property. This follows from a more general principle that also applies to relations:

Extreme Anti-essentialism For any distinct objects $x_1, ..., x_n$ and any qualitative relation R, if it is metaphysically possible for R to be instantiated by distinct objects, it is metaphysically possible for R to be instantiated by $x_1, \ldots, x_n$ in that order.[9]

Extreme Anti-essentialism entails that it is possible for Adam to bear to Noah the conjunction of all qualitative relations that are in fact borne by Noah to Adam, and thus that each could instantiate that conjunction of all the qualitative properties in fact instantiated by the other. Of course, one could accept the Adam-Noah conclusion without going as far as Extreme Anti-essentialism: less extreme versions of anti-essentialism might allow for a few 'categorial boundaries' within the domain of objects, such that objects can freely swap roles within but not across these boundaries.[10] Far from regarding Chisholm's argument as paradoxical, proponents of views like Extreme Anti-essentialism might see it as providing significant abductive support for these views.

But whether one finds the conclusion that Adam and Noah could have swapped qualitative roles repugnant or congenial, there remains the question whether

[9] If we dropped the clause requiring $x_1, \ldots, x_n$ to be distinct, Extreme Anti-essentialism would conflict with the necessity of identity, since it would entail that distinctness (a possibly instantiated, qualitative relation) is possibly instantiated by any x_1 and x_2, including the case where $x_1 = x_2$. This is ruled out by our basic modal logic. On the other hand, it would be consistent with our basic modal logic to strengthen Extreme Anti-essentialism by requiring merely that R be possibly instantiated rather than possibly instantiated by distinct objects. If we did this, the resulting principle would entail that any two objects could be identical (and thus conflict with the necessity of distinctness).

[10] Mackie (2006) defends a form of anti-essentialism perhaps a bit less extreme than Extreme Anti-essentialism. Mackie claims that Aristotle could have been a poached egg or a centipede, but does not hold that he could have been an event or a number (for example). Whether this requires weakening Extreme Anti-essentialism will depend on whether one takes events and numbers to be objects, as opposed to treating talk about them as some sort of disguised higher-order quantification, analogous to how we have been treating talk of properties. See §7.4 for more on this question.

Chisholm's sketch of a Tolerance Argument for that conclusion can in fact be fleshed out in such a way that its premises are prima facie plausible. It is far from clear that it can. In particular, it is far from clear that the details can be filled in in such a way that Non-contingency could be motivated by a Security Argument of the sort discussed in §3.3, rather than merely resting on the problematic 'small differences can't matter' intuition that gives prima facie appeal to all Sorites reasoning. Chisholm encourages us to imagine the complete qualitative profiles of Adam and Noah as big conjunctions of determinate properties that can be varied one at a time: first *age at death*; then *name*; then maybe *ratio of animals named to animals saved from drowning*, or *proportion of life spent naked*.... But this is misleading: a complete qualitative role is a super-detailed relational property that can't be decomposed into conjuncts in any natural way. Instead, we should be imagining tracing a path through the space of possible qualitative profiles, starting with Adam's actual profile and ending with Noah's. But not just any old path will do, no matter how small the steps along the path might be. For example, the argument won't be very compelling if some of the profiles we visit along the way entail *being a poached egg*, or *being a stack of miscellaneous body parts*, or anything like that. In a world where Adam existed with one of the qualitative profiles well into such a sequence (so that he was at best a borderline person) the only obvious motivations for thinking that he could have been one step still further along the sequence are the suspect, Soritical ones. Our "ordinary modal practices" don't tell us that monstrous borderline-people could still have concretely existed while being even more monstrous; still less, that they could still have been people in that case. Even some profile in the sequence is such that if Adam had instantiated it, we would have judged falsely that he could have instantiated the next profile after that in the sequence, this counterfactual error does not threaten the epistemological bona fides of our actual judgement that Adam could have instantiated the next profile in the sequence (the second one), since that actual judgement isn't just based on a suspect Soritical intuition, but on much more ordinary modal sensibilities.

So, if we want to construct a compelling, non-Soritical Tolerance Argument for the claim that one person could have instantiated the qualitative role actually instantiated by someone else, we will need to construct a path from the first qualitative role to the second that isn't just gradual, but also steers clear of the kind of "monstrousness" which would make counterfactual errors about Tolerance unsurprising, and unthreatening to the security of our actual practice. This means, inter alia, that we need to stay well clear of weird cases where something is so far from the paradigms of personhood that we might understandably feel more hesitant in our judgements about its capacity to tolerate qualitative differences than we do in ordinary cases. And it is by no means clear that this can be done in the case of two people. Let's set the biblical Adam aside (since his non-biological origin already makes it natural to suspect that he is rather special in his modal properties)

and consider two real people, Philip IV of Spain and his contemporary Charles I of England. Among Philip's qualitative properties was that of having parents who were second cousins (the Holy Roman Emperor Charles V was both his mother's father's father's father and his father's father's father's father). Charles I lacked this qualitative property. Because of this, any sufficiently gradual path from Philip's qualitative profile to Charles's will necessarily contain something that entails being close to the borderline of the property, which means something quite weird—e.g. a property that entails having developed from an egg whose genetic material was derived from several sources, so that a small difference in the relevant proportions would have been enough to make a difference as regards which of two people was one's (biological) parent. This isn't exactly *monstrous*—after all, in the actual world, there was in fact a child born in 2016 from an egg whose mitochondria had been exchanged with mitochondria from an egg derived from a different person (Zhang et al. 2017). One might (not very plausibly) think that the label "three-parent baby" is in fact a biologically accurate description of this child, in which case it might in fact be close to the borderline of *having parents who were second cousins.* Our "morphing sequence" could perhaps visit odd qualitative profiles that entail existing in worlds where the technology for such mitochondrial exchange was developed much earlier. Still, if there are people whose origin involved a process of transferring genetic material between eggs that would only need to have been slightly different to have resulted in a baby with a different set of parents, our ordinary modal practices don't make it *so* implausible to think that those people are "living on the edge", in that some mild quantitative variations in their originating process would be enough to take us to a possible situation where those very people are never born. Insofar as our actual confidence that people *aren't* living on the edge isn't just manifesting misguided Soritical intuitions, it may be based in part on our tacitly assuming that the relevant techniques have not in fact yet been developed to the point that there are borderline cases for biological parenthood. The task of constructing a Tolerance Argument that Philip could have instantiated Charles's qualitative profile for which there is a forceful, security-based case for Non-contingency thus looks quite difficult.

Once we are careful not to fall back on Necessitated Tolerance premises whose plausibility is merely Soritical, we can see that there is nothing like a general template for producing a plausible Tolerance Argument for the conclusion that some given object could have played the qualitative role of some other given object. That should dampen any ambition to use Tolerance Arguments to support some radical principle like Extreme Anti-essentialism. But that isn't to say that there is no way to give a forceful Tolerance Argument for a Hypertolerance conclusion that conflicts with Weak Anti-haecceitism. Indeed, in a later paper (1973), Chisholm suggests a different example involving two tables, where the obstacles are much less severe. In Chapter 6, we will be presenting some such arguments ourselves. But as we have said, the conflict with Weak Anti-haecceitism does not strike us as

a very forceful reason for rejecting Hypertolerance in these cases. So let's look for some other theoretical considerations which might provide weightier reasons for rejecting Hypertolerance.

5.2 From Sufficiency to Overlap Essentialism

In discussions of Tolerance Arguments involving original composition, principles such as the following have often been cited as the primary reason not to simply accept Hypertolerance:

Overlap Essentialism For every table x and collection of atoms C: if C originally composed x, then necessarily x is not originally composed by a collection of atoms with no members in common with C.

Given there are in fact two non-overlapping chemically matching collection of atoms, one of which originally composes a table, Overlap Essentialism conflicts with the version of Hypertolerance where K is 'table', the modality is atomic possibility (see §2.4), and the closeness relation is our standard origin-closeness relation.

If we were talking about *metaphysical* rather than atomic possibility, the conflict between Hypertolerance and Overlap Essentialism would not be immediate. By giving up the Necessity of Distinctness for metaphysical necessity, one could consistently accept both claims (along with the claim that there are two non-overlapping matching collections of atoms, one of which originally composes a table). One could, for example, accept the following thesis: whenever C originally composes a table, some atom in C and some atom in C' are such that it is metaphysically possible for them to be identical, so that it is metaphysically possible for C and C' to overlap. And one could maintain that in each such case, while it is possible for the table in fact composed by C to be composed by C', it is necessary that if this happens C and C' overlap, so that it is impossible for the table to be composed by a collection not overlapping C. However, when we switch from metaphysical possibility to atomic possibility, this kind of reconciliation is no longer available. Atomic necessity is defined in such a way that truths of the form $x \neq y$ when x and y are atoms automatically get to be atomically necessary whether or not they are metaphysically necessary, and given this, it follows that the same is true of truths of the form *C and C' do not overlap* where C and C' are collections of atoms.[11] And as we noted in §2.4, replacing metaphysical necessity with atomic

[11] Suppose: $\forall x(Cx \rightarrow \forall y(C'y \rightarrow \Box x \neq y))$, where C and C' are rigid. By the inextensibility of C' we have $\forall x(Cx \rightarrow \Box\forall y(C'y \rightarrow x \neq y))$, so by the inextensibility of C, $\Box\forall x(Cx \rightarrow \forall y(C'y \rightarrow x \neq y))$, i.e. it is impossible for C and C' to overlap.

necessity does not make a significant difference to the plausibility of the premises of the relevant Tolerance Arguments: there is just no interesting motivation for a view where the metaphysical-possibility Tolerance premise is true but the atomic-possibility Tolerance premise is false.

So, embracing Hypertolerance in our original-composition Tolerance Arguments will require giving up Overlap Essentialism. And why not just do that? There is a widely discussed strategy for arguing for principles like Overlap Essentialism from premises stating putative *sufficient* conditions for the creation of a given object. In the recent literature, the canonical source for this strategy is Saul Kripke. In a famous footnote, Kripke offers "something like a proof" of a necessity of originating matter principle for wooden tables:

> Let '*B*' be a name (rigid designator) of a table, let '*A*' name the piece of wood from which it actually came. Let '*C*' name another piece of wood. Then suppose *B* were made from *A*, as in the actual world, but also another table *D* were simultaneously made from *C*. (We assume that there is no relation between *A* and *C* which makes the possibility of making a table from one dependent on the possibility of making a table from the other.) Now in this situation $B \neq D$; hence, even if *D* were made by itself, and no table were made from *A*, *D* would not be *B*. (Kripke 1972: n. 56)

In an important discussion of this passage, Salmón (Salmon 1979, 1981: ch. 7) points out that the most promising ways of turning it into a valid argument turn not only on the necessity of distinctness (which Kripke highlights) but on some kind of *sufficiency of origin* premise, such as the following:

Sufficiency For any *x* and *y*: if it is possible that *x* is a table originally made of hunk *y*, then necessarily any table originally made out of *y* is identical to *x*.[12]

Without something like this, it is obscure how one would rule out the following elaboration of Kripke's example: it is possible for Belinda the table to be made out of Angus, the hunk of which Belinda is actually made, while Charlie (a hunk not overlapping Angus) is made into a different table, Daphne; but it is *also* possible for Belinda to be made of Charlie while Angus isn't made into any table. The combination of Sufficiency and ND rules this out: we can infer from the possibility of making Daphne out of Charlie that necessarily any table made of Charlie is

[12] Following Kripke, we are assuming that pieces of wood are hunks of matter. As Barnett (2005) points out, it is not obvious that this is true, since one might think that pieces of wood cannot survive being scattered whereas hunks of matter can; those who agree with this can pick whichever they prefer, or follow our own preference and do everything with collections of atoms.

identical to Daphne, and hence (given ND) that it is impossible for Belinda to be made of Charlie.[13]

Something like Kripke's general strategy of arguing from sufficiency principles to claims that rule out Hypertolerance can also be seen much earlier. Hobbes (1655: ch. 11) uses his elaboration of the Ship of Theseus scenario (already discussed in Chapter 4) to argue against the temporal Hypertolerance claim according to which the Ship survives all the plank-replacements:

> [I]f (for example) that Ship of Theseus…were, after all the Planks were changed, the same Numerical Ship it was at the beginning; and if some Man had kept the Old Planks as they were taken out, and by putting them afterwards together in the same order, had again made a Ship of them, this without doubt had also been the same Numerical Ship with that which was at the beginning; and so there would have been two Ships Numerically the same, which is absurd.

Hobbes is plausibly appealing to the following sufficient condition for the appearance of a certain object at different times: if a ship s is made at t of a certain collection of planks C, and C are "in the same order" at t' as at t and compose a ship at t', then the ship composed by C at t' is s. Or putting the idea using tense operators: for any sometime ship s, sometime collection of planks C, and arrangement A: if *sometimes* C instantiates A and composes s, then *always*, every ship composed by C while C instantiates A is identical to s. This is structurally parallel to Sufficiency. And as Hobbes points out, it is inconsistent (given the evidently true premise that no object can simultaneously be composed of two distinct collections of planks) with the relevant temporal Hypertolerance claim,

[13] Salmón (Salmon 1979: n. 11; 1981: §25.6), crediting Fine and Stalnaker, discusses an alternative reconstruction of Kripke's argument, which does not involve any sufficiency premise; he reports that in conversation, Kripke said that this reconstruction matched his intention (Salmon 1981: 214, n. 10). The argument requires ND, Origin Uniqueness (see below), and an extra premise along the following lines:

Super-strong Compossibility For any non-overlapping hunks y and y' and possible tables x and x': if it is possible for x to be a table made of y, and it is possible for x' to be a table made of y', then it is possible for x to be made of y while x' is made of y'.

Setting $x' = x$, Super-strong Compossibility immediately implies that if a table could be made of each of two non-overlapping hunks, it could be made of both of those two hunks, which means by Origin Uniqueness that it is possible for them to be identical, which is ruled out by ND since they are not in fact identical. But as Salmón points out, Super-strong Compossibility is very strong; once we have noticed that it covers the case where $x = x'$, we may naturally feel that it is too close to Overlap Essentialism to provide substantive dialectical support for it. Rohrbaugh and deRosset (2004) nevertheless defend roughly this argument, deriving Super-strong Compossibility from a hard-to-pin-down "locality of prevention" thesis; for discussion, see Robertson and Forbes 2006; Roca-Royes and Cameron 2006; Rohrbaugh and deRosset 2006, and Ballarin 2013.

i.e. that (in his version of the thought-experiment) the original ship is eventually composed by a collection of planks not overlapping those it began with.[14]

But let's get back to Kripke-style arguments for Overlap Essentialism. The combination of Sufficiency and ND isn't enough all by itself: one could in principle suppose that it is possible to make Belinda out of Angus, possible to make Belinda out of Charlie, and impossible to make any table distinct from Belinda out of either Angus or Charlie. However, this would require the failure of one of two plausible principles:

Origin Uniqueness Necessarily, no table is originally made of two distinct hunks.

Weak Compossibility For any two non-overlapping hunks, if each is such that it is possible for there to be a table originally composed of it, it is possible that both of them originally compose tables.

Origin Uniqueness looks unproblematic so long as we are clear that '*x* is originally made of *y*' requires all of the matter in *y* and no other matter to be part of *x* at the beginning of its career.[15] Weak Compossibility, meanwhile, seems pretty plausible. One might challenge it with certain esoteric possibilities: for example, someone might propose that for something to be a table, there have to be some people around to make it or use it, and that each of two hunks which jointly exhaust the matter of the universe is such that it could compose a very large table while some of the rest of the matter was made into people. But if one were worried about these kinds of examples, one could easily strengthen 'overlap' throughout the argument to bypass such concerns: while the conclusion will then be weaker than Overlap Essentialism, it will be just as good for the purposes of arguing against Hypertolerance.[16]

Overlap Essentialism—or rather its analogue for hunks—follows from Origin Uniqueness, Weak Compossibility, Sufficiency, and ND.[17] Suppose for reductio that *x* could have been a table originally made from *y* and also could have been a table originally made from *y*′, where *y* and *y*′ are non-overlapping hunks. By Sufficiency, it is necessary that any table originally made from *y* is identical to *x*, and also necessary that any table originally made from *y*′ is identical to *x*. But by Weak Compossibility, it is possible that both *y* and *y*′ are such as to originally

[14] Hobbes of course wants to draw the stronger lesson that even in a case where the old planks were destroyed, the Ship would still not have survived all the plank-replacements. That inference raises further questions which need not concern us here.

[15] In fact, the argument will still be valid if we weaken Origin Uniqueness by replacing 'distinct' with 'non-overlapping'.

[16] One simple option is just to replace '*x* overlaps *y*' throughout with 'it is not possible that both *x* and *y* originally compose distinct tables', or the conjunction of this with overlap: then there is no need for any premise like Weak Compossibility.

[17] We can make do with the restriction of ND to hunks.

compose some table. Hence it is possible that x is a table originally composed of both y and y', which by Origin Uniqueness implies that it is possible that $y = y'$, contradicting ND.[18]

When presenting his favourite version of Kripke's argument, Salmón (Salmon 1979, 1981) mentions Weak Compossibility (his 'II' on p. 201), but actually employs a stronger premise ('IV' on p. 203) along the following lines:

Compossibility For any possible table x and non-overlapping hunks of matter y and y', if it is possible that x is a table originally composed of y, and it is possible that some table is originally composed of y', then it is possible that: x is a table originally composed of y and some table is originally composed of y'.

As we have seen, the additional logical strength of Compossibility is not needed in arguing from Sufficiency to Overlap Essentialism.[19] However, with Compossibility in hand, one can make do with a weaker sufficiency premise which deletes the possibility operator from the antecedent:

Weak Sufficiency For any x and y: if x is a table originally made of hunk y, then necessarily any table originally made out of y is identical to x.[20]

For suppose for reductio that a table x is in fact originally composed by a hunk y and could have been composed by a non-overlapping hunk y'. Then by Compossibility, x could have been a table originally composed of y' while some table was originally composed of y. But by Weak Sufficiency, it is necessary that any table originally composed of y is identical to x; so we can conclude that it is

[18] Formally, letting '$T(x, y)$' mean 'x is a table and y is a hunk of matter of which x is originally composed' and $O(y, y')$ mean 'y overlaps y'', we have:

Sufficiency $\forall x \forall y(\Diamond T(x, y) \rightarrow \Box \forall x'(T(x', y) \rightarrow x' = x))$

Origin Uniqueness $\Box \forall x \forall y \forall y'(T(x, y) \wedge T(x, y') \rightarrow y = y')$

Weak Compossibility $\forall y \forall y'(\neg O(y, y') \wedge \Diamond \exists x T(x, y) \wedge \Diamond \exists x' T(x', y') \rightarrow \Diamond \exists x \exists x'(T(x, y) \wedge T(x', y')))$.

Given ND, these jointly entail that for any two non-overlapping hunks, there is no possible table that could have been a table originally composed of either of those hunks:

Strong Overlap Essentialism $\forall x \forall y \forall y'((\Diamond T(x, y) \wedge \Diamond T(x, y')) \rightarrow O(y, y'))$

By the T axiom we can drop the first $\Diamond$, yielding something close to Overlap Essentialism. In an S5 modal logic, with the extra premise that overlap is persistent (as it is if we replace hunks with collections of atoms), Strong Overlap Essentialism is equivalent to the necessitation of this weaker version.

[19] Since Salmón is interested in arguing for the necessitation of Overlap Essentialism, he uses necessitated versions of various premises; in this context, our point is that the necessitated version of Weak Compossibility will be able to do the relevant work.

[20] Cf. Gendler and Hawthorne 2000 (296), claim (20).

possible that x is both originally composed of y and originally composed of y'. So by Origin Uniqueness, it is possible that $y = y'$, contradicting ND.[21]

Weak Sufficiency avoids some challenges to Sufficiency. Consider a variant on the bucket example from Chapter 2: suppose a table Chunky, shaped like an inverted, thick-walled bucket, is made out of a hunk of wood composed in turn of two non-overlapping sub-hunks, Inner and Outer. It's tempting to think that Chunky could have been made thin enough so that only one of the two sub-hunks was required, so that it is possible for Chunky to be made from Inner and also possible for Chunky to be made from Outer. But it is plausible (and moreover follows from Weak Compossibility and Origin Uniqueness) that there could have been two distinct tables made respectively of Inner and Outer. If so, we have a counterexample to Sufficiency, but so far we have no counterexample to Weak Sufficiency.

True, if one endorses S5, this case will be a counterexample to the *necessary* truth of Weak Sufficiency, since it is either possible that (Chunky is made of Inner but a table distinct from Chunky could have been made of Inner), or possible that (Chunky is made of Outer but a table distinct from Chunky could have been made of Outer). But as Chapter 3 made clear, it does not go without saying that anything true in this vicinity is necessarily true, and indeed in Chapter 11 we will be sympathetically exploring views where certain principles in the vicinity of Weak Sufficiency are true but not necessary. And moreover, if like Salmón one does not accept an S5 modal logic, then the case just presented is compatible with the necessary truth of Weak Sufficiency. If the 5 axiom (that everything possible is necessarily possible) fails, then we could say that while it is in fact possible for Chunky to be made of Outer, it would not be possible if Chunky had been made of Inner.[22]

But as Salmón (Salmon 1981: 210) notes, there are other problems with Sufficiency, which carry over to Weak Sufficiency. Consider a possible situation in which the modernist designer Franz Erlich makes a table out of one of the hunks

[21] Formally:

Weak Suffiency $\forall x \forall y (T(x,y) \rightarrow \Box \forall x' (T(x',y) \rightarrow x' = x))$

Compossibility $\forall x \forall y \forall y' (\neg O(y,y') \wedge \Diamond T(x,y) \wedge \Diamond \exists x' T(x',y') \rightarrow \Diamond (T(x,y) \wedge \exists x' T(x',y')))$

[22] Since Salmón denies the 4 axiom but accepts the T axiom for metaphysical possibility, he is committed to denying at least the *necessary* truth of the 5 axiom as well, since 4 follows from the combination of T and $\Box$5. (If $\Diamond\Diamond p$, then $\Diamond\Box\Diamond p$ by $\Box$5 ($\Box(\Diamond p \rightarrow \Box\Diamond p)$), so $\neg\Box\Diamond\neg\Diamond p$, so $\neg\Diamond\neg\Diamond\Diamond p$ by 5, so $\Box\Diamond p$, so $\Diamond p$ by T.) In fact, Salmón's approach to Tolerance Puzzles involves actual failures of 5 along with actual failures of 4.

Apart from using his sufficiency principle to argue against Hypertolerance, Salmón also has a different application for it, namely motivating premise 6 of the "Four Worlds Paradox" which is his central argument against Iteration. For this, he really does need Sufficiency (or something structurally like it). But as we pointed out in note 27 in Chapter 2, Salmón doesn't really need this premise to run his argument against Iteration, since it can be dispensed with by making a minor modification of his other premises which leaves their motivation unaltered.

that actually was made into a table by Thomas Chippendale using one his characteristic plans. It seems tendentious at best to claim that the table made by Erlich in that possible situation is the very same table actually made by Chippendale. The literature has, in various ways, attempted to remedy things by finding ways to weaken the sufficiency principles that still leave in place an argument for Overlap Essentialism (or some nearby principle). Finding such a principle, however, is not an easy task. In the rest of this section, we will give a brief taste of the difficulties involved.

Salmón's response (Salmon 1979: 716; 1981: 211) to examples like the Chippendale/Erlich example is to weaken Sufficiency by mentioning the "plan" according to which the table is constructed:

Plan Sufficiency If possible table x could have been made out of a hunk y according to a plan P, then necessarily, any table made out of y according to P is identical to x.

We can derive Overlap Essentialism from Plan Sufficiency, Origin Uniqueness, ND, and two further premises:

Plan Necessity Necessarily every table is made according to some plan.

Weak Plan Compossibility If y and y' are non-overlapping hunks of matter and P and P' are plans, and it's possible that y is made into a table according to P, and it's possible that y' is made into a table according to P', then it's possible that: y is made into a table according to P and y' is made into a table according to P'.[23]

If you were convinced by the earlier argument about Chunky, you could weaken Plan Sufficiency to 'Weak Plan Sufficiency' by replacing 'could have been' with 'is', while strengthening Weak Plan Compossibility to a 'Plan Compossibility' parallel to Compossibility.

The notion of 'plan' in these principles could be spelled out in many different ways. It could just be a matter of the table's shape and mode of construction, but we could also allow it to have *de re* aspects so that, e.g., 'make the top from this portion of the hunk and the legs from that portion of the hunk', or 'get the wood by chopping down the Liberty Tree', or 'model the table after one made by Chippendale' could count as (parts of) plans.[24] The more we build into plans, the

[23] Formally, we can just reinterpret $T(x, y)$ in the previous argument as 'y is an ordered pair of a hunk h and a plan P such that x is made into a table according to P', and $O(y, y')$ as 'y and y' are ordered pairs whose first co-ordinates are overlapping hunks'.

[24] Salmón seems to allow for some *de re* elements, since immediately prior to introducing plans in Salmon 1981 (211) he suggests that making a table from the same matter but swapping the matter used for the top and the legs might not be sufficient to make the same table as in the actual world. However,

more plausible Plan Sufficiency will be; but if we build too much in—e.g. if we allow plans like 'make a table from the three most expensive planks in the world'—Weak Plan Compossibility will no longer be plausible.

Although Salmón initially characterizes Plan Sufficiency as 'exceedingly plausible, almost to the point of being indubitable' (Salmon 1979: 716; 1981: 211), a few pages later (Salmon 1979: 721; 1981: App. I) he recognizes that it is subject to apparent counterexamples, assuming that plans are unspecific enough that the same matter can be used twice over to make tables according to the same plan. Some of the counterexamples are quite simple. Suppose that Erlich fashions a *Blumentisch*, smashes it into pieces, and then uses the same hunk some years later to make a *Blumentisch* according to the same plan. Plan Sufficiency implies that the original table has been, so to speak, reincarnated. This is implausible on its face, and even worse when combined with the view that tables can survive total Theseus-style replacement of matter. Suppose that Erlich fashions the *Blumentisch*, then slowly replaces the wood chips in it with new ones, and finally uses the old wood chips to make a second *Blumentisch*. If the first *Blumentisch* survives the gradual change, the second *Blumentisch* must be distinct from the first, so Plan Sufficiency is in trouble.[25]

Salmón's response to the objection (Salmon 1979: 721, 1981: 230) is to weaken Plan Sufficiency further by replacing the relation *table x is made out of hunk y according to plan P* with the relation *table x is the only table ever made out of hunk y according to plan P*. As a matter of logic, this relation cannot hold both between x, y, and P and between x', y, and P unless $x = x'$. However, as Robertson Ishii (Robertson 1998) points out, this weakened principle has other problems. Suppose that in the actual world Belinda is made of Angus according to P, and no other table is ever made from Angus. Let Angus* be a hunk of matter differing from Angus only by a few wood chips. Given Salmón's embrace of Tolerance, it's possible that Belinda is made of Angus* according to P. But according to the weakened sufficiency principle, the following is not possible: first Belinda is made of Angus* according to P, and then later a different table is made of Angus according to P. But if it's possible to make Belinda out of Angus* at all, surely doing so doesn't preclude later using Angus to fashion a different table according to the same plan. (As Robertson Ishii makes vivid, this is particularly evident if we think that tables can survive Theseus-style part replacement, since then we could use Angus to make a new table while Belinda is still around, by having previously extracted all of

as we saw in §2.4, he also uses the word 'plan' in stating Tolerance premises in a way that makes obvious trouble for one or both of Tolerance and Non-contingency if plans were allowed to be *de re* in certain of these ways.

[25] Salmón (Salmon 1979: 720) denies that Plan Sufficiency would entail the problematic identity claim in a case like this, on the grounds that it is a modal rather than a temporal principle, and focuses instead on a more complicated counterexample based on a case from Chandler 1975. But although he does not mention a change of mind, the corresponding passage in Salmon 1981 (228) makes the opposite claim that Plan Sufficiency *does* imply the problematic claims of identity over time.

the Angus parts of Belinda by gradual replacement.) But as Salmón (Salmon 1981: App. VI) notes in response, one could avoid this sort of problem by going for an even more drastic weakening of Plan Sufficiency: e.g. we replace the relation *x is the only table originally made of hunk y according to plan P* with the relation *x is a table originally made of hunk y according to plan P, and no other table is ever made out of any hunk overlapping y*.[26]

These weakenings of the sufficiency principles do not leave the argument for Overlap Essentialism unscathed. If we replace the relation *x is a table made from y according to P* with some stronger relation *R* throughout the argument, then instead of Overlap Essentialism, the conclusion of the argument will be that no table that bears *R* to *y* and *P* could have borne *R* to some y' and P' where y' does not overlap *y*. For example, if bearing *R* to *y* and *P* requires that no other table is ever made of any hunk overlapping *y*, then this weakening of Overlap Essentialism will be completely silent as regards the modal behaviour of any tables which are such that at some point some other table is made of a hunk largely overlapping the hunk from which they were made. Moreover, even when *x* is the only table ever made from any variant of its originating hunk *y*, the weakened principle does not rule out the possibility of *x* being made of some completely non-overlapping hunk y', so long as some other table is at some time made from a hunk overlapping y'. It is not at all obvious how we could argue our way back to the original Overlap Essentialism from the weaker versions of it that can be motivated from the weaker sufficiency principles.[27] However, even the weakenings of Overlap Essentialism will rule out Hypertolerance in some variant Tolerance Arguments, e.g. Tolerance Arguments based on a family of properties of the form *being originally composed of hunk y and such that no distinct table is ever originally composed of any hunk overlapping* y'. So if we are primarily concerned to argue against Hypertolerance considered as a general strategy for addressing Tolerance Puzzles, the weaker versions of Overlap Essentialism may still be dialectically helpful.

[26] Gendler and Hawthorne (2000) suggest a somewhat less drastic weakening, in which the relevant relation is *x is a table originally made of hunk y according to plan P, and no other table is ever originally made according to P of a hunk largely overlapping y*. But this is still subject to a problem similar to Robertson Ishii's, assuming tables are at least mildly tolerant as regards their plan. Suppose Belinda is made from Angus according to *P*. Then Gendler and Hawthorne's principle will entail that it is not possible for the following to happen: first Belinda is made from Angus according to a slightly different plan P', and later a numerically different table is made from Angus according to *P*. As a compromise, we might use the relation *x is a table originally made of hunk y according to plan P, and no other table is ever originally made of any overlapping hunk* y' *according to any plan* P' *such that x could have been made from* y' *according to* P'.

[27] Robertson Ishii (Robertson 1998) suggests that one could recover Overlap Essentialism from the weaker claim of Overlap Essentialism corresponding to Salmón's 'only table' weakening of Plan Sufficiency by appealing to the premise that any possible table that could be made from a hunk could be the only table made from that hunk. But as noted in Robertson 1998 (737) and Gendler and Hawthorne 2000 (294), premises of this sort are hard to consistently combine with the relevant weakenings of Plan Sufficiency. (There is, however, less trouble combining them with the corresponding weakenings of Weak Plan Sufficiency.)

Overall, though, we do not think that sufficiency-of-origin principles provide a powerful bulwark against Hypertolerance. For unless one crafts an extremely rich notion of "plan", even the weakest sufficiency principle we considered will look quite tendentious. Consider once more our table Belinda; suppose Belinda was made of Angus according to a certain medieval-inspired design *D*, and no other table is ever made of any hunk overlapping Angus. In a certain possible world *w*, Angus was earlier formed into a wooden throne at the behest of Henry VIII; after the destruction of the throne, collectors tracked down all of Angus's parts, and re-fashioned them into a table according to *D* (wrongly thinking that Henry had had Angus made into a table using that design), selling the table to a museum. It seems quite tendentious to claim that our boring actual table Belinda is a table at *w*. One could avoid being committed to this by coming up with a conception of 'plan' that goes well beyond anything of interest to carpenters, so that the plan used for Belinda in the actual world counts as different from the one used for the table made from Angus at *w*.[28] But it is unclear how to specify an interpretation of 'plan' rich enough to avoid worries along these lines. Even if we did, the resulting sufficiency principle would likely be so complicated that it would be hard for us to be confident that it was counterexample-proof, in view of the history of failures of simpler sufficiency principles. And finally, as we pointed out above, if we make the notion of plan *too* rich, then the compossibility premise required to get results anything like Overlap Essentialism will start to become more problematic.

We have looked at various failed attempts to state plausible sufficiency principles that could plug in to a Kripke-style argument for something like Overlap Essentialism. However, we do nevertheless think that the arguments are on to something interesting. Even if the kinds of factors the sufficiency principles try to list—originating matter, artificer, design, previous history of the matter . . .—don't settle whether a given table-making process results in the creation of a given possible table, the facts about which tables emerge from which such processes are not brute and inexplicable. When a particular table *x* gets made, there will be *some* collection of facts about the underlying details of how everything unfolded that necessitates that *x* was made. And even if this does not establish Overlap Essentialism, this thought still puts pressure on the general strategy of responding to Tolerance Puzzles by embracing Hypertolerance. In Chapter 6 we will discuss what seems to us to be the most promising way of fleshing out this thought, where the role of sufficiency principles will be played instead by physicalistic supervenience claims.

First, however, we will consider a different strategy for arguing against Hypertolerance. Even if one sets little store in sufficiency arguments for Overlap

[28] Note that building the time of construction and the artificer responsible into the sufficiency principle alongside or as part of the "plan" won't help, since the story can be filled in in such a way as to hold these constant. And the intentions of the carpenter need not vary in any interesting way either, since we can imagine that at *w* the collectors don't tell the carpenter what's significant about Angus.

Essentialism, one might still regard Overlap Essentialism as highly plausible in its own right, and take the conflict between Hypertolerance and Overlap Essentialism to constitute a weighty objection for Hypertolerance. In the next section, we will raise some doubts about Overlap Essentialism which, though not decisive, suggest that it is not nearly as compelling as it would need to be to play this dialectical role.

5.3 Doubts about Overlap Essentialism

Let's look again at Overlap Essentialism.

Overlap Essentialism For every table *x* and collection of atoms *C*: if *C* originally composed *x*, then necessarily, *x* is not originally composed by any collection of atoms with no members in common with *C*.

Offhand, this already looks rather plausible, so one might feel that even without some further argument, the need to reject it is a problem for Hypertolerance. However, as we will see in this section, a variety of challenges to Overlap Essentialism come into view when we start looking at examples. These challenges are quite independent of the doubts that might be prompted by Tolerance Arguments, and suggest that Overlap Essentialism is by no means on such firm ground that giving it up would be a weighty problem all by itself for Hypertolerance. These challenges to Overlap Essentialism make additional trouble for the sufficiency principles considered in the previous section, though we will leave it as an exercise for the reader to draw out these implications. Moreover, they can be adapted to support the conclusion that at least some tables are hypertolerant, in the sense that they could have been originally composed of any collection of atoms chemically matching those of which they were in fact originally composed.[29]

First: P. F. Strawson suggests that while it is false of a particular table 'of some fairly standard kind' that it could have been made of (entirely) different materials, the situation is different with 'elaborate human constructions', including certain buildings and ships:

> [T]he Old Bodleian ... might have been built of (composed of) stone from quarry A instead of being built, as it was, of stone from quarry B. Will someone say: then it would not have been this building, but another just like it? The retort seems insufficiently motivated. Before the building existed, there existed a plan: a plan for a building on this site, for this purpose, to be constructed of materials of

[29] At least setting aside collections spanning distant galaxies and the like, which as pointed out in §2.4 already make trouble for the relevant Tolerance premise.

> such-and-such sort according to such-and-such architectural specifications. Here we have all the prehistory we need. The building, this building, is begotten of a particular project rather than a particular scattered part of the earth's material. If someone said: 'The QE II, you know, might have been built of quite a different lot of steel from that which it was actually built of'—and gave his reasons—would it not be absurd to reply: 'In that case it wouldn't have been the QE II at all—the QE II wouldn't have existed—it would have been a different ship of that name.'? (Strawson 1979: 235)[30]

We find these examples quite compelling. Unlike, say, the mass-produced tables churned out by some factory, the objects in Strawson's examples are distinguished from others of the same kind by salient properties that have to do with their origins, but have nothing to do with their originating *matter*. These properties provide a natural response to the question 'What could make a *K* made of entirely different matter be that very *K*?'[31]

Since the objects in Strawson's examples are not tables, they pose no immediate threat to Overlap Essentialism. But they do suggest that it might not be so strange to think that certain tables whose origins are sufficiently distinctive in respects other than originating matter are counterexamples to Overlap Essentialism. Suppose, for example, that IKEA adopted a new system where shoppers are given an empty box and a printed list of parts at the door and must travel around to different stations to collect the planks, screws, and so on. It's tempting to think that a shopper who only bought a single table could have ended up with the same table even if they had filled their box with numerically different parts. Although describing the table as an 'elaborate human construction' seems a stretch, this is also a case where the object's origin is distinctive in many ways that don't involve the matter—the identity of the shopper, the box, the shop, the shopper's purposes . . . —and this makes it easy to get into the mood where denying Overlap Essentialism feels natural.

Indeed, if you can get this far, you might even convince yourself that the table in question is *hypertolerant* (i.e. that it could have been originally made from *any* collection of atoms chemically matching those from which it was in fact originally made). By analogy, if the Old Bodleian could have been made of stone from Quarry A or from stone from Quarry B, it is natural to think that it could have been made by *any* chemically matching collection of atoms, if the collection was first arranged

[30] Strawson makes these points in the course of arguing that 'might have' functions as a past-tense epistemic modal; but he actually makes his point using counterfactuals, which seem pretty good quite independent of Strawson's larger project.

[31] Other examples where originating matter is plausibly pretty unimportant: the A34 could have arguably been made of entirely different tarmac and stone, so long as it still followed its actual route from Winchester to Salford; John's kitchen floor could have been made of linoleum instead of ceramic tiles; a certain tsunami could plausibly have taken place even if all the water in the relevant ocean were different, so long as it had the same triggering earthquake.

so as to constitute a shipment of stone and surreptitiously substituted for the actual stone during its journey from the quarry; maybe the IKEA parts work in the same way.

We are not suggesting that any of these judgements are so strong as to present a forceful objection to some more general theory having Overlap Essentialism as a consequence, if such a theory were available; but as we saw in the previous section, the dominant way of trying to provide such a theory seems highly problematic. Moreover, since Overlap Essentialism taken by itself is quite weak, its own theoretical payoffs are unlikely to be great enough to warrant setting aside plausible counterexamples in the spirit of Strawson's remarks.

Second, consider a plastic table Plasticky that was printed, upside down, by a 3D printer. It is not obvious at what point in the printing process Plasticky started to be inside the printer. But it is pretty natural to say that this was true almost immediately: Plasticky started out as a very thin rectangle of plastic, then gradually became fatter, then acquired short legs, which gradually became longer until finally Plasticky had assumed its final form. (Note that this doesn't require saying that Plasticky was ever a very thin and legless *table*—it might have been concrete before being a table, just as we were concrete before being philosophers.) If we equate '*x* is originally composed by collection of atoms *C*' with '*x* is composed by *C* at the beginning of its concrete existence', this means that Plasticky's originating collection of atoms is some quite small collection—one that by the end of the process composes just a very thin layer at Plasticky's top surface. But it seems quite natural to think that Plasticky could also have been made starting somewhere else, e.g. with its bottom rather than the top surface coming first, in which case there are at least *some* collections from which it could have originated despite the fact that they do not overlap its actual originating collection. So, Plasticky seems prima facie to be a counterexample to Overlap Essentialism.[32]

Indeed, if we accept this, it seems pretty plausible that Plasticky is *hypertolerant*, i.e. that *every* collection of atoms matching the thin collection of atoms that first composed Plasticky could have been Plasticky's originating collection, by being inserted into the 3D printer as the first layer, followed by the rest of Plasticky's actual atoms.[33]

One might hope to circumvent this sort of counterexample by adopting a less straightforward definition of 'originally compose' on which Plasticky counts as

[32] One might try to get out of this by insisting that no table was concrete before being table-*shaped*; but this won't help with tables that are made by first coating the inside of a mould with a thin layer of plastic and then pouring more molten plastic to fill up the cavity inside the mould.

[33] There are many other ordinary cases where it is highly debatable when a given table was created. For example, we seem to be comfortable with the idea that IKEA tables concretely exist, and are tables, prior to being assembled—after all, one can have tables shipped from IKEA. But when else could the table have been created, or become a table? Was it when all of the parts were made, even though at that point the chance that those particular parts would end up in the same box was presumably quite low? Or was the grouping of the parts into a box the crucial thing? There seems to be a lot of vagueness here, and perhaps also context-sensitivity. Chapter 11 will present a framework which can comfortably accommodate such vagueness.

being originally composed by some rather large collection of atoms despite the fact that many of these atoms were not incorporated into Plasticky until Plasticky had already been around for a while. One might for example say that an object is originally composed by a collection just in case it is composed by that collection when it is first *finished*. (This would take some getting used to, given that some tables are never finished and hence by this definition aren't originally composed by any collection of atoms. But for the purposes of raising a puzzle, the operative interpretation of 'originally compose' need not match its ordinary meaning.[34]) But even according to this new definition, it is still not so implausible that some ways of making a table make tables that are counterexamples to Overlap Essentialism, and perhaps even hypertolerant. Suppose a table is made by starting with a flimsy shell which is then filled in and sanded away, so that none of the atoms that were parts of it at the beginning are around when it is finished. It is not absurd to think that the same table could have been made in a circumstance where the initial flimsy shell is the same but the matter that is filled into the shell is entirely different.[35]

A third independent argument against Overlap Essentialism involves paying attention to some of the larger parts of tables, in addition to their atoms. As Barnett (2005) points out, the popular view that endorses Hypertolerance in certain *temporal* Tolerance Arguments (like the Ship of Theseus) also provides the resources for making many originating-matter Hypertolerance claims seem palatable. Suppose that a table, Woody, was made in such a way that at the time of its creation, it was composed of four planks $p_1 - p_4$ and four legs $l_1 - l_4$, which were in turn composed of a certain collection of atoms C. It seems fairly plausible that prior to being put together into a table, $p_1 - p_4$ and $l_1 - l_4$ could each have been subjected to a Ship-of-Theseus-style process whereby their original atoms (those belonging to C) were gradually replaced by atoms not belonging to C.[36] If the Ship of Theseus can survive the replacement of its planks, then it is plausible that at the end of such a process, $p_1 - p_4$ and $l_1 - l_4$ would still be around, but composed of some collection of atoms C' not overlapping C. Suppose that they are at this point assembled into a table, in much the same way as at the actual world. It is easy to believe that such a process would—or at least *could*—produce Woody. The reason

[34] Note that even the earlier definition also allows that there could be objects that have not always been concrete but weren't originally composed by any collection of atoms. If the set of times at which an object concretely exists is open towards the past, there won't be a first moment of its concrete existence; moreover, there may be no collection of atoms that composes it at some time and is such that no distinct collection composes it at any earlier time.

[35] Similarly, one could imagine a case where one wants to end up with a fine ebony table, but since the ebony will be shipped over a long period, one decides to build a table from pine and gradually replace the pine with ebony as it comes in. Plausibly, one could have built the same table if one had all the ebony right at the beginning.

[36] This would be even easier if, instead of being assembled from legs and planks, Woody had been carved directly from a tree, since living organisms are in fact prone to change their atoms over time. Gendler and Hawthorne (2000: n. 23) imagine a table carved from the "Liberty Tree", and point out that it might be considered a counterexample to Overlap Essentialism.

we are initially resistant to the idea that it is possible for Woody to be originally composed of some collection with no atoms in common with its actual originating atoms is, at least in part, that we can't see what could explain its being Woody rather than some other table that was created when those atoms were assembled into a table. But when we focus on the special Theseus-style way of making a table from some atoms, such queries seem far from unanswerable.[37]

Note that if this argument is good, then it can plausibly be adapted to show that Woody is not only a counterexample to Overlap Essentialism, but hypertolerant (with respect to its originating atoms), and indeed hypertolerantly a table. For if the planks and legs are able to survive *some* gradual replacement of their atoms, it seems plausible that they could survive the gradual replacement of their atoms with *any* other atoms of the same types. It would be odd to think that there are some atoms which could be cycled in without destroying the planks and legs, but other atoms of the very same sort which could not, just as it would be odd to think that among all the world's planks of the right sizes, some differ from others in that the Ship of Theseus could have survived having its original planks gradually replaced by them.[38] If this is right, then we can contrive for Woody to be originally composed by *any* collection of atoms matching its actual originating atoms, by arranging for those atoms to be Theseused into the planks and legs before their assembly.

By contrast with the earlier arguments in this section (the Strawson cases and the 3D-printed table), the Barnett-style argument has some prospect of motivating the claim that *all* tables are hypertolerantly tables. The more resistant cases will be ones where there is no intermediate stage of disassembled parts available for Theseus operations. Consider for example an austere modernist coffee table that is stamped in one go, along with several other tables, out of a large sheet of iron, by a machine that simultaneously folds down the ends of the stamped rectangle. But even here, one might possibly convince oneself that the table in question was hypertolerantly a table by considering a possible case where instead of stamping the table out in one go, one stamps out three separate rectangles (the top and

[37] Barnett uses a similar kind of argument to argue against a principle like Overlap Essentialism that refers to collections of planks and legs instead of collections of atoms. For, arguably, we could assemble Woody out of some collection of four other planks and four other legs by first Theseusing Woody's actual atoms into those planks and legs. There is no obvious tension in holding that Woody could both have been composed of the same atoms and completely different planks and legs, and also could have been composed of completely different atoms and the same planks and legs. The view that both of these scenarios are possible does lead to an interesting puzzle about what would happen if, having Theseused different atoms into the planks and legs, we assembled the planks and legs into one table while assembling the original atoms (Woody's actual atoms) into a different table. One might think that Woody could not be made of the same atoms while a different table was made of the planks and legs; or vice versa; or one might think that both are possible.

[38] As discussed in §2.4, one might want to make an exception for atoms in distant parts of the universe that could only have ended up in the right neighbourhood by being moved in some miraculous or remarkable way.

two sides), performs Theseus operations on the three rectangles, and only then welds them together. There are a variety of moods one can get into here, but one of them seems to be quite friendly to the claim of Hypertolerance even in this case. Certainly Hypertolerance is not *obviously* false.

The counterexamples to Overlap Essentialism are prima facie compelling, but there may be some still weaker principle in the vicinity that is true. For example, one could try defining some notion of an atom being "involved in the origin of" a table, where the atoms involved in a table's origin include not only those that originally compose it, but those that become parts of it later during its construction; those that originally compose its intermediate parts such as legs, cells, and molecules; those that originally compose certain larger objects from which the table is made, such as a sheet of metal or a tree; those that originally compose intermediate parts that those intermediate parts had before being assembled into the table; and perhaps even those that at various times compose the carpenters, saws, stamping machines, and other relevant bits of the environment. With such a notion in hand, one could weaken Overlap Essentialism to merely rule out the possibility that a given table is such that none of the atoms involved in its origin are among those that are actually involved in its origin.[39] (Perhaps something along these lines will take care of Strawson-style cases; or perhaps we will need to address them separately by restricting the principle to mass-produced tables.) But we have little appetite for the project of weakening Overlap Essentialism to avoid the putative counterexamples of this section, just as we had little appetite for the project of weakening Sufficiency to avoid the counterexamples in the previous section. In both cases, the problem is that even if the principle that results of the modifications is both defensible and incompatible with some interesting Hypertolerance claims, it will likely be too complex to be of much use as a *starting point* for arguing against such claims.

So far, then, things are looking pretty good for the strategy of dealing with Tolerance Puzzles by accepting Hypertolerance. The dominant argument against Hypertolerance in the literature turns on a conflict between it and Overlap Essentialism, a principle which seems counterexample-prone, and the primary arguments for which turn on dubious sufficiency premises. Moreover, as we saw in this section, some origin-theoretic Hypertolerance claims can be motivated quite independently of Tolerance Arguments, e.g. via Barnett-style arguments.

Nevertheless, there are many particular origin-theoretic Tolerance Arguments where Hypertolerance retains an air of absurdity, even in the light of all these considerations. For example, we could consider a family of properties of the form

[39] Note that there seems to be little prospect of being able to derive this weakening of Overlap Essentialism from some sufficiency argument. For one thing, given how permissive the concept of "originating atoms" is, it will plausibly be possible for two simultaneously produced tables to have exactly the same collection of originating atoms, whereas Kripke's argument turns crucially on the premise that no two tables originate in the same hunk of matter.

being one of four tables stamped all in one go from such-and-such portion of this circular metal sheet, with closeness requiring extensive overlap of the relevant portions. Hypertolerance here will entail that each of the four tables could have emerged via the same process from any of the relevantly sized portions of the sheet, including those that in fact formed the other tables. This Hypertolerance claim still seems absurd, though so far we have not succeeded in finding any general *theoretical* grounds for resisting it. In the next chapter, we will try to do better.

6
Hypertolerance and Supervenience

6.1 Supervenience Principles

In the previous chapter, we considered "sufficiency principles" along the following lines: 'Whenever a table could be made of some matter according to a certain plan, it is necessary that any table made of that matter according to that plan is identical to that table'. As we saw, this principle and various of its refinements are subject to counterexamples. Perhaps some descendants of these principles will avoid all clear-cut counterexamples; but these descendent principles will likely be too complex to prompt much confidence.

Nevertheless, there seem to be the beginnings of a good idea underlying the impulse to search for a counterexample-proof sufficiency principle. The motivating thought is that being identical to a given table (pyramid, painting . . .) is not brute and inexplicable: when a particular table *x* gets made, there will be *some* collection of more basic facts about the details of how everything unfolded which "makes it the case that *x* was made", in a sense that at least requires that it is metaphysically necessary that if those facts all obtain, *x* is made. The intuitive resistance we feel to various Hypertolerance claims seems at least in part based on a thought along these lines. To explain our sense of absurdity, we want to say things like 'What could make a thimble-sized pyramid be the *Great Pyramid*?'; 'What could make a painting with the spatial arrangement of colours characteristic of *The Scream* be the *Mona Lisa*?'; 'What could make a table in whose history none of these atoms played any role at all be *this very table*?' As we consider the counterexamples which doom the simpler sufficiency principles, we learn that the factors which determine which tables (pyramids, paintings) get made are more complex than we might have originally thought, so that the collections of relevant "underlying" facts may need to be quite large. But there is no reason this discovery should shake our confidence in the motivating vision about how things like tables fit into reality.

To our minds, though, the best way of making good on this idea is to relinquish the strategy of adding a long list of qualifying clauses to a sufficiency principle built around some specific word of interest like 'table', but to articulate the background motivating thought in a more general way.[1]

[1] Salmón (Salmon 1981: 280) also suggests that some of the argumentative role of the sufficiency principles could be taken over instead by a supervenience principle pretty much like the one we will introduce below. However, the argumentative role he considers is that of motivating premise (6) in

The Bounds of Possibility: Puzzles of Modal Variation. Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri, Oxford University Press. © Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri 2021.
DOI: 10.1093/oso/9780192846655.003.0007

One might worry that the driving thought will turn out to just be our old friend, anti-haecceitism (see §5.1). Weak Anti-haecceitism entails that any truth about a particular object—for example, the truth that a particular table was created—is necessitated by some qualitative truth; so it would certainly license the quest for a nontrivial Weak Sufficiency principle that would tell us something informative about *which* qualitative truth does the job in any particular case. And the stronger principle Anti-haecceitism entails that any *possible* proposition about a particular object—for example, the proposition that a certain table was originally formed from some particular collection of atoms that it was not in fact originally formed from—is necessitated by some possibly true qualitative proposition; so it would likewise license the quest for a nontrivial Sufficiency principle. An implicit or explicit commitment to one or other form of anti-haecceitism could thus provide a motivation for the intuition that the fact that a particular table was created should not be "brute and inexplicable". From this standpoint, this is just a special case of the general principle that *facts about particular objects* are not "brute and inexplicable", but rather are explained by—in a sense that requires being necessitated by—the qualitative facts.

But sufficiency principles like those considered in §5.2 aren't of the right sort to be motivated by anti-haecceitism, since the postulated sufficient conditions include facts about particular hunks of matter which are, prima facie, just as non-qualitative as the facts about particular tables which they allegedly necessitate. And the intuition about tables that seems to lie behind these principles does *not*, in fact, seem to generalize to just any object. It does not seem absurd to suppose that when two photons interact and an electron and positron are emitted, the fact that this particular electron and positron are emitted *is* brute, in the relevant sense—if you like, God is free to choose to have any possible electron he pleases pop up on this occasion. What makes the corresponding view about the construction of tables absurd is our conviction that tables are not *fundamental* objects; electrons, by contrast, are not obviously non-fundamental.[2]

Still, the general *forms* of the anti-haecceitism principles seem promising. They ascribe a special status to the property of being a *qualitative* proposition: Weak Anti-haecceitism says that every truth is necessitated by a truth that has it; Anti-haecceitism says that every possible truth is necessitated by a possible truth that has it. In trying to articulate the way in which facts about which particular tables are created seem to depend on something more basic, we could consider principles of a similar form, but which give some property of propositions other than being

his "Four Worlds Paradox"—a premise which (as we saw in note 27 of Chapter 2) is redundant given the motivations for the other premises. He does not explicitly discuss how the supervenience principle could help to motivate Non-hypertolerance (and thus his premise (3)).

[2] The success of quantum field theory tends to suggest that electrons are in fact rather analogous to waves in the sea, and hence non-fundamental; but the point is just that we are open to there being a category of fundamental objects that work very differently from tables in the relevant respects.

qualitative a starring role. We will choose the property of being a proposition *about microphysics*—for short, a *microphysical* proposition—so that the resulting principles capture a specifically physicalistic picture of reality, for which we have some sympathy. This gives us the following two principles:

Weak Microphysical Supervenience Every true proposition is necessitated by some true microphysical proposition.

Microphysical Supervenience Every possibly true proposition is necessitated by some possibly true microphysical proposition.

The modality here is metaphysical.[3]

The logical relationship between these principles is parallel to the relationship between their anti-haecceitist counterparts. Given the plausible principle that when every proposition in a certain collection of propositions is microphysical the proposition that some member of that collection is true is also microphysical, Microphysical Supervenience is equivalent to the claim that every proposition is necessarily equivalent to a microphysical proposition, and thus implies Weak Microphysical Supervenience.[4] Given S5 and the claim that every microphysical proposition is necessarily microphysical, Microphysical Supervenience is equivalent to the necessitation of Weak Microphysical Supervenience, and thus also to its own necessitation; but given weaker modal logics, the necessitation of Weak Microphysical Supervenience does not imply Microphysical Supervenience.[5]

[3] Given the principles (which we accept) that negations of microphysical propositions are microphysical and that *every proposition in C is true* is microphysical whenever *C* is a collection of microphysical propositions, Weak Microphysical Supervenience is equivalent to the claim that every possible world that differs from the actual world on some proposition differs from it on some microphysical proposition, and Microphysical Supervenience to the claim that any two possible worlds that differ on some proposition differ on some microphysical proposition. We use the labels 'weak' and 'strong' because of the analogy with weak and strong anti-haecceitism (§5.1) and weak and strong sufficiency principles (§5.2). There is no connection to Kim's (1984) use of 'weak supervenience' and 'strong supervenience' to refer to two different relations between sets of properties.

[4] See the corresponding discussion of Anti-haecceitism in §5.1.

[5] In particular, a denier of the 5 axiom could combine the denial of Microphysical Supervenience with the necessary truth of Weak Microphysical Supervenience by claiming that any two possible worlds that agree on all microphysical propositions but disagree on some other proposition are impossible relative to each other. To get a feel for how this might be motivated, consider the Chunky example from §5.2, where in the actual world a thick bucket-shaped table, Chunky, is made by combining two hunks of wood, Inner and Outer. It is very natural to think that Chunky could have been made just of Inner and could also have been made just of Outer. But it also seems odd to imagine that there are any hunks such that Chunky could only be made of them if certain other hunks are not made into tables. Putting these thoughts together, we get that it is possible for Chunky to be made of Inner while a different table is made of Outer, and also for Chunky to be made of Outer while a different table is made of Inner. And while we could imagine that the two sorts of possible worlds would inevitably differ microphysically in some way or other, it seems hard to think of a principled basis for this. While this line of thought can of course be resisted in several ways, it might convince some friends of Weak Microphysical Supervenience that the proposition that Chunky is made of Inner while a different bucket is made of Outer is a counterexample to Microphysical Supervenience.

We will not attempt to define what it means for a proposition to be microphysical. The rough idea we have in mind is that a microphysical proposition is one that could in principle be expressed in a language whose only nonlogical predicates are names for possible microphysical objects (elementary particles, spacetime points, non-concrete possible particles and spacetime points, maybe even atoms and molecules; but not paintings and tables) and predicates expressing qualitative microphysical properties and relations (such as geometric and field theoretic relations). But we will assume that when every proposition in some collection is microphysical, the propositions that every member of that collection is true and that some member of that collection is true are both microphysical as well (even if the collection is infinite), so this gloss needs to be taken with a grain of salt.[6]

We do not claim that either of these supervenience principles fully captures the idea that everything "obtains in virtue of" the microphysical. The last two decades have seen a burgeoning literature devoted to theorizing about the notion of *grounding* in play in such claims, a recurring theme in which is that modal principles like Microphysical Supervenience are not strong enough to capture the most important thought in the vicinity.[7] But whether or not one of the supervenience claims fully captures that idea, their truth has (as we will see) some interesting upshots for Hypertolerance; and moreover, we are not aware of any important further lessons for Tolerance Puzzles to be learned by paying attention to those grounding-theoretic claims that stay closest to the motivating "what could make such-and-such be the case?" intuitions.

Even without grounding-theoretic add-ons, both of our microphysical supervenience principles are deeply controversial. While we find them appealing, a full-dress, abductive defence of them will not be needed. As we will discuss in §6.4 below, many of the objections they face are not really relevant to the dialectic about Tolerance Puzzles, since the objections also suggest ways of weakening them which would could equally well play the dialectical role of making trouble for (certain) Hypertolerance claims. But before we discuss the objections and weakenings,

[6] Our supervenience principles articulate the idea that the description of the whole *world* in microphysical terms settles everything about the world. If one helps oneself to the notion of a microphysical *property*, one can formulate stronger principles which capture the idea that the description of any given *object* in microphysical terms settles everything about that object:

Weak Property Microphysical Supervenience Every instantiated property is necessitated by some instantiated microphysical property.

Property Microphysical Supervenience Every possibly instantiated property is necessitated by some possibly instantiated microphysical property.

These principles are subject to certain objections which do not apply to Microphysical Supervenience. For example, one might worry that a statue and the lump of clay that makes it up, or the song 'My Generation' and the novel *Bleak House*, share the same microphysical properties (e.g. *being such that there are electrons*), although they do not share all properties. Such objections are far from decisive, but responding to them would require saying a lot more about the notion of microphysicality as it applies to properties, especially properties instantiable by non-microphysical objects.

[7] See Fine 2012 (41).

we should first see what can be done with the supervenience principles in their current form.

As a warmup to seeing how Microphysical Supervenience might be relevant to Hypertolerance claims, let us consider the following radical thesis which we discussed in §5.1:

Extreme Anti-essentialism For any distinct objects $x_1, \ldots, x_n$ and any qualitative relation R, if it is metaphysically possible for R to be instantiated by distinct objects, it is metaphysically possible for R to be instantiated by $x_1, \ldots, x_n$ in that order.

Extreme Anti-essentialism is a natural position to adopt for someone whose general strategy for handling Tolerance Puzzles is to embrace Hypertolerance. This strategy will require accepting a host of prima facie surprising *de re* possibility claims, which would arguably add up to a weighty abductive case for some simple generalization like Extreme Anti-essentialism under which they could all be subsumed.

But on the assumption that there are only finitely many microphysical objects, and at least two qualitatively discernible non-microphysical objects (e.g. David Lewis and a poached egg), Weak Microphysical Supervenience plausibly entails that Extreme Anti-essentialism is false. Let a and b be two non-microphysical objects, and let F be a qualitative property instantiated by b but not by a. Let the microphysical objects be $x_1, \ldots, x_n$, and let R be a qualitative relation instantiated by $x_1, \ldots, x_n$ which necessitates every other qualitative relation they instantiate. Now consider the relation *being* $y_1, \ldots, y_{n+1}$ *such that* $R(y_1, \ldots, y_n)$ *and* $F(y_{n+1})$. It is qualitative (since R and F are) and it is instantiated by distinct objects, namely $x_1, \ldots, x_n, b$. So Extreme Anti-essentialism implies that it is possibly instantiated by the distinct objects $x_1, \ldots, x_n, a$: possibly, $R(x_1, \ldots, x_n)$ and $F(a)$. But $R(x_1, \ldots, x_n)$ is true. Moreover, since $x_1, \ldots, x_n$ are the only microphysical objects, it is very plausible that every microphysical proposition is necessarily equivalent to one predicating some qualitative relation of $x_1, \ldots, x_n$, in which case $R(x_1, \ldots, x_n)$ necessitates every true microphysical proposition. So Microphysical Supervenience implies that $R(x_1, \ldots, x_n)$ necessitates every truth, including the truth that $\neg F(a)$: contradiction.

(The assumption that there are only finitely many microphysical objects is of course quite dubious.[8] But the need for this assumption is an artefact of the weakness of our formulation of Extreme Anti-essentialism, which tells us about the qualitative roles that could be played by *finite* sequences of objects, but says nothing analogous about infinite sequences. The thought behind Extreme

[8] Especially since we are allowing non-concrete possible electrons to count as microphysical objects.

Anti-essentialism seems to generalize to infinite as well as finite sequences. If we could find a way of articulating that thought, it should conflict with Microphysical Supervenience without the need for any finitude assumption.[9])

6.2 Supervenience and Tolerance Puzzles

Even if we have jettisoned all intuitive resistance to the idea that certain objects could have been radically different from the way they actually are, Microphysical Supervenience (whether in its weak or strong form) imposes some principled limits. But none of the Hypertolerance claims we have considered so far is strictly inconsistent with Microphysical Supervenience. In this section, we will take up the task of formulating Tolerance Arguments whose Hypertolerant conclusion *is* inconsistent with Microphysical Supervenience, and whose premises are as plausible as possible. These Tolerance Arguments will involve "fine-grained" families of properties, each of which inter alia includes a complete specification of the microphysical lay of the land. Proponents of Microphysical Supervenience will need to develop some strategy for resisting such fine-grained Tolerance Arguments. And once they have done so, they may well find that the same strategy comfortably extends to Tolerance Arguments for Hypertolerant conclusions that are not strictly inconsistent with Microphysical Supervenience.

In formulating fine-grained Tolerance Arguments, the following definitions will be useful:

> Property F *incorporates* proposition p = F necessitates *being such that p*.
>
> p is a *microstory* = p is a possibly true microphysical proposition that necessitates every microphysical proposition or its negation.
>
> F is *microspecific* = F incorporates some microstory.
>
> F and G are *conspecific* = there is a microstory that both F and G incorporate.[10]

It follows from Microphysical Supervenience that if F and G are incompatible but conspecific, then anything which could be F couldn't be G.[11] Similarly, it

[9] A natural higher-order surrogate for a non-repeating infinite series of objects is a rigid binary relation R that is *well ordered*. (That is: Rxz whenever Rxy and Ryz; Rxx whenever Rxy or Ryx; and whenever there is a y such that Xy and Ryy, then there is a y such that Xy and for all z, if Xz and Rzz, then Ryz.) There is also a higher-order regimentation of the relation *having the same order-type*. Given this, the desired generalization of Extreme Anti-essentialism is as follows: for any rigid well-orderings R and S of objects that have the same order type, and any qualitative property Q of binary relations, if QS, it is possible that QR.

[10] Note that given the existence of a microstory, any impossible property is automatically micro specific, and any two such properties are automatically conspecific; so conspecificity does not entail consistency.

[11] For suppose F and G are incompatible but both incorporate a microstory p, and x could be F. By Microphysical Supervenience, there is a possibly true microphysical proposition q that necessitates that

follows from Weak Microphysical Supervenience that if *F* is microspecific and instantiated, then the only objects that could be *F* are those that are in fact *F*.[12]

So, if we want to rely on Microphysical Supervenience, our strategy will be to look for a closeness relation that holds between microspecific properties, such that some microspecific property instantiated by some object *x* is *ancestrally close* to two incompatible but conspecific properties. In any Tolerance Argument constructed using such a closeness relation, metaphysical modality, and a "kind" predicate *K* applying to some such *x*, Hypertolerance will imply that any such *x* could have each of the two incompatible but conspecific properties, which is ruled out by Microphysical Supervenience. Meanwhile, if we only want to rely on Weak Microphysical Supervenience, we will need a closeness relation that holds among microspecific properties such that for some *K* object *x*, a microspecific property not instantiated by *x* but instantiated by some other object is ancestrally close to one that *x* does instantiate. In a Tolerance Argument constructed using such a closeness relation, Hypertolerance will imply that any such *x* could have had an instantiated microspecific property that it doesn't in fact have, which is ruled out by Weak Microphysical Supervenience.

Let's set about trying to construct a compelling Tolerance Argument involving a closeness relation on microspecific properties with the features required to generate the conflict between Hypertolerance and Microphysical Supervenience. We had better make sure that our closeness relation isn't such as to make Tolerance already incompatible with (the relevant form of) Microphysical Supervenience: thus, the closeness relation can't directly connect a property instantiated by some *K* to two conspecific but incompatible properties, or to a microspecific and instantiated but not coinstantiated property.[13] And we also want to make sure that Non-contingency can be motivated by security considerations of the sort canvassed in §3.3, rather than merely by dubious Soritical "small differences can't matter" intuitions. This will require specifying the closeness relation in such a way

x is *F*. Since *F* incorporates *p*, *q* necessitates *p*. But *p*, being a microstory, necessitates either *q* or not-*q*. It can't necessitate not-*q*, since then *q* would necessitate not-*q* which conflicts with the possible truth of *q*. So *p* necessitates *q*, and thus also necessitates that *x* is *F*. By the incompatibility of *F* and *G*, it follows that *p* necessitates that *x* is not *G*. Thus *G*, which incorporates *p*, necessitates *being such that x is not G*. So it is necessary that if *x* is *G*, *x* is not *G*; hence it is impossible for *x* to be *G*.

[12] For suppose *F* incorporates a microstory *p*, *x* is *F*, but *y* is not *F*. By Weak Microphysical Supervenience, there is a true microphysical proposition *q* that necessitates that *y* is not *F*. Since *x* is *F*, *x* is such that *p* is true, so *p* is true. Since *p* is a microstory and doesn't necessitate not-*q* (since *q* is true), it necessitates *q*, and thus necessitates that *y* is not *F*. So, necessarily, if *y* is *F* *y* is not *F*; so it is impossible for *y* to be *F*.

[13] So, for example, it won't work to pick a closeness relation where *being originally composed of atoms C while microstory p is true* is close to *being originally composed of atoms D while microstory q is true* whenever *p* is sufficiently similar to *q* and *C* sufficiently overlaps and chemically matches *D*. Since every microstory is sufficiently similar to itself, Tolerance interpreted using this closeness relation entails that if a table *x* is originally composed of *C* while a microstory *p* is true, *x* could instead have been originally composed of any sufficiently overlapping, and chemically matching *D* while *p* was true, which is impossible since by Weak Microphysical Supervenience *p*, being a true microstory, necessitates that *x* is originally composed by *C*.

that none of the properties along the path of interest involves being "monstrous" in a way that would undermine the security-based case for Non-contingency.

For reasons that we already encountered in our attempt to flesh out Chisholm's Adam / Noah argument in §5.1, finding a route that steers clear of monstrousness is often a rather tricky matter. Many pairs of objects in the actual world, even if they are both paradigmatic members of some fairly specific kind (e.g. pairs of people), are such that any sufficiently gradual route from a highly specific property instantiated only by the first to a highly specific property instantiated only by the second will involve properties that entail being monstrous, or at least somewhat non-paradigmatic, in ways that might (if we knew about them) reasonably shake our confidence in the truth of the relevant Tolerance premise. But in some cases, we may be able to avoid such security-undermining monsters by taking moderately sized rather than tiny steps. And in some other cases, we can make the route as gradual as we please without encountering anything even remotely monstrous along the way.

Here is an example of the latter sort, modelled on a case of Williamson's (1990: 127f.). A certain craftsman simultaneously makes two semicircular earrings by cutting a circular metal disc in half 'with a device that simultaneously punches an attachment for the ear out of each half'. Suppose that the cutting tool is a T-shaped device: the stem of the T is a wooden handle and the crossbar is a blade, sharpened at the top. In the actual world, a craftsman places the disc flat on a workbench, grasps the cutting tool in his fist, blade downwards, and pounds it down onto the disc, with one end of the blade pointing towards him and the other pointing away from him (see Figure 6.1).[14] He thereby makes two earrings: Lefty on his left and Righty on his right. To be painfully exact: suppose that one end of the blade is so close to the craftsman that the radius that bisects Lefty is, to the nearest degree, 90° of the way clockwise around the disc from the point on the disc closest to the craftsman, while the radius that bisects Righty is, to the nearest degree, 270° clockwise around the disc from that point. Let p_0 be a true microstory; let E_0 be the property of being an earring that at its creation was centred 90° clockwise from the craftsman to the nearest degree and such that p_0 is true, and let F be the property of being an earring that at its creation was centred 270° to the nearest degree clockwise from the craftsman and such that p_0 is true. In fact, Lefty has E_0 and lacks F and Righty has F and lacks E_0. Given Weak Microphysical Supervenience, p_0 necessitates that this is the case. Hence it is metaphysically impossible for Lefty to have F.

Our task is to connect E_0 to F via a series of intermediate properties in which each step is close to its predecessor, in a sense of 'close' that makes Persistent

[14] In Williamson's version of the story, the craftsman dies before making the cut. We will discuss this case and the morals Williamson draws from it in §7.2.

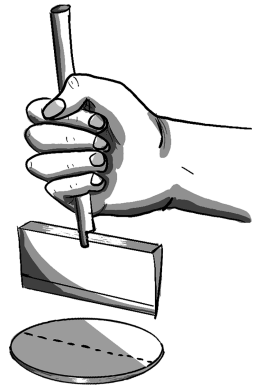

Fig. 6.1 The earrings about to be made.

Closeness unproblematic and makes both Tolerance and Non-contingency plausible. To do so, we consider a series of possible alternative microstories $p_1, \ldots, p_{180}$ similar in all respects up to the time of the cut to p_0, except for the orientation of the cutting tool in the craftsman's fist, which is rotated by 1° clockwise at each step.[15] We let E_i be the property of being an earring that at its creation was centred $i + 90°$ clockwise from the craftsman, to the nearest degree, while p_i is true. So, halfway along, we have a microspecific property E_{90} that entails being an earring made by a craftsman holding a cutting tool with one end of the blade pointing to his left and the other pointing to his right, on the opposite side of the blade from the craftsman's body. The final property E_{180} is similar to F in that it entails being an earring made towards the right of the cut by a craftsman holding a cutting tool with one end of the blade pointing towards him and the other pointing away from him. Unfortunately, we are still not done: E_{180} isn't exactly the same as F, since its microstory p_{180} differs from p_0 as regards the positions of the microspecific atoms composing the cutting tool. For example, when F entails being such that a certain atom a is located in the part of the blade further from the craftsman, E_{180} will entail being such that a is located in the part of the blade closer to the craftsman. But we can bridge the gap from E_{180} to F by another series of microspecific properties $G_1, \ldots, G_n$, where at each step we swap an atom on one side of the cutting tool with an atom on the other side of the cutting tool until we have finally moved

[15] Since our goal is to avoid any sort of "monstrousness" which would make counterfactual errors about Tolerance unthreatening to the epistemic standing of our actual Tolerance judgements, we prefer that all of $p_1, \ldots, p_{180}$ should be physically possible. Thus we can't keep *all* the microphysical facts except for the locations of the microscopic parts of the cutting tool *exactly* the same. For example, the pattern of photons bouncing off the cutting tool will need to be varied in tandem with its position, and if there are air molecules around, we will need to move them to make way for the cutting tool.

each atom into the position it needs to be in for p_0 to be true.[16] (If the cutting tool was perfectly symmetric, the microstories incorporated by $G_1, \ldots, G_n$ can all be qualitatively isomorphic, differing merely as regards which microphysical objects play which roles. If not, we can gradually reverse any minor qualitative asymmetries in the cutting tool at the same time as we swap the atoms.)[17]

The closeness relation we are interested in is the relation of being one step away in the series of properties we have just described. Let's consider the Tolerance Argument using this closeness relation, with K as 'earring' and metaphysical modality: call it the Earring Argument, and its premises Earring Tolerance, Earring Non-contingency, etc.[18] Earring Tolerance follows from two claims: first, every earring made centred $n°$ clockwise could have instead been made centred $n + 1°$ clockwise while the microphysical facts were similar in all respects up to the time of the cut, except that the cutting tool was rotated 1° clockwise in the craftsman's fist. Second, every earring made centred $n°$ clockwise could have still been made centred $n°$ clockwise while any microstory accurate in every way except for the locations of two atoms in the cutting tool was true. Even though this Tolerance judgement is somewhat stronger than the ones we have been focusing on up to now, it certainly seems pretty alien to suppose that any earrings are living on the edge in the peculiar way they would have to be for it to be false. Moreover, the sense of bizarreness is exactly the same when we change our perspective from the actual world of the story to, say, a world where Lefty instantiates E_{90}. Earrings made by craftsmen holding cutting tools perpendicular to their body are no less paradigmatic than those made by craftsmen holding cutting tools with one end of the blade pointing towards them; the security-based motivation for Earring Non-contingency is thus as strong as it ever is.

The case of the earrings was carefully designed to provide a gradual path from a microspecific property instantiated by one object to a microspecific property instantiated by some other object (and not by the first object), without bringing in security-undermining oddities along the way. It is helped along by certain approximate symmetries in the setup. But some of the symmetries aren't really needed. For example, it is fine if the earrings aren't semicircles, but are, say, rectangles, made by cutting a metal square in half. Even if we insist that all the properties in the field of our closeness relation should entail *being rectangular*—perhaps because we think those earrings couldn't possibly have been triangular—we can modify the rotational part of the sequence so that not only the cutting tool, but also the metal square beneath it, are gradually rotated through 180°. At the

[16] If in tracing the path to p_{180} we also ended up with different positions for air molecules and the like, we can swap those back in the same way.

[17] We could instead have imagined a series of cases where two kinds of change—rotation of blade and swaps of atoms within the blade—happen at once.

[18] Everything we will say about this argument will also apply to the analogous *de re* Tolerance argument based on the singular term 'Lefty'.

end of this phase, we have a microspecific property that entails being made on a craftsman's right, out of (approximately) the same atoms that Lefty is made of in the actual world. To get from this microspecific property back to one actually instantiated by Righty, we can now start swapping the atoms back to their positions in the actual world, simultaneously reversing any slight asymmetries between the two sides of the square. The only difference between this and the original earring case is that before we only needed to swap the atoms of the cutting tool, since we didn't rotate the disc along with the cutting tool.

One lesson of these examples is that if we could successfully carry out the task Chisholm seemed to be attempting—of constructing a compelling Tolerance Argument in which Hypertolerance is inconsistent with modal anti-haecceitism—then it should be straightforward to adapt this argument to make its conclusion be inconsistent with Microphysical Supervenience. If it is possible for a certain *K* object *x* to still be *K* while instantiating the qualitative role in fact instantiated by some other object, then there must be some microstory *M* such that it is possible for *x* to be *K* with that qualitative role while *M* is true. Whatever this *M* might be, it must agree with the true microstory as regards qualitative propositions. So, we can get from *M* back to the true microstory by a sequence of tiny changes, where at each change we simply swap the roles played by two microphysical objects. While Tolerance judgements according to which any *K* playing a certain qualitative role could still have been a *K* playing that role while any two microphysical objects swapped roles aren't exactly *everyday* ones, they certainly seem quite appealing, and their appeal is not just a manifestation of the temptations of Sorites-style reasoning. Given that ordinary tables, earrings, and people are evidently capable of still being tables, earrings, or people even when the microphysical facts are quite substantially different in any of many different ways, it is hard to believe that there are certain *minute* changes, involving just two microscopic objects, which would suffice to prevent some actual member of any of these kinds from belonging to that kind. And insofar as we find this thought compelling, the security considerations will in the usual way push us towards the idea that it is necessarily true if true.

So far, we have been considering Tolerance Arguments whose conclusions clash with Microphysical Supervenience where we try to make Tolerance as hard to resist as possible by using a very demanding closeness relation that involves tiny microphysical steps. But we can also make arguments of this sort that involve much bigger steps. Let's fill out the bucket scenario from §2.1 by supposing that there was an indeterministic microphysical process in the brain of the bucket-maker, at a time *t* shortly before the creation of Flimsy, which could easily have led to any of four outcomes—making two flimsy buckets, one from *A* and one from *B*; making a flimsy bucket from *A* and discarding *B*; making a flimsy bucket from *B* and discarding *A*; or making a single thick-walled bucket using both *B* and *A*. The old closeness relation was just the relation of being adjacent in the following sequence of five properties:

H_1 Being a bucket made of A while a different bucket is made of B.
H_2 Being a bucket made of A while B is discarded.
H_3 Being a fat bucket made of both A and B.
H_4 Being a bucket made of B while A is discarded.
H_5 Being a bucket made of B while a different bucket is made of A.

Whereas the field of the old closeness relation comprised just these five properties, the field of the new one will comprise all of the enormously many microspecific properties derived from them by conjoining them with *being such that p* for some microstory p. We can define it as follows:

> G is close in the new sense to F := For some properties F' and G' such that G' is close in the old sense to F, and some microstories p and q which match up to t, F is *being* F' *and such that p is true*, and G is *being* G' *and such that q is true*, and both F and G are such that it is nomologically possible for them to be instantiated.

Call the Tolerance Argument based on this closeness relation (with metaphysical modality and the singular term 'Flimsy') the Fine-grained Bucket Argument. Fine-grained Bucket Hypertolerance is inconsistent with Microphysical Supervenience for the reason explained above: a property actually instantiated by Flimsy, namely *being* H_1 *while p is true* where p is a true microstory, is ancestrally close in the new sense to another property not instantiated by Flimsy but instead by the bucket made of B, and incorporating the same microstory, namely *being* H_5 *while p is true*. Fine-grained Bucket Tolerance says that if Flimsy has H_i (where $i \in \{1, 2, 3, 4\}$), then it could not only have had H_{i+1}, but could have had it compatibly with any nomologically microstory matching the actual microstory up to t, so long as that microstory is compatible with *something* having H_{i+1}. Given that in fact Flimsy has H_1, this means that every nomologically possible microstory that matches actuality up to t and is compatible with there being a bucket made of A while B is discarded is compatible with *Flimsy* being made of A while B is discarded.[19] This rules out various deeply implausible claims that were compatible with the old, coarse-grained Tolerance premise, such as that Flimsy could be made while B was discarded with the bucket-maker's left hand but not made while B was discarded with the bucket-maker's right hand (assuming these are both necessitated by some microstories matching actuality up to t). Reflecting on the implausibility of such suggestions, it seems that if we were persuaded of Tolerance in the old argument, and we fill out the stipulated details of the example in such a way

[19] Given Microphysical Supervenience, any proposition compatible with a microstory is also necessitated by that microstory. But we should be careful not to just presuppose Microphysical Supervenience in explaining what Tolerance says.

that the differences between the different nomologically possible microstories where a given one of $H_1 - H_5$ is instantiated are not too great, we will find Fine-grained Bucket Tolerance hard to resist. And insofar as this is compelling, Fine-grained Bucket Non-contingency can be supported by the same kind of security considerations as its coarse-grained predecessor.

The strategy can be applied to many Tolerance Arguments whose conclusions are consistent with Microphysical Supervenience. What we will do is fine-grain the properties in the relevant family by making them microspecific, motivating Tolerance for the new family by appeal to the thought that if the relevant object could have had one of the original unspecific properties, it could have had it consistently with all the relevant microphysical realizations of that unspecific property. This kind of thought can be compelling in a particular case even when we are not in a position to write down a counterexample-proof general sufficiency principle which exhaustively specifies which aspects of the microphysical supervenience base are the relevant ones. However, it is a thought that might be resisted; we will have more to say about what such a resistance strategy would look like in §6.4 below.

6.3 Coarse-Grained Hypertolerance

If we hold on to Microphysical Supervenience, we will have to reject at least one premise in each fine-grained Tolerance Argument whose conclusion conflicts with it. Such a rejection is compatible with accepting a wide variety of coarser-grained Tolerance Arguments as sound, such as those considered in Chapter 2. And this is not a terribly unprincipled suggestion. After all, as we have seen, many of the coarse-grained Tolerance premises require little to no philosophical extrapolation from ordinary modal practice; the fine-grained Tolerance premises, by contrast, turn on more theoretically motivated thoughts about the irrelevance of certain fine microphysical details. Nevertheless, the combination of coarse-grained tolerance with a lack of fine-grained tolerance takes some getting used to. We shall be exploring its prospects in this section.

Given Microphysical Supervenience and the assumption that no two earrings can be composed by the same atoms, there is a function that takes a collection of atoms, and a complete microstory that necessitates that the atoms in that collection compose an earring, and returns the possible earring that they necessarily compose if that microstory is true.[20] Given Origin Hypertolerance for earrings, we know that for every possible earring that could be composed by atoms with a given distribution of chemical formulae, and every collection with that

[20] We also need to assume the Barcan Formula.

distribution, there is some microstory such that the function delivers that possible earring given that collection and microstory as inputs. There is very little prospect for a principled account from which we could read off the behaviour of the function. Suppose Lefty and Righty are two earrings which are made together by a Williamson-style disc-cutting process and happen to contain equally many atoms with each atomic number, and *C* is some distant collection of atoms with the same distribution of chemical formulae whose physical relations to *A* Lefty and Righty *B* are roughly symmetric. Some microstories will necessitate that *C* composes Lefty *A*, while others necessitate that *C* composes *B* Righty, and still others necessitate that *C* composes some earring distinct from both Lefty *A* and *B* Righty (since each of the enormously many possible earrings that could be composed by a collection with *C*'s distribution of chemical formulae must, by Origin Hypertolerance, be possibly composed by *C*). There is no prospect of telling a principled story from which we could read off which of these categories a given microstory belongs to. The view will have to have it that the identity of the earring composed by *C* depends on the microphysical facts in ways that are largely inscrutable.

But perhaps this lack of a principled account is something we can learn to live with. When it comes to vagueness, we have learnt to make our peace with speeches that seem to commit us to arbitrary and inscrutable dependencies, such as 'There is a number *n* such that necessarily, anyone with fewer than *n* hairs is bald, and possibly someone with exactly *n* hairs is not bald'. And it wouldn't be surprising if the kinds of moves that might help us overcome our initial resistance to those speeches extend to the kind of arbitrariness just contemplated. After all, in characterizing the relevant function we used several words that are plausibly vague, most notably the word 'earring'. So one should be careful before dismissing the view simply on the grounds of arbitrariness.

The "coarse hypertolerance" view can also be fleshed out in some ways that avoid certain particularly bizarre kinds of dependencies. The bare bones of the view are consistent with the thought that certain possible earrings depend for their existence on phenomena that common sense immediately recognizes as irrelevant. For example, it is consistent with the view that there is some particular earring which can only be composed by some particular collection of atoms in a world where a certain particular asteroid hits the moon at some time or other. But the story can be developed in a way that rules out such weird dependences, endorsing not just Origin Hypertolerance, but "Macro-hypertolerance", according to which, roughly speaking, the proposition that a certain collection of atoms composes one possible earring rather than another places no special constraints on the *macroscopic* facts. More precisely:

> For any earring *e* originally composed by a collection of atoms *C*, and any collection of atoms *C′* chemically matching *C*, and any *macroscopic p*, if it is possible that *p* is true while *C′* composes an earring, it is possible that *p* is true while *C′* composes *e*.[21]

For each possible earring *e* and chemically appropriate collection of atoms *C′*, the set of possible microstories that necessitate that *C′* originally composes *e* is sprinkled like a fine dust across the space of microstates compatible with each relevant macrostate, inextricably intermingled with microstories that entail that *C′* originally composes some earring other than *e*. There are further constraints you could play with too. You could hold that the question which earring is made when an earring is made out of a certain collection of atoms at some time is entirely settled by the microphysical history of the world *up to that time*. You could even maintain that it is entirely settled by the *intrinsic* microphysical facts about the relevant collection of atoms at the time of creation (in combination with the fact that an earring is made out of that collection at that time).

The coarse hypertolerance picture involves giving an intuitively weird significance to questions of microphysical detail. Can this intuitive resistance be turned into a rigorous argument? Such an argument would presumably rely on some more general principle about the ways in which microphysical differences can and can't matter to the identities of macroscopic objects of the relevant kind. But formulating a defensible generalization which can play this role turns out to be a tricky matter. Focusing on the case of earrings, here is one initially tempting thought: the microphysical facts would need to be *substantially* different to make a substantial difference to any of the physical properties (location, shape, atomic composition, concrete existence or lack thereof…) of any earring, or its earringhood. More carefully:

Earring Gradualness For any earring *e* originally composed by a certain collection of atoms *C*, and any *approximately* true microstory *p*, *p* is compatible with *e* being an earring originally composed by some collection largely overlapping *C*.

This principle does secure Earring Tolerance (the fine-grained premise from §6.2), assuming that for any pair of microstories incorporated by properties that are "close" in the sense at issue, the truth of the one counts as entailing the "approximate" truth of the other. But it faces several glaring problems. One,

[21] This is the Hypertolerance claim in a quantified Tolerance Argument where K is *earring* and *G* is close to *F* iff for some *C*, *C′*, and *q*, *C* and *C′* overlap by at least 90 per cent and chemically match, *q* is a conjunction of macroscopic propositions, *F* is *being composed of C and such that q*, and *G* is *being composed of C′and such that q*.

by now familiar, sort of worry concerns situations where some process creates something that is so poorly suited to being attached to people's ears that a tiny microphysical difference would take us to the situation where there is no earring at all anywhere in the vicinity. (Considering Sorites sequences is enough to establish the *possibility* of such borderline cases, and thus to refute the *necessitation* of Earring Gradualness; but since it is not unlikely that the actual world already contains some such borderline earrings, the problem also arises for Earring Gradualness itself.) A less familiar worry involves situations where a tiny microphysical change would constitute a dramatic increase or decrease in the size of some particular earring. For example, suppose an earring weighing five grams contains a diamond weighing three grams that gradually works its way loose until it falls off. It is plausible that at the end of the process, the original earring is still around but weighs only two grams. And the decrease in weight seems to be discontinuous—there was no time at which the earring weighed three or four grams. Now, consider some time when the earring still weighed five grams but would very soon weigh two grams: it seems that there should be possibilities where the earring already weighs only two grams at that time even though the microphysical situation at and before that time is only slightly different. And plausibly, there are similar cases where a small change would take us to a situation where a certain diamond was *never* part of some earring, so that an earring that in fact sometimes weighed five grams would instead have never weighed more than two. But here, again, the small microphysical differences that make a substantial difference to the physical properties of some particular earring also seem to make a similarly substantial difference to the general facts about which atoms compose earrings. When the earring goes from being big to being small, the small collection of atoms (not including the ones in the diamond) begins composing an earring, even though previously it didn't even "largely overlap" any collection of atoms that composed an earring.

Both of these concerns can be handled by retreating to:

Weak Earring Gradualness For any earring *e* originally composed by a certain collection of atoms *C*, and any approximately true microstory *p*: if *p* is compatible with there being *an* earring originally composed of most of *C*, then *p* is compatible with *e* being an earring originally composed of most of *C*.

This still secures Earring Tolerance, since it is clear that if in fact an earring is originally composed of *C*, then every microstory that differs only by a 1° rotation of the cutting tool or a rearrangement of two of its atoms is compatible with *an* earring being originally composed of *most* of *C*.

However, even Weak Earring Gradualness still faces some potential counterexamples of a more esoteric sort. Suppose that certain earrings are made by attaching a tiny but very sparkly diamond to a large, plain setting. Although the vast majority of the atoms that compose such an earring are in the setting, it is

not unnatural to think of the diamond as essential and the setting as inessential. Consider now a series of microstories where one tiny diamond slowly becomes detached from a setting, while another simultaneously becomes attached to a different part of the same setting. Given the essentialist way of thinking about the diamond just alluded to, such a series might be thought to contain a counterexample to the above principle: a pair of microstories such that the truth of the one necessitates the approximate truth of the other, which differ as regards which of the diamonds is sufficiently attached to the relevant setting, and which hence differ as regards the identity of the earring composed by most of the atoms in the setting. The lesson here is that when some atoms are much more load-bearing than others as far as the identity of an object is concerned, considerations about what *most* of the atoms are doing might be a poor guide to the facts about what that object is doing.[22] Of course, in the original earring example, there is no temptation to start assigning a similarly crucial role to any of the few atoms that would be affected by a small rotation in the angle of cut.

It thus seems quite hard to find defensible generalizations which capture the intuitive thought about the irrelevance of microscopic details, at least ones that aren't watered down with ceteris paribus clauses, restrictions to "normal" objects of the relevant sort, and so on. This is a bit disappointing for friends of fine-grained Tolerance premises like Earring Tolerance, and somewhat encouraging for proponents of the coarse-grained hypertolerance strategy which turns on rejecting those premises.

Further encouragement might be drawn from some exotic cases where Microphysical Supervenience generates pressure all by itself to posit some sort of inscrutable fine-grainedness in the manner in which the macro-facts supervene on the micro-facts. Consider two ships that have existed from eternity, where every year a plank from one ship is swapped with the corresponding plank from the other ship. It would be quite strange to deny that each ship could have existed without the other ship ever concretely existing. But since there is no such thing as the *originating* matter of either ship, it is hard to come up with any principled-looking story under which the coarse-grained facts at some world where the planks only ever compose one ship are enough to determine which of the two ships it is. If we hold Microphysical Supervenience fixed as a constraint, any conceivable story about how the identity of the ship supervenes on the microphysical facts in a one-ship world will have the feeling of the very sort of inscrutability that the coarse-grained hypertolerantist wants to posit all over the place.

[22] For other kinds of object, counterexamples to the analogue of Weak Earring Gradualness may be more widespread. Restaurants, for example, seem much jumpier than earrings—by signing a legal document, one can get a restaurant to move suddenly across a city. This suggests that there may be some microphysically close cases where some legal document that swaps the location of two restaurants is close to completion in one but complete in another, which differ dramatically as regards which restaurant is where, although they agree as regards where there is and isn't a restaurant.

So far, then, coarse-grained hypertolerance isn't looking too bad. But there is a further pressing worry for it which leads us to think that it cannot, or at least cannot all by itself, be the key to a satisfactory resolution of the puzzles of tolerance. The problem is that ordinary modal judgements support not merely claims of *tolerance* with regard to various families of coarse-grained properties, but also claims of *robustness* with regard to the same families of properties. Focusing on the case of the earrings, the coarse-grained tolerance premise says that wherever an earring is made, it is possible for it to be made any small number of degrees from where it is in fact made. The corresponding robustness claim is that it is hard for the same earring *not* to be made, so long as the cut is made only a few degrees from where it is in fact made. More generally, what we are calling robustness claims are claims to the effect that it would be hard for the same objects to fail to exist (in roughly the same location…) without things being substantially different in relevant underlying respects. As we have seen, exotic cases (like the very sparkly diamonds) make it challenging to write down premises that capture this kind of robustness thought and hold in full generality. Nevertheless, there are many perfectly everyday settings in which our ordinary judgements require some more localized robustness. Such judgements include, for example, claims about easy possibility like 'I couldn't easily have prevented the craftsman from making that table' (which would seem obviously true if my influence was limited to delaying the shipment of one of the planks), and counterfactuals like 'If the craftsman's hand slipped a little bit, this earring would almost certainly have been made a little further away from him than it actually was.'

Perhaps the most useful way of regimenting the language of hardness and easiness that we used in introducing robustness is to cash it out in terms of high and low objective chance. For example, it seems very plausible, and very much of a piece with other ordinary judgements, to think that before the craftsman decided where to cut the disc, there was a high objective chance of Lefty the earring being made on the left of the cut, conditional on the cut being made approximately where it is actually made. Indeed, we might eventually convince ourselves that the conditional chance was not merely high but *exactly* 1. (This would follow from the even stronger claim that it is metaphysically impossible for Lefty not to be made on the left in a world with matching history and where the cut was made just a few degrees from its actual position.) But since retreating from 'exactly 1' to 'close to 1' wouldn't seriously disrupt our ordinary thinking, the 'exactly 1' thought is less appropriate as a starting point, and more in need of a philosophical argument.

Given how entrenched robustness thoughts are in ordinary practice, a systematic rejection of them would be a steep price to pay. Moreover, it seems arbitrary and unprincipled to be deeply respectful of ordinary tolerance thoughts while dismissive of ordinary robustness thoughts.

Fortunately, although coarse-grained hypertolerantists are under pressure to reject very strong robustness thoughts such as those that require the relevant

conditional chances are *exactly* 1, they can agree that they are *close* to 1, which seems to be an adequate vindication of ordinary practice. In the case of the earrings, the picture would be that *most* of the microstories close to the actually true microstory entail that Lefty is made on the left, although a few of them don't, since according to the coarse-grained hypertolerantist it is possible for Lefty to be made on the right while the cut is made almost exactly where it actually is.[23]

Unfortunately, this picture is unstable in a subtler way. If *most* of the nearly true microstories where the cut is made near its actual position entail that Lefty is made on the left, but a few of them entail instead that Lefty is made on the right, then those few also seem to entail that our ordinary robustness judgements are badly mistaken. Those worlds are freaks, unrepresentative of their modal neighbourhood. In them, when we reason from 'There was a only a small chance of the cut not being approximately where it actually was' to 'There was only a small chance of any earring not being made approximately where it actually was', we would seem to be making a mistake. And the fact that there are nearly true microstories which entail that we are mistaken in this way seems epistemologically disturbing, as regards the robustness judgements we in fact make. Even setting questions about knowledge and luck aside, it just seems faintly absurd to hold that reality in fact provided us with some nice robust earrings, although we could have been unlucky and been served up instead with some super-fragile lookalike earrings that we would have wrongly assumed to be robust.

Granted, this kind of argument is not irresistible. Perhaps we should learn to stick to our guns in the face of nearby error, comforted by the thought that such error is very unlikely (and perhaps also "abnormal" in a way that goes beyond mere statistical unlikelihood). Or perhaps (as we will suggest in Chapter 11) we should resist the inference from the claim that earrings at the relevant close worlds fail to be robust to the claim that our robustness judgements at those worlds are erroneous. But either of these strategies for resisting the Security Argument applied to robustness judgements will work just as well for resisting the Security Argument applied to tolerance judgements. If we are willing to appeal to these strategies, we will have no stable motivation for accepting Non-contingency in the coarse-grained Tolerance Arguments, and hence no motivation for accepting coarse-grained Hypertolerance in the first place.

So, the coarse-grained Hypertolerance package looks dialectically unstable. To motivate it (without relying on bad Soritical intuitions), we need to accept the Security Argument for Non-contingency. But if we do, we should be equally impressed by the Security Argument against the existence of nearby worlds where the relevant objects fail to be robust.

[23] Here 'most p-worlds are q-worlds' can be understood as equivalent to 'the objective chance of q conditional on p is high'.

We shall be returning to the idea of robustness in two places later in this book. In Chapter 9 we will give a more careful development of the argument, and a defence of the Iteration-like premises about objective chance required for it to go through. And in Chapter 11 we will be exploring some metasemantical ideas which, while helping coarse-grained hypertolerantists deal with the robustness objection, also undercut much of the motivation for the view by blocking the Security Argument for Non-contingency.

6.4 Denying Microphysical Supervenience

Finally, let's examine the option of denying Microphysical Supervenience in order to preserve a much more thoroughgoing hypertolerance, extending even to fine-grained Hypertolerance claims like the conclusions of the arguments discussed in §6.2. Microphysical Supervenience is, as we emphasized, a highly controversial thesis, so those who reject it in response to Tolerance Arguments will not lack for company. The problem is that many of the other arguments against Microphysical Supervenience also suggest ways of weakening it which would be just as good for our current purposes (i.e. for arguing against Hypertolerance in certain otherwise compelling Tolerance Arguments). For example, many philosophers (such as Chalmers 1996) believe that facts about *qualia*—"what it is like" to be a certain subject of experience—fail to supervene on the microphysical facts. According to them, there are possible worlds where all the same microphysical propositions are true that are true at the actual world, but where, e.g., certain people's qualia are spectrally inverted with respect to how they actually are, or where everyone is a "zombie" with no qualia at all. But even if they are right, their arguments will obviously not refute "Microphysical-cum-qualia Supervenience", a thesis just like Microphysical Supervenience except that the supervenience base contains propositions about qualia in addition to propositions about microphysics. And that thesis is just as good for the purposes of ruling out certain Hypertolerance claims, since we can treat minor variations of the qualia in exactly the same way as minor variations of the microphysical facts.[24] Similar remarks apply to many other putative counterexamples to Microphysical Supervenience, such as facts about goodness, or moral obligation, or Cartesian minds and their thoughts, or God and his angels. None of these candidate expansions of the supervenience base seem to

[24] Having qualia in the supervenience base will open up the possibility of a variant of the coarse-grained hypertolerance package from §6.3 in which, e.g., Lefty and Righty could have swapped their total microphysical roles but not their total overall patterns of qualitative relations to microphysical objects: perhaps particles can feel pain, and getting Lefty and Righty to switch requires a modification in their pain levels. But everything we said in §6.3 about the form of coarse-grained hypertolerance designed to preserve Microphysical Supervenience will apply mutatis mutandis to this sort of view: in particular, the problem with robustness judgements will arise in the same way.

get in the way of describing an intuitively gradual path connecting two properties each of which incorporates a proposition that is maximally specific about the entire supervenience base, but which are instantiated by different objects, where the objects in question do not themselves figure in the supervenience base (e.g. the earrings Lefty and Righty). Given such a path, one can construct a Tolerance Argument whose conclusion is incompatible with the new supervenience thesis, and whose premises enjoy similar plausibility to those of a corresponding Tolerance Argument constructed with an eye to Microphysical Supervenience.

An alternative strategy for accommodating certain kinds of putative counterexamples to Microphysical Supervenience, such as those involving zombies, would be to retreat from the thesis that the microphysical propositions provide a supervenience base for *all* propositions to the thesis that they provide a supervenience base for all propositions *of the sort at issue in the relevant Tolerance Argument*—e.g. about what earrings they are and where they are when they are made. Salmón goes for a weakening along these lines: 'Perhaps some facts are underdetermined by the totality of material facts, but surely the question of whether a given actual table *a* is constituted by a certain hunk of matter *h″* must be so determined' (Salmon 1981: 281). But this really doesn't seem at all obvious. Even if there could be zombies, one might think that this very artificer couldn't have been a zombie; and it would not be absurd to think that these earrings couldn't have been made otherwise than by this artificer. For this reason, we think the strategy of expanding the supervenience base to include, e.g., qualia facts is more promising than that of narrowing the supervening class of propositions to exclude them.[25]

It is also worth pointing out that among those philosophers who take qualia and the like to provide counterexamples to Microphysical Supervenience, many have no problem with the analogues of Microphysical Supervenience for certain modalities narrower than metaphysical necessity. For example, Chalmers (1996), the most influential current proponent of the view that qualia-facts do not supervene with metaphysical necessity on physical facts, argues that they do nevertheless supervene with *nomological* necessity. But most of the Tolerance Arguments we have been concerned with have premises which remain quite plausible even when all the modal operators are interpreted using nomological necessity, or something even narrower than nomological necessity, such as having chance 1 at some appropriately chosen time *t*. So insofar as the doubts about Microphysical Supervenience are specific to the metaphysical-necessity version, we can bypass them just by focusing entirely on these narrower modalities.

A different challenge to Microphysical Supervenience (which also affects Microphysical-cum-qualia Supervenience) involves the claim that it is possible for

[25] Moreover, there are certain kinds of candidate failures of supervenience for which the latter strategy would make no sense. For instance, if facts about whether some atoms compose something don't supervene on microphysics—a view discussed, though not endorsed, by Markosian (1998: 216)—there is no prospect of thinking that such facts don't matter when it comes to tables and earrings.

all the microphysical facts to be the same while some "alien", non-microphysical, fundamental properties (or relations) are instantiated—where 'alien' means 'not belonging to the collection of all actually instantiated fundamental properties'. Again, we could bypass this by focusing on a narrower modality. Lewis (1983*a*) introduces the notion of the "inner sphere", comprising all possible worlds where no alien fundamental properties are instantiated, and suggests that these like Microphysical Supervenience should be weakened in such a way that only inner-sphere worlds matter. Without worldspeak, we can say that *p* is inner-sphere necessary just in case there is a collection that contains all instantiated fundamental properties, such that *p* is metaphysically necessitated by the proposition that this collection contains all instantiated fundamental properties.[26] Since the Tolerance Arguments we will be concerned with could, if they worked, easily be modified so as to support Hypertolerance claims involving inner-sphere possibility, questions about what goes on at possible worlds outside the inner sphere are not really relevant.[27] However, it is not obvious that such a retreat is in any case called for, even if one admits the possibility of alien non-microphysical fundamental properties. We could instead try to get around this problem by giving the property of being *fundamental* (i.e. fundamental in some particular type—a fundamental object, a fundamental monadic property of objects, etc.) the same status we implicitly gave to *logical* properties and relations—i.e. counting propositions as "microphysical" when they are expressible in a language that includes not only standard logical constants and constants for microphysical objects, properties, and relations, but also predicates for fundamentality. If so, then so long as it is *in fact* true that all instantiated fundamental properties are microphysical, truths of the form *Every instantiated fundamental property is one of* $F_1 \ldots F_n$ will count as microphysical, and thus metaphysically possible worlds

[26] We can do something similar for, e.g., alien fundamental binary relations (among objects). But there is a problem: in our higher-order regimentation we have no way of making sense of talk of collections that include entities of arbitrary types, so it is not obvious how to make the definition of 'inner-sphere possible' demanding enough to ensure that it is inner-sphere impossible for any alien fundamental properties or relations, of any type, to be instantiated. One way around this expressive difficulty is to include some semantic vocabulary in our object language: we can define '*p* is an alien-excluder' to mean 'For some type σ, the formula $\ulcorner \exists X(\text{Rigid}_{\langle\sigma\rangle}(X) \wedge \nu = \forall y(\text{Fundamental}_{\sigma}(y) \rightarrow X(y)))\urcorner$ is true relative to the assignment that maps the variable ν to *p*', and then define '*p* is inner-sphere necessary' as '*p* is metaphysically necessitated by some collection of true alien-excluders'.

[27] An alternative approach would just replace metaphysical necessity in Microphysical Supervenience with "microphysical necessity", defined as metaphysical necessitation by the proposition that all instantiated fundamental properties are microphysical. (See previous note for discussion of how this could be generalized to other types.) Unlike inner-sphere necessity, microphysical necessity is not metaphysically necessarily factive, which means that this modification of Microphysical Supervenience is compatible with all kinds of intuitively non-physicalist goings-on. But the conjunction of it with the proposition that all instantiated fundamental properties are microphysical could play a similar role to Microphysical Supervenience in our argument: if in fact all microphysically necessary propositions are true, various Tolerance premises involving microphysical necessity will be appealing, so the falsity of the corresponding Hypertolerance conclusions has broader morals that will carry over to Tolerance Arguments involving other modalities.

where alien fundamental properties are instantiated will not count as agreeing with the actual world on all microphysical propositions.[28]

So, if we are going to deny Microphysical Supervenience in order to be able to accept some Hypertolerance claim inconsistent with it, our denial must be a truly radical one, attacking not just the specifics, but the underlying picture of the world on which there are certain "basic" facts from which all the facts flow, in a sense that requires being metaphysically necessitated by them.[29, 30]

The idea that facts about middle-sized objects enjoy this surprising kind of autonomy is not entirely unprecedented: there are other ways of putting pressure not only on the letter of Microphysical Supervenience, but on the general vision of the supervenience of all truths on some privileged ground floor. Here is one putative counterexample, adapted from Hawthorne 2006*a* (282). Suppose the history of the world is one of one-way eternal recurrence, infinite in the pastwards direction. The same atoms have existed for all eternity. In each epoch, they are first assembled into the form of a statue, and then disassembled again. Eventually, the backwards-infinite series of intrinsically indistinguishable epochs comes to an end, with a giant explosion that scatters the atoms far and wide. Plausibly, a new statue is made in each epoch. And it is tempting to think that the explosion could have happened in exactly the same way but one epoch earlier, so that what was in fact the second-last statue would have been the last statue, what was in fact the third-last statue would have been the second-last statue, and so on; what was in fact the last statue would never have been made at all. But the situation we have just described seems to involve exactly the same totality of microphysical facts as the original history: the same collection of atoms exists in both histories, instantiating the same pattern of microphysical properties and relations. If these

[28] Another potential restriction is to worlds that don't contain any "new" objects—objects that do not belong to the collection that in fact contains every object. (Such a restriction is pointless for proponents of BF). This modification of Microphysical Supervenience may be needed if one thinks that there are possibilities where "new" microphysical objects—say, new electrons—are added while the distribution of the "basic" microphysical relations over the microphysical objects there actually are remains the same, and understands 'microphysical proposition' in such a way that this counts as something that could happen without any true microphysical proposition becoming false.

[29] For a full-throated defence of this radical form of supervenience-denial in the context of Tolerance Puzzles, see McKay 1999 (although McKay does not go for Hypertolerance).

[30] Note that rejecters of Microphysical Supervenience might still accept the following principle, which is compatible with all manner of Hypertolerance claims, and even with Extreme Anti-Essentialism:

Qualitative Microphysical Supervenience Every qualitative truth is necessitated by a qualitative microphysical truth.

One might even think that this was all along the intuitively compelling principle in the vicinity. We are quite sceptical about this last claim. The intuitive force of the thought that something must "explain what makes it be the case that *P*", when *P* is about macro-objects, seems to arise in exactly the same way whether or not *P* is qualitative. Still, Qualitative Microphysical Supervenience does also seem quite attractive; we will have more to say in Chapter 14 about its potential relevance to certain Tolerance Puzzles.

are a genuine pair of possibilities, Microphysical Supervenience must be false (at the world in question).

The argument just presented requires the tendentious assumption that the atoms and their parts are the only objects that instantiate microphysical properties and relations at the relevant worlds. This would be denied by spacetime substantivalists, who would contend that there are particular spacetime points (or regions) that are also among the relata of microphysical relations: examples might include geometric and field-theoretic relations, which hold solely among points, or relations like *occupation* which hold between points and other microphysical objects such as particles. Against this background, there are clear differences in the truth values of microphysical propositions between the two scenarios, since different particular spacetime points witness explosions in the two stories. Analogous remarks hold for the more old-fashioned view that reifies *instants of time*.[31]

It is worth noting, however, that there is a different kind of "shift argument" against Microphysical Supervenience that acquires a certain amount of plausibility in the context of these substantivalist approaches. Here, we start with a world of *two-way* eternal recurrence, e.g. one where the same atoms exist throughout, and in each duplicate epoch, a statue is assembled and then disassembled. Assuming again that a new statue is created each time, it is natural to think that the entire history could have played out in a Leibniz-shifted way, so that each statue inhabited a region of spacetime (or time) one epoch backwards or forwards from the region it in fact inhabits. While this judgement is more theoretically loaded than the earlier judgement about exploding statues (since it relies on judgements about relations to spacetime points, objects which are not familiar to common sense), it does have some intuitive force. Indeed, the possibility of this kind of shift, conditional on the existence of times, was Leibniz's key reason for refusing to reify times.

Despite these objections to Microphysical Supervenience, we are rather drawn to the thesis: for us, the guiding idea that the macro-facts are determined by the micro-facts carries more weight than intuitions about such exotic possibilities. Note that one way for proponents of Microphysical Supervenience to respond to these examples would involve positing fine-grained sensitivity to microphysical detail in the truth about which macro-things there are. For example, while accepting that it is impossible for the explosion to occur one epoch earlier while the microphysical history was *exactly* the same, one might allow that it could have occurred one epoch earlier while the microphysical history was *approximately* the same—perhaps some tiny change in the distance between two particles would be enough to make what was in fact the last statue be permanently non-concrete, what was in fact the second-last statue be the last statue, etc. This is reminiscent

[31] Dorr (unpublished *a*) discusses ways of combining the spacetime substantivalist ontology with an "A-theoretic" view of the tense-operators as analogous to modal operators.

of the sensitivity to detail required by coarse-grained hypertolerance, although it need not be combined with coarse-grained hypertolerance. And while it might seem absurdly arbitrary to suppose that a particular microphysical lay of the land determines one ship rather than the other, the points about vagueness that we made in explaining why such arbitrariness is not fatal for the coarse-grained hypertolerance package also apply here.

We will end by mentioning one final challenge to the package that combines a widespread endorsement of Hypertolerance with a denial of Microphysical Supervenience. As we have seen, there is dialectical pressure for proponents of this package to accept many Hypertolerance claims even involving rather narrow modalities like objective chance 1, since prima facie it looks like any strategy for blocking Tolerance Arguments couched in terms of the narrower modalities will extend easily to their analogues for metaphysical modality.[32] But when developed in this way, the hypertolerance package will struggle to accommodate some ordinary "robustness" judgements that we introduced in §6.3 as a difficulty for the competing coarse-grained hypertolerance package. Cashed out in terms of objective chance, the relevant robustness judgement says, for example, that there was a high chance of Lefty being made on the left, conditional on the cut being made roughly where it actually was. So far that is no problem for the hypertolerantist, since 'high chance' doesn't mean 'chance 1'. However, if we accept the relevant chance-1 Tolerance Argument, we must think that there was at least a small chance of Lefty being made on the right, conditional on the cut being made roughly where it actually was.[33] And prima facie, it looks like worlds where this happens are ones where our ordinary robustness judgements are radically false. There, we are just making a mistake when we say things like 'The earring he made on the right is one that he couldn't easily have failed to make on his right without significantly adjusting the cutting tool.' As we explained in §6.3, such a chance of error raises a security worry about our robustness judgements in the actual world. As in that setting, the worry may not be a knockdown one: one might firmly stick to one's epistemic guns in the face of the relevant chances of error, or one might—using ideas to be explored in Chapter 11—resist the inference from there being a chance of the relevant proposition being false to there being a chance of our making a mistake. But these responses also tend to undermine the security-based motivation for Non-contingency. So for our part we can't see a dialectically stable way to assemble a compelling hypertolerantist package, with or without Microphysical Supervenience.

[32] The most promising-looking way of trying to drive a wedge between the chance-1 Tolerance Arguments and the metaphysical necessity versions would be to deny Iteration for chance 1, while accepting it for metaphysical necessity. However, in Chapter 9 we will argue that there are strong reasons to accept Iteration for chance 1.

[33] The denier of the chance-1 version of Microphysical Supervenience will presumably also think that there was a nonzero chance of Lefty being made on the right conditional on the complete truth about microphysics. But this claim isn't so relevant for ordinary robustness judgements.

7
Rejecting Iteration

In one of the earliest discussions of our puzzles, Chandler (1976), focusing on issues of tolerance with respect to originating matter, noticed the fact that Necessitated Tolerance and Iteration together imply Hypertolerance, and argued that since Necessitated Tolerance is true and Hypertolerance is false, Iteration must be false. His central example is a bicycle which, he supposes, is necessarily such that whichever parts it is created with, it could have been created with one third of those parts being different, but could not have been created with two thirds of those parts being different. The proposition that it is created with a certain specific roster of parts which in fact only include one third of its actual parts is thus only contingently impossible: it is possibly possible, but not possible. Other proponents of this solution include Salmón (Salmon 1981, 1986*a*) and Lewis (1986*a*: §4.4).[1]

For some notions of possibility, such as the one expressed by 'It could easily have been that . . . ', Iteration doesn't seem especially compelling to begin with. As regards metaphysical possibility, on the other hand, many philosophers have felt the Iteration-denying position to be in tension with their basic grip on what metaphysical possibility comes to. We are sympathetic, and will develop this line of thought in the next chapter. But in this chapter, we will bracket the question whether Iteration holds for metaphysical possibility, and consider other kinds of puzzles which may still arise even on the assumption that it does not, and which might convince one that denying Iteration for metaphysical possibility isn't a sufficiently general strategy for handling Tolerance Puzzles. §7.1 will discuss Tolerance Puzzles involving tense operators. §7.2 will discuss a Tolerance-like puzzle due to Williamson (1990) for which it is not obvious how Iteration-denial could help. §7.3 will discuss Tolerance Puzzles involving the notion of *ancestral* metaphysical possibility, for which Iteration is guaranteed to hold in our basic modal logic; we will argue that certain higher-order identities make Hypertolerance problematic even in that case. Finally §7.4 will address a natural, but ultimately unsatisfactory, strategy for resisting the relevant higher-order identities.[2] The topic of

[1] Lewis's version is accompanied by some further distinctive ideas about what it is to be possibly a certain way, which we will discuss in Chapter 10.

[2] Chapter 9 will discuss a further range of Tolerance Puzzles having to do with chance, where again the denial of Iteration for metaphysical modality does not help in any obvious way.

The Bounds of Possibility: Puzzles of Modal Variation. Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri, Oxford University Press. © Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri 2021.
DOI: 10.1093/oso/9780192846655.003.0008

higher-order identity which we broach in this chapter will also be important to Chapter 8's discussion of metaphysical necessity.

7.1 Iteration and Tense Operators

As we saw in §2.3, Tolerance Arguments can be run using a wide variety of different operators. Among them are tense operators like 'It will always be the case that . . . ' and 'It always is the case that . . . ', for which Iteration seems especially compelling.[3] If it's going to be the case that it's going to be the case that *P*, how could it not be going to be the case that *P*? The suggestion seems baffling.

We can sharpen the point by considering so-called "metrical tense operators". If in twenty years' time Juhani will be twenty years from retirement, it seems utterly obvious that Juhani is now forty years from retirement. Likewise, if thirty years ago a certain ship was such that thirty years earlier it was made of certain planks, it is similarly obvious that sixty years ago it was made of those planks. But since anything that has ever been true was true some number of years ago and anything that was true some number of years ago was once true, these additive principles about the metrical tense operators imply Iteration for the ordinary past and future tense operators.[4]

Nevertheless, some Tolerance Arguments based on tense operators generate real puzzles. While many philosophers have got quite comfortable with the idea that the Ship of Theseus can survive any number of successive plank-replacements, even they may not be so comfortable with the thought that a painting could go by successive dabs and deletions from looking like the *Mona Lisa* to looking like *The Scream*, or that a language could go by small steps from working the way Latin did during the Roman Empire to working the way Italian currently does. Williamson (1990: 142) suggests that the implausibility of Iteration-denial as a solution to the temporal Tolerance Puzzles undermines the plausibility of Iteration-denial as a

[3] As we noted in §2.3 (note 17), there is an influential "B-theoretic" tradition of theorizing about time on which such expressions are not taken to express genuine operations (properties of propositions), but rather to work analogously to quantifiers. B-theorists might quibble with the classification of the relevant puzzles as Tolerance Puzzles, but they will still need a solution to their alternatively regimented versions of the puzzles, and it seems especially unpromising to think that that solution will be in any way analogous to Iteration-denial.

[4] The only examples we know of of philosophers who reject Iteration for tense operators are temporal counterpart theorists like Sider (2002: §5.8). On Sider's view, it could be true that in one hundred years you will be such that you have one hundred years left to live, although in fact you only have 137 years to live. Sider adopts a semantics on which the truth of Iteration would require a certain relation of "temporal counterparthood" to be transitive, and he denies that the relation is in fact transitive. His main argument against transitivity is based on premises to do with "what matters" when it comes to personal identity. But it seems to us that the case that the tense operators obey Iteration is overwhelmingly stronger than the required premises about "mattering". One general theme of our discussion of counterpart theory in Chapter 10 is that it provides no good excuse for giving up compelling principles of modal and tense logic.

solution to the modal paradoxes. The temporal paradoxes 'involve the failure of some other assumption; it will have a modal analogue; why should we suppose that the latter does not fail, and blame the S4 principle instead?'

Williamson's challenge is an important one. It is not answered merely by noting that there are some specific families of properties (e.g. those having to do with originating matter) for which there are modal paradoxes with no temporal variants, since the properties in question are necessarily had always if ever. If some initially appealing thing has to be sacrificed to solve a temporal Tolerance Puzzle without giving up temporal Iteration, that might be expected to shake one's confidence in the corresponding claim in any modal Tolerance Argument against Iteration.

On the other hand, schematic similarities between Tolerance Puzzles may mask important differences. Even within the domain of Tolerance Puzzles based on metaphysical possibility, there isn't *that* strong a presumption in favour of a uniform resolution to such disparate examples as that of chess-to-Twister and originating matter. And in the case of Hypertolerance-based solutions, there are special reasons for concern in the modal domain that don't carry over to typical temporal puzzles. As we saw in the previous chapter, there is a plausible general principle, namely Microphysical Supervenience, which rules out Hypertolerance in some modal Tolerance Puzzles involving highly specific families of properties, in a way that arguably generates problems for Hypertolerance claims involving less specific families. By contrast, Microphysical Supervenience does not pose a problem in the same way for Hypertolerant solutions to temporal Tolerance Puzzles. So a split treatment, combining Iteration-denial in the modal case with an acceptance of Hypertolerance in many temporal cases, is far from entirely unprincipled, whatever other objections it may face.

True, there are some very special temporal Tolerance Puzzles where Microphysical Supervenience does rule out Hypertolerance. These puzzles are set in possible worlds where the microphysical facts are subject to eternal recurrence, so that the complete microstory that is true now will be true again in the future. Suppose, for example, that at every time there are just two ships, each made of one hundred planks, circling one another in otherwise empty space. (They are wooden spaceships.) Every day, one plank becomes detached from each ship, and the two detached planks swap positions. The swaps always occur in the same order, so that the exact pattern of microphysical relations among physical particles is repeated every hundred days.[5] Considering this world, we can run a temporal Tolerance Argument where the kind *K* is *ship*, and the closeness relation is the one that

[5] We assume here that the particles are the only things that instantiate microphysical relations. This is somewhat tendentious even from an A-theoretic point of view, but highly dubious from a B-theoretic perspective: a B-theoretic fundamental ontology will need to include something like spacetime points, or instants of time, or instantaneous particle-slices. We see no way of running the current argument against temporal Hypertolerance in a B-theory-friendly way.

F bears to *G* iff *F* is *being composed of C* and *G* is *being composed of C′* for some collections of planks *C* and *C′* which differ by at most one plank and are both arranged shipwise at some time. Hypertolerance here entails that each ship is at some time composed by each of the 200 collections of planks that are ever arranged shipwise and. But this is ruled out by Microphysical Supervenience. For given our stipulations, every microphysical proposition that is ever true is true again one hundred days later. Given Microphysical Supervenience and the principle that metaphysically necessary propositions are always true (Dorr and Goodman 2020), it follows that *every* proposition that is ever true is true again one hundred days later, and thus that for each ship, there are at most one hundred collections of planks that ever compose it. So given Iteration for the tense operators, the Tolerance premise cannot be permanently true. At some time, a ship composed of one hundred planks must be such that after the next plank-exchange, it will not end up composed of ninety-nine of those planks and one other plank. (Perhaps it will, instead, cease to be concrete. Or perhaps it will undergo a sudden jump, coming to be composed by an almost entirely non-overlapping collection of planks.)

This is certainly a surprising result, and it would not be *crazy* to react to it by becoming markedly less confident that ships in the actual world survive single plank-replacements. However, given the extremely far-fetched character of the thought-experiment, it would also not be crazy to quarantine the surprise, by holding that worlds of eternal recurrence are very special as regards the behaviour of things like ships, and that ships at worlds that lack eternal recurrence are always tolerant of plank-replacements even if that means eventually coming to be composed of entirely different planks.[6]

7.2 Williamson's Earrings

Williamson also presents a different line of thought that he takes to diminish the plausibility of Iteration-denying solutions to modal Tolerance Puzzles. He considers an example like the earring case we described in §6.2, with the crucial twist that no earring is in fact made: 'in the actual world of the story, the craftsman is struck dead before he can cut the disc; had he not been, he would have chosen

[6] One might object to this kind of special treatment on the grounds that if ships are short-lived in worlds of eternal recurrence, they should be similarly short-lived in worlds where a tiny microphysical asymmetry spoils perfect eternal recurrence. But from an A-theoretic point of view (and as noted above, the argument is a non-starter from a B-theoretic point of view), there is an enormous gulf between perfect eternal recurrence and everything else. Perfect eternal recurrence involves the very same propositions that are now true being true again in the future, including propositions about who is alive, what wars are being fought, etc. By contrast, with anything short of perfect eternal recurrence, the complete truth about how things are will never be true again, and so one is not forced to posit widespread reincarnation.

a diameter at random' (Williamson 1990: 128). Williamson argues that for each point x on the circumference of the disc, there is a particular possible earring that *would* have been made centred on x, if the disc had been cut in such a way that some earring or other was made centred on x (i.e. cut along the diameter connecting the two points separated from x by 90°).[7] As he points out, it is tempting to think that when two points on the circumference are very close—say, at most 1° apart—the possible earrings that would have been made centred on each of those points, if the disc had been cut along the corresponding diameters, are identical. But this tempting thought leads by the transitivity of identity to the absurd consequence that even when two points are 180° apart, the possible earrings corresponding to these points are identical. This is absurd, since this would require that if the disc had been cut along any diameter, a single bilocated earring would have been created. The tempting thought must therefore be rejected: whether or not one accepts Iteration, one must accept that there is some pair of close points x and y such that the possible earring that would have been made if an earring had been

[7] One possible motivation for this view would appeal to two disputed principles of counterfactual logic:

Conditional Excluded Middle (CEM) $(P > Q) \vee (P > \neg Q)$

Either (if it were that P it would be that Q) or (if it were that P it would be that not Q)

Conditional BF $\forall x(P > Fx) \rightarrow (P > \forall xFx)$

If each thing would be F if it were that P, then everything would be F if it were that P

From these, together with some less controversial principles of counterfactual logic, we can derive the following schema:

Witnessing $\exists x(\exists yFy > Fx)$

(Suppose Witnessing were false; then $\forall x\neg(\exists yFy > Fx)$; so by CEM, $\forall x(\exists yFy > \neg Fx)$; so by Conditional BF, $\exists yFy > \forall x\neg Fx$; but then orthodox counterfactual logic lets us deduce $\exists yFy > Q$ for any Q we please, including Fx for any given x.) Taking F to be 'earring created centred on such-and-such point', Witnessing would suffice for Williamson's "One Earring Per Point" claim.

Williamson would accept Conditional BF for the same reasons he accepts BF for metaphysical necessity; however, he is not a proponent of CEM or of Witnessing. His motivation is something more like the sufficiency principles we discussed in §5.2: '[i]f there were two possible pairs of ear-rings either of which in the circumstances he might equally well have made by cutting along d [a particular diameter], a curious kind of indeterminism would follow: the same process in the same circumstances could lead to the creation of either pair of ear-rings, and nothing in the period up to the moment of creation would determine which pair was created. Indeterminism does not come so cheap' (Williamson 1990: 128). It is not obvious what exactly Williamson has in mind by 'the circumstances'. On one possible interpretation, 'the circumstances' stands for some very detailed proposition which is such that the conjunction of it with the proposition that the craftsman cuts along any given diameter d metaphysically necessitates a completely detailed *microphysical* characterization of history up to the time of the cut, and which would have been true no matter which diameter the craftsman had cut along. But without CEM, it is unclear why one would take there to be any such proposition: there might be multiple microstories such that if the craftsman had cut along d, one of them would have been true, although none of them is such that *it* would have been true in that case. On an alternative interpretation, 'the circumstances' stands for something weaker, and Williamson is appealing to the kind of sufficiency principle that coarse-grained hypertolerantists deny (although he is concerned with 'nearby' possibility rather than 'logical' possibility). However, determinism as standardly understood would not be threatened by the view that there are distinct pairs of earrings that are made by cuts along the same diameter at nearby worlds where the circumstances are the same in this coarse sense.

made centred on *x* is distinct from the possible earring that would have been made if an earring had been made centred on *y*.

Williamson's case provides a dramatic illustration of the disastrous consequences of the Sorites-generating thought that sufficiently small differences as regards "underlying" facts can't possibly make a difference to some higher-level facts. But of course we didn't need Williamson's case to convince ourselves of this: since we can get from a world where a given artefact exists to one where it doesn't by a sequence in which adjacent worlds differ only slightly in underlying respects, there must be some small changes that make a difference. If the appeal of Necessitated Tolerance rested entirely on the Sorites-generating thought, there would be no interesting case against Iteration. But as we discussed in Chapter 3, there is at least one other source of support for Necessitated Tolerance claims that might remain in place even after we have rooted out the impulses that lead to the Sorites paradox, namely, the combination of a case for Tolerance rooted in ordinary modal judgements with the Security Argument for Non-contingency. The case of the unmade earrings does nothing, as far as we can see, to undermine the argument against Iteration based on Necessitated Tolerance claims (motivated in some such non-Soritical way) and Non-hypertolerance claims (motivated in whatever way).

There are several interestingly different options for developing an account of Williamson's case consonant with the claims that motivate denying Iteration. One choice point is whether to accept Williamson's 'One Earring Per Point' thought: that for each point *x* on the diameter there is a unique possible earring *e* such that *e* could be made centred on *x* "in the circumstances". A second choice point is whether to accept the converse 'One Point Per Earring' thought, i.e. that for each possible earring *e*, there is a unique point *x* such that *e* could be made centred on *x* "in the circumstances".[8]

You might have thought that One Point Per Earring was inconsistent with Necessitated Tolerance, which entails that necessarily, any earring made centred on a point *x* could have been made centred on other points a few degrees away from *x*. This entails that when *x* and *y* are close points, and *e* is a possible earring that could be made centred on *x*, it is *possibly* possible for *e* to be made centred on *y*. But if we are giving up on Iteration, this doesn't imply that it is *possible* for *e* to be created centred on *y*.[9]

[8] Goodman (unpublished *e*) develops an Iteration-denying model of Williamson's case that accepts both One Earring Per Point and One Point Per Earring.

[9] Some Iteration-deniers might be tempted by the principle that if it could easily have been that it could easily have been that *P*, it is metaphysically possible that *P*: two steps of "easy" accessibility can never take you somewhere metaphysically inaccessible. But cases like Williamson's suggest that they should not accept this. For some close *x* and *y*, there's a possible earring *e* such that it could easily have been that *e* was made at *x* and could easily have been made at *y*, although it isn't even metaphysically possible for *e* to have been made at *y*.

You might have thought that denying One Earring Per Point would, as Williamson suggests, involve an objectionable failure of determinism.[10] But an Iteration-denier who denies One Earring Per Point is free to maintain that necessarily, if some earring e is made centred on some point x, then it is necessary that if "the circumstances" obtain and some earring is made centred on x, e is made centred on x. Perhaps, although for each point there are multiple possible worlds where distinct earrings are made centred on that point, none of these worlds are possible relative to one another.[11] While the existence of distinct possibilities in which different earrings are created at the same point is prima facie in tension with Microphysical Supervenience, it is consistent with the necessary truth of Weak Microphysical Supervenience (see §6.1). Iteration-deniers might maintain that that is good enough.

If (like Williamson) we accept One Earring Per Point and reject One Point Per Earring, we will be led to a picture where the circumference of the disc is partitioned into several distinct segments, where the partition is generated by the following equivalence relation (call it R):

> being points x and y such that for some possible earring e: necessarily, if a single earring is created centred on x in the circumstances, e is, and necessarily, if a single earring is created centred on y in the circumstances, e is.

Thus all the points within any one segment correspond to the same possible earring, while any two points from different segments correspond to distinct possible earrings.[12] One thing to note is that since earrings can't be bilocated, R cannot hold between opposite points. When x and x' are opposite points, it is impossible in the circumstances for an earring to be made centred on x without a distinct earring being made centred on x', so there can't be a single possible earring associated with both x and x'. Thus there must be at least two segments. On the most lax version of the view, there are *exactly* two segments, and thus only two possible earrings that could have been made in the circumstances: no matter where we cut, we would get the same two earrings, though there are special boundary points such that when we shift the cut from one side of the boundary

[10] To make trouble for Microphysical Supervenience we would need to understand "the circumstances" to be some extremely detailed proposition such that it, together with the proposition an earring is made centred on some point x, settles the truth value of every microphysical proposition.

[11] What's crucial here is that the failure of Iteration also involves failures of the 5 axiom: it can be possible, but not necessarily possible, for e to be an earring made centred on x. Given that the T axiom is uncontroversial for metaphysical necessity, anyone who rejects Iteration must also reject 5.

[12] The segments here are just sets of points on the circumference that would generate the same earring: it is natural but not inevitable to think that such sets correspond to continuous (i.e. topologically connected) sections of the circumference.

to the other, the earrings swap sides. Remember that, from the Iteration-denying perspective, this is consistent with Necessitated Tolerance: if points x and y are close but in different segments, Necessitated Tolerance means that if an earring had been made at x it *would* have been possible for that earring to be made at y; but this only gives us a claim of possible possibility, not one of possibility. The same remarks apply to other less lax views, that posit more than two segments and more than two possible earrings, including Williamson's favoured view on which there are exactly four possible earrings.

Such views might initially seem bizarre, since it seems impossible to give any principled, non-arbitrary answer to the question where the boundaries between distinct segments are located. But as Williamson notes, one should surely maintain that some of the expressions involved in the definition of R above are vague, so one can deploy whatever tools one normally deploys to handle apparently arbitrary vague boundaries.[13]

In sum, while Williamson's case presents Iteration-deniers with an interesting menu of options, it does nothing to undermine their case against Iteration, so long as they are supporting Non-contingency by appeal to a Security Argument rather than some bad Soritical thought.

7.3 Ancestral Hypertolerance and Higher-Order Identity

In evaluating the prospects of denying Iteration for metaphysical necessity, the notion of *ancestral necessity*, introduced in §1.4, will be helpful. Intuitively, a proposition is ancestrally necessary if it is metaphysically necessary, metaphysically necessarily metaphysically necessary, metaphysically necessarily metaphysically necessarily metaphysically necessary, and so on *ad infinitum*. More rigorously, we defined ancestral necessity using a variant of Frege's technique for defining ancestrals:

> Operation O_2 is a *finite iteration* of operation $O_1 :=: O_2$ has every property of operations X such that $\lambda p.p$ has X and whenever O has X, the result of composing O with O_1—i.e. $\lambda p.O(O_1 p)$—also has X.

[13] As Williamson (1990: 133) puts it: 'The symmetry of the original disc makes ignorance a scarcely credible option: what invisible lines could dictate that one pair of cuts a degree apart would yield the same ear-rings, while another pair a degree apart would yield distinct ear-rings? It is far more plausible to assume suppose that judgements without uncontroversial truth-values are indeterminate in truth-value.' (Presumably he would want to put things a bit differently given the views about vagueness he developed in the later Williamson 1994.) An interesting further choice point for the Iteration-denier who accepts One Earring Per Point but not One Point Per Earring is whether to pin the vagueness in the definition of R on the word 'earring' or instead on the word 'possible'. Salmón (Salmon 1981: 275–6) goes the latter route, taking accessibility among worlds to be a vague matter (though he treats such vagueness using a framework of truth value gaps that is prima facie in conflict with classical logic and thus outside the scope of this work).

> p is *ancestrally necessary* := p has every finite iteration of metaphysical necessity.

As we establish in Appendix B, whenever a modal operator $\Box$ obeys our basic modal logic H_{KR}, its ancestral version $\Box^*$ obeys the stronger logic $\mathsf{H}_{\mathsf{S4R}}$. Thus we automatically secure Iteration for $\Box^*$: even if some metaphysically necessary truths are not metaphysically necessarily metaphysically necessary, any ancestrally necessary truth is ancestrally necessarily ancestrally necessary.

This is quite intuitive if we conceive of ancestral necessity as the infinite conjunction of all of the finite iterations of metaphysical necessity. It's easy to see that when p is ancestrally necessary, the proposition that p is necessary is also ancestrally necessary: the result of applying any finite iteration of necessity to the latter is also the result of applying some finite iteration of necessity to the former. The same reasoning applies to the proposition that p is necessarily necessary, and more generally to every proposition that results from p by the application of any finite iteration of necessity. But since each finite iteration of necessity commutes with conjunction, ancestral necessity does too; so the conjunction of all these propositions—i.e. the ancestral necessitation of p—must also be ancestrally necessary. That said, since we don't have the resources to define ancestral necessity directly in a list-like fashion, the official definition is the quantificational one above, and for that reason the derivation in Appendix B needs to rely crucially on the rigidity axioms introduced in §1.5.

The fact that Iteration holds for ancestral necessity makes it an interesting status to think about for those who advocate solving certain Tolerance Puzzles by denying Iteration for metaphysical necessity (while accepting the remaining premises). Giving up Iteration will not be an option for the parallel puzzle about ancestral necessity. Since everything metaphysical possible is ancestrally possible, the truth of Tolerance in the original puzzle suffices for its truth in the parallel puzzle, so giving that up will also not be an option for such theorists either. And the definitions of 'close' required to vindicate the metaphysical-necessity version of Persistent Closeness will typically also vindicate its ancestral-necessity analogue. So, in the interesting cases, the choice will be between rejecting Ancestral Non-contingency and accepting Ancestral Hypertolerance.

Chapter 3's Security Argument for Non-contingency carries over fairly directly to ancestral modality. We can bring out the security-based motivation for Ancestral Non-contingency by focusing not on our judgement that Tolerance is *true*, but on judgements about its modal status. Consider, for example, the judgement that Tolerance is metaphysically necessary—a judgement that is central to the original argument that motivated the rejection of Iteration for metaphysical necessity. Suppose that while Tolerance is necessarily true, it could have failed to be necessarily true. Then it seems that it could *easily* have failed to be necessarily true: all that would have been required would be for certain planks to have been

cut in moderately different positions. But in that case, the truth of our judgement that it is necessarily true would seem to be fragile in a way that is prima facie worrying. Insofar as one was moved to accept the *necessary* truth of Tolerance in order to avoid having to treat our original judgement that Tolerance is true as a risky hostage to empirical fortune, one should be equally moved to accept the *necessary* necessary truth of Tolerance in order to avoid having to treat our judgement that Tolerance is *necessarily* true as a risky hostage to fortune, in exactly the same sense. And this kind of reasoning can be replayed for the judgement that Tolerance is necessarily necessarily true, for the judgement that it is necessarily necessarily necessarily true, and so on. Although the pressure to say these things is not irresistible, the most promising place for resistance seems to be right at the beginning, in the transition from Tolerance to Necessitated Tolerance, rather than somewhere further down the line.

So, the most natural option for someone who wants to use some Tolerance Argument involving metaphysical modality as an argument against Iteration is to accept Ancestral Non-contingency and thus also Ancestral Hypertolerance. And this is indeed what Salmón—the philosopher who has done most to develop the Iteration-denying perspective—does, for the puzzles having to do with originating matter which are his central focus. After introducing his Tolerance premise concerning tables (see §2.4), he asserts without argument that it 'is such that if it is true at all, then it is necessary that it is necessarily true, and it is necessary that it is necessary that it is necessarily true, and so on' (Salmon 1986*a*: 77)—i.e. he endorses Ancestral Non-contingency, thereby implicitly committing himself to Ancestral Hypertolerance.[14]

If ancestral necessity is distinct from metaphysical necessity, it is not a familiar status, so direct "appeals to intuition" will have little force against Ancestral Hypertolerance claims. But even *metaphysical* necessity is not all *that* familiar, and we have in any case not been inclined to put much stock in "appeals to intuition" as an objection to Hypertolerance claims involving metaphysical possibility. Rather, we suggested in Chapter 5 that the weightiest considerations against Hypertolerance are those based on sweeping supervenience claims like Microphysical Supervenience. In the present section, we will argue that there is a strong theoretical case for the ancestral-modality analogue of Microphysical Supervenience:

[14] For any proposition that is ancestrally but not metaphysically possible, we can ask what minimum number of metaphysical possibility operators need to be applied to it to turn it into a truth. You might naturally think that the answers will vary wildly depending on the proposition in question. But in conversation, Jeremy Goodman has entertained a view on which although Iteration fails for metaphysical necessity, anything that's ancestrally possible is at most two steps away. His thought is that objects become maximally unfussy when they are non-concrete, so that (e.g.) if Woody the table were non-concrete, it would then be possible for it to be a table made of any given table-parts, or indeed to be a poached egg.

Ancestral Microphysical Supervenience For every ancestrally possible *p*, there is an ancestrally possible microphysical *q* such that it is ancestrally necessary that if *q* then *p*.

Ancestral Microphysical Supervenience creates problems for Ancestral Hypertolerance claims parallel to those that Microphysical Supervenience creates for their metaphysical-possibility analogues. Its truth would thus be bad news for the approach that treats Tolerance Puzzles as arguments against Iteration for metaphysical possibility.

§6.4 discussed various objections to Microphysical Supervenience, e.g. one based on claims about qualia and one based on claims about "alien" fundamental properties; similar objections can be raised to Ancestral Microphysical Supervenience. Since the possible fallbacks and refinements which we considered in that section can also be used as responses to the objections to Ancestral Microphysical Supervenience, we will ignore these worries here.

As we saw in §6.3, Microphysical Supervenience does not completely blow Hypertolerance-accepting strategies out of the water: there are principled "coarse-grained Hypertolerance" packages that hold onto Microphysical Supervenience while endorsing a wide range of Hypertolerance claims concerning properties that are not too specific. Similarly, Ancestral Microphysical Supervenience does not blow Ancestral Hypertolerance-accepting strategies out of the water: one could coherently accept Ancestral Hypertolerance for a wide range of coarse-grained families of properties, while rejecting Tolerance for certain super-specific families. However, Ancestral Microphysical Supervenience does not combine comfortably with the argument against Iteration based on the rejection of the corresponding coarse-grained Hypertolerance theses for metaphysical necessity. Suppose for example that one takes it to be ancestrally possible but metaphysically impossible for Woody to have been originally composed by a certain collection of atoms *C*—say, one does not include any of those that in fact composed Woody. Given Ancestral Microphysical Supervenience, there is some ancestrally possible microstory *p* that ancestrally necessitates that Woody is originally composed by *C*. Since everything ancestrally necessary is metaphysically necessary, *p* must be metaphysically impossible. One will thus be committed to the existence of counterexamples to Iteration even within the domain of the microphysical. This isn't completely out of the question: for example, one might think that when two atoms are sufficiently far apart, it is possibly possible but not possible for them to be close together. But it is bizarre to think that the only ancestrally possible worlds where Woody is made of *C* are ones where certain *microphysical* objects are stretched beyond their possible limits, in such a way that their pattern of microphysical relations is merely ancestrally possible. If Woody's being made of *C* is ancestrally possible at all, it is much more promising to think that it is ancestrally compossible with some *metaphysically* possible microstories (which in fact metaphysically necessitate that *C* composes a table distinct from Woody).

Thus, proponents of the Iteration-denying strategy are under a lot of pressure to reject Ancestral Microphysical Supervenience. But why would anyone accept it? To get a feel for what's compelling about it, it will be helpful to begin with some much less ambitious claims of *de re* ancestral necessity. Consider the Tri-State Area, which comprises New York, New Jersey, and Connecticut.[15] There is a strong case, quite independent of Iteration, that the relationship between these four objects includes certain ancestrally necessary connections. For example:

(1) It is ancestrally necessary that an object is in the Tri-State Area just in case it is in New York, New Jersey, or Connecticut.

While one could be led to endorse (1) by many different arguments, the argument that we find gripping appeals to the following plausible *higher-order identity* claim:

(2) To be in the Tri-State Area is to be in New York, New Jersey, or Connecticut.[16]

On the reading we are interested in, claims like (2) are higher-order analogues of regular first-order identity claims like 'Hesperus is Phosphorus'. Given the policy for using words like 'property' which we announced in §1.2, we could express the same thing by saying, 'The property of being in the Tri-State Area is the property of being in New York, New Jersey, or Connecticut.' But sentences like (2) show that English has the resources to articulate the relevant kind of claim without going in for the kind of reification that requires a steady hand with the Fregean grains of salt.[17]

The argument from a higher-order identity like (2) to a corresponding ancestrally necessitated biconditional like (1) is straightforward. Uncontroversially, it is ancestrally necessary that an object is in the Tri-State Area just in case it is in the Tri-State Area. Given (2), (1) follows immediately from this claim by (higher-order) Leibniz's Law. We can do the same thing with any predicates F and G: since $\Box^*$ obeys our basic modal logic, we have $\Box^*\forall x(Fx \leftrightarrow Fx)$, so by Leibniz's Law we can infer $F = G \rightarrow \Box^*\forall x(Fx \leftrightarrow Gx)$.[18]

[15] We are aware that other things are also called "the Tri-State Area".

[16] Take 'in' here to mean *partially in* rather than *entirely in*.

[17] Dorr (2016*b*: §§1–2) argues that English sentences like (2) have a salient reading as predications of higher-order identity, while acknowledging certain other possible interpretations.

[18] In the official logical system of Chapter 1, this Leibniz's Law inference has to be mediated by two applications of Eβ, as follows:

1.	$F = G$	Premise
2.	$\Box^*\forall x(Fx \leftrightarrow Fx)$	Premise
3.	$(\lambda X.\Box^*\forall x(Fx \leftrightarrow Xx))(F)$	2, Eβ
4.	$(\lambda X.\Box^*\forall x(Fx \leftrightarrow Xx))(G)$	1, 3, LL
5.	$\Box^*\forall x(Fx \leftrightarrow Gx)$	4, Eβ

While we find (2) plausible, we do not mean to suggest that it is obviously true. §7.4 will extensively discuss one kind of argument against it (based on a "structured" theory of higher-order identity). Another, less promising, argument against it appeals to considerations involving propositional attitudes. For example, it seems that someone might want to be in the Tri-State Area without wanting to be in New York, New Jersey, or Connecticut (owing to a mistaken belief about the composition of the Tri-State Area): prima facie, (2) seems to rule out such a possibility by Leibniz's Law. But we should be wary of arguments against identity claims which turn on apparent substitution-failures in propositional attitude contexts. For it also seems that someone might want to visit Hesperus without wanting to visit Phosphorus; but Hesperus is Phosphorus, so there must be *something* wrong with the argument that takes that appearance at face value and then invokes Leibniz's Law to conclude that Hesperus is not Phosphorus. It is not obvious what the problem is. Maybe 'Someone could want to visit Hesperus without wanting to visit Phosphorus' is simply false (Salmon 1986*b*; Soames 1987; Braun 1988); maybe it is context-sensitive and false on all uniform resolutions of its context-sensitivity, though true on a salient non-uniform reading (Dorr 2014*b*; Goodman and Lederman 2021); or maybe it is true (on a uniform interpretation), and this is another of the many cases where the logic of English is more complicated than that of our formal language (see §1.3).[19]

As well as setting aside arguments against (2) based on attitude ascriptions, we should also set aside a certain argument *for* (2) which would be dialectically ineffective in the present context. This argument derives (2) from Intensionalism, the principle that metaphysically necessary coextensiveness suffices for higher-order identity, together with the premise that it is metaphysically necessary that an object is in the Tri-State Area iff it is in New York, New Jersey, or Connecticut. As noted in §1.4, Intensionalism entails Iteration (since it entails that there is only one necessary truth), so Iteration-deniers will have no patience for arguments that assume it.

But the appeal of (2) is by no means tied to Intensionalism. It just looks plausible on its face! If its appeal isn't obvious to you, consider the compelling, if obscure, thought that the Tri-State Area is "nothing over and above" its three constituent states. (2) seems at least to begin to capture that thought. By contrast, it does not seem to be captured at all by the mere claim that it is metaphysically necessary that an object is in the Tri-State Area iff it is in New York, New Jersey, or Connecticut.

Finally, if one or both of F or G are complex predicates like 'be in the Tri-State Area'—formally, $\lambda y.\,\mathrm{In}(y, \mathrm{tsa})$—we can add a final step of β-reduction under the $\Box^*$, e.g. replacing $\Box^*\forall x(\lambda y.\,\mathrm{In}(y, \mathrm{tsa})(x) \leftrightarrow (\lambda y.\,\mathrm{In}(y, \mathrm{ny}) \vee \mathrm{In}(y, \mathrm{nj}) \vee \mathrm{In}(y, \mathrm{ct})(x)))$ with $\Box^*\forall x(\mathrm{In}(x, \mathrm{tsa}) \leftrightarrow \mathrm{In}(x, \mathrm{ny}) \vee \mathrm{In}(x, \mathrm{nj}) \vee \mathrm{In}(x, \mathrm{ct}))$.

[19] It's also not obvious that the problem with the bad argument that Hesperus isn't Phosphorus is the same as the problem with the bad argument against (2). Perhaps (as Dorr (2016*b*: n. 24) suggests, following Soames (2002: ch. 10)) Leibniz's Law is valid in English for syntactically simple expressions like 'Hesperus', but not for complex expressions like 'be in the Tri-State Area'.

(2) is only about one relation that things might bear to the Tri-State Area, namely the *in* relation. But clearly there is nothing special about that relation. Insofar as (2) is plausible, it is plausible that for every qualitative $n+1$-ary relation R, there is an qualitative $n+3$-ary relation R', such that to stand in R to the Tri-State Area is to stand in R' to New York, Connecticut, and New Jersey.[20] This gives us a localized ancestral supervenience claim: every ancestrally possible predication of a qualitative $n+1$-ary relation of some $x_1, \ldots, x_n$ and the Tri-State Area is ancestrally necessitated by (since it is identical to) some ancestrally possible predication of a qualitative $n+3$-ary relation of $x_1, \ldots, x_n$, New York, New Jersey, and Connecticut.

Similar relationships can be found in many other domains. For example, it seems plausible that the intimate relationship between a particular molecule m of diatomic oxygen and its two constituent oxygen atoms a_1 and a_2 includes higher-order identities such as (3):

(3) For m to be an oxygen molecule is for a_1 and a_2 to compose an oxygen molecule.

Nor is the plausibility of such identities confined to cases where the objects on one side *compose* the object on the other, as we can see from two more examples:

(4) Muhammad Ali's left fist and Muhammad Ali's left hand are such that: to be hit by the former is to be hit by the latter while it is clenched.

(5) Plato's beard is such that: to be identical to it is to be a beard of Plato.

We don't want to overstate the standing of sentences like (2)–(5). Plausibly, some of them are context-dependent or polysemous, and true only on some suitable resolutions of their context-sensitivity or polysemy. For example, there are different ways of talking about beards: we sometimes say things like 'I missed your beard and I'm glad you've grown it back', but in a different mood we might say things like 'Your new beard is much greyer than the one you had as a young man.' This plausibly involves different interpretations of 'beard'; for (5) to be defensible, we need to interpret it in the first way. But for the point we are making, it suffices that the relevant identities are true in the context we are making salient.[21]

[20] For example, if we take R to be *identity*, it follows that there is a certain four-place relation such that to *be* the Tri-State Area is to bear that relation to New York, New Jersey, and Connecticut. Plausibly this is something along the lines of *being a territorial unit composed of*, but we might not have a word that picks it out exactly. Another interesting value of R is the instantiation relation $\lambda Xy.Xy$ (of type $\langle\langle e\rangle, e\rangle$): in the presence of full Beta-conversion (see note 13 in Chapter 1), the truth of the claim for this R would imply its truth for every other R.

[21] Even on the favourable interpretation of 'beard', (5) does not look very plausible unless we hold some form of "animalism" about people: if one thought that moving Plato's brain into a new body would lead to Plato having a new body, one should probably not think that his beard would follow him into that new body (even if it was also bearded). A different kind of worry stems from certain odd possibilities

If identities like (2)–(5) are true, they suggest a general moral: non-fundamental objects—not only areas and molecules and fists and beards, but tables, pyramids, and so on—are linked to more fundamental objects by a rich array of higher-order identities. And one upshot of these identities will be a similarly rich array of ancestrally necessary truths, tying properties and relations among less fundamental objects to properties and relations among more fundamental objects. Given our working physicalistic orientation, reflecting on these examples inclines us strongly towards accepting Ancestral Microphysical Supervenience.[22]

Indeed, to our minds, some version of this picture provides the most compelling theoretical basis for Microphysical Supervenience. The gripping thought pushing us towards that view is that the non-fundamental objects are "nothing over and above" the fundamental microphysical ones. And higher-order identities of the sort we have been considering provide a relatively clear and satisfying way of making sense of such "nothing over and above" claims.

But even setting aside the hermeneutics of 'nothing over and above' slogans, it seems deeply suspicious to think that there are two quite different categories of non-fundamental entities: those that are tied to the ground floor by higher-order identities, and others that float free from the ground floor as far as identity-theoretic facts are concerned, although they nevertheless enjoy some other kind of intimate tie to the ground floor (e.g. via Finean essence or grounding), thanks to which facts about them still supervene with metaphysical necessity on the ground-floor facts.

One might be led to posit such a contrast by noting that our examples—the Tri-State Area, the oxygen molecule, the fist, the beard—have been cherry-picked: in these cases there are relatively plausible and easy-to-write-down identities connecting the relevant objects to more fundamental ones. By contrast, in the case of objects like tables, it is a lot harder to actually write down defensible candidates to be the higher-order identities playing this role. One might initially be tempted to

in which Plato might be thought to have more than one beard—e.g. if he grew a second bearded head, or a second bearded face on the back of his head. However, variants of (5) are available that avoid these objections—e.g. 'Plato's beard and Plato's face are such that: to be identical to be the former is to be a beard of the latter.'

[22] The most obvious route to Ancestral Microphysical Supervenience from an enthusiasm about the relevant kinds of higher-order identities would go by way of the thesis that *every proposition is microphysical*. Of course, convincing oneself of that thesis would also require being open to controversial kinds of property-identities not having to do with particular objects. For example, if one denied that *being an atom that is part of a city* is identical to any microphysical property, then one might reject Ancestral Microphysical Supervenience on the grounds that there are ancestrally possible worlds where the microphysical facts are the same but there are no cities. But the suspicions that follow about the view that there are two drastically different categories of nonfundamental *objects* carry over to the view on which there are two drastically different categories of nonfundamental *properties*. There may also be other routes to Ancestral Microphysical Supervenience that don't involve the thesis that all propositions are microphysical: for example, one could derive it from the thesis that every proposition is *about* the collection of all microphysical objects (in a sense of 'about' that will be explored in Chapter 15), together with the thesis that all propositions about that collection ancestrally supervene on microphysical propositions.

think that something of the form of (6) is true (where $a_1, \ldots, a_n$ are the molecules that in fact originally composed Woody, and k is a number considerably bigger than 1 but considerably smaller than n):

(6) To be identical to Woody is to be the first table-shaped object to be originally composed of at least k of atoms $a_1, \ldots, a_n$.

But (6) is too simple to withstand even a cursory search for counterexamples. Surely a table made out of k of $a_1, \ldots, a_n$ by a different artificer at a very different time according to a very different plan need not have been Woody (and its parts need not have been parts of Woody), even if Woody could under more favourable circumstances have been made of k of $a_1, \ldots, a_n$. Of course, we can formulate more complicated variants of (6) that attempt to take these further factors into account in appropriately nuanced ways. But any variant of (6) simple enough for us to write down will feel artificial and over-committal, even if it manages to avoid clear counterexamples.

However, we don't think *this* contrast is indicative of some deep metaphysical contrast between Woody and objects like the Tri-State Area, since the relevant sort of difficulty in spelling things out crops up wherever there is vagueness. For example, it is plausible that there is some real number x such that to be pretty full is to be at least x% full, even though for each particular choice of cutoff x, the sentence 'To be pretty full is to be at least x% full' is too committal to be assertible. The difficulty in singling out any specific candidate x certainly doesn't stop us thinking that there is an x such that being at least x% full is metaphysically necessary and sufficient for being pretty full. And if we accept that there is such an x, it seems plausible that this claim of metaphysically necessary equivalence can be strengthened to a higher-order identity: to be pretty full is to be at least x% full.[23] In a non-Intensionalist setting, the task of formulating a nontrivial plausible identity of the form 'To be pretty full is to be …' may require making further arbitrary-seeming choices owing to the multiplication of intensionally equivalent properties (suppose for example that we thought that *being at most 20 per cent empty* is distinct from *being at least 80 per cent full*); but here again we can naturally appeal to vagueness rather than concluding that being pretty full is distinct from all the candidates. With most vague words, the situation is even messier: for example, whether someone is bald seems to depend not just on how many hairs they have, but on various facts about what those hairs are like and how they are distributed on the scalp. But in none of these cases does the folly of putting forward any

[23] At least, this seems plausible given the kind of broadly linguistic theory of vagueness that we favour. On a non-linguistic account of vagueness like that of Bacon (2018*b*), there is more prospect of a principled position that accepts the existentially quantified metaphysical necessity claim but denies the corresponding existentially quantified identity claim. In Chapter 8 we will have a bit more to say about Bacon's picture.

particular identity with the vague word in question on one side and some complex expression in more precise language on the other side show that that disjunction of all such identities is incorrect. And we submit that the same diagnosis is in order when it comes to facts about particular tables, pyramids, and so on. One should be quite suspicious about views on which, e.g., the identities about the Tri-State Area and the oxygen molecule are true, but *being a region that contains Woody* is not bearing any relation to any microphysical objects. The suspicion is that such views are reading too much metaphysical significance into the contrast between the simple cases where reasonably plausible higher-order identities can be written down and more complex cases where considerations of vagueness (and indeed plain old ignorance) make this impossible.

Of course, there is an argument that *containing Woody* is not identical to any microphysical property which has no parallel for the Tri-State Area, namely the argument for Ancestral Hypertolerance (with respect to an appropriately fine-grained family of properties, based on Tolerance, Ancestral Non-contingency, and Ancestral Iteration). One person's Modus Ponens is another's Modus Tollens! But we think some wariness is in order here, just as with the use of Tolerance Arguments to argue against Microphysical Supervenience. While the other premises of the relevant Tolerance Arguments are prima facie plausible, in aggregate they do not seem so very forceful as to constitute a refutation of the sort of sweeping metaphysical vision of the relationship between the everyday world and the reality described by physics that these supervenience claims aspire to.

7.4 Structure

We conclude that there is pressure on friends of Ancestral Hypertolerance to find some principled basis for rejecting identities like (2)–(5), as opposed to accepting them while insisting that nothing relevantly similar is true when it comes to the objects that star in Tolerance Paradoxes. And indeed, many readers will have regarded these identities as hopeless from the beginning. For opponents of Intensionalism often endorse, or implicitly assume, a "structured" theory of higher-order entities, which would rule out the above identities as non-starters. One characteristic claim of such structured theories is that the result of predicating a property of one entity (of some given type) is distinct from the result of predicating any other property of any entity, or the result of predicating the same property of a different entity. Contrapostively:

Structure If for x to be F is for y to be G, then x is y and to be F is to be G.
$$\forall F \forall G \forall x \forall y (Fx = Gy \rightarrow (x = y \wedge F = G))^{24}$$

[24] This is a schema with one instance for each type σ: x and y are variables of type σ, and F and G are thus variables of type $\langle\sigma\rangle$.

Structure already rules out a wide range of higher-order identities of the sort whose plausibility we have been advocating. For example, since Muhammad Ali's fist is not identical to Muhammad Ali's hand (since, e.g., the latter touched some things that the former never touched), Structure rules out:

(7) For Muhammad Ali's fist to be touching a flower is for Muhammad Ali's hand to be clenched while touching a flower.

Alternatively we could rule out the combination of Structure and (7) by appeal to the fact that touching a flower isn't the same as being clenched while touching a flower.

Structure is restricted to identities between monadic predications, but it has natural generalizations that will rule out our other example identities. The picture we have in mind will endorse (the universal generalizations of) principles like the following:

$$Rab = Scd \rightarrow (R = S \wedge a = c \wedge b = d)$$
$$\lambda x.Rax = \lambda x.Sbx \rightarrow (R = S \wedge a = b)$$
$$Fa \neq Rbc$$
$$Fa \neq Gb \quad \text{(where } F \text{ and } G \text{ are of different types)}$$

These can be subsumed under the following general schema (here $k \geq 0$, $m, n \geq 1$, and all the variables can have any types that make sense):

Generalized Structure (a) If $(\lambda x_1 \ldots x_k.Rx_1 \ldots x_k y_1 \ldots y_n) = (\lambda x_1 \ldots x_k.Sx_1 \ldots x_k z_1 \ldots z_n)$ then $R = S$ and $y_1 = z_1$ and $\ldots$ and $y_n = z_n$.

(b) $(\lambda x_1 \ldots x_k.Rx_1 \ldots x_k y_1 \ldots y_n) \neq (\lambda x_1 \ldots x_k.Sx_1 \ldots x_k z_1 \ldots z_m)$ where $n \neq m$ or any y_i is of a different type from z_i.[25]

Here R and S are variables standing, respectively, for $n + k$-ary and $m + k$ ary relations, where the first k relata are of the same type. Thus $\lambda x_1 \ldots x_k.Rx_1 \ldots x_k y_1 \ldots y_n$ and $\lambda x_1 \ldots x_k.Sx_1 \ldots x_k z_1 \ldots z_m$ are the k-ary relations that result, respectively, from plugging entities $y_1, \ldots, y_n$ into R's last n argument-places and plugging entities $z_1, \ldots, z_m$ into S's last m argument places. Structure is the special case of (a) where $k = 0$ and $n = 1$. Generalized Structure will rule out some more of our example identities, for example the claim that to be identical to Plato's beard is to be

[25] Formally, part (a) of Generalized Structure is a schema with one instance for any two sequences of types $\sigma_1, \ldots, \sigma_k$ (the types of $x_1, \ldots, x_k$) and $\sigma'_1, \ldots, \sigma'_n$ (the types of $y_1, \ldots, y_n$ and $z_1, \ldots, z_n$). Part (b) is a schema with one instance for any three sequences of types for $x_1, \ldots, x_k, y_1, \ldots, y_n$ and $z_1, \ldots, z_m$, so long as the latter two sequences are distinct.

a beard of Plato. Formally this would be $(\lambda x.x = \mathrm{b}) = (\lambda x.\,\mathrm{BeardOf}(x, \mathrm{p}))$, which given Generalized Structure implies the obviously false $\mathrm{p} = \mathrm{b}$ (Plato is identical to the beard) and $(=) = \mathrm{BeardOf}$ (*being identical to* is *being a beard of*).

Although Generalized Structure is quite strong in some ways, it is still not strong enough to rule out all the identities in our examples, since many of them correspond to formulae in our formal language that are not substitution instances of $\lambda x_1 \ldots x_k.Rx_1 \ldots x_k y_1 \ldots y_n = \lambda x_1 \ldots x_k.Sx_1 \ldots x_k z_1 \ldots z_m$ (for any types of R and S). We could formulate further generalizations of Structure that would rule them out. But the task of exploring generalizations of Structure is moot, since as it turns out, Structure is already inconsistent in our classical background logic. For recall that the variables x and y in Structure need not be of type e (the type of objects). They can be of any type σ, in which case the predicates F and G must be of type $\langle\sigma\rangle$ (the type of properties of type-σ things). In particular, x and y can be variables of type $\langle\rangle$, the type of propositions (in which F and G will be of type $\langle\langle\rangle\rangle$, the type of monadic operators). The classical inconsistency of Structure for these particular types is a corollary of the Russell-Myhill theorem (Russell 1903: appendix B; Myhill 1958).[26] The proof is not difficult; we give it in a footnote.[27]

The Russell-Myhill theorem is, we think, a result of enormous and under-appreciated importance for metaphysics, especially since it has been common for metaphysicians to talk about properties and propositions in ways that seem to presuppose something like Structure.[28] But while Structure is ruled out by the theorem, there are weakenings of Structure (and Generalized Structure) which

[26] For recent discussions, see Uzquiano 2015 (384–6), Hodes 2015 (63–4), Dorr 2016*b* (§6) Fritz 2017*b*, Goodman 2017, and Robertson Ishii and Salmón 2020.

[27] Following the version in Dorr 2016*b*, pick some arbitrary proposition p; let O abbreviate

$$\lambda q.\exists X((q = Xp) \wedge \neg Xq)$$

and let q be Op. In the terminology of Robertson Ishii and Salmón 2020, an O-proposition is a "stone-caster": it attributes to p some property that it itself lacks. Here is the derivation.

1.	$O(Op) \leftrightarrow \exists X(Op = Xp \wedge \neg X(Op))$	Eβ
2.	$\forall X(Op = Xp \rightarrow X(Op)) \rightarrow O(Op)$	UI, Ref
3.	$\neg O(Op) \rightarrow O(Op)$	(1, 2)
4.	$O(Op)$	(3)
5.	$\exists X(Op = Xp \wedge \neg X(Op))$	(1, 4)
6.	$\exists X(Op = Xp \wedge O \neq X)$	(4, 5, LL)

[28] Admittedly, philosophers talking in these ways (e.g. Salmon 1986, 1989*a*, 1993; Soames 1985, 1987, 1989, 2010; King 2007) often work in a first-order setting where properties are "just more objects". Analogues of the Russell-Myhill theorem are still apt to arise in first-order theories of properties (as argued in Deutsch 2008). However, such theories must (unless they reject classical propositional logic) already somehow restrict Naïve Property Conversion—their analogue of Eβ (see §1.2)—in order to block Russell's paradox. If one is in any case in the market for some restriction, one might hope that the restriction can be made strong enough to also block (the analogue of) the Russell-Myhill theorem and preserve (the analogue of) Structure. However, the required restrictions seem much more drastic than those needed to block Russell's paradox, and we predict that they will seriously diminish the capacity of such theories to sustain the ordinary ways of reasoning that make property-talk so expressively useful.

are consistent in our background logic, but would still suffice to rule out many of the higher-order identities which we appealed to in our case for Ancestral Microphysical Supervenience. One such weakening is what Dorr (2016*b*: §6) calls 'Atomic Structure': this is just the special case of Structure where the variables x and y are of type e (i.e. they "stand for objects") and F and G are thus of type $\langle e \rangle$ (they "stand for properties of objects"). This is already enough to rule out identity (7), since 'Muhammad Ali's fist' and 'Muhammad Ali's hand' stand for objects.[29] And if we likewise restrict Generalized Structure by requiring the variables $y_1, \ldots, y_n$ and $z_1, \ldots, z_m$ to be of type e—call the result of doing this 'Generalized Atomic Structure'—we can similarly rule out some of the other example identities without falling into inconsistency.[30]

While Atomic Structure is consistent in our background logic, it is inconsistent with certain other quite attractive general principles which we might want to add. The problem can be seen by considering the following instance of Atomic Structure:

$$(8) \qquad (\lambda x.Axc)(c) = (\lambda x.Acx)(c) \rightarrow (c = c \wedge (\lambda x.Axc) = (\lambda x.Acx))$$

Taking c to be a name for Cleopatra and A to mean 'admires', this says that if for Cleopatra to be an admirer of Cleopatra is for Cleopatra to be admired by Cleopatra, then Cleopatra is Cleopatra and to admire Cleopatra is to be admired by Cleopatra. The consequent of (8) is clearly false: it is not the case that to be an admirer of Cleopatra is to be admired by Cleopatra, since there could have been (and probably were) people who admired Cleopatra but were not admired by her. But the antecedent of (8), i.e. (9), seems very plausible:

$$(9) \qquad (\lambda x.Lxc)(c) = (\lambda x.Lcx)(c)$$

It is natural to think that for Cleopatra to be an admirer of Cleopatra is just for Cleopatra to admire Cleopatra, and that's also what it is for Cleopatra to be admired

[29] For the consistency of Atomic Structure, and various other restrictions of Structure to specific carefully selected types, see Goodman unpublished *d*.

[30] The situation is a bit delicate. Generalized Atomic Structure rules out some claims in the vicinity of (2), such as:

$$(\lambda x.\,\mathrm{In}(x, \mathrm{tsa})) = (\lambda x.(\lambda yzuv.\,\mathrm{In}(y, z) \vee \mathrm{In}(y, u) \vee \mathrm{In}(y, v))(x, \mathrm{ny}, \mathrm{nj}, \mathrm{ct}))$$

But it doesn't rule out (2) itself, which in our formal language would be

$$(\lambda x.\,\mathrm{In}(x, \mathrm{tsa})) = (\lambda x.\,\mathrm{In}(y, \mathrm{ny}) \vee \mathrm{In}(y, \mathrm{nj}) \vee \mathrm{In}(y, \mathrm{ct}))$$

We could infer the falsity of this from the falsity of the previous identity by appeal to Beta-conversion (see below); but this is little help since Atomic Structure is inconsistent with Beta-conversion. The package that includes Beta-conversion together with Generalized Qualitative Atomic Structure, a Beta-conversion-compatible weakening of Generalized Atomic Structure to be introduced below, is actually a more powerful tool for ruling out the kinds of identities we relied on in making the case for Ancestral Microphysical Supervenience.

by Cleopatra. (9) is not only intuitive on its face, but follows from an attractive general schema, which upgrades the biconditional in Extensional Beta-conversion to a propositional identity:

Beta-identity $(\lambda v_1 \ldots v_n.Q)(a_1, \ldots, a_n) = Q[a_1/v_1, \ldots, a_n/v_n]$

$(\lambda x.Axc)(c) = Acc$ and $(\lambda x.Acx)(c) = Acc$ are both instances of Beta-identity; by Leibniz's Law, they entail $(\lambda x.Axc)(c) = (\lambda x.Acx)(c)$.[31]

While the plausibility of identities like (9) clearly cries out to be subsumed under *some* kind of formal generalization, there are possible generalizations weaker than Beta-identity. For example, some philosophers who are dubious about Beta-identity, such as Dorr (2016*b*) and Goodman (unpublished *d*), accept the following weaker version:

Non-vacuous Beta-identity $(\lambda v_1 \ldots v_n.Q)(a_1, \ldots, a_n) = Q[a_1/v_1, \ldots, a_n/v_n]$
where all of $v_1, \ldots, v_n$ have free occurrences in Q.

This avoids some objections to Beta-identity, but is still sufficient to imply (9) (and hence to rule out Atomic Structure).[32] The same is true of other principled restrictions of Beta-identity—for example, a restriction that required all the variables $v_1, ..., v_n$ to occur *exactly* once in Q (see Bacon unpublished).

The arguments based on instances of Beta-identity also provide a useful supplement to the argument against Structure from the Russell-Myhill theorem. While our own view is that the intuitive and theoretical attractions of classical logic are far greater than those of Structure, less stalwart souls may have been leaning towards taking the Russell-Myhill theorem as a reductio not of Structure but of the background classical logic used in the theorem. But some of them may not have noticed that Structure also requires giving up such obvious-looking claims as that for Cleopatra to admire Cleopatra is for Cleopatra to admire herself; and when they do notice this, it may do something to dislodge any impression they may have had that Structure is *obviously* true, or that identities like (2) which it rules out are *obviously* false.

However, it is possible to weaken Atomic Structure further, in ways that make it consistent with Non-vacuous Beta-identity and even with full Beta-conversion, while still sufficing to rule out the kinds of higher-order identities we are interested

[31] Beta-identity follows in turn from the full Beta-conversion schema: see note 13 in Chapter 1. The latter seems stronger than Beta-identity: we see no way of getting in the background logic from Beta-identity to instances of Beta-conversion like $F(\lambda x.(\lambda yz.Rzy)xx) \leftrightarrow F(\lambda x.Rxx)$.

[32] The restriction in Non-vacuous Beta-identity may seem off-puttingly ad hoc. But Dorr (2016*b*) argues that this is just an artefact of our current notation—the analogous principle in a language where vacuous lambda-abstracts were disallowed as ill-formed would not look in any way ad hoc, and those who accept Non-vacuous Beta-identity while rejecting Beta-identity should regard such languages as superior for the purpose of stating simple true theories.

in. In the monadic case, the key move for achieving consistency with Non-vacuous Beta-identity is to require F and G to be *qualitative* (see §5.1 and §14.1):

Qualitative Atomic Structure If F and G are both qualitative properties of objects, and for x to be F is for y to be G, then x is y and to be F is to be G.

This will still suffice to rule out our candidate identity (7), given the further plausible premise that *touching a flower* and *being clenched while touching a flower* are qualitative properties. And, as it turns out, Qualitative Atomic Structure is consistent with Non-vacuous Beta-identity, though it would lead to absurdity when combined with full Beta-identity.[33] If we want a principle consistent with full Beta-conversion, we can do so by requiring that F and G be not just qualitative but *non-forgetful*, where a property F is forgetful iff it is $\lambda x.p$ for some p. Call this *Weak Qualitative Atomic Structure*.

Formulating a polyadic generalization of Qualitative Atomic Structure which preserves its compatibility with [Non-vacuous] Beta-identity requires a few further modifications beyond the restriction to qualitative relations. Here is one that does the job:

Generalized Qualitative Atomic Structure (a) Whenever R and S are qualitative, objects $y_1, \ldots, y_n$ are all distinct, objects $z_1, \ldots, z_n$ are all distinct, and $\lambda x_1 \ldots x_k.Rx_1 \ldots x_k y_1 \ldots y_n = \lambda x_1 \ldots x_k.Sx_1 \ldots x_k z_1 \ldots z_n$, there is some permutation $\pi_1, \ldots, \pi_n$ of the numbers $1, \ldots, n$ such that y_1 is z_{π_1} and … and y_n is z_{π_n} and S is $\lambda x_1 \ldots x_k u_1 \ldots u_n.Rx_1 \ldots x_k u_{\pi_1} \ldots u_{\pi_n}$.[34]

(b) If R and S are qualitative, then for all objects $y_1, \ldots, y_n, z_1, \ldots, z_m$ $\lambda x_1 \ldots x_k.Rx_1 \ldots x_k y_1 \ldots y_n \neq \lambda x_1 \ldots x_k.Sx_1 \ldots x_k z_1 \ldots z_m$ (where $m \neq n$).

This rather complicated principle is the result of making three further changes to the original Generalized Structure in addition to requiring $y_1, \ldots, y_n$ and $z, \ldots, z_n$

[33] For the consistency result see Goodman unpublished *d*. To see the problem with Beta-identity, let p be some qualitative proposition (e.g. $\forall y(y = y)$). Then $\lambda z.p$ (being such that p) is a qualitative property, but Beta-identity implies that $(\lambda z.p)(x) = p = (\lambda z.p)(y)$ for any x and y, which in conjunction with Qualitative Atomic Structure has the absurd consequence that there is only one object.

One might wonder whether the result of weakening Structure by requiring F and G to be qualitative, but still allowing x and y to be of any type, would be consistent. The answer is no, given some very basic principles about qualitativeness: the Russell-Myhill derivation can be conducted entirely within the domain of the qualitative, by choosing the target proposition p to be something qualitative, and defining O as *attributing to p some qualitative property that one lacks* ($\lambda q.\exists X(\text{Qual}(X) \wedge q = Xp \wedge \neg Xq)$). O must itself be qualitative given that p, qualitativeness, and the logical constants are qualitative and that qualitativeness is preserved under application. But then we can rerun the derivation to prove that there must be some *other* qualitative O' such that $Op = O'p$.

[34] Formally, we don't really need to quantify over permutations here: the principle is a schema with one instance for every pair of a number n and a sequence of types for $x_1, \ldots, x_k$, and the consequent is a disjunction with one disjunct for each of the $n!$ permutations of the numbers $1, \ldots, n$.

to be of type *e* (the restriction that yielded Generalized Atomic Structure). First, we require *R* and *S* to be qualitative: that restriction has already been motivated. Second, we require that $y_1, \ldots, y_n$ be all distinct and that $z_1, \ldots, z_n$ be all distinct: this is needed to make sure that the principle doesn't conflict with Beta-identities like $(\lambda x.Lxx)c = Lcc$ (where *L* is qualitative). And third, we allow that $y_1, \ldots, y_n$ might be the result of applying some permutation to $z_1, \ldots, z_n$, rather than being the same things taken in the same order: this is needed to make sure that the principle doesn't conflict with Beta-identities like $(\lambda xy.Lyx)ac = Lca$ (where *L* is qualitative and *a* and *c* are distinct objects, e.g. Antony and Cleopatra). Finally, if we want, we can weaken the principle further by requiring *R* and *S* to be *non-forgetful* in every argument[lace; the resulting principle is consistent with full Beta-identity.[35]

Generalized Qualitative Atomic Structure suffices, given plausible premises about the qualitativeness of various properties and relations (*being in*, hitting, identity, instantiation), to do as much work as Generalized Structure would have done in ruling out the kinds of identities that featured in the argument of the previous section. Moreover, if we combine Generalized Qualitative Atomic Structure with Non-vacuous Beta-identity, we can rule out a lot more, by applying Beta-conversion on both sides of some candidate identity to turn it into one of the form targeted by Generalized Qualitative Atomic Structure. All of our example identities can be ruled out in this way. Consider (2), for example, which would be rendered in our formal language as follows:

$$(\lambda x.\,\mathrm{In}(x, \mathrm{tsa})) = (\lambda x.\,\mathrm{In}(x, \mathrm{ny}) \vee \mathrm{In}(x, \mathrm{nj}) \vee \mathrm{In}(x, \mathrm{ct}))$$

Given Non-vacuous Beta-identity, this implies

$$(\lambda x.\,\mathrm{In}(x, \mathrm{tsa})) = (\lambda x.(\lambda yzuv.\,\mathrm{In}(y, z) \vee \mathrm{In}(y, u) \vee \mathrm{In}(y, v))(x, \mathrm{ny}, \mathrm{nj}, \mathrm{ct}))$$

But given that the 'in' relation is qualitative, the denotation of the complex four-place predicate on the right is also qualitative (since it is defined in terms of 'in' and disjunction); since we also have the distinctness of the three states, Generalized Qualitative Atomic Structure rules out the truth of this identity (thanks to the '$n = m$' clause).

Our identities (2)–(5) are very plausible. Nevertheless, if Structure were consistent, its power and elegance might tempt us to relinquish them, so as to avail ourselves of the clean and complete picture of higher-order identity and distinctness that Structure seems to promise. By contrast, while less general principles like

[35] Goodman (unpublished *d*) develops a theory that entails Generalized Qualitative Atomic Structure, and shows how it can be derived from certain simpler and initially weaker-seeming axioms. Bacon (2020: n. 48) states principles structurally like Generalized Qualitative Atomic Structure, modulo replacing 'Qualitative' with 'Pure' and replacing type-*e* quantifiers with quantifiers of arbitrary type restricted by a 'Fundamental' predicate; he also gives simpler axioms from which this can be derived.

Generalized Qualitative Atomic Structure could also be used to justify rejecting these identities, and have the advantage of consistency, their theoretical virtues do not seem nearly so compelling. These weakenings involve singling out type *e*—the type of objects—for some kind of special treatment: they have in common the idea that objects are special, in being less apt than entities of other types to enter into interesting identity-theoretic relationships with one another.[36] However, it is far from clear what basis there is supposed to be for treating type *e* as special in this way. Considerations of theoretical virtue would need to be quite dramatic in order to prop up the view against apparent counterexamples like (2–5). But we see nothing like this level of theoretical virtue on offer.

Those who do not agree with us about the plausibility of (2–5) may be more convinced by apparent counterexamples to (Generalized) Qualitative Atomic Structure involving *abstract* objects. Consider ordered pairs, such as the ordered pair of Paris, France and Paris, Texas—call it Parispair:

(10) For Parispair to be such that its first element is bigger than its second element is for Paris, France to be bigger than Paris, Texas.

In the case of ordered pairs, the "nothing over above" thought is especially gripping: the practice of talking about ordered pairs involves a kind of freedom in going back and forth between attributions of binary relations and attributions of properties to ordered pairs that is strongly suggestive of identities like (10). But given that *being bigger than* is a qualitative binary relation and *being an ordered pair whose first element is bigger than its second element* is a qualitative property, (10) is ruled out by Generalized Qualitative Atomic Structure, which entails that it never happens that for two distinct objects to stand in some qualitative binary relation is for a single object to instantiate some qualitative property.[37]

Another widely discussed case where higher-order identities involving abstract objects are especially plausible is that of numbers. For example:

(11) For zero to number the children of Wittgenstein is for Wittgenstein not to have any children.

[36] Perhaps certain other types are special in the same way: Goodman (unpublished *d*) shows that Atomic Structure can consistently be extended to certain larger sets of types, though generalizations of the Russell-Myhill theorem show that none of these sets can include $\langle\rangle$ or any type built up from $\langle\rangle$.

[37] Some might reject (10) on the grounds that the proposition that Parispair is a pair whose first element is bigger than its second element "presupposes the existence of ordered pairs", whereas the proposition that Paris, France is bigger than Paris, Texas does not. If "presuppose" is understood in an epistemic way, this thought looks bad in the same way as 'It can't be that for there to be water in this glass is for there to be H_2O molecules in this glass, since the proposition that there are H_2O molecules in this glass presupposes the existence of oxygen.' On a more metaphysical understanding of "presuppose", the thought is more defensible. Notice, however, that Generalized Qualitative Atomic Structure also rules out the claim that for Parispair to be such that its first element is bigger than its second element is for Paris, France and Paris, Texas to form an ordered pair whose first element is bigger than its second, in which both sides might be thought to "presuppose the existence of ordered pairs".

On the plausible assumption that the binary relation *numbering the children of* and the property *not having any children* are both qualitative, Generalized Qualitative Atomic Structure rules out the truth of (11), along with its more complicated analogues for other natural numbers. But such identities seem to capture something quite important about the connection between numbers and counting facts.[38]

These examples involving abstract objects would of course not rule out a yet weaker version of Generalized Qualitative Atomic Structure in which all the relevant objects ($y_1, \ldots, y_n$ and $z_1, \ldots, z_m$) are required to be non-abstract. Since regions, molecules, fists, and beards presumably count as non-abstract, this weakening would still rule out the likes of (2)–(5). But whatever one might think about Generalized Qualitative Atomic Structure itself, *this* restriction feels badly ad hoc. The initial plausibility of identities like (2)–(5) suggests that we are dealing with a phenomenon with nothing special to do with the abstract-concrete distinction (whatever *that* is). If one is going to be restricting Generalized Qualitative Atomic Structure in any case, it seems more plausible and principled to think that the relevant restriction involves the notion of *fundamentality* rather than that of non-abstractness. But this restriction would be perfectly compatible with our examples, and with Ancestral Microphysical Supervenience.[39]

Some defenders of Generalized Qualitative Atomic Structure might concede that examples like (10) and (11) do capture true higher-order identities, but deny that they are counterexamples, by maintaining that despite superficial syntactic appearances, the relevant expressions—-e.g. 'Parispair' and 'Zero'—do not really denote *objects*, but should instead be understood as standing for something higher up in the type hierarchy. We already have a precedent for this sort of attitude: in Chapter 1, we stipulated that for the purposes of this book, sentences that seem on the surface to be about special objects called "properties", "relations", and "propositions" should be read as shorthand for claims that could be more perspicuously regimented using higher-order terms and quantifiers of various types. We did not take a stand on whether the relevant sentences in ordinary English (or less technical-sounding variants involving expressions like 'redness', 'some characteristic', 'having something in common', etc.) in fact work that way.

[38] Higher-order identities like (11) are central to the view Rayo (2013) defends under the name 'Trivialist Platonism'. But since Rayo's overall package also includes (a contingentism-friendly variant of) Intensionalism, one might get the impression from him that acceptance of identities like (11) goes hand in hand with Intensionalism; whereas it is crucial to the point we are making that (11) is plausible even assuming the falsity of Intensionalism.

[39] Moreover, whereas the abstract/concrete contrast only has nontrivial application within the domain of objects, the fundamental/nonfundamental distinction seems to make sense in all types. Bacon (2020) discusses principles analogous to Generalized Qualitative Atomic Structure which apply in arbitrary types, with a restriction to fundamental entities in the relevant type; his principles involve a demanding notion of *purity* in place of qualitativeness, which is appropriate since qualitativeness also seems to be a notion with a special relation to type *e*.

But there is much to be said for the claim that they do (see Prior 1971: chs. 1 and 3). And insofar as one takes this attitude to ordinary discourse about properties, relations, and propositions, it is extremely natural to extend it to certain other superficially first-order domains of discourse, including discourse about numbers and ordered pairs.

A range of workable higher-order reconstructions are available for these cases. For example (following a suggestion considered by Frege (1884: §55) and developed by Whitehead and Russell (1910)), quantification over natural numbers could be reconstructed as (restricted) quantification into type $\langle\langle e\rangle\rangle$, the type of properties of properties of objects. 'Zero' is thus treated in the same way that we would treat the expression 'the property of being an uninstantiated property', both being assimilated semantically to the quantifier 'nothing'. Likewise 'One' is treated like our 'the property of being a property that has an instance that is identical to all its instances'; the *numbering* relation is simply instantiation. Quantification over ordered pairs of objects could be understood as restricted quantification into type $\langle\langle e, e\rangle\rangle$, the type of properties of binary relations among objects: 'the ordered pair of Paris, France and Paris, Texas' is then treated like our 'the property of being a binary relation that Paris, France bears to Paris, Texas'. But there are several other natural regimentations, using different types. And there is plausibly a fair amount of vagueness as regards which regimentation is correct, since many structures can satisfy the job descriptions set out by ordinary discourse.[40]

Such "concealed higher-order quantification" views strike us as attractive, although they face some serious challenges which we will not attempt to answer here. Insofar as they are tenable, generalizations like Generalized Qualitative Atomic Structure start looking more defensible, since some of the putative counterexamples can be set aside as not really involving *objects*. If one could provide principled grounds for thinking that higher-order "reconstruction" is only appropriate for discourse (apparently) about *abstract* objects, one could set aside putative counterexamples like (10) and (11), but one would still have to bite the bullet of rejecting identities like (2)–(5). Alternatively, one might admit higher-order reconstruction when it comes to discourse about regions, beards, fists, and the like, but hope to provide principled grounds for resisting such reconstruction when it comes to certain other macroscopic objects, such as tables, for which Tolerance Puzzles can be raised. However, we see no such principled grounds in either case. Insofar as the project of explaining higher-order identities apparently in conflict with Generalized Qualitative Atomic Structure by appeal to "concealed higher quantification" is plausible for things like numbers and ordered pairs, it

[40] Benacerraf (1965) argues that the identification of numbers with sets should be rejected on the grounds that any particular identification seems arbitrary; a similar argument could be made against the the higher-order regimentations currently under discussion. But Benacerraf provides no reason to think that the kind of arbitrariness he is worried about is any different from the kind of arbitrariness that is associated with vagueness.

also seems plausible for things like fists and beards. And the project of drawing a principled line between, e.g., fists and tables seems even more unpromising. As we discussed earlier, there are differences here having to do with how easy it is to write down simple proposals that are proof against obvious counterexamples; but as before, such contrasts should not be taken as indicative of a deep metaphysical contrast. Insofar as we go in for higher-order reconstruction and don't try to draw lines in these unpromising places, we will be led to a picture on which the real type *e*—the domain of objects—is quite sparse, corresponding roughly to the old notion of a *fundamental* object. Those like us who were previously inclined to look to physics as a guide to fundamentality will now use physics as a guide to "ontology" (understood as the investigation of what there is, on a type-*e* reading of the quantifier). But the programme does not require this physicalist outlook. Russell (1918–19), for example, was happy to entertain higher-order reconstruction for all kinds of apparent objects, such as Socrates and Piccadilly (that's essentially what he means when he says that these objects are 'series of classes'), while taking sense data to be among the genuine objects.[41]

The question *how* discourse about ordinary objects like tables should be reconstructed in higher-order terms is partly bound up with the questions about what the genuine type-*e* objects are like; insofar as our ignorance about ultimate physics makes us cautious about answering the latter questions, we should also be cautious about the former. But to fix ideas we will sketch a couple of simple strategies, both of which go along with the slogan 'ordinary material objects are really properties'—in other words, quantification over tables and the like is reconstructed as quantification into type $\langle e \rangle$. On one possible view (inspired by a natural interpretation of field theories in physics), the only type-*e* objects are *spacetime points*. Given this, it would be natural to say that the points that *instantiate* any given ordinary object (understood as a property) are those at which that object is *located*. Physical properties of and relations among ordinary objects would then be explained in terms of properties of and relations among the spacetime points that instantiate them: to take a simple example, *being spherical* might be *being such that for some distinct points x and y, one is instantiated by all and only the points that are at least as close to x as y is*.[42] On another possible

[41] Just as the "higher-orderist" take on ordinary 'property'-talk might be categorized as a form of nominalism, if 'nominalism' is interpreted as a thesis expressed using a genuinely first-order quantifier, the "higher-orderist" take on 'material object'-talk might be categorized as a form of mereological nihilism, if 'mereological nihilism' is regarded as a thesis perspicuously expressed using a genuinely first-order quantifier. (Russell (1918–19: 18): '[T]hat thesis, if it can be maintained, will dissolve Piccadilly into a fiction'.) On the other hand, if nominalism and nihilism are conceptualized as *error* theories according to which ordinary people express or believe falsehoods when they say such things as 'there are prime numbers' or 'there are expensive tables', the kinds of views we are describing will not count as nominalist or nihilist. For a non-error-theoretic conception of the nihilist programme see Dorr 2005*b*; a version of nihilism with a more error-theoretic flavour is discussed in Dorr and Rosen 2002.

[42] The same treatment can be extended to various not-so-ordinary objects, e.g. spatial and spatiotemporal regions, subatomic particles, etc.

view, the type-*e* objects include certain *particles*: in this setting, it might be more natural to think that the properties that are ordinary objects are instantiated only by these particles, and (for example) to say that to be instantiated by Woody is to be a particle that is part of Woody.

We want to remain noncommittal about this general project of higher-order reconstrual of apparently first-order quantification. In the case of numbers, for example, there is a long tradition of resistance to higher-order reconstruction, going back to Frege's (1884: §57) insistence that every number is a "self-subsistent object". Those who are moved by the kinds of considerations that moved Frege in the case of numbers will certainly be all the more resistant to higher-order reconstruction when it comes to tables and so on. What we mainly want to insist on is that if one takes higher-order reconstruction seriously as an option in some of our warmup cases, one should also take it seriously as an option for huge swathes of ordinary discourse. If it is to be rejected across the board, Generalized Qualitative Atomic Structure looks badly counterexample-prone; if it is a serious option across the board, Generalized Qualitative Atomic Structure looks defensible, but becomes useless for the purposes of ruling out higher-order identities involving ordinary objects. Either way, identities like (2)–(5) remain plausible, and continue to encourage a picture on which Ancestral Microphysical Supervenience is true and underwritten by higher-order identities. And as we have remarked, the truth of Ancestral Microphysical Supervenience will be bad news for any view that goes in for Ancestral Hypertolerance as part of a general strategy for responding to Tolerance Puzzles.

8

Iteration for Metaphysical Necessity

This chapter will take up the question the last chapter bracketed: whether Iteration holds for metaphysical necessity. Many have had the sense that an argument for Iteration will somehow fall out from a proper understanding of what metaphysical necessity is. One way to try to make this happen is to set down some principles connecting metaphysical necessity with some other metaphysical categories. For example, Fine (1994: 9) suggests that metaphysical necessity can be "reduced" to a certain operator expressing a fine-grained conception of essence. And many philosophers, including Williamson (2007), have tried to shed light on metaphysical necessity via principles connecting metaphysical modality to counterfactuals. But as far as we can see, neither of these programmes has much effect on the dialectic about Iteration. One can write down principles about the Finean essentialist operator or about counterfactual conditionals which putatively correspond to Iteration, but the principles in question do not seem to be any more immediately compelling or theoretically virtuous than Iteration itself.[1]

A more dialectically helpful strategy is to appeal to the thought that metaphysical necessity is in some important sense *absolute*, or maximally strong. In this chapter, we will defend this thought and show how it can be used to develop a case for Iteration.

[1] In some cases they seem *less* immediately compelling. For example, given Williamson's claim that 'it is metaphysically necessary that *P*' is tantamount to 'If not-*P* it would be that ⊥', Iteration is equivalent to the following schema: if (if *P* it would be that ⊥), then (if *Q*, it would be that if *P*, it would be that ⊥). While this might be correct, compared to other principles of counterfactual logic it is far from being well entrenched in ordinary reasoning. In Fine's case, Iteration comes down to a complex Iteration-like principle for his essentialist operator (Fine 1995: 247, II.iii) which certainly does not wear its truth on its sleeve.

You might think that some other kind of connection between metaphysical necessity and essence would be more promising. Leslie (2011) maintains that Iteration-denial is 'not consistent with the notion of essence' since 'a thing's essence could not have been different than it is'. Iteration-deniers would have to disagree with this claim if they identified an object's essence with the collection of all the properties it has necessarily. But Fine (1994) influentially rejects this kind of identification, and the motivations for Iteration-denial may provide further reasons to doubt it. In general, we worry that the ideology of essence—and particularly the noun 'essence' as opposed to the adverb 'essentially'—is too laden with disparate associations picked up over its long and tortuous history to be dialectically useful in the debate about Iteration. However, some of the ideas about the connection between metaphysical necessity and *identity* which we will be exploring below could reasonably be billed as having to do with essence, given the historical importance of identity-theoretic glosses on essence (such as 'what it is to be' a given thing).

The Bounds of Possibility: Puzzles of Modal Variation. Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri, Oxford University Press. © Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri 2021.
DOI: 10.1093/oso/9780192846655.003.0009

8.1 The Broadness of Metaphysical Necessity

Metaphysical necessity is widely characterized in ways that emphasize its strength. Being metaphysically necessary is equated with being "unrestrictedly" or "absolutely" necessary; it is 'necessity in the highest degree—whatever that means' (Kripke 1972: 99); 'necessity in the widest sense' (Stalnaker 2003*a*: 203). Metaphysical possibility, correspondingly, is 'possibility without qualification' (van Inwagen 1998: 72). Formulae like these played a central role in putting metaphysical necessity on the map as a topic for philosophical debate in the twentieth century, and continue to be central for introducing the topic to new generations of students. And they lead to a prima facie worry for the view that Iteration is false for metaphysical necessity. For whatever it means for one operation $\Box_1$ to be at least as strong as (less restricted than, at least as high in degree as, at least as broad as) another operation $\Box_2$, it certainly requires that the extension of $\Box_1$ is contained in that of $\Box_2$, i.e. that $\forall p(\Box_1 p \rightarrow \Box_2 p)$. But if Iteration fails for metaphysical necessity, then there is another status—metaphysically necessary metaphysical necessity—which is less inclusive than metaphysical necessity. As we will discuss, this status also seems to have a good claim to count as a kind (or "degree" or "sense") of necessity, since it is guaranteed to play the characteristic logical roles of a necessity operation by the fact that metaphysical necessity does. But if it does count, and some metaphysical necessities fail to enjoy it, that will contradict the claim that metaphysical necessity is at least as broad as every other necessity operation.[2]

Introductions to the expression 'metaphysically necessary' (or to the interpretation of 'necessary' that came to be labelled 'metaphysical') often also involve philosophers giving examples of truths they take to be metaphysically necessary. Many of the standard stock of examples come from Kripke (1972):

(1) Hesperus is Phosphorus.

Every golden thing is made of atoms containing seventy-nine protons.[3]

Nixon is not an inanimate object.

Lectern *L* is not made of ice.

[2] Following Bacon (2018*a*), we will use 'broader than' as our label both for the relation among kinds of necessity and for the corresponding relation among kinds of possibility. This use of 'broad' is inspired by possible-worlds semantics: roughly, the broader the operation, the more worlds are relevant to it. Of course, if you were thinking 'the broader the operation, the more propositions it applies to', you would need to systematically replace all talk of 'broader than' with 'narrower than' in the case of necessity operations.

[3] We use the adjective 'golden' rather than the mass noun 'gold' in order to sidestep questions about the semantics of mass nouns, which in our view have led to unhelpful distractions from Kripke's central points. We of course mean 'golden' in the sense of interest to metallurgists and alchemists, not the sense having to do with the colour of surfaces.

If these truths were *obviously* not necessary in the broadest sense, then we couldn't take Kripke's remarks suggesting that the status he is claiming for them is the broadest necessity operation at face value, on pain of convicting him of denying the obvious. Many philosophers have regarded the idea that these examples are necessary in the broadest sense as a non-starter, while remaining open to the idea that metaphysical necessity is some distinct interesting status that they do have. But as we will go to some lengths to explain in this section, it is just *not at all* obvious that Kripke's examples fail to be necessary in the broadest sense. Indeed it strikes us as pretty plausible that all of them are. The primary arguments that seem to have convinced people that they are not necessary in the broadest sense are simply bad arguments.

Of course, there may ultimately be some interesting, non-obvious theoretical reasons to think that some or all of Kripke's candidates are not necessary in the broadest sense. But even if we discovered such reasons, that would not be a good basis for rejecting the thesis that metaphysical necessity is the broadest form of necessity, since the examples were put forward in the spirit of discovery rather than stipulation. Indeed, in many cases, Kripke's central concern is not so much to convince his readers of the specific necessity claims that he gives as examples (although he does think they are true), but rather to correct what he rightly saw as a pervasive tendency to assume that the relevant propositions *can't* be necessary in the sense he was concerned with, based on mistaken ideas linking necessity to topics like a priori knowledge and linguistic convention. As we will make clear, those who think that the examples are obviously not necessary in the broadest sense are making precisely the sort of mistake that Kripke was trying to purge.

Why might one think it obviously false that the Kripkean examples are necessary in the broadest sense? One thought we can quickly set aside is that a broadest necessity operation would have to be one that applied to nothing whatsoever. Someone could be led to this by the thought that for any property F of worlds, *being true at all F worlds* counts as a necessity operation, combined with a liberal account of worldhood like the one we presented in §1.6, on which for every collection of propositions that includes every proposition or its negation, there is a world at which exactly those propositions are true.[4] Then in particular *being true at every world* will be a necessity operation, even though nothing is true at every world. But it seems completely in the spirit of the characterization of metaphysical necessity as the broadest necessity operation to treat certain kinds of logical well-behavedness as criterial for anything to count as a necessity operation. For example, one might plausibly require that if X is a necessity operation, then $\forall pX(p \rightarrow p)$ should be true. More generally, one might require that every closed theorem of some basic modal logic such as our H_{KR} should come out true when $\Box$ is interpreted as expressing X.

[4] Salmón (Salmon 1989) favours a view along these lines, and points out that it makes *truth at all worlds* an uninteresting status; see also Nolan 2011 (317).

Such criteria will rule out uninstantiated properties like *being true at every world* on the liberal understanding of worlds.[5]

But even if such logical constraints on the category of necessity operations are in place, many philosophers still regard the thesis that the Kripkean examples are broadly necessary as a non-starter. One common thought is that *logical* necessity is a form of necessity that the examples obviously lack.[6] Here is a representative statement from Clarke-Doane (2021):

> But this thesis [that metaphysical possibility is 'the most inclusive non-epistemic, non-deontic notion of possibility'] is false. Consider logical possibility. In particular, consider the notion of logical possibility corresponding to some fixed S5 system of quantified modal logic with contingent identity . . . . Logical possibility in the present sense is dramatically more inclusive than metaphysical possibility. It is logically possible in the present sense that you could have had different parents, that Hesperus ≠ Phosphorus, and that there are not infinitely-many prime numbers—even assuming the actual facts about you, Hesperus, and the numbers.

Given any system of logic, there is indeed a notion of *consistency* in that logic which can apply to sentences in the relevant language. But Clarke-Doane uses 'logically possible' as a *sentential operator* ('it is logically possible that…'), and thus seems to elide the important distinction between operators and predicates of sentences.[7] Perhaps he is assuming that for every predicate F of sentences, there is a corresponding operation—in his terminology, a "notion"—such that when a sentential operator O is interpreted as expressing it, $\ulcorner O\varphi \leftrightarrow F\langle\varphi\rangle\urcorner$ will come out true for every sentence φ. But this assumption is false. For example, there is no operation that corresponds in this way to the predicate 'has exactly four words', since the application of that predicate turns on facts about sentences that don't have anything to do with what they mean. And prima facie, 'is consistent in such-and-such logical system' is like 'has exactly four words' in this respect. After all,

[5] Deontic logicians discuss a status sometimes called *deontic necessity*, supposed to be closely related to more familiar topics like obligation and what ought to be the case. Some influential deontic logicians have argued against the view that it is deontically necessary that $p \vee \neg p$, for every p: e.g. von Wright (1951) accepts a "principle of deontic contingency" on which it is never *necessarily* deontically necessary that $p \vee \neg p$, and McNamara (2019) suggests that nothing was deontically necessary when there were no agents. If so, so-called deontic necessity will also fail to count as a necessity operation by our logical criterion.

[6] See, e.g., Fine 2002. Salmón (Salmon 1989*b*: 13) remarks that logical necessity shows that 'Metaphysical necessity is definitely *not* an unrestricted limiting case', and offers mathematical necessity as another example of a modality less restrictive than metaphysical necessity.

[7] Similarly, Priest (2021) thinks that 'necessity' has multiple senses (denying that metaphysical necessity is one of them) but treats them all as predicates of sentences, e.g. 'The analytic necessities are the things true simply in virtue of the meanings of words.' Even Plantinga (1974: 1), though he explicitly takes necessity to be a property of propositions, says that "the sense of necessity in question is wider than that captured in first-order logic"; he does not explain in what sense a property of propositions could be captured in first-order logic.

the relevant definitions in logic are sensitive to various syntactic distinctions, such as the difference between sentences with two occurrences of the same term and sentences with two different terms, that do not seem to require any distinctions at the level of meaning.

Our point is not that anyone who uses 'logically possible' in such a way that ⌜It is logically possible that φ⌝ is treated as synonymous with ⌜'φ' is logically consistent⌝ is thereby automatically making a mistake.[8] We are free, if we wish, to institute special semantic conventions concerning a certain expression, whereby certain sentences involving it (and no quotation marks) are to be treated as semantically equivalent to certain other sentences that do involve quotation marks. If our new conventions are sufficiently systematic, they will allow us to treat the new expression as belonging to some familiar *syntactic* category (e.g. that of sentential operators), by specifying a meaning for every sentence that would be well formed when it is placed in that category. But if we do this, we have to be cautious about extending logical schemas that were valid in the old language to to the new language. For instance, suppose we introduce the expression 'Efree' into the language by a stipulation that guarantees that for any type-*e* expression α, ⌜Efree(α)⌝ will be a closed sentence, true if α does not contain the letter 'e'.[9] Then since '$\forall x$(Efree(x))' is vacuously true (since the letter 'x' does not contain the letter 'e'), while 'Efree(Hesperus)' is false, the universal instantiation schema

$\mathbf{UI_{sub}}$ $\qquad \forall xP \rightarrow P[a/x]$

will fail in the expanded language.[10] Similar caution is required as regards higher-order quantification, since although 'Efree(Phosphorus) $\wedge$ ¬Efree (Hesperus)' is true, '$\exists X(X$(Phosphorus) $\wedge$ ¬X(Hesperus))' is still false (it doesn't contain the expression that triggers the new convention). In the same way, if we institute a convention whereby ⌜It is logically possible that φ⌝ is synonymous with ⌜'φ' is logically consistent⌝, there is no reason to expect higher-order

[8] This way of using 'logically possible' must be sharply distinguished from the convention that 'logically possible' means *not expressed by any contradictory sentence* (cf. Salmon 1986*b*, app. B, on 'logically necessary'). Insofar as the relevant *expressing* relation is in good standing, this is a perfectly fine operation. But on this interpretation, it is not automatic that ⌜It is logically possible that φ⌝ is true ⌜'φ' is logically consistent⌝ is, since a proposition expressed by a contradictory sentence may also be expressed by one that isn't contradictory.

[9] Note that even if successful this stipulation leaves various questions about the semantics of the expanded language unsettled, e.g. what ⌜Efree(α)⌝ *means*.

[10] It is not immediately obvious that we really have to say that $\mathrm{UI_{sub}}$ is *invalid* in the new language. For languages that admit ambiguity and context-sensitivity, a sensible notion of validity needs to invoke some notion of uniform resolution of ambiguity and context sensitivity (see Dorr 2014*b*). But the relevant notion of uniformity is naturally understood in terms of facts about the semantics of subsentential constituents, about which the relevant stipulations are silent since they only concern the meanings of sentences. Thus it isn't obvious where there is and isn't uniformity failure in the new language.

existential generalization into operator position to remain valid in the newly extended language.[11]

Armed with an appropriate caution about the way in which quotation-like behaviour—including apparent failures of familiar logical laws—can arise even when actual quotation marks are nowhere to be seen, we can see that arguing that there are necessity operations that don't apply to the Kripkean examples is harder than one might have thought. Consider the following schematic argument against the thesis that it is broadly necessary that Hesperus is Phosphorus (using Nec for 'is a necessity operation', a predicate of type $\langle\langle\langle\rangle\rangle\rangle$):

P1 $\mathrm{Nec}(O)$

P2 $\neg O(\text{Hesperus} = \text{Phosphorus})$

C $\exists X(\mathrm{Nec}(X) \wedge \neg X(\text{Hesperus} = \text{Phosphorus}))$

There is a wide range of expressions one might consider substituting for '*O*' in this argument, including expressions in the vicinity of 'it is logically necessary that', 'it is knowable a priori that', 'it was epistemically necessary during the time of the Babylonians that', 'no perfectly rational person could dispute that', and 'if Hesperus were distinct from Phosphorus it would be the case that'.[12] But none of these arguments are any good. The problem is that if we were to grant that both premises are true (on a uniform interpretation that is also available for the conclusion), we would then have good reason to think that the argument was invalid. For all these values of *O*, P2* is also very plausible:

P2* $O(\text{Hesperus} = \text{Hesperus})$

But if P2 and P2* are both true, then the existential generalization schema

$\mathrm{EG}_{\mathrm{sub}}$ $P[a/x] \rightarrow \exists x P$

[11] Note that if we want to fix a meaning for every sentence that would be well formed when 'it is logically possible that' is treated as operator, we will also need to make some stipulations about sentences where it occurs as an argument of some higher-order predicate, as well as those where it occurs with an argument of its own. We could just stipulate that every such sentence is to be synonymous with 'Snow is white.' Or, less arbitrarily, we could stipulate that when it occurs as an argument, 'it is logically possible that' should be interpreted as we would have interpreted 'λp. it is logically possible that p', i.e. as the vacuous lambda term 'λp.'p' is logically consistent'. Since the sentential variable *p is* in fact a logically consistent formula, this will make for surprising failures of Eβ: $\ulcorner(\lambda p.\text{it is logically possible that } p)(\varphi)\urcorner$ will be true for every φ although $\ulcorner$it is logically possible that $\varphi\urcorner$ is false for some φ. But as we have just pointed out, some surprising failures of the usual logical rules are to be expected in any case.

[12] The argument is given in a semi-formal language which imports the relevant bits of English into higher-orderese. This requires some finesse, since the relevant strings of words aren't generally regarded as syntactic units in English. But if one were concerned about that fact, one could replace P1 with something like $\mathrm{Nec}(\lambda p.$it is knowable a priori that $p)$.

must fail in some cases where O plays the role of a.[13] For by Leibniz's Law, since Hesperus = Phosphorus, we have

C* $\neg\exists X(X(\text{Hesperus} = \text{Hesperus}) \wedge \neg X(\text{Hesperus} = \text{Phosphorus}))$

Importantly, the argument for C* only requires applying Leibniz's Law within the pure language of logical constants and the names 'Hesperus' and 'Phosphorus', so its (apparent) failures for sentences containing O are not relevant here.[14]

Note that blocking the argument in this way does not require denying P1. Our point does not turn on coming up with some refined understanding of the relevant sense of 'necessity operation' that precludes *being knowable a priori* and the like from counting as necessity operations. We are trying to avoid relying on the kind of language—hard words like 'objectivity' and 'worldliness' and 'metaphysics'—that might be invoked in such a refinement.[15] For reliance on such expressions will leave the door open to Iteration-deniers to complain that the only nontrivial sense they can make of them renders them tantamount to 'applies to every metaphysical necessity', so that the claim that double metaphysical necessity is a necessity operation will just be a restatement of Iteration, rather than part of an interesting argument for it. Our point is that quite irrespective of the truth of P1, the application of existential generalization required for the argument to work is invalid if P2 and P2* are both true. Meanwhile, if P2* is *false* for our chosen O, we will then be on firm ground denying P1. As we explained above, it is dialectically legitimate to take it for granted that necessity operations are *logically* well behaved in certain ways—e.g. that $\forall X(\text{Nec}(X) \rightarrow \forall y X(y = y))$. *That* kind of clarification does not trivialize the characterization of metaphysical necessity as the broadest necessity, or undermine the prospects of deriving Iteration from this characterization.[16]

[13] Recall from §1.3 that in endorsing the base logic H_0 we are not committed to the claim that analogous generalizations hold when constants are added to the language by "naïvely formalizing" English expressions such as 'it is knowable a priori that…'. In H_0, EG_{sub} can be derived from the combination of EG proper ($Fa \rightarrow \exists F$) with Eβ. Those who reject EG_{sub} for some expanded language thus have a choice whether to blame the failure on EG or Eβ. For the expansions at hand, the latter option seems more natural to us, but we will not pursue the matter here.

[14] The Leibniz's Law argument from Hesperus = Phosphorus to C* will be rejected by those like Caie, Goodman, and Lederman (2020) who advocate giving up our background logic (even within the purely logical signature) for reasons relating to the behaviour of attitude verbs. On their picture, identity is not the strongest reflexive relation: there *is* a strongest such relation, namely Leibniz equivalence ($\lambda xy.\forall Z(Zx \leftrightarrow Zy)$), but it does not hold between Hesperus and Phosphorus, for example. See §1.3 for our reasons for rejecting this view.

[15] Williamson (2016) characterizes metaphysical modality as the "maximal objective modality". Rosen (2006) characterizes it as the strictest "real" modality.

[16] Given that we are understanding 'necessity operator' in this way, we are perhaps not really in disagreement with some philosophers who speak of varieties of necessity that do not apply to all metaphysical necessities. For example, Fine (2002) suggests that being a mathematical necessity is just being a metaphysical necessity that is also a mathematical truth. Given the fine-grained setting he is assuming, he would probably not count the proposition that Hesperus is Hesperus (or that all dogs

The idea that the usual quantifier laws fail when applied to certain sentences involving propositional attitude words should not be surprising. Anyone who holds the popular view that the substitution of 'Hesperus' for 'Phosphorus' inside attitude verbs fails to preserve truth (within a single context) is under pressure to posit failures of quantifier rules, in order to block arguments like the following four-step argument:

(i) Hesperus = Hesperus ∧ the Babylonians believed that Hesperus was visible in the evening.

(ii) $\exists x$(x = Hesperus ∧ the Babylonians believed that x was visible in the evening).

(iii) $\forall x$(x = Hesperus → the Babylonians believed that x was visible in the evening).

(iv) Phosphorus = Hesperus → the Babylonians believed that Phosphorus was visible in the evening.

If one thought that $\mathrm{EG}_{\mathrm{sub}}$, $\mathrm{UI}_{\mathrm{sub}}$, and the universally quantified version of $\mathrm{LL}_{\mathrm{sub}}$ were valid for the language expanded to include 'believe', one could use the first to derive (ii) from (i), the second to derive (iv) from (iii), and the combination of the three to derive (iii) from (ii). Proponents of the Classical Opacity view discussed in §1.3 will resist the step from (ii) to (iii). For them, the claim that something identical to an object has a property is weaker than the claim that the object itself has the property, which is in turn weaker than the claim that everything identical to the object has the property. But setting this radical option aside, those who hold that (i) is true and (iv) is false (in the same context) must think that one or both of the inferences from (i) to (ii) and from (iii) to (iv) fails. Those who embrace substitutivity-failures and do not go the Classical Opacity route are thus already in the business of giving up the standard quantifier rules; it is thus unsurprising that, as we have seen, this view also requires failures of those rules when the attitude

are dogs, etc.) as a mathematical truth; if so, being mathematically necessary will not be a necessity operation in our sense. Notice also that given the kinds of logical well-behavedness required, it is not obvious that the epistemic operations we considered do in fact qualify as necessity operations. For example, it is not so obvious that each and every particle and spacetime point is such that it is knowable a priori that it is self-identical: one might deny this, on the grounds that some of them are such that it would be impossible for anyone to think about them at all, and a fortiori to know them to be self-identical. It is also debatable whether 'it is knowable a priori that' satisfies consequences of the K schema like $\forall F \forall y \forall p(O(Fy \rightarrow p) \rightarrow O(Fy) \rightarrow Op)$; on account of the fact that (so to speak) one might might know a priori Fy under one guise and $Fy \rightarrow p$ under another, but not know p a priori under any guise. (For example, perhaps Quine's Ralph (Quine 1956) knows a priori that Ortcutt is a spy iff the person visible on the beach is a spy, and knows a priori that if Ortcutt is a spy then a leading citizen is a spy, but not that if the person visible on the beach is a spy then a leading citizen is a spy.) And depending on how tightly one takes the epistemic interpretation of 'must' to be tied to 'know', it too may fail the logical test for being a necessity operator.

operator itself (e.g. 'it is knowable a priori that'), rather than something in its scope, is the expression being substituted for a quantified variable.[17]

For many of the candidate operators *O*, premise P2 of the above argument is in fact very controversial. Many authors have argued that Leibniz's Law inferences involving the substitution of proper names are valid even when the names occur in attitude reports. Most influentially, Salmón (Salmon 1986*b*), Soames (1987), and Braun (1988) have argued that sentences like 'The Babylonians knew that Hesperus = Phosphorus' are simply true, though they may be bad to assert because of their tendency to communicate falsehoods. Similar considerations extend to 'It is possible to know a priori that Hesperus = Phosphorus.'[18] Others (Dorr 2014; Goodman and Lederman 2021) have defended a contextualist picture where there is no single context where the substitution of 'Hesperus' for 'Phosphorus' in an attitude report would affect its truth value, though such substitutions often do affect which resolutions of context-sensitivity are most natural for a given sentence. These views block the above style of argument that existential generalization fails for 'it is knowable a priori that'. Of course, they also block the argument that it is not broadly necessary that Hesperus is Phosphorus. However, they leave it open that some of the other Kripkean examples could be used in place of that identity. For example, one might argue as follows:

P1 Nec *O*.

P2′ $\neg O$(every golden thing is made of atoms containing seventy-nine protons).

C′ $\exists X$(Nec $X \wedge \neg X$(every golden thing is made of atoms containing seventy-nine protons)).

Even if you were convinced by the Salmón-Soames argument that it is knowable a priori that Hesperus is Phosphorus, you might still think that it is obviously not knowable a priori that every golden thing is made of atoms containing seventy-nine protons: Salmón and Soames certainly don't intend their arguments to

[17] The view that the standard quantifier laws fail for the language supplemented with ("naïvely formalized") attitude verbs naturally suggests the project of providing a systematic translation-scheme from this logically unruly language into a well-behaved one. One possible model for such a translation was provided earlier, when we discussed a translation from a language where 'it is logically possible that' appears in the syntactic role of an operator into a language without this expression, that works by replacing formulae involving this operator with predications of quoted sentences. (Quotation could in turn be eliminated in favour of descriptive structural characterizations of linguistic items.) However, the idea that attitude ascriptions should ultimately be understood in terms of some relations between people and natural-language sentences is not very promising, so we shouldn't expect that model to be particularly useful for the attitudinal case, though it might serve as a starting point for something more defensible. A further and more ambitious project is that of providing a systematic *semantic theory* for the logically unruly language (in a well-behaved language): such a theory will also need to address questions about the semantic properties of subsentential expressions.

[18] See Salmon 1986*b* (app. B.1) and Soames 2002 (ch. 1). Of course, you might not even think that it is possible to know a priori that Hesperus is Hesperus, on the grounds that you can't know a priori that Hesperus is identical to something. This motivates Salmón (Salmon 1986*b*) to opt for the more cautious claim that it is knowable a priori that *if Hesperus exists* Hesperus is Phosphorus.

generalize in that way.[19] But when we bear in mind that identities can be higher order as well as first order, we can see that switching from 'Hesperus = Phosphorus' to examples like those above does not make a difference to the dialectic. For example, the following higher-order identity strikes us as very plausible:

(2) To be golden is to be made of atoms containing seventy-nine protons.

Given (2), we can argue as before that there is no substitution of *O* that makes both premises true and the argument valid. Just as our basic logic licenses the inference from 'Hesperus = Phosphorus' and the unproblematic

(3) $\forall X(\mathrm{Nec}(X) \rightarrow X(\text{Hesperus} = \text{Hesperus}))$

to

(4) $\forall X(\mathrm{Nec}(X) \rightarrow X(\text{Hesperus} = \text{Phosphorus}))$

it likewise licenses the inference from (2) and the unproblematic

(5) $\forall X(\mathrm{Nec}(X) \rightarrow X(\text{Every golden thing is golden}))$

to

(6) $\forall X(\mathrm{Nec}(X) \rightarrow X(\text{Every golden thing is made of atoms containing seventy-nine protons}))$

[19] Soames (2002: 272, 276) "presumes" that it is not knowable a priori that "for all *x*, *x* is a drop of water iff *x* is a drop of a substance molecules of which contain two hydrogen atoms and one oxygen atom". His strategy for distinguishing this case from the case of "Hesperus is Phosphorus" is not based on any special contrast between singular terms and expressions in other grammatical categories, but on the contrast between simple expressions and compound expressions (cf. Salmon 1986*b* (72) on "single-word terms"). He thus accepts that it is knowable a priori that all and only woodchucks are groundhogs (278) (since 'woodchuck' and 'groundhog' designate the same "natural kind"). Dorr (2016*b*: n. 25) likewise suggests a view where $\mathrm{LL}_{\mathrm{sub}}$ instances involving attitude verbs are valid when the arguments of = are both simple expressions, but not otherwise.

However, there is a forceful argument that proponents of this package should accept that 'We can know a priori that every golden thing is made of atoms containing seventy-nine protons' is true at least in some contexts. Let's introduce a new simple predicate "seventyninish" by stipulating that the identity 'To be seventyninish is to be made of atoms containing seventy-nine protons' should be true. Surely 'We can know a priori that every seventyninish thing is made of atoms containing seventy-nine protons' is true, at least in some contexts. But since 'golden' and 'seventyninish' are both simple and 'to be golden is to be seventyninish' is true (by (2) and Leibniz's Law), the restricted form of Leibniz's Law licenses the inference from this to 'We can know a priori that every golden thing is made of atoms containing seventy-nine protons.' (Robertson Ishii (2018) considers, but ultimately rejects, a closely related argument where instead of a new predicate we use quantifying in, along the lines of 'being made of atoms containing seventy-nine protons is such that we can know a priori that everything that does it is made of atoms containing seventy-nine protons'.) For more relevant discussion, see Yli-Vakkuri and Hawthorne forthcoming b.

So, if you want to insist that it is obviously *not* broadly necessary that every golden thing is made of atoms containing seventy-nine protons, you had better think that the identity (2) is obviously false. But it is obviously *not* obviously false! The investigation of identities like (2) seems to play a central role in science. Something has gone terribly wrong if we start dismissing them out of hand based on philosophical arguments having to do with a priori knowledge and the like.[20]

Just as with the earlier 'Hesperus is Phosphorus' argument, there are two possible diagnoses of the failure of the above argument from P1 and P2′ to C′. One diagnosis maintains that the substitution of identicals is valid, even for propositional attitude verbs, for higher-order as well as first-order identities. On this view, the truth of (2) guarantees that (7) is true on every uniform literal interpretation:

(7) If it is knowable a priori that every thing made of atoms that contain seventy-nine protons is made of atoms containing seventy-nine protons, it is knowable a priori that every golden thing is made of atoms containing seventy-nine protons.

Those who think (7) is true on every uniform literal interpretation might think that its *consequent* is also, surprisingly, true on every uniform literal interpretation (Stalnaker 1978), or they might think that its antecedent is, surprisingly, false on some uniform literal interpretations (Goodman and Lederman 2021). The second diagnosis holds that despite the truth of (2), substituting 'made of atoms containing seventy-nine protons' for 'golden' fails to preserve truth in attitude ascriptions (holding context fixed), so that (7) is false on at least some uniform literal interpretations. But in that case, higher-order existential generalization must fail when attitude verbs are around, both for terms in the scope of the verb and for the verb itself.

As a last resort, one might try to preserve $\mathrm{EG}_{\mathrm{sub}}$ for the target operator O by claiming that it licenses both the substitution of 'Hesperus' for 'Phosphorus' and the substitution of 'made of atoms containing seventy-nine protons' for 'golden', while still offering a version of the P1–P2 argument to refute the broad necessity of some other Kripkean example, e.g. that Nixon is not an inanimate object. But this is completely unpromising. If we could make our peace with the initially bizarre idea that we can know a priori that every golden thing is made of atoms containing

[20] It is worth being clear that even the Structure principle from §7.3 does not rule out the truth of (2), except insofar as, being inconsistent, it rules out everything. The idea of Structure is that when the expressions on either side of a true identity are both complex, they must be complex in the same way, and identities involving the corresponding constituent expressions must also be true. A principle that ruled out identities with a simple expression on one side and a complex expression on the other would be much more obviously problematic. Many such identities, such as (2), are highly plausible. Moreover, unless one were willing to say that $\exists x(x = a)$ is false whenever a is complex, such a principle would have to be restricted to closed terms; but it is hard to see what the principled basis for such a restriction could be. What is to stop us from introducing a new simple term to denote something already denoted by some complex term?

seventy-nine protons, it should at that point be totally unobvious whether we can also know a priori that Nixon is not an inanimate object. The same point applies to the other candidate values of *O*.

We conclude that it is *not* obviously false that Kripke's examples (1) are broadly necessary, so that the need to avoid attributing obvious falsehoods does not preclude us from taking at face value his remarks suggesting that the status he is claiming for them is the broadest necessity. While there are seductive arguments that they are not broadly necessary based (inter alia) on claims about a priori knowledge, these arguments should already have been ringing alarm bells for anyone familiar with *Naming and Necessity*. Our discussion has confirmed that such arguments can be independently seen to be problematic. And for what it's worth, we think that Kripke's celebrated warnings about inferring possibility claims from various kinds of epistemic status are most needed when the topic is broad possibility. For a restricted modality, it is a lot harder to see why anyone would have been tempted to make such inferences in the first place. For example, there wouldn't have been much point in writing a book inveighing against inferring *nomological* possibility from premises about lack of a priori knowability. Interpreting Kripke as talking about some narrower modality not only doesn't fit with his explicit invocations of 'absoluteness' and fails to attend to the crucial dialectical role of Leibniz's Law inferences, but also makes his central theses much less interesting. We suggest that those who still think that there is a good argument from failures of a priori knowledge (or analyticity, or logical truth . . .) to broad possibility haven't taken to heart the central lessons of *Naming and Necessity*.[21]

Even if you concede that it isn't obvious that Kripke's examples aren't broadly necessary, you might doubt that we can *know* that they are broadly necessary, and you might on these grounds be tempted to think that Kripke must have been talking about some restricted necessity that is easier to know about. But we see little basis for these sceptical worries. The arguments of this section turned on one important way of arguing for ascriptions of broad necessity, from first- or higher-order identities. But it would be a mistake to draw the lesson that one could only know something to be broadly necessary by deriving it from identities in this way. For example, it seems that one could know that it's broadly necessary that there are no round squares even if one would flounder embarrassingly when asked to give a helpful answer to the question what it is to be round or square. More generally, we are happy to think that the kinds of psychological procedure that led to Kripke's modal judgements—methods that don't give pride of place to

[21] Of course, there may be some other interesting theoretical argument that certain of the examples are not broadly necessary. But as we have been at pains to point out, such an argument should be construed as an argument that the examples lack the status Kripke was claiming for them, and not as an argument that he was claiming a lesser status than that of broad necessity.

propositions about higher-order identities in the order of knowing—are generally good ways of acquiring knowledge about what is and isn't broadly necessary.

8.2 The Broadest Necessity in the Setting of Classicism

We haven't offered any definition of 'necessity operation'. Nor have we argued that there is a necessity operation that applies only to propositions to which all necessity operations apply, or that if there is such an operation X, the operation of applying X twice over (i.e. $\lambda p.X(Xp)$) is itself a necessity operation. Some readers might, understandably, still be pessimistic as to whether there is any way of getting a grip on the category of necessity operations in such a way as to make both these claims plausible, without an antecedent commitment to Iteration for metaphysical necessity. For example, if the only reasonable interpretation of 'necessity operation' made it equivalent to 'operation applying to all metaphysical necessities', the thesis that double metaphysical necessity is a necessity operation is just a restatement of, rather than an interesting argument for, Iteration. Alternatively, one might think one does have some independent fix on the category of necessity operations, but see no reason to preclude the thesis that for each such operation $\Box_1$ there is another one $\Box_2$ such that $\exists p(\Box_2 p \wedge \neg\Box_1 p)$ (see Fritz 2017*a*; Rayo 2020; Roberts 2019).

To help dispel such pessimism, this section and the next will consider how the questions whether there is a broadest necessity operation and whether Iteration holds for it look against the background of a controversial but attractively simple theory about the fineness of grain of higher-order reality which a number of theorists have been tempted to adopt. The theory in question is dubbed "Classicism" by Bacon and Dorr (forthcoming). Its basic idea is that logical equivalence (in our basic classical higher-order logic H_0, supplemented with the full Beta-conversion and Eta-conversion schemas: see note 13 in Chapter 1) is sufficient for higher-order identity. Classicism can be axiomatized by extending H_0 with every instance of Beta-conversion, Eta-Conversion, and the following schema:

Logical Equivalence $\lambda x_1 ... x_n.P = \lambda x_1 ... x_n.Q$ whenever $P \leftrightarrow Q$ is a theorem of H_0 (where $n \geq 0$).[22]

Classicism's simplicity and naturalness become more apparent when one considers the range of different ways in which it can be axiomatized. Bacon and Dorr

[22] Classicism is a strengthening of the theory Dorr (2016*b*: §7) calls *Booleanism*: Booleanism can be axiomatized by replacing Logical Equivalence with "Tautological Equivalence", which is the same except that $P \leftrightarrow Q$ has to be a theorem of classical *propositional* logic. Classicism is an extremely natural generalization of Booleanism, since it extends to the other logical constants (identity and the quantifiers) a treatment that Booleanism, somewhat arbitrarily, restricts to the truth-functional connectives.

(forthcoming) discuss some axiomatizations which use just a tiny, simple subset of the instances of Logical Equivalence.[23] They also show that the Classicism is the result of closing H_0 under the following rule, according to which provable claims of coextensiveness can always be "upgraded" to identities:

EQUIV If $\vdash Fx_1...x_n \leftrightarrow Gx_1...x_n$, then $\vdash F = G$ (for F, G of type $\langle \sigma_1, ..., \sigma_n \rangle$, $n \geq 0$)

It follows that Classicism also includes $\lambda x_1...x_n.P = \lambda x_1...x_n.Q$ whenever $P \leftrightarrow Q$ is a theorem of Classicism but not of H_0.

We can explicitly define, in purely logical terms, an operation with an excellent claim to be the broadest necessity on the assumption that Classicism is true. Let $\top$ be some arbitrarily chosen theorem of Classicism, say $\forall p(p \vee \neg p)$; and let $\Box_\top$ abbreviate $\lambda p.p = \top$. Then it follows from the closure of Classicism under EQUIV that $\Box_\top$ obeys not only our basic modal logic H_{KA}, but the stronger logic H_{S4}. To show this, it suffices (see Appendix A, Corollary A12) to show that Classicism includes the $\Box_\top$-versions of the K, 4, and T axioms and is closed under necessitation. For the K axiom, note that $q \leftrightarrow (\top \rightarrow q)$ is a theorem since $\top$ is, so $q = (\top \rightarrow q)$ is a theorem by EQUIV, so $(p = \top) \rightarrow (q = (p \rightarrow q))$ is a theorem by LL; finally, another application of LL gives us $(p \rightarrow q) = \top \rightarrow (p = \top \rightarrow q = \top)$, which by E$\beta$ yields $\Box_\top(p \rightarrow q) \rightarrow \Box_\top p \rightarrow \Box_\top q$. For the T axiom, note that since $\top$ is a theorem, $(p = \top) \rightarrow p$ is too by LL. For the 4 axiom, note that since $\top \leftrightarrow \Box_\top \top$ is a theorem, $\top = \Box_\top \top$ is too by EQUIV, so $(p = \top) \rightarrow (\Box_\top p = \top)$ is by LL, and thus $\Box_\top p \rightarrow \Box_\top \Box_\top p$ is by Eβ. For necessitation, suppose that P is any theorem of Classicism. Then $P \leftrightarrow \top$ is also a theorem, so by EQUIV, $P = \top$ is, so by Eβ, $\Box_\top P$ is.[24]

While it's a tall order to come up with a defensible and nontrivial account of what it is to be a necessity operation that is completely neutral on questions of grain, the task is more tractable once we assume Classicism. In this context, there are several candidate definitions of 'necessity operation' that can be spelled out in terms of various constraints on logical role. Here are six worth considering:

$\mathrm{Nec}_1\, X := X\top$, i.e. $\forall p(\Box_\top p \rightarrow Xp)$.

$\mathrm{Nec}_2\, X := X^*(\mathrm{Nec}_1\, X)$, i.e. $Y(\mathrm{Nec}_1\, X)$ for every finite iteration Y of X.

$\mathrm{Nec}_3\, X := \Box_\top(\mathrm{Nec}_1\, X)$, or equivalently, $\Box_\top(\mathrm{Nec}_2\, X)$.[25]

[23] Dorr (2016*b*: n. 59) gives another fairly minimal axiomatization of Classicism using certain "Adjunction" principles. The present section closely follows the discussion on pp. 68–70 of that paper.

[24] Bacon (2018*a*) shows that the strongest *propositional* modal logic for $\Box_\top$ that Classicism guarantees is *exactly* S4.

[25] *Proof*: since H_0 implies that every operation is a finite iteration of itself, Nec_2 entails Nec_1 and so $\Box_\top(\mathrm{Nec}_2\, X)$ entails $\Box_\top(\mathrm{Nec}_1\, X)$. In the other direction, if $\Box_\top(\mathrm{Nec}_1\, X)$ then we have $(\lambda p.p)(\mathrm{Nec}_1\, X)$ by the factivity of $\Box_\top$ and $\forall Y(Y(\mathrm{Nec}_1\, X) \rightarrow YX(\mathrm{Nec}_1\, X))$ which implies that $Y(\mathrm{Nec}_1\, X)$ for every finite iteration Y of X.

$\mathrm{Nec}_4\, X :=$ $\mathrm{Nec}_1\, X \wedge \forall p \forall q(X(p \to q) \to Xp \to Xq)$.[26]

$\mathrm{Nec}_5\, X :=$ $X^*(\mathrm{Nec}_4\, X)$.

$\mathrm{Nec}_6\, X :=$ $\Box_\top(\mathrm{Nec}_4\, X)$, or equivalently, $\Box_\top(\mathrm{Nec}_5\, X)$.[27]

Of course, there are other variations: for example, one might wish to give the T axiom $\forall p(Xp \to p)$ the same status that the last three definitions give the K axiom.

Bacon (2018*a*) calls necessities$_1$ "weak necessities" and necessities$_3$ "necessities". But even the category of necessities$_3$ seems to let in too much: for example, broad *possibility*, i.e. $(\lambda p.\top \neq \neg p)$, is a necessity$_3$. The necessities$_4$ do better on this front, but X's being a necessity$_4$ does not even require that $XX(p \vee \neg p)$, which seems to be just as integral as the K axiom to the kind of logical role we are trying to articulate. The category of necessities$_5$ is much more promising, since it captures the idea of "obeying the basic modal logic H_{KA}" in a pretty natural way: whenever P is a theorem of H_{KA} (not containing the constant $\Diamond$), $\mathrm{Nec}_5 \Box \to P$ is a theorem of Classicism.[28] However, it seems somewhat odd to count operations which just *happen* to be necessities$_5$ as necessity operations. For example, perhaps Satan could decide on a whim to write down a list of sentences in such a way as to make *expressed by some sentence written by Satan* count as a necessity$_5$, but it seems strange to suppose that such a decision would make that a necessity operation. So assuming Classicism, our favourite option of these six is to identify *being a necessity* with *being a necessity*$_6$.

But for the purposes of arguing for Iteration, it doesn't matter whether you agree with this judgement. For all the candidates above, and all the other natural candidates we can imagine, have two things in common: $\Box_\top$ counts as a necessity operation, and $\forall p(\Box_\top p \to Xp)$ (i.e. $X\top$) is true for every necessity operation X. Thus, no matter which candidate we pick, it will turn out that '$\Box_\top$ is an extensionally minimal necessity operation' is true.[29]

Granted, one might balk at the idea that being a necessity operation is just a matter of playing the right kind of logical role, and thus worry that all the definitions canvassed above are inadequate. They all have the consequence that

[26] This is equivalent given Classicism to $X\top \wedge \forall p \forall q(X(p \wedge q) \leftrightarrow (Xp \wedge Xq))$. *Left-to-right*: suppose $X\top \wedge \forall p \forall q(X(p \to q) \to Xp \to Xq)$. Then since $\top = ((p \wedge q) \to p) = ((p \wedge q) \to q) = (p \to (q \to (p \wedge q)))$, we have $X(p \wedge q) \to Xp$, $X(p \wedge q) \to Xq$, and $Xp \to Xq \to X(p \wedge q)$. *Right-to-left:* Suppose $X\top \wedge \forall p \forall q(X(p \wedge q) \leftrightarrow (Xp \wedge Xq))$. Then since $((p \to q) \wedge p) = (p \wedge q)$ by Logical Equivalence, we have $(X(p \to q) \wedge Xp) \to X(p \wedge q)$ and hence $X(p \to q) \to Xp \to Xq$.

[27] *Proof*: as in note 25.

[28] *Proof*: this is immediate for axiom $\Box^*$K. For axiom $\Box^*$N, note that when P is a theorem of H_0, Classicism proves $P = \top$ and so $\Box P = \Box\top$. Also, since H_0 proves $\forall v_1 ... v_n \Box\top \leftrightarrow \Box\top$, Classicism proves $\forall v_1 ... v_n \Box\top = \Box\top$, and so $Y\forall v_1 ... v_n \Box\top \leftrightarrow Y\Box\top$. Hence, if $Y\Box\top$ whenever Y is a finite iteration of $\Box$, then $Y\forall v_1 ... v_n \Box P$ whenever Y is a finite iteration of $\Box$.

[29] If we choose Nec_3 or Nec_6, something stronger will be true, namely that $\Box_\top$ *entails* every necessity operation (i.e. $\Box_\top \forall p(\Box_\top p \to Xp)$ whenever X is a necessity operation). But that stronger claim, while plausible, is not needed for the purposes of establishing Iteration.

if they start with any necessity operation X and proposition p, the operation $X_p := \lambda q.((p \wedge X(p \rightarrow q)) \vee (\neg p \wedge X(\neg p \rightarrow q)))$, derived from X by "holding fixed the truth value of p", is also a necessity operation. X_p is guaranteed to apply either to p or to its negation, even when p is some proposition like *I had breakfast today*, to which we would not ordinarily be tempted to apply the word 'necessary'. This has led some to embark on a project of coming up with further tests for "genuine modal force" that go beyond the sorts of logical features invoked in the above definitions, and that will rule out these cooked-up operations.[30] But it is hard to imagine how $\Box_\top$ could end up failing any reasonable further tests that might emerge from such a project.[31]

All of this creates a strong presumption that if Classicism is true, $\Box_\top$ counts as a necessity operation whose extension is contained in that of every other necessity operation, and hence that metaphysical necessity is identical to or at least coextensive with $\Box_\top$.[32] This presumption might be defeated if there was a sufficiently obvious case, assuming Classicism, against the $\Box_\top$-necessity of many of the standard Kripkean examples. But there is no such case: everything that we said in the previous section about the defensibility of the thesis that the Kripkean examples are broadly necessary applies to the thesis that they are identical to $\top$.

Indeed, even if we completely bracket the question whether $\Box_\top$ coincides with metaphysical necessity, there is a strong case given Classicism for the identity to $\top$ of most of the usual examples. For example, it is very plausible that (8) is true if Classicism is:

(8) $\Box_\top$ Everything golden is composed of atoms with seventy-nine protons.

[30] Rosen (2006: 33) insists that there is '*no sense whatsoever*' in which he but not Tony Blair has his location necessarily. For similar thoughts see Fine 2002, Lange 2005, and Kment 2018 (ch. 2). In general we are sceptical about the project of finding a *natural* way of demarcating the class of necessities that excludes the "breakfast-necessity" operation but still lets in the operations expressed by everyday uses of modal words, as in 'I couldn't finish my muesli this morning' or 'I had to scramble the eggs (because the yolk broke).'

[31] Lange (2005) and Kment (2018) endorse counterfactual tests on which, if X is a genuine necessity, the counterfactual $p > q$ must be true whenever $\neg X \neg p$ and Xq. Assuming Classicism, $\Box_\top$ plausibly passes this test. Indeed, $\top$ might be thought to enjoy a unique kind of counterfactual robustness: whereas '$\top$ would be the case no matter what' (i.e. $\forall p(p > \top)$) is surely true on all readings, it is natural to think that for any p distinct from $\top$, we can at least get into a context in which 'p would be the case no matter what' (and hence $\neg p > p$) is false.

[32] Given Beta and Eta-conversion, the combination of Classicism with the view that $\Box_\top$ is metaphysical necessity is equivalent to Intensionalism (see §1.4). To see how Intensionalism follows from that combination, note that since $(Fx \wedge \forall x(Fx \leftrightarrow Gx)) \leftrightarrow (Gx \wedge \forall x(Fx \leftrightarrow Gx))$ is a theorem of H_0, the corresponding identity $(\lambda x.Fx \wedge \forall x(Fx \leftrightarrow Gx)) = (\lambda x.Gx \wedge \forall x(Fx \leftrightarrow Gx))$ is an instance of Logical Equivalence. By LL this implies $\forall x(Fx \leftrightarrow Gx) = \top \rightarrow (\lambda x.Fx \wedge \top) = (\lambda x.Gx \wedge \top)$. But Classicism proves $Fy \leftrightarrow (\lambda x.Fx \wedge \top)y$ and hence also $F = (\lambda x.Fx \wedge \top)$ by EQUIV, and similarly for G, so we can simplify this to $\forall x(Fx \leftrightarrow Gx) = \top \rightarrow F = G$. (A parallel proof works for polyadic or propositional F and G.) To derive Logical Equivalence from Intensionalism, we need only note that when $P \leftrightarrow Q$ is a theorem of H_0, $\forall x_1 \ldots \forall x_n((\lambda x_1 \ldots x_n.P)x_1 \ldots x_n \leftrightarrow (\lambda x_1 \ldots x_n.Q)x_1 \ldots x_n)$ is too, so H_K includes its necessitation, which implies $\lambda x_1 \ldots x_n.P = \lambda x_1 \ldots x_n.Q$ given Intensionalism. Finally, Intensionalism implies $\Box = \Box_\top$, since it implies $\Box \forall p(\Box(p \leftrightarrow \top) \leftrightarrow (p = \top))$ and hence also $\Box \forall p(\Box p \leftrightarrow \Box_\top p)$.

For given Classicism, (9) is true:

(9) $\Box_\top$ Everything composed of atoms with seventy-nine protons is composed of atoms with seventy-nine protons.

And (8) follows from the combination of (9) with an identity which is plausible quite independently of Classicism:

(2) To be golden is to be composed of atoms with seventy-nine protons.

Moreover, once we have taken the truth of (8) on board, it would be quite weird if any of the following were false:

(10) a. $\Box_\top$ There are no round squares.
b. $\Box_\top$ Everything scarlet is red.
c. $\Box_\top$ Every ham is made from a pig.
d. $\Box_\top$ Every carrot is a root.
e. $\Box_\top$ Penne is a type of pasta.

We find it very hard to take seriously a view that accepts Classicism and (8) but rejects these less scientific examples. True, in the case of the examples in (10) we are not in such a good position to lay our finger on a higher-order identity that can play the argumentative role that (2) played with respect to (8), i.e. it both implies the relevant $\Box_\top$ claim given Classicism, and is one that even proponents of theories of higher-order identity much finer-grained than Classicism have good reason to accept.[33] But the absence of that very specific sort of argument is certainly not any kind of reason for a *Classicist* to doubt that the relevant propositions are identical to $\top$.[34]

[33] Of course, if we assume Classicism, it's as easy as can be to find plausible higher-order identities that imply the ones in (10): for example, 'To be scarlet is to be scarlet and red' implies (10b). However, unlike (2), there are principled non-Classicist reasons for rejecting this identity: e.g. it is ruled out by the 'Only Logical Circles' principle from Dorr 2016*b*.

[34] One interesting kind of view that is compatible with Classicism, and explored by Bacon (2020), is that there is a category of *fundamental* entities (objects, properties, relations, operations . . .) which obey some form of *combinatorialism* with respect to $\Box_\top$. On this view, in a language all of whose nonlogical constants stood for distinct fundamental entities, $\Box_\top P\phi$ is true for closed sentences ϕP only when $\forall v_1 \ldots v_n \phi P^*$ is, where ϕP^* is the result of replacing all the nonlogical constants in ϕP with distinct variables $v_1, \ldots, v_n$. This picture provides an interesting route from attributions of fundamentality to claims of higher-order distinctness. For example, it allows one to argue that if *exact location*, *being a particle*, and *being a point* are all fundamental, then it's false that for no particle to be exactly located at more than one point is for it to be the case that $\top$. If one combines this with the further premises that it is metaphysically necessary that no particle is exactly located at more than one point and that exact location, being a point, and being a particle are fundamental, one can conclude that metaphysical necessity is not the broadest necessity. Sociologically, we suspect that some of the resistance to the claim that metaphysical necessity is the broadest necessity is driven by this sort of thinking. But we find it

As we discussed in the previous chapter, there is a coherent position on which *objects* are "less prone to enter into interesting patterns of higher-order identity" than things of other types. Some versions of this position are consistent with Classicism, and would provide a principled basis for denying that $\Box_{\mathsf{T}}$Nixon is not an inanimate object, that $\Box_{\mathsf{T}}$Lectern L is not made of ice, etc. But that position is highly tendentious, since it also rules out identities like 'To be in the Tri-State Area is to be in New York, New Jersey, or Connecticut' and 'To be the first element of Parispair is to be identical to Paris, France', which are plausible quite independently of Classicism. For those who are willing to take identities like these on board, the *de re* $\Box_{\mathsf{T}}$ claims become similar in plausibility to the examples in (10).

So, against the background of Classicism, the $\Box_{\mathsf{T}}$-attributions corresponding to standard Kripkean claims about metaphysical necessity are not only defensible but prima facie plausible. Of course, they are not *irresistible*, but our case that metaphysical necessity is coextensive with $\Box_{\mathsf{T}}$ doesn't require them to be irresistible, since as we emphasized in §8.1, the corresponding attributions of metaphysical necessity are not irresistible either. The upshot is that we have as yet seen nothing to defeat the presumption that $\Box_{\mathsf{T}}$ is what the expression 'metaphysically necessary' has been getting at all along, or at the very least, that the two statuses are coextensive.[35]

As noted above, Classicism implies that $\Box_{\mathsf{T}}$ obeys the modal logic H_{S4}, which includes $\Box\forall p(\Box p \rightarrow \Box\Box p)$. The thesis that metaphysical necessity is *identical* to $\Box_{\mathsf{T}}$ thus immediately implies that the Iteration principle for metaphysical necessity is not only true but metaphysically necessary. But while there is much to be said in

hard to sympathize with someone is who simultaneously confident in the required way of attributions of fundamentality, attributions of metaphysical necessity, and broad combinatorialism. Fundamentality in particular was never supposed to be the sort of status one could easily discern. And in the particular case of exact location, the claim of fundamentality seems especially suspect: one might for example think that *exact* location is not fundamental but *location* is fundamental, and to be exactly located at a point is to be located at that point and not at any other point.

The combinatorialist picture is sometimes combined with the idea that *objects* are all fundamental (and thus freely recombinable), so that something like Generalized Qualitative Atomic Structure from Chapter 7 is true. Unless it is combined with a programme of reconstruction whereby ordinary singular terms are not really of type e, this will immediately yield rampant failures of Microphysical Supervenience for $\Box_{\mathsf{T}}$. Even if we were tempted by such a radical view, we would take it as reason to endorse correspondingly radical claims about metaphysical possibility (including Extreme Anti-essentialism: see §5.1) rather than a reason for distinguishing $\Box_{\mathsf{T}}$ from metaphysical necessity.

[35] To make a case for identity rather than mere coextensiveness, one could bear down further on what 'broader than' means in the context of formulae like 'metaphysical necessity is the broadest necessity'. As Bacon (2018*a*) points out, if we adopt a definition of 'necessity operator' that implies Nec_3 above, and define 'X is broader than Y' means $\Box_{\mathsf{T}}\forall p(Yp \rightarrow Xp)$, it follows that $\Box_{\mathsf{T}}$ is at least as broad as every necessity operation, and moreover that there are no two distinct operations each of which is at least as broad as the other. Bacon's account of 'at least as broad as' seems natural given a definition of 'necessity operator' such as Nec_3 or Nec_6, but would be unpromising if one favoured, say, Nec_5. But the whole question seems a little academic: the extreme naturalness of $\Box_{\mathsf{T}}$ creates considerable pressure to think that 'metaphysical necessity' refers to it irrespective of the interpretation of 'broader than', and the interesting reasons for resisting this pressure are also reasons to think that metaphysical necessity isn't even coextensive with $\Box_{\mathsf{T}}$.

favour of that identity, the weaker claim that metaphysical necessity is *coextensive* with $\Box_\top$ is perfectly adequate for the purposes of arguing that Iteration is *true*. We don't even need to argue that double metaphysical necessity is itself a necessity operation; all we need is the completely uncontroversial claim (in the setting of Classicism) that $\top$ is doubly metaphysically necessary. If metaphysical necessity is coextensive with $\Box_\top$, so that $\top$ is the *only* metaphysical necessity, this immediately implies that everything metaphysically necessary is doubly metaphysically necessary, i.e. that Iteration is true.[36]

8.3 Is There Only One Metaphysically Necessary Truth?

Bacon (2018*a*) accepts Classicism, and argues that given Classicism, $\Box_\top$ is the broadest necessity operation.[37] But he denies that $\Box_\top$ is coextensive with metaphysical necessity. He gives several arguments. One appeals to the claim that not all metaphysical necessities can be known a priori, which we have already discussed. Another appeals to the thesis that some propositions—e.g. that it is raining if it is raining now—are metaphysically necessarily true but not *always* true. Since one of us has co-written a whole paper arguing against this thesis (Dorr and Goodman 2020), we will not discuss it further here. A third argument appeals to the phenomenon of vagueness, which Bacon regiments using an 'it is determinately the case that' operator Δ: Bacon suggests for example that for some n, it is metaphysically necessary but not determinately the case that all and only those with at least n hairs are bald. But as Bacon recognizes (2018*b*: ch. 4), the orthodox view that vagueness is a linguistic phenomenon suggests that if we are going to insist on using something with the superficial syntax of a sentential operator here, we should expect it to be logically ill-behaved in the manner of operators involving hidden quotation.[38] One consideration in favour of the orthodoxy—or at least, the logical ill-behavedness of Δ—is that there seem to be cases where a disjunction of identities is true even though all of its disjuncts are vague. If one uses a sentential operator Δ, this means we should regard the following as consistent:

$$(11) \qquad (a = b \vee a = c) \wedge \neg\Delta(a = b) \wedge \neg\Delta(a = c)$$

[36] We owe special thanks here to Andrew Bacon and Peter Fritz for helpful discussion.

[37] Modulo the replacement of a functional with a relational type system (see note 6 in Chapter 1), Classicism is equivalent to the result of adding the "Modalized Functionality" axiom (Bacon 2018*a*: 779) to Bacon's system HE.

[38] Dorr (2010: §1) discusses how one might cash out sentences using Δ in terms of predicates applicable to sentences and open formulae. As with logical necessity (see note 8), another option would be to define Δp as something like 'p is expressed by a precise sentence'; but then if we think the same p can be expressed by both precise and vague sentences, we won't be able to infer $\ulcorner\neg\Delta\varphi\urcorner$ from $\ulcorner$'φ' is (extensionally) vague'$\urcorner$.

Lewis (1988) gives an example of this form where *a* is 'Princeton', *b* is 'Princeton Borough', and *c* is 'Princeton Township'. Another example using higher-order identities has *a* be 'mass', *b* 'relativistic mass', and *c* 'rest mass' (Field 1973). But as Lewis points out, for (11) to be consistent, Δ must be implicated in failures of Leibniz's Law, in which case (as we have explained) it will also be ill-behaved with respect to the quantifier laws. And it seem to us to be quite integral to our practice of thinking about vagueness that it gives rise to cases which will have the form of (11), when represented using the Δ operator. The vagueness of a word *always* involves there being multiple "candidates", so that when the word *a* (of any type) is vague, $\forall x \neg \Delta(x = a)$ is true.[39]

However, Bacon has one more strategy for arguing against the identification of metaphysical necessity with $\Box_\top$ which we find much more challenging. This argument appeals to the premise that ND ($\forall x \forall y(x \neq y \rightarrow \Box x \neq y)$) holds when $\Box$ is interpreted as metaphysical necessity, but fails when it is interpreted as $\Box_\top$.[40] We discussed some possible motivations for thinking that metaphysical necessity obeys ND in §4.2. The argument that we found most persuasive is Williamson's (1996) argument from certain principles about the interaction between metaphysical necessity and the "actually" operator; we suggested that it might be possible to run a similar argument using resources of "modal anaphora" closer to those found in natural languages. Bacon (2018*a*) suggests one argument that $\Box_\top$ does not obey ND which we do not find convincing: this argument both relies on a non-linguistic account of vagueness and invokes a notion of "conceptual truth" that we find obscure. But there are other motivations that we find weightier for at least taking the failure of ND for $\Box_\top$ as a live hypothesis. Bacon (2020) shows that ND for $\Box_\top$ is incompatible with a certain strong but attractively simple "logical combinatorialist" picture articulating the idea that *fundamental* entities are all independent of one another as regards broad necessity.[41] Bacon and Dorr

[39] Bacon has an interesting and highly developed theory (Bacon 2018*b*) that rejects both the consistency of sentences like (11), and the orthodox picture of vagueness as a linguistic phenomenon (although he is happy to allow that propositions about linguistic relations like reference can suffer from vagueness in the same way as propositions about other matters). Space precludes a proper exploration of Bacon's view, or of the fallbacks he might suggest as substitutes for (11). But we note that while Bacon opposes the idea that vagueness is a *linguistic* phenomenon, his positive theory turns on the idea that there are intimate connections between vagueness (and the Δ operator) and notions like *rational credence*. In developing these connections, Bacon assumes that operators like '*a* has credence *x* that…' are logically well behaved with respect to the quantifiers. But if one accepted any of the putative counterexamples to universal instantiation and existential generalization for attitude verbs which we discussed above, one might expect Δ to suffer from the same logical unruliness given Bacon's picture of it as belonging to the same family.

[40] ND is a really a schema with an instance for every type. The argument works so long as there is *any* type for which ND is true for metaphysical necessity but false for $\Box_\top$.

[41] See note 34 above. Relatedly, Goodman (unpublished *d*) discusses a principle in the neighbourhood of Generalized Qualitative Atomic Structure (see previous section), but compatible with Classicism, on which the uncontroversial facts that distinctness and mutual self-identity (i.e. $(\lambda x^e y.x = x \wedge y = y)$) are both qualitative and distinct from one another would entail that we never get the same proposition by predicating these relations of the same two objects. As we explained in the

(forthcoming) explore related ideas, some of which lead to failures of ND for $\Box_\top$ without any need to mention fundamentality.

But even if you find the arguments for accepting ND for metaphysical necessity and rejecting it for $\Box_\top$ convincing enough to warrant rejecting (or suspending judgement about) the identification of metaphysical necessity with $\Box_\top$, that will not do much to undermine our case for Iteration. For beginning with any of the candidate definitions of 'necessity operation' considered in the previous section, we can readily define a more demanding category of "ND-respecting" necessity operations. These refined definitions simply give ND the same criterial status that definitions Nec_4–Nec_6 gave to the K axiom.[42] And as we show in Appendix D, there is a specific operation $\Box_{\neq}$, defined like $\Box_\top$ in purely logical terms, for which Classicism proves that $\Box_{\neq}$ is extensionally minimal within each of the relevant categories of ND-respecting necessity operations. Moreover, $\Box_{\neq}$ provably obeys H_{S4}, just as $\Box_\top$ does.

The spirit of the slogan 'metaphysical necessity is broadest necessity' seems entirely consistent with this *logical* way of narrowing down the category of operations within which metaphysical necessity is supposed to be extensionally minimal. For the interesting arguments that metaphysical necessity obeys ND, like Williamson's, are not specific to metaphysical necessity, but are equally compelling for any interpretation of 'necessary' allowable in English. Insofar as you find such an argument persuasive, you should thus regard the principles required to run it—e.g. 'Necessarily, if a proposition is actually not true, it is not actually true'—as capturing an important part of the characteristic "logical role" of the operations expressible by modal words in natural languages. In that case, a narrower interpretation of 'necessity operation' that includes some form of respect for ND should also seem reasonable.[43]

There is a further narrowing that could be motivated on similar grounds. If $\Box_\top$ fails to obey ND, then it might also fail to obey the Barcan Formula (BF: $\forall x \Box Fx \rightarrow \Box \forall x Fx$).[44] Moreover, the considerations that might lead one to take failures of ND for $\Box_\top$ seriously seem likely to also count against BF.[45] But BF for *metaphysical*

previous section, our view is that theories like Goodman's which single out type *e* for this kind of special treatment rule out too many plausible identities, unless they are combined with a programme of reconstructing a lot of *apparent* type-*e* quantification in other terms.

[42] Plausibly, then, the properties of operations corresponding to the refined definitions are only a little less *natural* (Lewis 1983*a*; Dorr and Hawthorne 2013) than those corresponding to the earlier definitions.

[43] Note that even if we reject the identity of metaphysical necessity with $\Box_\top$ on the grounds that ND is true for the former and false for the latter, that does not undermine the reasons for thinking that lots of familiar Kripkean examples are $\Box_\top$. For example, if to be golden is to be made of atoms with seventy-nine protons, the proposition that everything golden is made of atoms with seventy-nine protons is $\Box_\top$ and not merely $\Box_{\neq}$.

[44] By contrast, in the setting of Classicism, the necessitation of ND for type $\langle\rangle$ ($\Box_\top \forall p \forall q (p \neq q \rightarrow \Box p \neq q)$) is equivalent to the necessitation of the 5 axiom ($\Box_\top \forall p (\Diamond_\top p \rightarrow \Box_\top \Diamond_\top p)$), so adding it will secure the full modal logic H_{55} and hence BF.

[45] For example, the "maximalist" view discussed in Bacon and Dorr forthcoming requires failures of both principles.

necessity can, like ND, be derived from some prima facie plausible principles about the logical interaction between metaphysical necessity and 'actually'. Consider the following argument for BF:

1.	$\forall x \Box Fx \rightarrow \Box @ \forall x \Box Fx$	Premise
2.	$\Box \forall y @ (\forall x \Box Fx \rightarrow \Box Fy)$	Premise
3.	$\forall x \Box Fx \rightarrow \Box \forall y @ \Box Fy$	1, 2
4.	$\forall x \Box Fx \rightarrow \Box \forall y Fy$	3

Line 1 can be justified by the widely accepted principle $\forall p (p \rightarrow \Box @ p)$, which we already encountered as part of the derivation of ND in §4.2. Line 2 can be justified by an appeal to the idea that the joint logic of metaphysical necessity and @ should include the result of closing our nonmodal logic H_0 under necessitation for both $\Box$ and @ (as well as GEN). The step from lines 1 and 2 to line 3 can be justified by appeal to the principle that @ obeys the K axiom with metaphysical necessity, i.e. $\Box \forall p \forall q (@(p \rightarrow q) \rightarrow (@p \rightarrow @q))$. Finally, the step from line 3 to line 4 can be justified by appeal to the somewhat less familiar, but still very tempting, principle $\Box \forall p (@ \Box p \rightarrow p)$: necessarily, whatever is actually necessarily the case is the case.[46]

While these principles are not irresistible, those who are convinced that metaphysical necessity "plays well with @" in the ways required to run Williamson's argument for ND will likely also think that it "plays well" in the further ways required by the above argument for BF. But insofar as one is convinced, one will also have a good motivation for further refining one's conception of "having the logical behaviour of a necessity operation" so to include some form of "respect for BF" along with respect for ND. (No further refinement would be required if $\Box_{\neq}$ already obeyed BF; but it is unclear what would motivate a view that accepts BF for $\Box_{\neq}$ but not for $\Box_{\top}$: why should the coming into being of new entities require any two old entities to collapse into one?) As we show in Appendix D, while there are several natural ways of articulating the "respect for BF" requirement, the differences don't matter much, since there is a logically definable operation $\Box_5$ for which we can prove, in Classicism, that it is extensionally minimal within each of the relevant classes of operations. Strikingly, Classicism proves that the logic of this $\Box_5$ is not only H_{S4} + ND + BF, but H_{S5}; indeed, it can equivalently be characterized as the broadest "S5-respecting" necessity operation. Thus, no matter what $\Box_{\top}$ might be doing, $\Box_5$ is guaranteed not only to obey ND and BF, but

[46] If you wanted to resist this argument, where should you resist? On the natural interpretation of @ as a rigid property coextensive with truth (see note 23 in Chapter 4), the problem will arise in the step from 1 and 2 to 3. Suppose $\forall x \Box Fx$ but not $\Box \forall x Fx$. Then while $\Box @ \forall x \Box Fx$ by Persistence, and $\Box \forall y @ (\forall x \Box Fx \rightarrow \Box Fy)$ since $\Box @ \top$ and $\Box \forall y ((\forall x \Box Fx \rightarrow \Box Fy) = \top)$, it is false that $\Box \forall y @ \Box Fy$: if there were a new, non-*F* object, the proposition that it is necessarily *F* would be new too, and thus not a member of @.

to be logically well behaved in every way that could reasonably be demanded of metaphysical necessity.[47]

So if $\Box_5$ is distinct from $\Box_T$, the thesis that metaphysical necessity is $\Box_5$ also strikes us as reasonable and principled, and compatible with the spirit of its characterization as the broadest necessity operation. This thesis implies that metaphysical necessity obeys Iteration. And even the weaker thesis that metaphysical necessity is *coextensive* with $\Box_5$ is bad news for Iteration. For insofar as there is reason to accept it, there is reason to think that metaphysical necessity obeys H_{KA} + ND + BF. But any reasons to think that ND and BF hold for metaphysical necessity will have natural analogues for double metaphysical necessity: in particular, the Williamsonian argument for ND and the analogous argument for BF lose little when $\Box$ is replaced throughout with $\Box\Box$. If so, then since $\Box_5$ is extensionally minimal amongst operations that obey H_{KA} + ND + BF, and double metaphysical necessity is such an operation, it cannot have a narrower extension than $\Box_5$. It follows that everything metaphysically necessary is metaphysically necessarily metaphysically necessary.

8.4 The Broadest Necessity in a Grain-Neutral Setting

The previous section considered the thesis that metaphysical necessity is an extensionally minimal necessity operation from the standpoint of Classicism, a relatively coarse-grained view in pure higher-order logic. While the inconsistency of Structure (see §7.3) should defeat any temptation to think that Classicism is *obviously* false, it is also not obviously true, and there has recently been some exploration of some consistent finer-grained alternatives. But it turns out that the basic argument of the last two sections can also be made using much weaker assumptions which may be acceptable to many of those who reject Classicism as excessively coarse-grained. It is sufficient if our theory T is such that, for a certain operator L of the form $\lambda p.P = Q$ (where p occurs free in one or both of P and Q):

(i) T implies that L obeys H_{K} (i.e. that whenever $\vdash_{\mathsf{H}_{\mathsf{K}}} \vdash P$, $\vdash_T P[L/\Box]$).

(ii) H_0 implies $\forall p(Lp \rightarrow p)$

Given these conditions, we automatically get a stronger version of (i), namely that T implies that L obeys H_{S4}: H_0 secures the T axiom $\forall p(Lp \rightarrow p)$, and the necessity

[47] One nice thing about the result that $\Box_5$ obeys H_{S5} is that, when combined with the actuality-based arguments for ND and BF, it provides a potentially dialectically effective motivation for the standard view that metaphysical necessity obeys S5. This seems more compelling than an argument that treats the sociological fact that S5 is widely assumed as in itself a weighty reason for the semantic gods to interpret 'metaphysical necessity' so as to validate S5 (as suggested in Dorr 2016*b*: 70).

of identity secures $\forall p(Lp \rightarrow LLp)$ via $\forall p((P = Q) \rightarrow L(P = Q))$. Moreover, if we add any other operator $\Box$ to our theory subject to the basic modal logic H_{KA}, the resulting combined theory will prove $\forall p(Lp \rightarrow \Box p)$. For since H_{KA} validates the necessity of identity, it implies $Lp \rightarrow \Box Lp$; since $Lp \rightarrow p$ is a theorem of H_0, we have $\Box Lp \rightarrow \Box p$ by necessitation and K, and hence $Lp \rightarrow \Box p$. L will thus have an excellent claim both to be the broadest necessity and to be metaphysical necessity, just as $\Box_{\mathsf{T}}$ does given Classicism. And if we take its claim to be defeated because we think ND or BF should hold, $L_{\neq}$ and L_5 are waiting in the wings as backups; L_5 in particular will be guaranteed to obey H_{S5} and thus meet every reasonable standard of "logical well behavedness".

These assumptions cover quite a wide range of theories. In Classicism, they hold with $L := \lambda p.p = \top$ (for any H_0-theorem $\top$). In the theory of Goodman (2018, unpublished *d*), they hold with $L := \lambda p.p = (p \vee \neg p)$.[48] In the theory developed in Dorr 2016*b*, they hold with $L := \lambda p.(p \wedge (p \vee \neg p)) = (p \vee (p \vee \neg p))$.[49] And Fine's recent work on "truthmaker semantics" (Fine 2016, 2017*a*) suggests a view on which they hold with $L := \lambda p.(p \wedge \forall q(q \vee \neg q)) = \forall q(q \vee \neg q)$.[50]

But we might want to proceed in a way that is completely neutral on questions about "fineness of grain" not settled by H_0. In that case, there will no longer be much prospect of arguing that some specific operation defined in purely logical terms (i.e. in terms of truth-functional operations, quantifiers, and identity) is an extensionally minimal necessity operation. Indeed, it is not obvious how one would even go about answering the question what it is to be a necessity operation, if one wants an answer that is defensible independent of any view about fineness of grain. If we help ourselves to some semantical vocabulary in the object language, we could try defining '*X* is a necessity' along the lines of 'Every closed theorem of such-and-such modal logic is true when $\Box$ it interpreted as expressing *X*'. But this sort of definition seems too weak, since it will allow *expressed by a sentence*

[48] Goodman's overall theory entails all instances of Generalized Qualitative Atomic Structure (see §7.3), which guarantees that Extreme Anti-essentialism (see §5.1) is true when $\Box$ is interpreted using this L (or the L_5 defined in terms of it). This doesn't of course mean that proponents of Goodman's theory should regard the relevant Kripkean necessities as obviously false, since they should realize that their theory is not obviously true.

[49] See the weak Commutativity, Distributivity, and Dissolution principles in Dorr 2016*b* (app. 8). In the relevant models, all instances of Logical Equivalence hold in the same weakened form, which implies that $\lambda p.Ap = Bp$ obeys H_{S4}.

[50] As Fine (2016*a*) points out, his "bilateral" truthmaker models validate a certain collection of axioms for propositional identity due to Angell (1977, 1989). These axioms are an interesting subsystem of the Boolean identities (see note 22), which many authors who would reject Booleanism as too coarse-grained have found metaphysically plausible (Fine 2017*a*; Correia 2010; Goodman 2018; Lovett 2020*a*). Angell's identities do not apply to the quantifiers, but there is a natural way of extending them to the quantifiers, assuming one thinks that quantifiers should behave in the relevant ways like conjunctions and disjunctions of their instances, as Fine (2017*a*: 639–40) and others (e.g. Lovett 2020*b*) have suggested (though Goodman disagrees). This will yield identities like $\forall xFx = (\forall xFx \wedge Fa)$ (analogous to $(p \wedge q) = ((p \wedge q) \wedge q)$) and $\forall x(p \vee Fx) = (p \vee \forall xFx)$ (analogous to $(p \vee q) \wedge (p \vee r) = (p \vee (q \wedge r))$). We conjecture that when all these identities are added to H_0, $\lambda p.(p \wedge \forall q(q \vee \neg q)) = \forall q(q \vee \neg q)$ will obey H_{S4}.

written down by Satan to count as a necessity if Satan writes down the right list of sentences.

Nevertheless, the fact that we were able to get an adequate fix on how *being a necessity operation* works on the assumption that Classicism is true helps to dispel the pessimistic thought that we have no grip on 'necessity operation' that isn't parasitic on 'metaphysically necessary', and hence useless for arguing against Iteration. And insofar as we are comfortable with the notion of a necessity operation, we may be in a position to see, even without having a definition, that the category of necessity operations obeys certain structural conditions from which it follows that there is an extensionally minimal necessity operation and that all extensionally minimal necessity operations obey Iteration.

Such a strategy is pursued by Williamson (2016).[51] He endorses four conditions, of which only two are important for our current purposes:

Closure Under Composition Whenever X and Y are necessity operations, the result of *composing* X and Y—i.e. $\lambda p.X(Yp)$—is also a necessity operation.

Closure Under Conjunction Whenever C is a collection of necessity operations, the conjunction of C—i.e. *having every member of* C, $\lambda p.\forall X(CX \rightarrow Xp)$—is also a necessity operation.

By Rigid Comprehension, there is a collection of all necessity operations: call the conjunction of all the members of this collection $\Box_{\text{min}}$. For any necessity operation X, we have $\forall p(\Box_{\text{min}}p \rightarrow Xp)$. Whenever $\Box_{\text{min}}p$, Xp is true for every necessity operation X. By Closure Under Conjunction, $\Box_{\text{min}}$ is a necessity operation, and thus an extensionally minimal one. And by Closure Under Composition Conjunction, the result of composing $\Box_{\text{min}}$ with itself (i.e. $\lambda p.\Box_{\text{min}}\Box_{\text{min}}p$) is also a necessity operation. By the extensional minimality of $\Box_{\text{min}}$, it follows that $\forall p(\Box_{\text{min}}p \rightarrow \Box_{\text{min}}\Box_{\text{min}}p)$, i.e. that $\Box_{\text{min}}$ obeys Iteration.

If we also make the reasonable further assumption that all necessity operations obey H_{KR}, so that $\mathsf{R}C \rightarrow CY \rightarrow Z\forall q((\lambda p.\forall X(CX \rightarrow Xp))q \rightarrow Yq)$ is true whenever Z is a necessity operation, we can also infer that if there is some other necessity operation X coextensive with $\Box_{\text{min}}$, it too obeys Iteration. For suppose Xp; then $\Box_{\text{min}}p$, so $\Box_{\text{min}}\Box_{\text{min}}p$; but by the further assumption, $\Box_{\text{min}}\forall q(\Box_{\text{min}}q \rightarrow Xq)$ and hence $\Box_{\text{min}}Xp$, and so XXp by the coextensiveness of X and $\Box_{\text{min}}$.[52]

[51] Williamson speaks of "objective" necessity operations rather than necessity operations simpliciter, but given that he assumes a coarse-grained (Booleanist) account of propositional identity and suggests that $\top$ is the only metaphysical necessity, he is clearly not thinking that the domain of operations also includes non-objective necessity operations that don't apply to all metaphysical necessities. Williamson's assumption of Booleanism is not required for the part of his project we are currently discussing.

[52] The assumption that all necessity operations obey H_{KR} also lets us derive the necessitated 4 axiom $\Box_{\text{min}}\forall p(\Box_{\text{min}}p \rightarrow \Box_{\text{min}}\Box_{\text{min}}p)$, setting Z as $\Box_{\text{min}}$ and Y as $\lambda p.\Box_{\text{min}}\Box_{\text{min}}p$.

Neither Closure Under Conjunction nor Closure Under Composition would be plausible if we adopted the kind of weak interpretation of 'necessity operation' on which Satan's writing down the right list of sentences would be enough for *expressed by a sentence written by Satan* to count as a necessity operation. For even if both Satan and Gabriel write down appropriate lists, there is no guarantee that either λp.*One of Satan's sentences expresses that one of Gabriel's sentences expresses* p or λp.*One of Satan's sentences expresses* p *and one of Gabriel's sentences expresses* p will meet even the most elementary necessary conditions for being a necessity operation. The former won't if none of Satan's sentences expresses that one of Gabriel's sentences expresses that snow is either white or not white, since $X(p \vee \neg p)$ must surely be true for any necessity operation X and proposition p. The latter won't if none of Satan's sentences expresses that the proposition that snow is either white or not white is both expressed by one of Satan's sentences and expressed by one of Gabriel's sentences, since $XX(p \vee \neg p)$ must surely be true for any necessity operation X and proposition p. But it is hardly a cost of these closure principles that they require an interpretation of 'necessity operation' that would exclude *expressed by a sentence written by Satan*, since categorizing such operations as necessities was never particularly plausible. Once we bear this in mind, and allow ourselves not to worry about the fact that we don't have in hand a grain-neutral *definition* of 'necessity operation' that excludes them, both closure conditions seem plausible.

We conclude that there is, overall, good reason to think that there is a broadest necessity operation (at least extensionally). And even though they don't come up all that often in everyday conversations, questions about what is broadly necessary are by no means unfathomable. We know how to think about them, and when the thinking goes well, it yields plenty of knowledge. That includes a lot of cases where the thinking isn't based on any worked-out arguments. (Recall our earlier example of someone who knows that it is broadly necessary that there are no round squares but who struggles to say anything helpful about what it is to be round or to be square.) And it includes a lot of cases where the thinking is a posteriori by any reasonable criterion: for example, by having a little experience with cats we can know that it is broadly necessary that no cats are robots, although we would have thought otherwise if presented with

Williamson gives arguments of a similar sort for the conclusion that the logic of $\Box_{\text{min}}$ is S5, i.e. that it includes the T, B, and 4 axioms and their necessitations. To get the T axiom, he uses the premise that the propositional truth operator $\lambda p.p$ is a necessity operation. (The following weaker premise would also do the job: there is a necessity operation X such that $\forall p(Xp \rightarrow p)$ and for any necessity operation Y, $Y\forall p(Xp \rightarrow p)$.) To get the B axiom, Williamson relies on the less clearly well-motivated premise that for every necessity operation X, there is a necessity operation Y such that for any p, p "entails" $X\neg Y\neg p$, defining "entails" in a way that is fit for purpose only on the assumption of Booleanism. (The following grain-theoretically neutral premise would also work: for every necessity operation X, there is a necessity operation Y such that for every necessity operation Z, $Z\forall p(p \rightarrow X\neg Y\neg p)$.)

misleading evidence suggesting that cats have all along been Martian spy-robots. Likewise for our knowledge that it is broadly necessary that Nixon wasn't a robot.

We are willing, of course, to engage with arguments that would suggest that broad necessity is much less widespread than we suppose. In addition to the arguments discussed in this chapter, such arguments also include Tolerance Arguments couched in terms of broad necessity, which can be used to support a wide range of broad-possibility Hypertolerance claims. Our strategy for responding to these arguments does not turn on our case that broad necessity is coextensive with metaphysical necessity, any more than our rejection of Hypertolerance for ancestral necessity in Chapter 7 turned on the claim that ancestral necessity is coextensive with metaphysical necessity. And indeed, our response to arguments for Broad Hypertolerance would largely parallel our response to arguments for Ancestral Hypertolerance.

If, despite our best efforts, you remain convinced by some argument that broad necessity is much rarer than we suppose, you might also think that there is some highly *natural* operation narrower than the broadest necessity, to which the kind of intuitive theorizing Kripke used is a more reliable guide.[53] If you are in this position, you should think that we and many other philosophers—plausibly including Kripke—have fallen into a state of *identity confusion*, taking what are in fact two distinct operations to be one and the same.[54] Looking in the other direction, we will of course think that you are a victim of the other kind of identity confusion, thinking (like the character Ralph in Quine 1956) that there are two things where there is in fact only one.[55] Seeing that the debate has this abstract structure doesn't help us see who is right: identity confusions of both

[53] "Natural" in the sense of Lewis 1983*a*; for an adaptation to the context of higher-order logic, see Dorr and Hawthorne 2013 and Dorr 2019.

[54] Alternatively, you might think that broad necessity is much rarer than we suppose and there is no particularly natural modality to which the kinds of methods we like is a good guide, although of course there will be various more gerrymandered operations which apply to the propositions we take to be broadly necessary and don't apply to the propositions we rightly take not to be broadly necessary. (For suggestions along these lines see Lewis 1986*a* (252) and Sider 2011 (ch. 11).) The diagnosis of identity confusion is less plausible on that view. If someone introduced a technical term 'inperforate' by explanatory remarks along the lines of 'inperforateness is being absolutely free of holes', but then went around applying this word to lots of objects that have holes too small to see, it would be odd to think that they falsely believed the property of being absolutely free of holes to be identical to the property of being such that all of one's holes are too small to see.

[55] Whichever form of confusion is afoot, banning the expression 'metaphysical necessity' is unlikely to make it go away. For example, if you think there is a natural operation narrower than the broad necessity and which is extensionally in keeping with Kripke-style essentialist remarks, and introduce a new word that you hope will refer to this operation, we will probably still end up thinking that your new word also refers to the broadest necessity. For our part, we have no inclination to ban 'metaphysical necessity': having multiple terms for the same thing is often useful, since it enables us to enter into arguments with opponents who falsely think that there are two things where there is only one, and enables these opponents to say some important true things.

kinds (false identity beliefs and false distinctness beliefs) are quite common in philosophy.[56] But we hope that we have managed to put the case for our own perspective in the strongest light. And insofar as you are inclined to agree with us that metaphysical necessity is the broadest necessity, you should think that the Iteration-denying response to Tolerance Puzzles involving metaphysical modality is on the wrong track.

[56] This pair of perspectives is characteristic of a number of important philosophical disputes. Consider for example the dispute between Meinongians who think that some mountains are made of gold (when 'some' is given its broadest sense) and more orthodox philosophers who say that none are. The Meinongian throws orthodoxy a bone by saying that there is a natural property, *existence*, which none of the golden mountains have, such that the methods the orthodox use to find out what there is are roughly reliable as guides to the realm of existence. They will plausibly think that the orthodox are misled by a false belief in the identity of this property with some universal property like *being identical to something*. The orthodox, by contrast, think that there is no natural property in the vicinity, other than some universal ones, and that Meinongians are being misled by a false distinctness belief.

9
Tolerance and Chance

The previous chapter focused on metaphysical possibility, a particularly undemanding modal status. In this one, we will consider a much more demanding modal status: that of having a positive *objective chance* at a certain time. In the first two sections, we will consider Tolerance Puzzles based on objective chance, and argue that there are some distinctive problems with the Iteration-denying strategy for solving these puzzles. As we will see, this generates further dialectical difficulties for Iteration-denial as an approach to other Tolerance Puzzles such as those involving metaphysical modality. (This part of our discussion has much in common with Kment 2018.) In the final section, we will show how chance can be used to precisify some thoughts about "robustness" which we invoked in our discussion of Hypertolerance in Chapter 6. The result will be a novel family of puzzles, in some respects even more challenging than Tolerance Puzzles.

9.1 Chance and Iteration

One interesting way to instantiate Chapter 2's general schema for Tolerance Arguments (Figure 2.3) is to read 'possible' as 'has positive objective chance at *t*' and 'necessary' as 'has objective chance 1 at *t*', for some chosen *t*. The result of making these substitutions is shown in Figure 9.1. Denying the validity of this argument seems a very unpromising option, as our basic modal logic H_{KA} seems well supported for objective chance 1. For many of the puzzling instantiations, Chance Tolerance arguably has an even firmer footing in ordinary practice than the corresponding Tolerance premise using metaphysical possibility. And likewise, as we discussed in §3.3, the security-based argument for Non-contingency seems especially forceful in the case of chance, since non-sceptics are accustomed to being more concerned about positive chances of certain kinds of error than about mere metaphysical possibilities of error. In many of the instantiations of interest, e.g. those where "closeness" is just adjacency on some finite list, Chance Persistent Closeness is unproblematic. Meanwhile, Chance Hypertolerance is often highly counterintuitive, and indeed at least as repugnant as its analogue for metaphysical modality. Even if one were willing to live with some anything-goes vision of metaphysical possibility such as Extreme Anti-essentialism (see §5.1), one might

The Bounds of Possibility: Puzzles of Modal Variation. Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri, Oxford University Press. © Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri 2021.
DOI: 10.1093/oso/9780192846655.003.0010

Chance Tolerance For every K object x and properties F and G such that x is F and G is close to F, the chance at t that x is K and G is positive.

Chance Non-contingency If Chance Tolerance is true, its chance at t is 1.

Chance Iteration For any proposition p, if p has a positive chance at t of having a positive chance at t, then p has a positive chance at t.

Chance Persistent Closeness Whenever G is close to F, the chance at t that G is close to F is 1.

Chance Hypertolerance For every K object x and properties F and G such that x is F and G is ancestrally close to F, the chance at t that x is K and G is positive.

Fig. 9.1 Chance Tolerance Argument.

well balk at the idea that there was a positive chance, prior to the construction of a pair of tables, of one of the tables being made out of the very materials that the other was in fact made out of; a positive chance, prior to the articulation of the rules of chess, of chess being played according to the rules of Twister; etc.

The remainder of this section will focus on the premise for which there are especially interesting and distinctive things to say on the chance interpretation, namely Chance Iteration. That principle follows immediately from the thesis that questions about what the chances are at a certain time are not themselves matters of chance at that time:

Chance Fixity For any proposition p, time t, and real number x: if the chance at t of p is x, then there is chance 1 at t that the chance at t of p is x.

In our view there is a strong case for Chance Fixity, making Iteration-denial unpromising as a solution to Chance Tolerance Puzzles.

Chance Fixity does not wear its plausibility on its sleeve: we are not used to asking about the chance at t of a proposition's having a certain chance at t. But it is widely accepted, and for good reasons.[1] When we look more closely at our practice of reasoning about objective chance, we can see patterns suggesting a quite deeply rooted commitment to a more general principle of which Chance Fixity can be seen as a special case:

[1] In the terminology of Lewis (1980*a*), Chance Fixity follows from the claim that facts about the chances at a time are "admissible" at that time.

History Fixity Any true proposition *p* entirely about the history of the world up to and including *t* has chance 1 at *t*.

For example: if an AI-written novel has never made the *New York Times* bestseller list, and there is now a 2 per cent chance of this occurring in the next decade, there is a currently 2 per cent chance of an AI-written novel making the *New York Times* bestseller list *for the first time* in the next decade. Likewise, if Maya is a mile from the exit to the maze and there is currently a 20 per cent chance of her reaching the exit within the next hour, there must currently be at least a 20 per cent chance of her travelling at least a mile within the next hour. Here we are implicitly treating the facts that an AI-written novel has never made the *New York Times* bestseller list and that Maya is currently a mile from the exit as having chance 1.[2]

In applying History Fixity, we have to be careful about what it means for a truth to be "about the history of the world up to and including *t*". If you smoked your last ever cigarette yesterday, it does not follow that the current chance of your smoking a cigarette today is zero: the fact that you smoked your last ever cigarette yesterday is not, in the relevant sense, "entirely about" the history of the world up to and including the present. But we seem to assume that when a proposition has a certain chance at a certain time, the fact that it has that chance at that time counts as a fact "about" that time (and about any interval including that time), in the sense relevant for History Fixity. For example: if there is no chance that the chance of Maya's escaping the maze will change between now and noon, and the chance of her escaping is 50 per cent now, it seems obvious that there is no chance of the chance being anything other than 50 per cent at noon. We do not countenance the possibility that although in fact the current chance of escape is 50 per cent, there is a 20 per cent chance that it is 60 per cent and will remain 60 per cent for the next hour. Rather, we treat the truth about Maya's current chance of escape as fixed, in the same way as the truth about Maya's current location. Thus, while we rarely directly consider questions about the current chances of propositions about the present and past (including the present and past chances), our commitment to their having current chance 1 if true emerges in our reasoning about the chances of propositions about how the future relates to the present and past (including the present and past chances).

In the literature on objective chance, a different argument for Chance Fixity has been widely discussed (and widely endorsed, at least as a good approximation). This argument appeals to the Principal Principle (Lewis 1980*a*), which concerns

[2] If we want to interpret chance-talk in a way that is useful even if determinism is true, we cannot treat *all* the truths about the microphysical state of the universe in the past, to infinitely many decimal places, as having chance 1: we will have to treat the domain of truths that automatically get chance 1 as restricted to "macroscopic" facts, or more generally to the facts about some limited domain of investigation. For defences of the propriety of this move, see Loewer 2001; Eagle 2011; Dorr 2016*a*; Gallow 2021.

the way in which rational people take propositions about chance into account in their credences (degrees of confidence). At a first pass, this principle says that one's rational *prior* credence in p, conditional on the chance of p being within a certain closed interval at t, should itself be within that interval if it is well defined.[3] This first-pass version isn't strong enough to recover many of the intuitively appropriate ways of taking chances into account in one's credences: for example, it doesn't tell us that a rational credence that it will rain would be conditional on a conjunction like *the chance at t that it will rain is between* 0.5 *and* 0.9 *and the chance at t that it will snow is between* 0.1 *and* 0.2. But if the first-pass version is correct, surely this conditional prior credence should be between 0.5 and 0.9. Here is a more careful version of the Principal Principle which delivers such results:

Principal Principle If I is any closed interval within $[0, 1]$, and C is a collection of propositions entirely about history up to and including t one of which is *the chance of p at t is in I*, then if one's prior credence that every member of C is true is positive, one's prior credence in p conditional on the proposition that every member of C is true should be in I.

It follows from this that you should *not* have a positive credence in any conjunction of the form *the chance of p at t is within I, but the chance that the chance of p is not within I is at least x*, where I is any closed interval and $x > 0$. For if one did, one's credence in *the chance of p at t is not within I* conditional on this conjunction would have to be $\geq x$ by the Principal Principle (with the interval $[x, 1]$), but it would also have to be 0 given that the probability of a conjunct conditional on a conjunction is 1. This means that if you have positive prior credence in a proposition of the form *the chance of p at t is within I*, then conditional on this proposition being true, you should have prior credence 1 that its chance at t is greater than $1-x$, for each $x > 0$. Given the (controversial) Countable Additivity principle, this comes to the same thing as having prior credence 1 that Chance Fixity holds as regards any given p: i.e. that if the chance at t of p is within I, the chance at t that it is within that interval is 1.[4]

[3] Here we can understand having conditional credence x of p on q in the usual way, namely as being such that x is the ratio of one's credence in $p \wedge q$ to one's credence in q. We'll also talk about credences conditional on collections of propositions: one's credence in p conditional on a collection C is one's credence in p conditional on the single proposition *every member of C is true* (i.e. the conjunction of C).

We count the singleton of any real number as a limiting case of an interval. By stating the principle in a form that also applies to intervals other than singletons, we bypass the worry that one's prior credence that the chance of p is *exactly* x will be zero for almost all x, so that the relevant conditional credence will almost always be ill-defined.

[4] Without Countable Additivity this fails, but we still get certainty that Chance Fixity holds to within any arbitrarily good approximation to 1, which will be sufficient (as §9.2 will explore) for the relevant applications of Chance Fixity.

Even with Countable Additivity, the result falls short of the claim that Chance Fixity itself deserves a prior credence of 1: one might have credence 1 that any given p and x are not a counterexample to Chance Fixity while having positive credence that there are counterexamples lurking somewhere.

However, we do not want here to rely on this argument from the Principal Principle, since the application of the Principal Principle to propositions about particular objects (which are our current concern) raises difficult issues having to do with the variety of "guises" or "modes of presentation" under which we can have credences in propositions about particular objects. Hawthorne and Lasonen-Aarnio (2009) imagine introducing a name 'Lucky' for whoever will win a certain fair lottery: we seem to justifiably be sure that Lucky will win the lottery despite being sure that the current chance that Lucky will win the lottery is low. Analogously "cheesy modes of presentation" might well arise in the present setting: before a table is assembled by some chancy selection of materials, the only ways we have of forming credences in propositions about what that particular table will be like will involve something analogous to singling out Lucky as the lottery winner (e.g. fixing the reference of 'Woody' as 'whatever table will be made'). So it is unclear how many of the cases of interest would fall inside the scope of a suitably qualified version of the Principal Principle.[5]

It is, however, worth noting a related argument that does not raise these worries about modes of presentation, and provides further strong motivation for Chance Fixity over and above our argument from History Fixity. This second argument appeals to a principle which says that current chances "defer to" future chances in the same sense that, according to the Principal Principle, rational prior credence functions defer to chances at any time:

Chance Reflection If I is a closed interval within $[0, 1]$, t and t' are times where t is no later than t', and C is a collection of propositions C which are all entirely about history up to and including t', one of which is *the chance of p at t′ is in I*: if the chance at t that every member of C is true is positive, then the chance at t of p conditional on every member of C being true is in I.[6]

For example: the chance at 11 am of Maya's escaping the maze, conditional on the proposition that the chance of escape will be 60 per cent at noon, is 60 per cent. This is also the chance of escape conditional on the proposition that at noon the chance of escape will be 60 per cent at noon and the chance of subsequently landing a book deal will be 59 per cent. A strong case can be made that this principle is implicit in a great deal of our reasoning about chance processes unfolding over time. Consider, for example, the way we calculate that the current chance of a later

But in a Chance Tolerance Argument where the relevant family of properties only has countably many members, this will not make trouble for the application of Chance Iteration needed for the argument.

[5] The guise-theoretic problems are not confined to guises of *objects*: properties too can be thought of under cheesy guises that make prima facie trouble for the Principal Principle, as can the relation *the chance of p at t is x* itself. Nevertheless, many applications of the Principal Principle seem immune to these worries.

[6] Here conditional chance is defined in terms of the ratio of unconditional chances, parallel to the case of conditional credence.

Tails outcome is 1/4 if there is 50 per cent chance of grabbing a fair coin and a 50 per cent chance of grabbing a double-headed coin. And just as the Principal Principle implies that we should have prior credence 1 that Chance Fixity holds in any particular case, Chance Reflection implies that the *chance* of Chance Fixity holding in any particular case is 1 at any given time.[7]

So, denying Chance Iteration looks unpromising as a way of blocking Chance Tolerance Arguments. But as we already noted, there are many instantiations where the remaining premises and the negation of the conclusion are pretty compelling. Chance Tolerance Puzzles thus present a particularly forceful instance of the general kind of puzzle with which this book is concerned.

Chance Tolerance Puzzles also present two particular challenges for those who want to treat Tolerance Puzzles as grounds for denying Iteration for metaphysical possibility. First, there is a dialectical challenge. Insofar as Chance Iteration is in good standing, they will need to choose some other strategy to escape Chance Tolerance Puzzles, and whatever they say, it will risk undermining their case for the corresponding premise in their argument against Iteration. Second, if their favoured way out of a Chance Tolerance Puzzle is to embrace Chance Hypertolerance, the following plausible principle will force them to accept the corresponding Hypertolerance claim for metaphysical possibility, thereby blocking their argument against Iteration:

Chance/Possibility Whenever *p* has a positive chance at any time, *p* is metaphysically possible.

Is Chance/Possibility plausible? It is standard to distinguish two kinds of interpretations of 'chance': the "objective" interpretations we have been focusing on, on which chance facts are a central topic investigated by physics and other sciences, and "epistemic" interpretations, on which the chances depend on the evidence available to some contextually selected agent or group of agents. Chance/Possibility would certainly be extremely tendentious if it were interpreted as having to do with epistemic chance. When the relevant agents are sufficiently ignorant, there may arguably be a positive epistemic chance of gold atoms having only fifty-seven protons, of Mars being Venus, of there being a barber in a village who shaves all and only those in that village who do not shave themselves, and so on. But Tolerance Arguments conducted in terms of *having positive epistemic chance* are not our focus

[7] Kment (2018) shows how certain somewhat weaker principles in a similar spirit to Chance Reflection, having to do with the *expectations* at earlier times of chances at later times, can be used to do similar work to Chance Fixity in generating Tolerance Arguments.

here.[8] Once we firmly fix our sights on objective chance, Chance/Possibility strikes us as overwhelmingly plausible.[9, 10]

Kment (2018) accepts Chance/Possibility (equivalent to his "Modal Chance Principle"), and uses it to argue against Chandler and Salmón's Iteration-denying approach to Tolerance Puzzles. His strategy is to argue from certain Tolerance claims involving metaphysical possibility to the corresponding claims involving positive chance, by invoking certain weak sufficiency-of-origins principles (much weaker than those discussed in Chapter 5). But even setting those principles aside, the package that systematically denies Tolerance claims in Tolerance Puzzles involving chance while accepting the corresponding claims of metaphysical possibility strikes us as too implausible to merit an extended treatment. As we see it, the central motivations for the metaphysical-possibility Tolerance premises are a range of ordinary judgements, which in many of these cases of interest provide an even more direct motivation for the corresponding chance-theoretic Tolerance premises. Many such judgements directly concern the positive chances of scenarios in which the same objects are around but have somewhat different properties.[11]

Responding to Kment, Salmón (2019) seems to reject Chance/Possibility. But it would be more accurate to characterize him as rejecting the very notion of objective chance. The only notion of chance he admits is epistemic, relativized to an agent (or "epistemic situation") rather than to a time. And his counterexamples

[8] One obvious complication those arguments raise is that it is dubious whether the relevant epistemic operators conform to our basic modal logic, for reasons explored in Chapter 8.

[9] If an argument is wanted, some might find it helpful to derive Chance/Possibility from the following two premises: (i) whenever p has positive chance at some time, p is nomically possible; (ii) whenever p is nomically possible p is metaphysically possible. We could also argue for Chance/Possibility from the premise (defended in Chapter 8) that metaphysical necessity is the broadest necessity, together with the premise that *having chance 1 at t* is a necessity operation.

[10] Chance/Possibility can also be used to further support Chance Fixity, at least if we also accept Ancestral Microphysical Supervenience (a thesis whose plausibility we urged in Chapter 6). Given Ancestral Microphysical Supervenience, every proposition p is ancestrally necessarily equivalent to some microphysical proposition q; given the necessary truth of Chance/Possibility, there is no chance of p and q differing in their chance at any time. So if any proposition is a counterexample to Chance Fixity, some microphysical proposition is also a counterexample to Chance Fixity. But physicists hope to formulate simple general laws which fix the chances of all microphysical propositions at a given time, given the microphysical history up to that time. It would be very strange to think that while these simple general laws are in fact true, they have a positive (and indeed quite substantial!) chance of being false.

[11] It's not always obvious whether an ordinary 'chance' judgement has to do with objective chance or epistemic chance, but it would be quite odd to think that the practice of talking about objective chance is confined to science. And there are plenty of ordinary judgements where an objective interpretation seems overwhelmingly natural, e.g. 'No-one knows just how likely it was on the morning of D-Day that it would be a success, but it certainly had a decent chance of succeeding.' Even setting aside ordinary judgements using the word 'chance', ordinary counterfactual judgements are also highly relevant. A counterfactual like 'If I had spent more time with the sander yesterday this table would have been smaller through the door' fits very poorly with a denial of Chance Tolerance. Plausibly, if the antecedent of a counterfactual has positive chance at a time while its consequent has zero chance at that time, then the counterfactual (at least on its most natural interpretation) also has zero chance at that time.

to Chance/Possibility generally rely on the assumption that when something is epistemically certain (or known), it has chance 1. For example, he thinks that relative to an epistemic situation in which a roulette wheel was spun out of sight of the agent and came up 26, the metaphysically impossible proposition that predicates *being prime* of the number 26 has a positive chance, since there is a positive chance that a prime number came up and it is certain that *that number* is the one that came up. He also argues that since propositions of the form *p iff actually p* have chance 1, any case where a false proposition *p* has a positive chance is also one where the metaphysically impossible proposition that *actually p* has the same positive chance.[12] All of this seems prima facie plausible as regards epistemic chance. But as we see it, Salmón is overlooking an important subject matter which is often, though not always, the target of ordinary uses of 'chance'. As evidence for this, we note that speeches like 'There was a chance that things would not turn out as well as they actually did' and 'There was a chance that no-one who was actually there would be there' are completely fine. We would also point out that chance-talk in the sciences seems to make pervasive use of the probability calculus in ways that are hard to make sense of Salmón's epistemic interpretation of 'chance'. For example, it is integral to most applications of the probability calculus that for each proposition there is at most *one* real number that is the chance of that proposition (at a given time), whereas on Salmón's picture, a single proposition will often have many different chances, corresponding to the different "guises" under which the relevant agents may entertain it.[13]

9.2 Chance Fixity, Humeanism, and Approximation

Opponents of Chance Iteration may take comfort from the fact that Chance Fixity has been rejected, on grounds entirely unrelated to Tolerance Puzzles, by

[12] Salmón derives this premise from the thesis that any proposition *p* is "logically equivalent" to *actually p*, and logically equivalent propositions have equal chances. It's true enough that in standard axiomatizations of propositional chance, truths of propositional logic get chance 1. But extrapolating from that to a principle that requires *p* and *actually p* to be interchangeable in a joint theory of chance and actuality seems as dubious as making an analogous extrapolation from standard necessitation principles in modal logic. We are also resistant to treating logical equivalence as a relation between propositions, for reasons explained in Chapter 8.

[13] Apart from various bald assertions to the effect that all chance is epistemic, and that chance is monotonic with respect to the relevant notion of logical consequence (on which 'actually *p*' is a logical consequence of *p*), Salmón also offers an argument that essentially depends on the thought that while many ordinary dice are fair, and the chance that any pair of fair dice come up on a composite number is exactly 7/12, it is false (or unintelligible) to claim that 7/12 is the *exact* degree to which any pair of ordinary dice are 'causally, physically, or metaphysically predisposed or inclined' to yield a composite number. In response, we think that if 'fair' means "precisely fair", it is not at all obvious that any real dice are fair. Wouldn't a physicist be making a perfectly good point if they showed that the dots on die-faces with higher numbers make those faces slightly lighter or heavier (depending on whether the dots are concavities or blobs of paint), and thus make the ideal of perfect fairness unattainable for real-world dice? Of course, 'fair' may often be used to mean 'roughly fair', just as 'equally tall' is often used to mean something like 'roughly the same height'.

proponents of "Humean" theories of chance. The basic idea of such theories is that chance-facts supervene on facts about the distribution of non-chance-theoretic, "categorical", properties and relations. According to the best-developed picture of how this supervenience works, Lewis's (1994) "best system" analysis, what it takes for a certain function to be the chance function is for it to have an optimal combination of simplicity and "fit", where fit is a matter of assigning high chances to truths.[14] For example, a world where there are billions of particles of a certain sort most of which decay quickly, and where the ones that decay slowly are scattered throughout the history in no special pattern, will be one at which the true chance function treats these events as independent and assigns each particle some high chance short of 1 of decaying quickly. A world where most decay slowly will be one in which the chance of decaying slowly is much higher. But this means that early in the history of the quick-decay worlds, there is some small but nonzero chance that most of the decays will be slow, and hence some small but nonzero chance that the chance function is the one that assigns the particles a high chance of decaying slowly. This conflicts with Chance Fixity: at the early time, in the world where most decays are quick, there is a small but nonzero chance that most decays are slow, and hence a small but nonzero chance that at that very time there is a high chance that most decays are slow.[15]

Examples of a similar sort can be brought to bear directly against Chance Iteration. For a very abstract example, imagine a world that is just a very long but finite sequence of instants that come in two types, '0' and '1'. The pattern is just the sort of random-looking one that would be produced by tossing a fair coin, with one exception: on each of the seventeen occasions when there are are one hundred or more '1's in a row, they are followed by an equal number of '0's. By Humean lights, this pattern would plausibly constitute the chance function's being one on which all sequences (of the relevant length) are initially equiprobable except that those in which one hundred or more '1's fail to be followed by equally many '0's have chance 0. If so, then at any time before the first run of one hundred '1's, there is a positive chance that there will never be a run of one hundred '1's. But if that happened and there were no other interesting patterns, then plausibly the true chance function would be the simpler one on which all the sequences are equiprobable. (Suppose that the sequence is long enough that not having any runs of one hundred '1's is only moderately unlikely.) So, at the relevant early times, although the chance of

[14] Simplicity is understood in terms of, or simply equated with, naturalness (see Dorr 2019: 4.6). As Elga (2014) points out, Lewis's suggestion that the "fit" of a candidate chance function is determined by the chance it ascribes to a proposition true only at the actual world founders on the fact that in all likelihood, such a world-proposition will have chance 0. Elga suggests an alternative account of "fit" on which what matters is assigning high values to *simple* (natural) truths.

[15] This case will also be a counterexample to Chance Reflection, since at the quick-decay worlds, the chance that most of the particles will decay quickly, conditional on the hypothesis that the chance function favours slow decay, is lower than the chance of most of the particles decaying quickly according to that hypothesis.

there being a run of one hundred '1's followed by a '0' followed by another '1' is zero, there is a positive chance of there being a positive chance of this happening.

The failure of principles like Chance Fixity has widely been regarded as a serious cost of Humeanism, by its proponents as well as its detractors. Lewis's response (1994: 486) is that the cost is tolerable, because the required deviations are so tiny as to be negligible. Indeed, when we are dealing with times fairly late in the history of a large universe, where we can be very confident that the earlier micro-history is completely representative by the lights of the true chance function, it is completely in the spirit of Humeanism to say that the current chance of every false proposition about the current chances is *astronomically* small.[16] This means that principles like Chance Fixity are close enough to being true that relying on them in reasoning about ordinary, fairly localized chance processes will never lead us astray.[17]

Against this background, it is important to see that the failures of Chance Fixity required by an Iteration-denying solution to chance-theoretic Tolerance Puzzles will be far more significant than the tiny departures motivated by Humeanism. To illustrate, consider the following toy model of how the chance-facts might play out on an Iteration-denying approach to a Chance Tolerance Puzzle. Suppose we have a circular arrangement of five small steel bars *abcde*, of which we pick three in a row to make a triangle (the musical instrument); each of the collections *abc*, *bcd*, *cde*, *dea*, *eab* has chance 1/5 of being chosen. In fact, *abc* was chosen, and we made a triangle, Steely. Chance-Necessitated Tolerance tells us that there was chance 1 that whichever triangle was made would have a positive chance of being made of each of the three collections that include at least two of its originating bars. Chance Non-hypertolerance tells us that at least one of the five collections is such that there was zero chance of Steely being made from it. For concreteness, let's flesh this out by saying that with chance 1, any triangle made of three bars would have zero chance

[16] This would not be the case on a "finite frequentist" view where, for example, the chance of an atom of a certain sort emitting a certain kind of particle when it decays is just the proportion of all decays of that atoms of that sort that emit that kind of particle (assuming there are only finitely many such decay events). On that view, there are true propositions about the current chances whose current chance is quite low. Even given finite frequentism, it is still plausible that there is currently a high chance that the current chances are *approximately* as they actually are; Lewis (1994: 488) focuses on approximations that can be recovered even given finite frequentism. But finite frequentism is very implausible (see Hájek 1996). And as Lewis (1994: 479) emphasizes, the more adequate "best system" form of Humeanism suggests that we can reasonably be confident that "nature is kind" in such a way that the best system is "robustly" best, meaning that something very unlikely by its own lights would have to happen for its aggregate simplicity-and-fit score to be outmatched by that of some other system. This will license the stronger approximation to Chance Fixity that we focus on in the text. As we will discuss in note 20 below, even the weaker approximation can still generate certain kinds of interesting Tolerance Puzzles.

[17] This attitude towards Chance Fixity is somewhat similar to the attitude towards the more general History Fixity required by the view that determinism is compatible with nontrivial chances (see note 2). Given that the deterministic dynamical laws have chance 1, the *fully detailed* specification of some past state of the world cannot have chance 1. But we can still ascribe chance 1 to a proposition specifying the course of history in enough detail as to rule out all but microscopic departures from actual history, and this will be enough to undergird all ordinary applications of History Fixity.

of being made conditional on only one of those bars being chosen, and chance 1 of being made conditional on two or three of those bars being chosen. Let *p* be the proposition that Steely is made of *cde* or *dea*. *p*'s chance is 0. But conditional on a triangle being made from *bcd*, there was chance 1 that *p*'s chance would be 1/5, since it would have been equal to the chance of a triangle being made of *cde*. For the same reason, *p* had chance 1 of having chance 1/5 conditional on a triangle being made of *eab*. So, although the chance of *p* was 0, it had a 2/5 chance of having chance 1/5. This is dramatically larger than the kinds of exponentially tiny failures of Chance Fixity and Chance Iteration that Humeans must learn to live with. It is far from negligible, and the case in which it arises is an entirely realistic one.

The pattern this example illustrates will be quite pervasive on the natural way of fleshing out an Iteration-denying approach to Chance Tolerance Puzzles. For any given object, there will be a range of properties in the relevant family such that the object has zero chance of instantiating any property not within that range, but has positive chance of instantiating each property within the range. But there will be quite a substantial chance that the object will have a property that is a good bit closer to one of the boundaries of the range than the one it in fact has. And in any such case, there would have been quite a substantial chance of the object's instantiating some property that it actually has no chance of instantiating.

In principle, a proponent of the Chance Iteration-denying strategy could avoid large failures of Chance Fixity by claiming that while (as Chance Tolerance claims) there is a positive chance of the object in question having each of the properties close to the one it in fact instantiates, the chance of this object instantiating any of those properties is minuscule. For example, in the Steely example, one could say that conditional on *bcd* or *eab* being chosen, the chance of Steely rather than some other triangle being made is very small (but not zero). Perhaps making Steely out of *bcd* or *eab* would require some further chancy microphysical parameters to take on certain very specific, and hence unlikely, values. Or perhaps the chance of Steely being made is small even conditional on any way of filling in the microphysical details of a scenario in which a triangle is made of *bcd* or *eab*.[18]

This proposal preserves the letter of Chance Tolerance, but is deeply out of keeping with the ordinary thinking that motivated Chance Tolerance in the first place. For this example and many others, our ordinary thinking supports not merely the judgement that a given object had a *nonzero* chance of instantiating each property close to one it instantiates, but that it had a fairly *substantial* chance of instantiating each of these properties. For example, in the Steely case, to the extent that it seems plausible that Steely had a nonzero chance of being made out of *bcd* or *eab*, it seems plausible that this chance is in the region of 1/5 in each case. Such strengthenings of Tolerance are well motivated in many other cases: for

[18] This latter option might naturally be combined with the denial of Microphysical Supervenience.

example, insofar as there is any pressure to think that a certain flimsy bucket had a nonzero chance of being a thick-walled bucket made with an extra portion of plastic, there is pressure to think that the chance in question was quite substantial, at least roughly equal to the chance of *any* bucket being made of both pieces of plastic.

Once Chance Tolerance is strengthened by replacing 'nonzero' with 'substantial', we can adapt the Chance Tolerance Argument to use a weaker, approximate version of Chance Iteration, which Humeans have no special reason to reject. While we are at it, we can also modify Chance Non-contingency so that it too allows for a small chance of failure rather than requiring the chance of the (strengthened) Tolerance premise to be exactly 1. This also makes the argument harder to resist, since this approximate version of Chance Non-contingency will be easier to motivate by a Security Argument of the sort discussed in §3.3: after all, substantial chances of error are more troubling than tiny chances of error. We can weaken Chance Persistent Closeness in a parallel way, though since Chance Persistent Closeness is already unproblematic in many interesting instantiations, the payoff of this isn't so interesting. The resulting argument is displayed in Figure 9.2. Given standard probabilistic principles for chance, the argument is valid so long as the threshold s (representing 'substantial chance') is greater than the sum of ε, ε', and ε''.[19] In the Steely example, for example, Strong Chance Tolerance is plausible with $s \approx 1/5$, so we have plenty of wiggle room. Persistent Closeness is automatic given the list-like definition of closeness, so even after we make allowance for Humeanism by choosing some tiny ε, we can afford to set $\varepsilon' \approx 1/5$ as well. So if we wanted to block the Security Argument for Loose Chance Non-contingency by sticking to our guns in the face of an acknowledged chance that the very same methods would lead us into error, we would have to learn to

[19] We just need the following theorem of the probability calculus (in addition to chance-1 obeying the basic modal logic):

Chance-MP $\quad \text{Chance}_t(Q) \geq \text{Chance}_t(P \to Q) + \text{Chance}_t(P) - 1.$

This implies a monotonicity schema: when $\vdash_{H_0} P \to Q$, we must have $\text{Chance}_t(Q) \geq \text{Chance}_t(P)$, since $\text{Chance}_t(P \to Q) = 1$.

First, apply MP to Strong Chance Tolerance and Loose Chance Non-contingency to deduce that the chance of Strong Chance Tolerance is at least $1 - \varepsilon'$. When G is close to F, the chance at t that G is close to F is at least $1 - \varepsilon''$ by Loose Chance Persistent Closeness, so by Chance-MP, the chance that G is close to F and Strong Chance Tolerance is true is at least $1 - (\varepsilon' + \varepsilon'')$. By monotonicity, this implies an approximate version of Stepwise Necessitated Tolerance:

$$\forall F \forall G \forall x (G \sim F \to \text{Chance}_t((Kx \wedge Fx) \to \text{Chance}_t(Kx \wedge Gx) > s) \geq 1 - (\varepsilon' + \varepsilon''))$$

We want to get from this to an analogue of Possibility Transfer: if $G \sim F$ and $\text{Chance}_t(Kx \wedge Fx) > s$, then $\text{Chance}_t(Kx \wedge Gx) > s$; Strong Chance Hypertolerance follows from this for the same reason that Hypertolerance follows from Possibility Transfer in the usual Tolerance Argument. So, suppose $G \sim F$ and $\text{Chance}_t(Kx \wedge Fx) > s$. By Chance-MP and the above displayed formula, this implies $\text{Chance}_t(\text{Chance}_t(Kx \wedge Gx) > s) > s - (\varepsilon' + \varepsilon'')$. Since $s > \varepsilon + \varepsilon' + \varepsilon''$ it follows that $\text{Chance}_t(\text{Chance}_t(Kx \wedge Gx) > s) > \varepsilon$. But Loose Chance Iteration says that the only way this can be true is if $\text{Chance}_t(Kx \wedge Gx) > s$; so we are done.

Strong Chance Tolerance For every K object x and properties F and G such that G is close to F, the chance at t that x is K and G is greater than s.

Loose Chance Non-contingency If Strong Chance Tolerance is true, its chance at t is $\geq 1 - \varepsilon'$.

Loose Chance Iteration For any proposition p, if there is a chance above ε at t that p's chance at t is above s, then p's chance at t is above s.

Loose Chance Persistent Closeness When properties are close, the chance at t that they are close is $\geq 1 - \varepsilon''$.

Strong Chance Hypertolerance For every K object x and properties F and G such that G is ancestrally close to F, the chance at t that x is K and G is greater than s.

Fig. 9.2 Approximate Chance Tolerance Argument.

live not only with *small* chances of such error, but with a chance in the region of 1/5. This does not seem promising.[20]

The Approximate Chance Tolerance Argument thus improves in at least two ways on the original Chance Tolerance Argument: the argument from Humeanism against Chance Iteration does not apply to Loose Chance Iteration, and the security considerations which motivate Chance Non-contingency even more strongly motivate Loose Chance Non-contingency. The only price we pay is the need to strengthen Chance Tolerance; but since any of the ordinary judgements that motivate claiming a nonzero chance for the relevant propositions will pretty much always motivate something considerably stronger, this isn't a serious cost.

In the next section we will turn to a new kind of chance-theoretic puzzle that's similar to, but even more challenging than, a Tolerance Puzzle. This new puzzle also turns on Chance Fixity, though it can (as we will briefly discuss) also be

[20] While Loose Chance Iteration is designed to be acceptable to Humeans, some Humeans will not accept it, either because they accept a view like finite frequentism on which even small variations in future history generally make for differences in the current chances (see note 16 above), or because they are more pessimistic than Lewis is about nature being "kind" in the ways that would make for a substantial current chance of the actually best system not being best. These Humeans might still accept a further weakening of Loose Chance Iteration that replaces the second occurrence of 'above s' with 'not significantly below s': while the *exact* value of the current chance of p may have a low current chance, its approximate value will still have a high chance. This weakening will invalidate the argument as stated, but in many instantiations, it will still enable us to derive weaker but still disturbing conclusions in the vicinity of Strong Chance Hypertolerance. For example, if G is only ten closeness steps away, and 'x is significantly below y' means '$x + \delta < y$', then we will be able to conclude that the chance of x being K and G is greater than $s - 10\delta$.

adapted so as to rely only on a Humean-friendly approximate version of that principle.

9.3 Chance and Robustness

The Approximate Chance Tolerance Argument turned on judgements to the effect that certain objects have substantial chances of existing with certain properties different from those they actually have. A very natural way of generating such judgements is to derive them from the premise that there is an appropriately substantial chance of *some object or other* having the relevant properties, together with the premise that the relevant objects have a high chance of having those properties *conditional* on anything having them. For example, we can derive the claim that the chance of Steely being made of *bcd* was close to 1/5 from the claim that the chance of *bcd* being assembled into a triangle was 1/5 (which is part of the stipulation of the case) in combination with the plausible claim that the chance of *Steely* being made of *bcd*, conditional on some triangle being made of *bcd*, was high. This judgement of high conditional chance is an example of what §6.3 called a *robustness* claim. Robustness claims about a particular object or range of objects say, intuitively, that it would be hard for those objects to *fail* to exist (with roughly their actual properties) without things being substantially different in the relevant underlying respects. Many such claims seem part and parcel of our ordinary modal practices. For example, if we learn that there was a significant chance that the carpenter who made Woody would pick approximately the same parts, make a table, and then spill ink all over it, we will unreflectively conclude that there was a significant chance of *Woody* being covered with ink. Our ordinary picture is emphatically not one on which when a certain plan, materials, etc. are selected, God then selects the table to be made from a giant urn of haecceities in a modally fragile way.

In §6.3 and §6.4, we suggested that such robustness judgements make trouble for various Hypertolerance-accepting packages. The intuitive thought was that, e.g., if there are multiple possible tables that could have been made from a certain hunk of originating matter, they can't *all* have had a high chance of being made conditional on some table being made of that matter. So prima facie it seems that there is a danger of our being in a scenario where a low-chance table is made, but we wrongly take it to be a high-chance table. Our treatment of this argument in Chapter 6 was somewhat impressionistic; now that we have said a bit more about the principles governing reasoning about chance, we can put it on a more rigorous footing.

Let's consider a chance-theoretic elaboration of the earring example from Chapter 6. At a certain t, shortly before stamping a cutting tool down onto a disc to create two semicircular earrings, the artificer gave it a random spin, so that for every $n < 180$, there was at t a 1/180 chance that earrings would soon thereafter be

made centred $n°$ and $n + 180°$ clockwise from the artificer (to the nearest degree). In fact, the cut was made with the tool pointing directly away from the artificer, and the earrings made, Lefty and Righty, were centred at 90° and 270°, respectively.

Say that an earring e, is *robust* just in case, for some n, e was made centred n degrees clockwise from the artificer (to the nearest degree), and there was a chance greater than 0.9, conditional on *some* earring being made centred $n + 1°$ clockwise from the artificer, that that earring would be e.[21] It seems obvious that both Lefty and Righty are robust. Moreover, it doesn't seem that we were just *lucky* to get a pair of robust earrings. The form of thinking that leads us to believe that we got robust earrings doesn't seem like it had any chance of leading to error. So, insofar as we are convinced by the Security Argument that provides the most solid motivation for Non-contingency in Tolerance Arguments, we will on similar grounds want to think that there was no chance at t of non-robust earrings being produced.

But now we are in trouble, assuming Chance Fixity. Say that a number n is "Lefty-friendly" iff the chance of Lefty being an earring centred at $n°$ conditional on there being an earring centred at $n°$ was greater than 0.9. (Thus Lefty is robust iff for some n, it is centred at n while $n+1$ is Lefty-friendly.) Suppose for reductio that for some $n < 360$, n is Lefty-friendly but $n+1$ is not. Given Chance Fixity, the truth about which numbers are Lefty-friendly has chance 1, so there was no chance that $n + 1$ would be Lefty-friendly. But since there was a positive chance of Lefty being an earring centred at $n°$, it follows that there was a positive chance of Lefty being a non-robust earring. Given the premise that there was no chance of any non-robust earrings being produced, we can conclude by reductio that $n + 1$ is Lefty-friendly whenever n is. Since Lefty is robust and centred at 90, 91 is Lefty-friendly; hence (by mathematical induction) every number from 91 to 359 is Lefty-friendly. In particular, 271 is Lefty-friendly. But this is absurd! The chance of an earring being made centred at 271° conditional on an earring being made centred at 91° is 1. So if both 91 and 271 are Lefty-friendly, that would mean that conditional on an earring being made centred at 91°, the chance that Lefty is an earring centred at 91° and the chance that Lefty is an earring centred at 271° are both > 0.9. This is inconsistent. When two propositions both have chance > 0.9 conditional on some third proposition, their conjunction must have chance > 0.8 conditional on that proposition. But the proposition that Lefty is both an earring centred at 91° and an earring centred at 271° is manifestly impossible, and thus must have chance 0, both unconditionally and conditional on any positive-chance proposition.

We can abstract from the above discussion a general argument-pattern that can be instantiated with an arbitrary closeness relation, an arbitrary predicate K in

[21] This is obviously only a necessary condition for the more intuitive notion of robustness: for example, the latter would require robustness in the anti-clockwise direction too. Indeed, it is tempting to think that there was *zero* chance of Lefty and Righty not being made (or not being made in their approximate actual positions, or not being earrings), conditional on the cut being made within a few degrees of its actual location; but the weaker claim is enough for our purposes.

Robustness For every K object x and properties F and G: if Fx and G is close to F, then $\text{Chance}_t(Kx \wedge Gx \mid \exists y(Ky \wedge Gy)) > h$.

Non-contingent Robustness If Robustness is true then its chance at t is 1.

Robustness Iteration For any propositions p and q, if there is a nonzero chance at t that the chance at t of q conditional on p is $> h$, then the chance at t of q conditional on p is $> h$.

Chance Persistent Closeness Whenever G is close to F, the chance at t that G is close to F is 1.

Hyperrobustness For every K object x and properties F and G: if Fx and G is ancestrally close to F, then $\text{Chance}_t(Kx \wedge Gx \mid \exists y(Ky \wedge Gy)) > h$.

Fig. 9.3 Robustness Argument.

place of 'earring', and an arbitrary threshold h in place of 0.9. The result, displayed in Figure 9.3, is structurally very like a Tolerance Argument. The schema can be instantiated with any closeness relation, kind-term K, and threshold h. Setting $h = 0$ gives something tantamount to the original Chance Tolerance Argument. With higher values of h, the premises do not *automatically* inherit the plausibility of the corresponding premises of the Chance Tolerance Argument, but for many choices of K and closeness relation they can be supported on similar grounds. And Robustness Iteration, just like Chance Iteration, is an obvious consequence of Chance Fixity.

Although *the chance is* $> h$ *that* obviously fails to play the role of $\Box$ in our basic modal logic, the argument can still be shown to be valid given that chances at a time obey some very elementary principles of probability theory. We have three intermediate lemmas:

Necessitated Robustness The chance of Robustness at t is 1.

Stepwise Necessitated Robustness For any F and G such that G is close to F, and any object x: there is chance 1 at t that $(Kx \wedge Fx) \rightarrow \text{Chance}_t(Kx \wedge Gx \mid \exists y(Ky \wedge Gy)) > h$.

Robustness Transfer For any F and G such that G is close to F, and any object x: if $\text{Chance}_t(Kx \wedge Fx \mid \exists y(Ky \wedge Fy)) > h$, then $\text{Chance}_t(Kx \wedge Gx \mid \exists y(Ky \wedge Gy)) > h$.

Necessitated Robustness follows by Modus Ponens from Robustness and Non-contingent Robustness. Stepwise Necessitated Robustness follows from Chance

Persistent Closeness and Non-contingent Robustness by the principle that when $\vdash_{H_0} (P \wedge Q) \rightarrow R$, if $\text{Chance}_t(P) > h$ and $\text{Chance}_t(Q) = 1$ then $\text{Chance}_t(R) > h$. Robustness Transfer follows from Stepwise Necessitated Robustness and Robustness Iteration by the principle that if $\text{Chance}_t(P \rightarrow R) = 1$ and $\text{Chance}_t(P \mid Q) > h$, $\text{Chance}_t(R \mid Q) > h$.[22] And finally Robustness Transfer implies Hyperrobustness for the same reason that Possibility Transfer implies Hypertolerance.

To assimilate the argument about the earrings to this schema, take K to be 'earring' and define 'G is close to F' as 'for some n, F is *being made centred n^o clockwise from the artificer to the nearest degree* and G is *being made centred $n + 1^\circ$ clockwise from the artificer to the nearest degree*'. In that particular instantiation, a knock-down refutation of Hyperrobustness was available, using the fact that $h > 1/2$ and there are two properties G and G' both of which are ancestrally close to a property x has such that there is zero chance that $\exists x(Kx \wedge Gx \wedge G'x)$ but a positive chance that $\exists x(Kx \wedge Gx) \wedge \exists x(Kx \wedge G'x)$. Similar grounds for rejecting Hyperrobustness will be available in many other cases where the family of properties in question is such that its members are pairwise incompatible, but there is a chance of several of them being instantiated: for example, the family of five properties that featured in the Bucket Argument (§2.1), or the family of properties of the form *being originally composed by C* where C is a collection of atoms.

Humean reasons for regarding Chance Fixity as a mere approximation will motivate rejecting Robustness Iteration in the above argument. But just as we did for the Chance Tolerance Argument in the previous section, we can adapt the Robustness Argument to rely only a weaker, approximate version of Robustness Iteration which Humeans have no special reason to reject, that replaces 'nonzero' with 'greater than ε' for some small ε. We can further sharpen the argument by making analogous weakenings to Non-contingent Robustness and Chance Persistent Closeness, replacing '$\ldots = 1$' in these two premises with '$\geq 1 - \varepsilon''$' and '$\geq 1 - \varepsilon''$', respectively. The compensating price we have to pay to get the argument still to be valid with these weakenings is that each relevant property F—i.e. every property F which is close to some property—has more than a tiny chance of being

[22] Suppose G is close to F and $\text{Chance}_t(Kx \wedge Fx \mid \exists y(Ky \wedge Fy)) > h$. By Stepwise Necessitated Robustness we have

$$\text{Chance}_t(\text{Chance}_t(Kx \wedge Gx \mid \exists y(Ky \wedge Gy)) > h \mid Kx \wedge Fx) = 1$$

So setting $P = Kx \wedge Fx$, $Q = \exists y(Ky \wedge Fy)$, and $R = \text{Chance}_t(Kx \wedge Gx \mid \exists y(Ky \wedge Gy)) > h$ in the principle, we have

$$\text{Chance}_t(\text{Chance}_t(Kx \wedge Gx \mid \exists y(Ky \wedge Gy)) > h \mid \exists y(Ky \wedge Fy)) > h$$

and hence by the definition of conditional chance,

$$\text{Chance}_t(\text{Chance}_t(Kx \wedge Gx \mid \exists y(Ky \wedge Gy)) > h) > 0$$

But by Robustness Iteration, this can only be true if $\text{Chance}_t(Kx \wedge Gx \mid \exists y(Ky \wedge Gy)) > h$.

instantiated by a *K* object. This is obviously true in the earring case, where each of the 360 relevant properties has chance 1/180 of being instantiated by an earring, and it will be similarly unproblematic in many other instantiations. The smaller our $\varepsilon, \varepsilon'$ and ε'', and the larger our h, the lower we can set the threshold for being "tiny": see the footnote for details.[23]

Robustness Arguments present a forceful puzzle, with fewer avenues for escape than Chance Tolerance Arguments. For in many Robustness Arguments (like the one with the earrings), Hyperrobustness is demonstrably false. By contrast, though Chance Hypertolerance is quite counterintuitive in many cases, it is not demonstrably false, and might have seemed like the best option overall (either in conjunction with the metaphysical-possibility version of Hypertolerance, or in conjunction with the denial of Chance/Possibility).

Robustness Arguments also shed dialectical light on the corresponding Chance Tolerance Arguments, since the considerations that support each premise in the latter can naturally be adapted to support the corresponding premise in the former. Let's consider the premises one by one, using the earring example.

(i) Chance Tolerance in that argument says both earrings had a positive chance of being made a degree clockwise from where they were in fact made. The ordinary patterns of judgement that support this suggest that the chance in question is not just positive, but close to 1/180, i.e. to the chance of the cut being made 1° clockwise from its actual location, as required by Robustness. There is just no interesting motivation for a picture on which using slightly different materials would *very likely* have led to the creation of different earrings, although there is a small chance of getting the same earrings.

(ii) The best motivation we have found for Chance Non-contingency comes from the Security Argument which we explored in Chapter 3 (where we already found a chance-theoretic sharpening helpful). This argument carries over

[23] Let the threshold for 'non-tiny' be n, so our extra "Non-tininess" premise says that $\forall F \forall G(G \sim F \rightarrow Ch_t(\exists x(Kx \wedge Fx)) > n)$. For the argument for Hyperrobustness to go through, our various thresholds need to satisfy the following condition: $hn > \varepsilon + \varepsilon' + \varepsilon''$. For the earrings instantiation, Non-tininess is true by stipulation with $n = 1/180$ and $h = 0.9$, and we can take $\varepsilon'' = 0$ since Chance Persistent Closeness is unproblematic, so ε and ε' don't even have to be astronomically small—their sum only needs to be less than 1/200. The argument works as before up to Stepwise Necessitated Robustness, which we now get in an approximate version:

$$\forall x \forall F \forall G(G \sim F \rightarrow \text{Chance}_t((Kx \wedge Fx) \rightarrow \text{Chance}_t(Kx \wedge Gx | \exists y(Ky \wedge Gy)) > h) > 1 - (\varepsilon' + \varepsilon''))$$

To derive Robustness Transfer, suppose $G \sim F$ and $\text{Chance}_t(Kx \wedge Fx \mid \exists y(Ky \wedge Fy)) > h$. By Non-tininess, $\text{Chance}_t(\exists y(Ky \wedge Fy)) > n$, so $\text{Chance}_t(Kx \wedge Fx) > hn$. By probability theory, whenever $\text{Chance}_t(P) > u$ and $\text{Chance}_t(P \rightarrow Q) > v$, $\text{Chance}_t(Q) > u + v - 1$. Setting $P = Kx \wedge Fx$, $Q = \text{Chance}_t(Kx \wedge Gx | \exists y(Ky \wedge Gy)) > h$, $u = hn$ and $v = 1 - \varepsilon'$, we get $\text{Chance}_t(\text{Chance}_t(Kx \wedge Gx | \exists y(Ky \wedge Gy)) > h) > hn - (\varepsilon' + \varepsilon'')$. Since $hn > \varepsilon + \varepsilon' + \varepsilon''$, this implies $\text{Chance}_t(\text{Chance}_t(Kx \wedge Gx | \exists y(Ky \wedge Gy)) > h) > \varepsilon$. By the approximate version of Robustness Iteration, this can only be true if $\text{Chance}_t(Kx \wedge Gx | \exists y(Ky \wedge Gy)) > h$, completing the proof of Robustness Transfer. Hyperrobustness follows as before.

straightforwardly to Non-contingent Robustness: a picture where we were just lucky not to be mistaken in our Robustness judgements is no more appetizing than one where we were just lucky not to be mistaken in our Tolerance judgements. And since Robustness is a strengthening of Chance Tolerance, the chance of Robustness being false is at least as high as the chance of Chance Tolerance being false. Given an appropriate "Chance Independence" premise (see §3.3), this will translate into a substantial chance of our mistakenly believing Robustness.

(iii) As we have discussed, the central motivations for Chance Iteration involve certain general patterns in our reasoning about chances which also entail Chance Fixity, and thus Robustness Iteration. Moreover, it is enough for both the Chance Tolerance and Robustness arguments that Chance Fixity be *approximately* true in the way that Humeanism permits.

(iv) Chance Persistent Closeness plays exactly the same role in both arguments.

But since the premises of the Robustness Argument imply Hyperrobustness claims which are, in many instantiations, obviously untenable, at least one of them must be resisted, thereby undermining the case for the corresponding premise of the Chance Tolerance Argument.

Robustness Arguments also carry important dialectical lessons for thinking about Tolerance Puzzles that aren't about chance, at least insofar as one's response to them involves keeping Chance Fixity (at least as a good approximation). Since denying Persistent Closeness has no prospect of providing a general solution, there are really just two options: denying Robustness and denying Non-contingent Robustness. But it is hard to see how the panoply of ordinary judgements that support Robustness premises has lesser standing than the panoply of similar judgements that support Tolerance premises: the only option there seems to be the kind of radical error theory which we already discarded in the Introduction. So there is a lot of pressure to reject Non-contingent Robustness. But it is natural to expect that any tenable way of doing this will justify a corresponding rejection of Non-contingency in our central Tolerance Arguments. After all, the primary obstacle to be overcome in both cases is the Security Argument, which works in more or less the same way in each setting. In Chapter 11 we will provide some tools for escaping this argument, which do indeed undercut both Non-contingency and Non-contingent Robustness. But first, we will take a chapter to discuss an influential cluster of ideas about modality which has not come up in our discussion so far, although it has seemed to many that it would offer some distinctive sort of help with Tolerance Puzzles.

10
Tolerance and Counterpart Theory

An important Iteration-denier whose view we have not yet explicitly engaged with is David Lewis. Lewis's view is developed within the context of his distinctive "counterpart-theoretic" account of metaphysical modality (Lewis 1968). This account has been quite influential since it was first unveiled; many authors who do not share the theoretical concerns that led Lewis to introduce counterpart theory have nevertheless seen the view as offering some distinctive and attractive strategy for defusing Tolerance Puzzles. For example, Stalnaker (1986: 117) suggests that a central benefit of "the counterpart move" is its capacity to dissolve various modal puzzles, by replacing 'the restrictive cross-world identity relation with a more flexible relation that permits intransitivities and asymmetries'. In this chapter we will consider the possible bearing of counterpart theory on our puzzles. We will conclude that when the view is developed in a way that avoids unpalatable failures of our basic modal logic, it doesn't really change the relative appeal of any of the options we have considered for handling Tolerance Puzzles, Iteration-denial included.

10.1 Counterpart Theory

At a first pass, counterpart theory says that there is a certain binary relation—*counterparthood*—such that for any property F (or at least, any *qualitative* property F), being metaphysically possibly F is having an F counterpart.[1] So understood, the view has many startling consequences: for example, that if it is metaphysically possible for Juhani to ride on a talking donkey, then Juhani has a

[1] For characterizations along these lines see Merricks 2003*b* and Sider, 2002 (194), 2009. Neither Merricks nor Sider restricts the claim to qualitative F, and nor does the word 'qualitative' occur in Lewis's early expositions of counterpart theory. But the restriction provides a natural response to a pressing objection to the unrestricted version. John is possibly Norwegian, but he has no Norwegian counterparts, since none of his counterparts are from Norway and everyone Norwegian is from Norway. (He does plausibly have counterparts from *counterparts* of Norway, but they are not Norwegian.) The qualitativeness restriction bypasses such concerns, since being Norwegian is plausibly not qualitative. While Lewis's own exposition of counterpart theory, in the form of a translation-scheme, does not involve any assertions of higher-order identity, similar problems arise if we apply the scheme to a language with syntactically simple but non-qualitative predicates like 'Norwegian', since the translation of $\Diamond Fx$ logically implies $\exists y(Cyx \wedge Fy)$. Dorr (unpublished *a*) argues that the qualitativeness-based formulation is preferable to Lewis's more complex official formulation.

The Bounds of Possibility: Puzzles of Modal Variation. Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri, Oxford University Press. © Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri 2021.
DOI: 10.1093/oso/9780192846655.003.0011

counterpart who rides on a talking donkey.[2] Since there have to *be* talking donkeys for Juhani to have a counterpart who rides on one, the view will thus attract "incredulous stares" (Lewis 1986*a*: §2.8) from those who are highly confident that it is metaphysically possible for Juhani to ride on a talking donkey, but far from highly confident that there are any talking donkeys. That covers most philosophers. Nevertheless, the view is an important one to think about, since it has struck some philosophers that if it were credible, it would help in some distinctive way with Tolerance Puzzles. Moreover, some philosophers have thought that they could keep something valuable in Lewis's theory while giving up the aspects responsible for the incredulous stares, although there is little agreement on the nature of the theory that might result from this process.[3] So for now, let's suspend disbelief as regards the existence of talking donkeys and so forth, and see what can be done with the view in the form in which Lewis presented it.

If one were focusing entirely on Tolerance Puzzles expressed in the language of possible worlds, with premises like 'For every *K* object *x* and property *G* close to one instantiated by *x*, there is a possible world at which *x* is *G*', it might seem obvious that counterpart theory would dissolve the puzzles. For one of Lewis's axioms says that each object is "in" at most one world, and that might seem to rule out there being any possible world *at which* some object has any property that it in fact lacks (see, e.g., Mackie and Jago 2017). This perhaps explains why counterpart theory comes up so often in discussions of the puzzles. But as we have emphasized from the outset, Tolerance Puzzles can be raised without mentioning worlds at all, and it is not so obvious how, or whether, it is relevant to the puzzles as we have set them up.

Some authors (including Kripke (1972: note 18) and Salmón (Salmón 1981: §41)) have understood counterpart theory as a programme for understanding certain modal claims about particular individuals as mere loose talk. On this interpretation, Juhani's having a counterpart who rides a talking donkey only secures that the *loose* truth of 'Juhani could have ridden a talking donkey'. For that sentence

[2] We are assuming for now that *riding a talking donkey* is a qualitative property.

[3] For example, Hazen (1979) says that that Lewis's proposals constitute a 'semantic theory', 'compatible with any of a wide range of metaphysical theories about possible worlds and the nature of possibility'. Sider (2009: 3) says that 'modal realism . . . is not obligatory for counterpart theorists'. Others who envisage disentangling counterpart theory from Lewis's modal realism include Stalnaker (1986), Forbes (1985), and Fara (2008).

One strategy for keeping Lewis's formal theory while avoiding its startling consequences would be to insist that even the *unrestricted* quantifiers of English correspond to the quantifiers in the "input" language to which Lewis applies his translation scheme, rather than to the lookalike quantifiers in the "output" language. On this very un-Lewisian approach, the truth of "output" sentences like $\exists x(\text{Talking}(x) \wedge \text{Donkey}(x))$ is compatible with there not being any talking donkeys (even unrestrictedly speaking), and likewise with there not being any physical universes other than the one we are part of. This view raises pressing questions what such sentences *do* mean, which might lead one to seek a reverse translation that maps every sentence in the "output" language back to one in the already-understood "input" language. This is what Sider has in mind when he speaks of "strategies for reducing possibilia".

to be *literally* true, Juhani would need to be "in" some world where is a talking donkey. Lewis's axiom that nothing is "in" multiple worlds will thus be taken to imply that literally speaking, it is not possible for anything to exist while having different properties.[4] Understood in this way, counterpart theory is a way of fleshing out the view we discussed in the Introduction, according to which the Tolerance premises that drive our puzzles are strictly speaking false but loosely speaking true. But as we pointed out there, the claim of strict falsehood is both implausible on its face, and not really helpful with the puzzles, since the allegedly "loose" mode of interpretation can also be applied to the remaining premises and the conclusion of any Tolerance Argument. It is also worth noting that there is no hint of this attitude to be found in Lewis. On his view, it is just a mistake to think that the literal truth of ordinary possibility claims requires objects to be "in" multiple worlds.[5]

A more interesting way in which counterpart theory might be thought to help with Tolerance Puzzles involves its potential vindication of Iteration-denial. At least on first-pass characterization of counterpart theory with which we began, failures of transitivity in the counterpart relation will lead to failures of the Iteration schema $\Diamond\Diamond P \rightarrow \Diamond P$.[6] More specifically: a sufficient condition for Iteration to fail is for there to be some items x, y, and z which are, respectively, the unique instances of some qualitative properties F, G, H, such that y is a counterpart of z and z is a counterpart of y, but z is not a counterpart of x. For in that case it will not be possible for x to be H (since x doesn't have any H counterparts), but it will be possibly possible for x to be H (since x's counterparts include y, which has the qualitative property *being possibly H* since its counterparts include z).

As we characterized it, counterpart theory is neutral as regards *which* relation plays the role of counterparthood. Lewis (1968) is less neutral: he proposes that x is a counterpart of y just in case x resembles y sufficiently closely, and nothing else in x's world resembles y (too much) more closely. As Lewis points out, this account makes it implausible that counterparthood is transitive. However we interpret 'resemble sufficiently closely', there will plausibly be cases where x resembles y sufficiently closely, and y resembles z sufficiently closely, but x does

[4] On the other hand, Kripke's (1972: note 18) envisaged use of counterpart theory does not rely on any such general axiom.

[5] Lewis does allow for pervasive context-sensitivity in modal claims correlated with context-sensitivity in 'counterpart'. Perhaps he would admit extreme contexts where 'counterpart' expresses a relation that no object bears to any object other than itself: this would generate an interpretation of 'possible' on which it is impossible for anything to have any qualitative property that it does not in fact have. But Lewis never singles out these contexts as having any special status, or suggests that all other counterpart relations generate interpretations that are nonliteral or otherwise second-rate.

[6] Since counterpart theory is generally developed in a strictly first-order framework, we focus on the schematic version of Iteration rather than the quantified version. If one did extend counterpart theory to higher-order quantification, there would be a question of whether to combine rejection of the schema with endorsement of the quantified version. This would require rejecting higher-order UI, but as we will in the next section, many versions of counterpart theory already induce UI-failure for first-order quantification.

not resemble z sufficiently closely. But if we thought there was a good argument for the truth of Iteration, it would be completely appropriate to turn this around, and conclude that if any relation plays the role of counterparthood, that relation must be transitive, and hence cannot take the specific similarity-theoretic form that Lewis proposes. For instance, one might propose that what plays the role is instead the transitive closure of Lewis's relation. Lewis doesn't seem to have any distinctive positive argument that the relation that plays the role of counterparthood has the similarity-theoretic form he favours. Of course, Tolerance Puzzles can be run as arguments against Iteration, and thus as arguments for the thesis that the relation that plays the counterparthood role is not transitive (see Lewis 1986*a*: 245–6). But the abstract structure of counterpart theory doesn't shift the balance of these arguments in any obvious way.

The dialectic with respect to Coincidence Puzzles (Chapter 4) is in some respects similar. Given Lewis's 1968 version of the theory, for the Necessity of Distinctness (ND: $\forall x \forall y(x \neq y \rightarrow \Box x \neq y)$) to be true, it would have to be true that no two objects in the actual world have a shared counterpart. And this is false on his similarity-theoretic account of counterparthood. But if one thought there was a good argument for ND, it would be completely appropriate to turn this around and conclude that if the translation-scheme is correct, the relation that plays the counterpart role is not Lewis's similarity-theoretic one, but something else that behaves in the way required for ND to come out true. Moreover, in later work Lewis (1986*a*: §4.4), following Hazen (1979), describes a variant of the original translation-scheme in which the translations of sentences where multiple free variables under under one modal operator involve a counterpart relation whose relata are ordered n-tuples. This theory makes it much easier for ND to be true: we need only claim that every counterpart of a pair of distinct objects is itself a pair of distinct objects. Of course, there are also independent arguments against ND, including the Coincidence Arguments themselves. But as with Iteration-denial, Lewis's counterpart theory in the abstract (as opposed to his specific proposal about the role of similarity) does not seem to shift the balance of these arguments in any decisive way.[7]

But we are getting ahead of ourselves. Lewis's overall package commits him not only to failures of Iteration, but to failures of (the first-order fragment of) the basic

[7] Sider (1996, 2002) has proposed a *temporal* counterpart theory, which stands to the tense operators as Lewis's theory stands to metaphysical necessity and possibility. (Similar ideas are also articulated in Hawley 2001.) Sider fills in the abstract structure of this view with some proposals about the nature of the relation that plays the counterparthood role that lead to some highly revisionary consequences with respect to standard tense logic. For example, intransitivities in the temporal counterpart relation induce the failure of the tense-logical Iteration principles (see note 4 in Chapter 7). Moreover, the oddities in Lewis's modal logic that we will discuss in §10.2 will carry over to Sider's tense logic. But just as there are versions of modal counterpart theory that are logically orthodox and preserve Iteration, there are versions of temporal counterpart theory that preserve standard tense logic: see Dorr unpublished *a* for such a theory.

modal logic which has so far been a fixed point in our discussion of Tolerance and Coincidence Puzzles. Other versions of counterpart theory have been similarly logically unorthodox, though sometimes in different respects. In the next section, we explain how such unorthodoxies arise, arguing that the derivation of the violations from premises about counterparthood does nothing to make them more palatable. However, there are coherent versions of counterpart theory that preserve the basic logic (while still rejecting Iteration and/or ND); we will turn to those in the final section. Our overall conclusion will be that the widespread impression that counterpart theory offers some distinctive path to defusing Tolerance Puzzles is an illusion.

10.2 Counterpart Theory and the Basic Modal Logic

Lewis's (1968) translation-scheme—unless combined with further postulates about counterparthood which Lewis in fact rejects, on the grounds that they are incompatible with the similarity-theoretic analysis he favours—predicts that certain theorem-schemas of the first-order fragment of our basic modal logic have false instances (in the "input" language with $\Box$ and $\Diamond$). There are two kinds of failures: first, some theorem-schemas from classical *nonmodal* logic have false instances; second, certain standard laws for $\Box$ and $\Diamond$ have false instances.

Let's start with the surprising failures of nonmodal schemas. The following version of Leibniz's Law is a valid schema of our basic nonmodal logic H_0 (and of its first-order fragment):

$\mathbf{LL_{sub}}$ $\qquad a = b \rightarrow P[a/x] \rightarrow P[b/x]$

But as Lewis notes, according to the 1968 translation-scheme, some instances of $\mathrm{LL_{sub}}$ will fail (i.e. have false universal closures) so long as there is an object in the actual world with two counterparts in some world. In fact, the translation of (1) (which is a universal closure of the $\mathrm{LL_{sub}}$-instance where a is x, b is y, and P is $a = x$) is logically equivalent to the claim that there is no such object:

(1) $\qquad \forall x \forall y(x = y \rightarrow \Box x = x \rightarrow \Box x = y)$

A second surprise—not noted by Lewis, but pointed out by Kripke in a letter to Lewis (Kripke 1969)[8]—is that Lewis's view also requires rejecting the following standard principle for quantification:

[8] In the letter Kripke predicts (understandably, but apparently wrongly) that Lewis will accept the failure of UI as a conclusive objection, and suggests a neat modification to Lewis's translation scheme which guarantees both $\mathrm{LL_{sub}}$ and $\mathrm{UI_{sub}}$ without the need for any new postulates about the counterpart relation. We thank Gary Ostertag for bringing this letter to our attention.

$\mathbf{UI_{sub}}$ $\forall xP \rightarrow P[a/x]$

For example (2), which is a universal closure of the UI_{sub}-instance where a is y and P is $\Diamond x \neq y$, is false if something the actual world has two counterparts in some world:

(2) $\forall y(\forall x \Diamond x \neq y \rightarrow \Diamond y \neq y)$

Lewis points out that on his similarity-theoretic account, objects do sometimes have multiple counterparts in a world; his response is simply to deny the validity of LL_{sub} (for the language with modal operators). He would presumably have said the same thing about UI_{sub}.[9]

These failures seem bad to us. it is not that we are fetishistically attached to LL_{sub} and UI_{sub}. As we discussed in Chapter 1, we are open to the idea that these might fail when we add words like 'believe' to the language and translate from English in the obvious way. So we shouldn't dismiss out of hand the suggestion that 'necessarily' and 'possibly' are similarly disruptive. But when we actually consider the English sentences that Lewis will have to reject if he wants to reject the "one counterpart per world" constraint, they seem so obviously true that it seems hard to imagine how any argument for the claim that some things in the actual world have multiple counterparts in the same world could outweigh the case for them. For example, it just seems blindingly obvious that if each person in the room could survive Juhani's being flattened by a steamroller, and Juhani is in the room, then Juhani could survive his being flattened by a steamroller. But for Lewis, this is false if, although no-one survives flattening, each person in the room, including Juhani, has a counterpart in some world who survives the flattening of some counterpart of Juhani in that world.[10]

[9] In H_0, the basic LL and UI schemes are $a = b \rightarrow (Fa \rightarrow Fb)$ and $\forall F \rightarrow Fa$, respectively; LL_{sub} and UI_{sub} are derived using Eβ. Since Lewis's original system doesn't include lambda terms, this derivation doesn't work in it. But there is an natural extension of Lewis's translation-scheme to a language with lambda abstraction on first-order variables: add the clause $[\lambda v_1 \ldots v_n.\varphi P]^w = \lambda v_1 \ldots v_n.[\varphi P]^w$; change the clause for predications to $[Fv_1 \ldots v_n]^w = [F^w]v_1 \ldots v_n$; and let $[F]^w = F$ when F is a simple predicate. When we do this, the failures of LL_{sub} and UI_{sub} will be traced to failures in Eβ. For example, although $\forall x \Box x = x$ is true, $\forall x((\lambda xy.\Box x \neq y)(x, x))$ will be false if something in the actual world has multiple counterparts in some world. LL and UI proper remain valid in this system.

[10] Someone might think that the theoretical payoffs of giving up LL_{sub} in blocking the standard arguments for claims like 'No statue is identical to a lump of clay' (see Chapter 4) are so large as to outweigh these objections: see Lewis 1986*a* (§4.5), and in the temporal case, Sider 1996 (§5). However, as we note below, Lewis's system validates the restriction of LL_{sub} to the case where a doesn't already occur free in P. Thus sentences like $\exists x \exists y(x = y \wedge \Diamond \text{Spherical}(x) \wedge \neg \Diamond \text{Spherical}(y))$ are inconsistent even for Lewis. Likewise '$\exists x \exists y(x =$ y $\wedge$ x will emerge from the left exit of a fission machine $\wedge$ y will not emerge from the left exit of a fission machine)' will be inconsistent in an analogous temporal counterpart theory. But insofar as there is any temptation to give up LL_{sub} to avoid coincidence, it is bound up with the temptation to accept claims like these.

Following Lewis (1986*a*: §4.5), one might license the assertion of such logical inconsistencies by positing context-sensitivity in 'counterpart', and appealing to mid-sentence context-shifts (cf. 'it is

There are other versions of counterpart theory where $\mathrm{LL}_{\mathrm{sub}}$ and $\mathrm{UI}_{\mathrm{sub}}$ (and more generally, orthodoxy in nonmodal logic) can be validated even without adding the 'at most one counterpart per world' axiom. One possibility is to modify Lewis's translation scheme so that each *occurrence* of a variable in the scope of $\Box$ or $\Diamond$ introduces a separate quantifier over counterparts.[11] This restores $\mathrm{LL}_{\mathrm{sub}}$ and $\mathrm{UI}_{\mathrm{sub}}$ at the cost of making the logic of the modal operators even odder: for example, we will no longer have $\forall x \Box x = x$ or even $\forall x \Box(Fx \rightarrow Fx)$.[12] More promisingly, one could adopt the "counterparts of *n*-tuples" variant mentioned above: in that case orderliness in the nonmodal logic can be restored by adding the natural principle that any *n*-tuple whose *i*th and *j*th elements are identical can only have other such *n*-tuples as counterparts.[13]

Apart from the odd failures of nonmodal principles like $\mathrm{LL}_{\mathrm{sub}}$ and $\mathrm{UI}_{\mathrm{sub}}$, Lewis's system also induces failures of a very basic principle of modal logic, namely the K schema:

K $\qquad \Box(P \rightarrow Q) \rightarrow \Box P \rightarrow \Box Q$

As various commentators have noted, K fails in Lewis's logic unless everything in the actual world has a counterpart in every world.[14] Suppose for example that *a* and *b* are both in the actual world, and every counterpart of *a* is *F*, while *b* has some non-*F* counterparts all of which are in worlds where there are no counterparts of *a*. Then given Lewis's translation-scheme, *a* and *b* will be a counterexample to

raining but it is not *raining*'). But this move is perfectly compatible with the claim that $\mathrm{LL}_{\mathrm{sub}}$ is valid in the sense of being true on every *uniform* interpretation (see Dorr 2014*b*). By contrast, Lewis (1971) presents the "multiple counterpart relations" idea in a very different way, that doesn't suggest that they involve any sort of context-shift: rather, he suggests that it is part of the "sense" of a particular term to "select" a specific counterpart relation. In an unpublished document (Lewis 1974), this treatment is formalized and extended from names to variables; Fara (2008) develops a similar idea in detail. In these systems, even the restricted versions of $\mathrm{LL}_{\mathrm{sub}}$ and $\mathrm{UI}_{\mathrm{sub}}$ fail. Indeed even $\forall x \Diamond Fx \rightarrow \forall y \Diamond Fy$ is not a theorem, because the variables *x* and *y* may select different counterpart relations. The failure of variables to be interchangeable strikes us as so radical as to call into question the utility of languages with variables as a tool for regimenting claims about the world. (Can such logics be naturally extended to allow for plural variables? Do they have any natural analogues for variable-free formalisms like combinatory logic?)

[11] For example, $\exists x \Diamond x \neq x$ is translated as $\exists x(Ix@ \wedge \exists w \exists y \exists z(Ww \wedge Iyw \wedge Izw \wedge Cyx \wedge Czx \wedge y \neq z))$.

[12] Forbes (1982, 1985) and Ramachandran (1989, 1990) develop variant counterpart theories which preserve $\mathrm{LL}_{\mathrm{sub}}$ and $\mathrm{UI}_{\mathrm{sub}}$ by taking an "occurrence-based" approach along these lines. Ramachandran's (1989) way of doing it also preserves necessitated tautologies like $\forall x \Box(Fx \rightarrow Fx)$. By adding a special-purpose clause for $=$, Ramachandran (1990) also manages to validate $\forall x \Box x = x$, but still must sacrifice, e.g., $\Box \forall x Rxx \rightarrow \forall x \Box Rxx$ and $\Box \forall x Rxx \rightarrow \forall x \Box(\exists y(y = x) \rightarrow Rxx)$, even for a simple *R* like 'is the same height as'.

[13] This is automatic in the 'counterpart functions' approach sketched by Hazen 1979. A less flexible variant, which Kripke suggests to Lewis in his letter (Kripke 1969), is to modify the 1968 translation-scheme to preserve NI by explicitly requiring the counterparts to respect any identities between the items of which they are counterparts.

[14] Kripke points this out in his letter (Kripke 1969): 'This isn't meant to be a damning objection … Probably you are already aware of this.'

$\forall x \forall y(\Box(Fx \rightarrow Fy) \rightarrow \Box Fx \rightarrow \Box Fy)$, the universal closure of a K-instance.[15] And the failure of K extends even to the principle that a conjunction is necessary only if the conjuncts are: the same *a* and *b* will also be a counterexample to $\forall x \forall y(\Box(Fx \wedge Fy) \rightarrow \Box Fx)$.[16]

By our lights, the failure of K is a decisive reason to either modify the counterpart-theoretic framework, or modify one's conception of the counterpart relation. This wasn't Lewis's attitude: he seems to regard counterpart theory as motivating extreme diffidence with regard to pretty much any general principle articulated in the "language of boxes and diamonds". We find this attitude mystifying. In general, theorists who are trying to analyse some familiar notions in other terms should be extremely concerned to make their analyses compatible with ordinary judgements involving those notions, and especially with deeply entrenched principles of the sort apt to be counted as part of their "logic". Once again, the lesson isn't that counterpart theory is simply untenable, but rather that we should prefer versions of counterpart-theory that preserve the basic modal logic, whether by plugging some suitably constrained account of counterparthood into an old framework, or modifying the framework itself in such a way as to make it more friendly to the basic modal logic (e.g. Dorr unpublished *a*; Yli-Vakkuri unpublished *a*).

The surprising failures of schemas like $\mathrm{LL}_{\mathrm{sub}}$, $\mathrm{UI}_{\mathrm{sub}}$, and K in Lewis's theory are not only problematic on their face, but also don't really open up any interesting avenues for dealing with Tolerance Puzzles. For his view validates restricted versions of all three schemas: instances of K in which every free variable in *P* is also free in *Q*, instances of $\mathrm{LL}_{\mathrm{sub}}$ in which neither *a* nor *b* already occurs free in *P*, and instances of $\mathrm{UI}_{\mathrm{sub}}$ in which *a* does not already occur free in *P* are all fine.[17] And the weaker versions of $\mathrm{UI}_{\mathrm{sub}}$ and K are all we need to run Tolerance Arguments. ($\mathrm{LL}_{\mathrm{sub}}$ is not needed at all.)[18]

[15] On the most straightforward version of the "*n*-tuple" translation-scheme, K will also fail in this case if all the non-*F* counterparts of *b* are in worlds where the ordered pair $\langle a, b \rangle$ has no counterpart.

[16] Since Tolerance Arguments rely on the basic modal logic, those who, like Lewis, are willing to reject it might consider responding to certain Tolerance Puzzles by simply accepting all the claims that are jointly inconsistent in the presence of the basic modal logic (i.e. Iteration, Tolerance, Non-contingency, Persistent Closeness, and Non-hypertolerance). However, it is hard to see how the factors that lead to K-failures in Lewis's theory could provide the basis for a general diagnosis of Tolerance Puzzles, since it seems possible to generate such puzzles even while confining one's attention to worlds in which Lewis would take all the relevant objects (tables, molecules, etc.) to have counterparts.

[17] At least, Lewis's view validates the *universal closure* of any UI-instance in which the substituted term is a variable that isn't already free in *P*. The translations of the open formulae themselves need not be logical consequences of the axioms of counterpart theory: for example, the $\mathrm{UI}_{\mathrm{sub}}$-instance $\forall x Fx \rightarrow Fy$ is translated by $\forall x(Ix@ \rightarrow Fx) \rightarrow Fy$. If we allowed for singular terms that denote objects not in the actual world, we could also get false *closed* $\mathrm{UI}_{\mathrm{sub}}$-instances involving those. See Yli-Vakkuri and Hawthorne forthcoming a for more on these qualifications, which do not affect our current point.

[18] Likewise, the modal formalization of the Coincidence Argument in Chapter 4 requires only the weak versions of K and $\mathrm{UI}_{\mathrm{sub}}$. Of course, the Coincidence Arguments we focus on mostly use counterfactuals, which are not treated in Lewis 1968. Natural extensions of the 1968 system to counterfactuals introduce many further unwelcome surprises (see Yli-Vakkuri and Hawthorne forthcoming a), but do not threaten the idea that for any *P*, the context 'If it were that *P* it would be that . . . ' obeys $\Box^*$N and the weakened version of K.

But other philosophers have devised variants of counterpart theory in which the basic modal logic fails in ways that look more interesting for our purposes, insofar as they actually invalidate some paradigmatic Tolerance Arguments, opening up the prospect of a strategy that simply accepts all the premises while denying the Hypertolerant conclusion, at least for some significant range of Tolerance Arguments.

The views we have in mind are the versions of counterpart theory developed by Forbes (1982, 1985) and Ramachandran (1989, 1990). In them, failures of transitivity in the counterpart relation do not lead to failures of Iteration; indeed $\Diamond\Diamond$ can freely be substituted in any formula with $\Diamond$, and $\Box\Box$ with $\Box$.[19] Moreover, all instances of K, and universal generalizations and necessitations thereof, are theorems. The surprises come instead in the form of failures of necessitations of theorems of classical nonmodal logic. In both systems, intransitivities in the counterpart relation lead to failures of the *necessitation* of $\mathrm{UI}_{\mathrm{sub}}$:

$\Box\mathbf{UI_{sub}}$ $\qquad\qquad$ $\Box(\forall xP \rightarrow P[a/x])$

For example, suppose that every F thing has a G counterpart in some world, and a has an F counterpart in some world, but no G counterparts. Then $\Box\forall x(Fx \rightarrow \Diamond Gx)$ is true, but a is a counterexample to $\forall x\Box(Fx \rightarrow \Diamond Gx)$, so the $\Box\mathrm{UI}_{\mathrm{sub}}$-instance

(3) $\qquad\qquad$ $\Box(\forall x(Fx \rightarrow \Diamond Gx) \rightarrow (Fa \rightarrow \Diamond Ga))$

is also false.[20]

$\Box\mathrm{UI}_{\mathrm{sub}}$-instances like (3) play a crucial role in our quantified Tolerance Arguments, so their failure provides a way of blocking the conclusion without denying any premises. For example, in Forbes's and Ramachandran's logics, we could consistently accept the Necessitated Tolerance claim (4) while denying (5):

(4) Necessarily, any two collections of atoms that overlap by at least 90 per cent are such that any table originally composed of the first could have been a table originally composed of the second.

[19] Forbes regiments counterparthood as a three-place relation *Cxyw*, '*x* is a counterpart of *y* relative to *w*' between two objects and a world. When we speak of "failures of transitivity", we really mean cases where *Cxyw* and *Cyzw′* but not *Cxzw*.

[20] In dramatic contrast with Lewis, Forbes's translation scheme includes replacing each atomic predicate with a new predicate with an extra argument for a world, so that the translation of 'There could be a talking donkey' is along the lines of 'there is a world which something talks-at and is-a-donkey-at'; it thus avoids incredulous stares, but risks attracting stares of incomprehension. So in his case, the assumption required for (3) to fail is really that everything that is F relative to any world has a counterpart relative to some world that is G relative to that world, and a has a counterpart relative to some world that is F relative to that world, but has no counterpart relative to any world that is G relative to that world.

(5) Every table x is such that necessarily, for any two collections of atoms that overlap by at least 90 per cent: if x is a table originally composed of the first, then x could have been a table originally composed of the second.

The argument from (4) (plus Iteration plus Persistent Closeness) to Hypertolerance depends crucially on the inference to (5), so blocking that inference will block the derivation of Hypertolerance.[21,22]

This piece of logical revisionism might seem relatively mild, since $\Box\text{UI}_{\text{sub}}$ has also been widely rejected for a different reason. Many philosophers are *contingentists*, and hold that certain objects are such that it could have been the case that there is nothing identical to them. Where a is such an object—Saul Kripke, say—they will accept

(6) $$\Diamond\neg\exists y(y = a)$$

But these philosophers still all think it necessary that everything is identical to something:

(7) $$\Box\forall x\exists y(y = x)$$

They thus reject the $\Box\text{UI}_{\text{sub}}$-instance

(8) $$\Box(\forall x\exists y(y = x) \rightarrow \exists y(y = a))$$

On similar grounds, they might deny that it is necessary that if everything is made of hydrogen atoms then Kripke is made of hydrogen atoms. But most contingentists endorse a weakening of $\Box\text{UI}_{\text{sub}}$ that inserts a proviso requiring the object in question to be identical to something:

Free $\Box\text{UI}_{\text{sub}}$ $$\Box(\forall xP \rightarrow \exists y(y = a) \rightarrow P[a/x])$$

For example: necessarily, if everything is made of hydrogen atoms then *if Kripke is identical to something*, Kripke is made of hydrogen atoms. But in the systems of

[21] In discussing Tolerance Puzzles, Forbes does not actually avail himself of this option, but relies instead on a nonclassical propositional logic, which we briefly mentioned in Chapter 1 (note 4). Ramachandran (2020), by contrast, embraces this validity-blocking option.

[22] Similarly, $\Box\text{UI}_{\text{sub}}$ (for the relevant operator, e.g. 'If A and B had been combined, then . . . ') is crucial for the validity of Coincidence Arguments as formalized in Chapter 4. Without $\Box\text{UI}_{\text{sub}}$, one could consistently combine the claim that if *A* and *B* had been combined they would not have composed any two distinct buckets with the claim that if *A* and *B* had been combined, Flimsy and Frail would have been distinct buckets both composed of *A* and *B*.

Forbes and Ramachandran, instances of Free $\Box\text{UI}_{\text{sub}}$ can fail for the same reason as instances of $\Box\text{UI}_{\text{sub}}$. For example,

$$(9) \qquad \Box(\forall x(Fx \to \Diamond Gx) \to \exists y(y = a) \to Fa \to \Diamond Ga)$$

will be true under the same assumptions that lead to the truth of (3).[23]

We think that the failure of $\Box\text{UI}_{\text{sub}}$ is already an important cost for contingentism; but the failure of Free $\Box\text{UI}_{\text{sub}}$ strikes us as dramatically worse than any of the main options for handling Tolerance Puzzles that are compatible with the basic modal logic. If it is necessary that every table is tolerant, how on earth could it be possible for Woody to be identical to something, and be a table, and yet not be tolerant?[24]

In this section we have surveyed a variety of ways in which counterpart theory might suggest overturning the logical presuppositions that have provided the backdrop for nearly all of our discussion. The survey has only reinforced our commitment to those presuppositions. All of the departures from orthodoxy require rejecting some *extremely* obvious-looking principles that are deeply entrenched in our ordinary reasoning about identity, quantification, and modality. Only one of the contemplated departures, namely the rejection of Free $\Box\text{UI}_{\text{sub}}$, would actually open up the possibility of a new strategy for solving Tolerance Puzzles or Coincidence Puzzles. But this strategy seems deeply unappealing in comparison with any of the main strategies compatible with our basic modal logic.

However, there are versions of counterpart theory that are compatible with the basic modal logic. One example is the result of adding both the "at most one counterpart per world" and the "at least one counterpart per world" axioms to Lewis's set of axioms (which will of course require revising his proposed similarity-theoretic definition of 'counterpart'). But this is by no means the only option, and there are others that might be considered superior on metaphysical grounds. Exploring the range of options for such theories is beyond the scope of this book:

[23] In Forbes's case, we need the additional assumption that being *F* relative to a world requires being "in the domain" of that world.

[24] Even if one were willing to invalidate the Quantified Tolerance Argument by rejecting Free $\Box\text{UI}_{\text{sub}}$, that would not help with the *de re* version of the argument (see §2.1). This argument (or rather its first-order, schematic version) is valid even in Forbes's and Ramachandran's logics. And since they accept (schematic) Iteration, if they want to accept Tolerance and reject Hypertolerance in this argument, they will have to reject Non-contingency, saying for example that while Woody is in fact tolerant, Woody could have failed to be tolerant. They will thus still need a response to Chapter 3's Security Argument for Non-contingency. Could we easily have falsely believed that Woody was tolerant? And if so, isn't that bad? it is hard to know how to even approach this question without extending Forbes's or Ramachandran's systems to include a belief operator. Whereas in the case of Lewis's translations this is routine, in Forbes's and Ramachandran's systems, the obvious approach will yield strange results: for example, 'Possibly, John believes Juhani is sleepy' will be translated along the lines of 'For some world *w*, John has a counterpart in *w* who believes that Juhani has a sleepy counterpart in *w*.'

see Dorr unpublished *a*.[25] We do, however, want to say something about whether counterpart theory could help with Tolerance Puzzles and Coincidence Puzzles in some way that doesn't turn on a rejection of their logical underpinnings.

10.3 Logically Orthodox Counterpart Theory

Once we have secured the basic modal logic, it will still be an open question whether we should also set things up in such a way as to secure Iteration or ND. The Iteration-denying strategy for handling Tolerance Puzzles is certainly still available in a counterpart-theoretic setting, as is the ND-denying treatment of Coincidence Puzzles. But the arguments for ND which we discussed in Chapter 4, and the arguments for Iteration which we developed in Chapters 7, 8, and 9, also work perfectly well in the context of a counterpart theory that preserves the basic modal logic. Let's briefly revisit them to check that this is true.

The main arguments for ND which we mentioned in §4.2 were those from the B axiom; in the tense-logical case, from the standard "mixing" axioms; from ordinary arguments involving counting under counterfactual suppositions; and Williamson's (1996) argument from the logic of 'actually' (or modal anaphora). Insofar as we can find a workable counterpart theory that preserves the basic modal logic, it is hard to see how it could blunt (or sharpen) the force of any of these considerations. True, developing a counterpart-theoretic treatment of the 'actually' operator that preserves the kinds of principles that feature in the Williamsonian argument involves some difficulties over and above those required to preserve the basic modal logic (Hazen 1979; Fara and Williamson 2005). But it can be done (Bacon 2014; Russell 2013; Dorr unpublished *a*). And the counterpart-theoretic framework does not undermine the reasons for wanting to do it.

Turning to Iteration, let's start with the arguments from Chapter 8. The framework of counterpart theory does nothing at all to undermine the reasons for thinking that 'metaphysically necessary' expresses an extensionally minimal necessity operation if there is one. True, Lewis himself held that not only the ordinary

[25] Since Lewis's counterpart theory is developed in a strictly first-order framework, readers might worry that in imagining a counterpart theory compatible with our higher-order modal logic, we are venturing into completely uncharted territory. But extending the basic idea of counterpart theory to higher types seems straightforward enough. Properties can have other properties as counterparts: if Cleopatra′ is a counterpart of Cleopatra, then *loving Cleopatra′* is plausibly a counterpart of *loving Cleopatra*. And a property possibly instantiates a qualitative higher-order property (such as *being instantiated*) iff one of its counterparts instantiates that higher-order property. On the natural way of developing this thought, *qualitative* properties will have no counterparts other than themselves, since when *F* is qualitative, *being identical to F* is a qualitative property that *F* has necessarily, hence one shared by all *F*'s counterparts. This may, however, require some rethinking of Lewis's specific brand of set-theoretic reductionism about things like properties and propositions: see Dorr 2005*a*.

That said, the arguments below do not depend on the full resources of higher-order logic: counterpart theorists who rejected all higher-order quantification as unintelligible are not automatically immune.

language 'necessary' but the special philosophical use of 'metaphysically necessary' were context-sensitive, being associated with different counterpart relations on different occasions of use. But this thesis is entirely orthogonal to the counterpart-theoretic framework. Counterpart theory might provide an interesting characterization of what all the necessity operations have in common, but doesn't diminish the reasons for placing a lot of interpretative weight on the standard characterizations of metaphysical necessity as "absolute".[26]

Might counterpart theory perhaps help to undermine the reasons for thinking that there *is* an extensionally minimal necessity operation, or that if there is such an operation it obeys Iteration? It certainly isn't obvious how it would do so. Lewis himself was willing to admit a "broadest" interpretation of modal operators, which he takes to correspond to a *universal* counterpart relation:

> In the broadest sense, all possible individuals without exception are possibilities for me. But some of them are *accessible* possibilities, in various ways, others are not. (Likewise for joint possibilities: in the broadest sense, all pairs of possible individuals are possibilities for any pair, but some are accessible possibilities in ways others are not.) (Lewis 1986*a*: 234)

Here 'joint possibilities' refers to the "counterparts of n-tuples" approach we mentioned earlier. If one wanted to modify that approach to preserve the basic modal logic, one would need to something slightly more restrictive than the universal relation on tuples of a given length, on which (e.g.) the counterparts of an identity pair are all the identity pairs rather than all the pairs. But this does not disrupt the reasons for thinking that the disjunction of all admissible counterpart relations will itself be an admissible counterpart relation, corresponding to a maximally demanding necessity operation.[27]

[26] Lewis, ironically, accuses *other* opponents of Iteration, like Salmón, of failing to do justice to the absoluteness of metaphysical necessity (Lewis 1986*a*: 246). What seems to be going on is that Lewis is envisaging (as suggested in Lewis 1968: §5) that the general version of counterpart theory applicable to all interpretations of modal operators has two independently context-sensitive moving parts: an *accessibility relation* (which only relates worlds) in addition to a counterpart relation (which relates objects in worlds). The "metaphysical" interpretations are absolute in that their accessibility relation is as weak as can be, but their counterpart relations can still differ in all sorts of ways. But it is completely obscure what could motivate treating this differential treatment of the two contextual parameters. Moreover, the use of two separate parameters seems rather artificial in any case: Lewis (1986*a*: 230–1) gestures towards a vision where they are unified into a single "possibility for" relation. The version of counterpart theory in Dorr unpublished *a* fleshes out this picture, in which Lewis's "worlds" turn out to play no distinctive role in the analysis of modality.

[27] If, instead, we preserve the basic modal logic by keeping Lewis's 1968 scheme while adding an "exactly one counterpart per world" constraint, the situation is a bit different, since there is no guarantee that the disjunction of a collection of relations which individually obey this constraint will itself obey the constraint. However, given some collection of modal operators $\Box_i$ corresponding to counterpart relations C_i which individually obey the constraint, there is no obvious bar to our introducing a new operator $\Box$ such that $\Box P$ is equivalent to $\bigwedge_i \Box_i P$. While this $\Box$ is not equivalent to the result of plugging $\bigvee_i C_i$ or any other counterpart relation into the 1968 translation scheme, it is hard to see what would

Let's turn next to the arguments from Chapter 7, where we argued that there is a Tolerance Puzzle about ancestral necessity that arises whether or not we embrace Iteration for metaphysical necessity. Might counterpart theory somehow help to undermine our case for Ancestral Microphysical Supervenience, thereby making the Ancestral Hypertolerance claims required by an Iteration-denying approach to Tolerance Puzzles more palatable? Sociologically, acceptance of counterpart theory has certainly tended to go along with claims that suggest that ancestral possibility should be something of a free-for-all. In discussing Chisholm's argument, Lewis suggests that we can get from the pair ⟨Adam, Noah⟩ to the pair ⟨Noah, Adam⟩ via a chain of pairs where each is a counterpart of its predecessor (Lewis 1986*a*: 243–4), so that it is ancestrally possible for Noah to bear to Adam any qualitative relations that Adam in fact bears to Noah. And if there is a necessity operation that works anything like the "broadest sense" of 'possibility' envisaged by Lewis in the passage quoted above, *it* will be a complete free-for-all. For example, it will not be not necessary in the broadest sense that everyone in the Tri-State Area is in New York, Connecticut, or New Jersey, since every quadruple is a possibility for the quadruple of these four objects, including quadruples whose first element contains people who aren't *in* their second, third, or fourth element. For similar reasons, if the actual world consists of only finitely many atoms $a_1, \ldots, a_n$ standing in a maximally specific qualitative microphysical relation R, it will not be necessary in the broadest sense that if $Ra_1 \ldots a_n$ then Woody is wooden, since the counterparts (in the broadest sense) of the $n+1$-tuple $\langle a_1, \ldots, a_n, \text{Woody}\rangle$ include $n+1$-tuples whose first n elements stand in R but whose last element is not wooden (such as $\langle a_1, \ldots, a_n, \text{Saul Kripke}\rangle$). So on the picture suggested by Lewis's remark, the statement of Microphysical Supervenience will be false if its modal operator is interpreted in the "broadest sense". In that case our arguments for Ancestral Microphysical Supervenience must be wrong, since they would apply just as well to the broadest necessity.

But setting sociology aside, it is hard to see how counterpart theory (once modified to preserve the basic modal logic, including Leibniz's Law) could actually do anything to blunt the force of the arguments for Ancestral Microphysical Supervenience that we gave in §7.3. As we observed, given Leibniz's Law, higher-order identities like the plausible-looking (10) will imply corresponding claims of ancestral necessity:

(10) To be in the Tri-State Area *is* to be in New York, Connecticut, or New Jersey.

warrant refusing to count it as a "necessity operation". Likewise, even though the composition $\Box_1\Box_2$ of two operations corresponding to constraint-obeying counterpart relations C_1 and C_2 need not itself correspond to any constraint-obeying counterpart relation, this also seems like a dubious basis for refusing to count it as a necessity operation.

If (10) is true, the requirements for a counterpart relation to correspond to a logically well-behaved necessity operation are more demanding than we might have thought. While we can secure the validity of first-order Leibniz's Law by imposing the constraint that no object has two counterparts in a single world, or that identity pairs only have other identity pairs as counterparts, securing higher-order Leibniz's Law requires a lot more, e.g. that every counterpart of ⟨the Tri-State Area, New York, Connecticut, New Jersey⟩ is such the objects in its first element are exactly those in its second, third, or fourth element.

We are not suggesting that identities like (10) are *obviously* true. As we saw in §7.4, there are interesting and consistent general principles which would rule out all such identities at a stroke, articulating a vision where *objects*, unlike entities of other types, all play interchangeable roles in the pattern of higher-order identities. We might conjecture that most counterpart theorists have been implicitly committed to this picture. Still, the abstract structure of counterpart theory certainly doesn't require it, and nor does counterpart theory (understood as an account of necessity and possibility) offer any obvious resources for undermining the plausibility of such identities, which involve no modal vocabulary. But as we argued in §7.4, once we countenance identities like (10), it is hard to draw a principled line between objects like the Tri-State Area and the ones that figure in standard Tolerance Puzzles.[28]

[28] It might be suggested that counterpart theorists should think of the higher-order identity connectives themselves as sources of context-sensitivity, analogous to the context-sensitivity introduced by modal operators. Thus, e.g., (10) could be true in some contexts and false in others, even when the contexts all agree on the interpretation of the names and the preposition 'in'. The problem we see with this view is that identity (higher-order as well as first-order) is provably coextensive with Leibniz-equivalence, which is defined in terms of unrestricted quantification and the truth-functional connectives. So for higher-order identity to be a source of context-sensitivity, unrestricted higher-order quantifiers would also have to be such a source. But that thought is hard to make sense of, since it seems that for any two interpretations of a quantifier, either one is a restriction of the other or there is a third interpretation of which both are restrictions, and in either case they are not both unrestricted.

A different idea is that counterpart theory might posit context-sensitivity in higher-order identities stemming not from the identity connective itself, but from the expressions flanking it. For example, you might think that the denotation of the predicate 'be in the Tri-State Area' depends on a contextually given counterpart relation: e.g., relative to counterpart relation *C*, it might be thought (speaking model-theoretically) to denote the function that takes each world *w* to the set of things in *w* that are in something that bears *C* to the Tri-State Area. This picture might be suggested by views like those of Ramachandran (1989), Hellie, Murray, and Wilson (2021), and Yli-Vakkuri and Hawthorne (forthcoming a), which invoke counterpart-theoretic ideology in the translations of, or model-theoretic clauses for, atomic sentences (as opposed to modal operators). On it, the dimension of context-sensitivity represented by the choice of a counterpart relation is present already in simple nonmodal sentences like 'Brooklyn is in the Tri-State Area'; any context-sensitivity that modal operators themselves might introduce is something quite separate. This kind of counterpart theory does not seem to do anything to support the Iteration-denying approach to Tolerance Puzzles. However, in Chapter 11 we will see how positing a surprising amount of context-sensitivity in simple nonmodal sentences can help overcome the main challenge for the strategy of denying Non-contingency, though our account of the source of this context-sensitivity does not fit well with the idea that it is driven by the choice of a counterpart relation.

Turning finally to the arguments for the truth of Iteration in chance-theoretic Tolerance Arguments which we presented in Chapter 9, we face the obstacle that there is no standard counterpart-theoretic account of sentences of the form 'The chance at *t* that *P* is *x*' to begin with. Kment (2012) considers some natural suggestions, e.g. that 'The chance at *t* that *a* is *F* is *x*' is equivalent to 'the chance-measure at *t* of the set of worlds that contain at least one *F* counterpart of *a* is *x*'. He points out that in the absence of the "exactly one counterpart per world" constraint, these natural suggestions lead to bizarre violations of the usual probability laws which would play havoc with our ordinary chance-theoretic reasoning: for example, the chance that *a* is *F* and the chance that *a* is not *F* could both be greater than 1/2. Order could be restored either by imposing the constraint, or by switching—as Kment suggests—to a different treatment, modelled on a different version of modal counterpart theory such as the one with counterparts of *n*-tuples. Once we do this, the Chance Tolerance and Robustness Arguments from Chapter 9 will be valid. The option of avoiding them by giving up Chance Fixity will still be on the table, but it looks no more appealing in the presence of counterpart theory than without it.

We conclude that contrary to what many philosophers have thought, counterpart theory is more or less inert when it comes to the relative merits of the main options for escaping Tolerance Puzzles.

11
Resolving the Puzzles: Plasticity and Plenitude

We have now scrutinized two important ways of responding to Tolerance Arguments: accepting Hypertolerance and giving up Iteration. We found good reasons to accept Iteration in the case of metaphysical modality (and various narrower modalities defined in terms of it such as atomic possibility), tense operators, and positive objective chance. Since our puzzles have generally been framed using these modal operators, we don't hold much hope for the Iteration-denying strategy. In the case of Hypertolerance, we saw reasons to doubt that it provides an adequately general strategy for defusing Tolerance Puzzles, though we didn't preclude its being the right response to some of them. Meanwhile, so long as we are careful in how we set up the puzzles, Tolerance will be on a very firm footing, while Persistent Closeness can be made unproblematic. Given this state of play, there is ample motivation for a second look at the case for Non-contingency.

11.1 Security and Semantic Plasticity

In Chapter 3 we considered a "Security Argument" for Non-contingency, a key premise of which was

Independence If Tolerance could easily have been false, we could easily have falsely believed it.

We also introduced a chance-theoretic variant of this argument, in which the role of Independence was played by the following premise:

Chance Independence If the chance at *t* of Tolerance being false is very substantial, the chance at *t* of our falsely believing Tolerance is substantial.

In discussing these arguments, we did our best to make the Independence premises seem intuitive. On the view that Tolerance is true but could easily have been false, all kinds of minor differences—e.g. the positioning of cuts in certain particular boards, or the random selection of materials from certain piles of interchangeable

The Bounds of Possibility: Puzzles of Modal Variation. Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri, Oxford University Press. © Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri 2021.
DOI: 10.1093/oso/9780192846655.003.0012

parts in distant portions of the world—can function as difference-makers between the actual world and worlds where Tolerance is false. Independence is then motivated by the thought that since the difference-makers aren't facts to which our belief-forming processes are sensitive, we are at their mercy.

Readers who are familiar with externalist themes in contemporary philosophy of mind may already have noticed the problem.[1] The sense in which we aren't "sensitive" to the relevant factors is physical: nothing of interest in the physical state of our brains, our dispositions to utter certain sentences, and so on is significantly correlated with the positioning of the cuts or the selection of the materials from the piles. But famously, the facts about which propositions we believe and assert are not determined by our brain states. Differences in parts of the world that are not interestingly causally related to our physical configuration might nevertheless correlate systematically with differences in which propositions we believe and assert. For example: we see a grabber pick one of many matching balls at random and we exclaim, 'That ball was picked!' We thereby assert of a certain ball *b* that it was picked. In nearby worlds where different balls are picked, we instead assert of balls other than *b* that they were picked. Our beliefs are different even though our physical state in some of those worlds is more or less the same, since there is no interesting causal correlation between our physical state and the physical process that determines which ball is picked.

So there is room in principle for resisting the Independence premises by appealing to the external determination of propositional attitude and speech act content. Let's see how this might work in connection with a *de re* Tolerance Argument, where Tolerance is stated using a demonstrative. In the actual world, we point to Woody the table and utter the sentence 'That is tolerant' (having earlier defined '*x* is tolerant' to mean '*x* could have been originally composed by any of the collections of atoms that chemically match the collection of atoms that in fact originally composed *x* and overlap it by at least 90 per cent'). In uttering our sentence, we assert the proposition that Woody is tolerant. Let's assume that this proposition is only contingently true, and indeed that it could quite easily have been false. For concreteness, imagine that the carpenter who made Woody very nearly picked two different legs, and that if she had used those legs, Woody would still have been made, but would then not have been tolerant. Given the lack of significant physical correlation between the carpenter's choice of legs and our physical state, we still utter the *sentence* 'That is tolerant' at many of these nearby worlds where Woody is not tolerant. To avoid nearby error, we must say that at these leg-switched worlds, we do not believe that Woody is tolerant, or assert that it is tolerant when we utter 'That is tolerant'. Instead, when we utter that sentence, we assert some other, true, proposition. Just like 'That ball was picked', the sentence

[1] See the introduction to Yli-Vakkuri and Hawthorne 2018 for a survey.

is subject to what we will call *semantic plasticity*: there is fine-grained variation within the space of nearby worlds as regards what it is used to assert.[2]

Another example of semantic plasticity that may provide a helpful analogy is the sentence 'The Standard Metre is exactly one metre long', familiar from Kripke's argument for the possibility of contingent a priori knowledge (Kripke 1972: 55). When we utter the sentence 'The Standard Metre is exactly one metre long', we assert a proposition that is not only contingent, but extremely fragile: differences in the distances between atoms that are far too tiny to be correlated significantly with our speech behaviour would suffice for the proposition to be false. But the fact that we speak the truth when we uttering that sentence is not similarly fragile (though of course it is still contingent). And as a result, the fragility of the truth of the proposition we express poses no threat to our claim to know it.[3]

But what *are* we asserting when we say 'That is tolerant' at nearby worlds where Woody isn't tolerant, and we utter the sentence without asserting anything false? One possibility is that we are asserting of some other thing, distinct from Woody, that it is tolerant. Another is that we are still referring to Woody, but attributing to it not the property of being tolerant, which it lacks, but some other property which it has. A third is that we are both using 'that' to refer to some object other than Woody *and* using 'is tolerant' to attribute some property other than tolerance. While all of these options are worth exploring, the first is particularly natural given the fact that demonstratives are so obviously prone to shift their reference across worlds owing merely to changes in their external environment, as illustrated by the grabber example. And indeed, this will be our preferred view of the matter.[4]

[2] We are not the first to suggest that semantic plasticity can help stabilize a Non-contingency-denying approach to Tolerance Puzzles. Williamson (1990: 134–5, crediting Eli Hirsch) suggests that the expression 'that ear-ring' might have an 'indexical component', such that 'when we point in the direction of material *m* and say "that ear-ring", we should be interpreted . . . so that no slight change of material would have made a different ear-ring in the relevant sense.' See also Hawthorne 2006*a*: 241, n. 8.

[3] Neither this sentence nor 'That is tolerant' is a great candidate for expressing a priori knowledge: for one thing, on the standard way of thinking about a priori knowledge, we are not in a position to exclude the sceptical hypothesis that all appearances of sticks and tables are hallucinatory. Material conditionals like 'If the Standard Metre has a length, it is exactly one metre long' are stronger candidates for expressing contingent a priori knowledge, though it is far from clear exactly what needs to be packed into the antecedent in order to control for all relevant sceptical scenarios. But if you're happy to think that some conditional of the form 'If *P* then the Standard Metre is exactly one metre long' is contingent a priori, we suggest that you should be equally happy to think that some conditional of the form 'If *P* then that is tolerant' has the same status. Thus insofar as semantic plasticity can provide a principled basis for rejecting the Independence premise in the Security Argument, it also undermines any putative path from claims of a priority to Non-contingency.

[4] Granted, there is an interesting difference between the kind of variation we need to posit in the reference of 'that' to preserve the truth of the nearby Tolerance speeches and the kind illustrated by the grabber example. In the latter case, the nearby worlds where we don't refer to ball *b* when we say 'that' are worlds where *b* isn't even in our line of sight. By contrast, to preserve Tolerance we will have to say that we don't refer to Woody by 'that' at the relevant nearby worlds, even though Woody is still sitting right in front of us while we make our speech. Notice, however, that there are fairly mundane cases where the referent of a perceptual demonstrative can shift without the original referent falling out of the asserter's line of sight. Suppose that in the actual world you are looking straight down on a

If we refer to some tolerant object when we say 'that' in the nearby worlds where Woody isn't tolerant, what do we refer to? Whatever it is, it must surely be similar to Woody in some salient respects: it is something table-shaped that's sitting right in front of us. (It would be completely implausible to think that in nearby worlds, we avoid error when we just say, 'That is tolerant', but fall into error when we say, 'That is tolerant and table-shaped and sitting right in front of us.') The reference-shift strategy thus turns crucially on the idea that the world of macro-sized objects is rich enough that when one sits down in front of a table, one will generally be sitting in front of a range of different objects which are very similar in physical respects—all of them are table-shaped; all of them are composed of atoms; etc.—but which may nevertheless be interestingly dissimilar in some modal respects (e.g. because some of them are tolerant and others are not).

It might initially seem radical, even outrageous, to suppose that whenever you sit down in front of one table-shaped object, you sit down in front of many table-shaped objects. But there are many precedents for this idea within philosophy. One widely discussed motivation (Geach 1962; Unger 1980; Lewis 1993) has to do with differences in exact spatial boundaries. Wherever there is a table, there are astronomically many different *collections of atoms* which are configured in a table-shaped arrangement, and which differ from one another only in that some include and others exclude certain atoms around the boundaries. It seems plausible that for every collection of atoms there is a material object composed of those atoms—e.g. the "aggregate" of the atoms, or the hunk of matter they comprise; it would be strange to suppose that only one or a few of the many collections of atoms is such as to compose any material object at all.

A second, even more widely discussed motivation for multiplying table-shaped objects is based on Leibniz's Law arguments for the conclusion that any table is distinct from the quantity of wood that composes it, analogous to the arguments for the non-identity of statues and lumps of clay which we considered at the beginning of Chapter 4. In a typical case, one can argue for distinctness on the grounds that the quantity of wood was around before the table was constructed, whereas the table wasn't around then. And even in unusual cases where the quantity of wood didn't predate and won't postdate the table (e.g. because both are eternal), there are still modal grounds for taking them to be distinct: for example, one but not the other could have existed in a world where there is no furniture.

However, the abundance of table-shaped objects required to sustain the response to the Security Argument that we are exploring cannot directly be

hemisphere. Wrongly assuming yourself to be looking down at a whole sphere, you utter the sentence 'That is spherical' and thereby assert, falsely, of a hemispherical object that it is spherical. In a nearby possible world, you and the hemispherical object you actually refer to are just where they actually are, but the flat side of the hemisphere is glued to that of a second hemisphere (which is entirely occluded from your point of view by the first hemisphere). Plausibly, when you utter 'That is spherical' in that world, you refer to something with both hemispheres as parts, and say nothing false.

subsumed under either of these models. Merely reflecting on the abundance of collections of atoms with slightly different spatial boundaries provides no grounds for confidence that the set of table-shaped objects approximately coincident with any one table will differ modally in the ways we need, i.e. with regard to tolerance. Moreover, in some Tolerance Arguments, such as those having to do with the range of possible atomic compositions for a particular molecule, there is no relevant unclarity about the spatial boundaries of the target objects. The contrast between a table and the quantity of wood that composes it gives us something a bit more like what we are looking for, since it involves a modal difference without any difference in atomic composition. But it is not at all plausible that in the nearby worlds where Woody isn't tolerant, our demonstrative refers to something like a *quantity of wood* rather than a table. In all relevant worlds, we are disposed to say things like 'That was made by a carpenter' and 'That would be destroyed if we dropped it from a ten-storey building.' And we wouldn't want to buy the truth of 'That is tolerant' at the expense of conceding the falsity of mundane speeches such as this.

If we want to undermine Independence by appealing to shifts in reference, we are thus badly in need of a general theory of material objects that guarantees, and explains, the existence of an abundance of objects that differ in the fine-grained modal respects needed to make the relevant 'That is tolerant' speeches true, without differing from one another in other ways that would introduce unwanted errors into other speeches we are disposed to make using the same demonstrative.

Many philosophers have defended "plenitude" principles that imply that wherever there is one material object, there is a dizzying variety of other, coinciding material objects. These objects include not just objects with familiar modal profiles like those of tables and hunks of wood, but hordes of other objects with bizarre modal profiles, e.g. an object that necessarily coincides with Woody if there is an odd number of lions and Woody is concrete and is otherwise not concrete. Different theoretical frameworks have been developed for articulating views in this spirit, and so long as they are consistent, our general strategy for solving Tolerance Puzzles is not particularly sensitive to the points that distinguish them.[5] But it is not a straightforward matter to state a principle that is that is precise, internally consistent, and capable of underwriting the intuitive picture of a plenitudinous reality.[6] Since the literature has often struggled in this

[5] Examples include Schlesinger 1985, Yablo 1987, forthcoming, Fine 1982, 1999, Hawthorne 2006*e*, Leslie 2011, Inman 2014, Jago 2016, and Goodman unpublished *c*. See Fairchild 2019 for a survey.

[6] Fairchild (2019) surveys a range of theories whose structure is modelled on the main proposal in Yablo 1987, which take roughly the form 'For any material object *y* and any collection *C* of "well-behaved" properties of *x* satisfying such-and-such logical constraints, there is an object *x* that coincides with *y*, has every property in *C* essentially, and has no other "well-behaved" property of *y* essentially.' She points out some logical pitfalls into which views with this structure can fall, including inconsistency and vacuity. Her own proposal, "Global Plenitude", is precise and consistent and takes this form, but

respect, it is worth spending some time exploring some carefully formulated theories which would do the job. The next section will thus serve as a mini-primer in plenitudinous metaphysics. Readers who don't mind working with the impressionistic statement of plenitude above can skip forward to §11.3.

11.2 Plenitude Principles

The first version of plenitude we will discuss (drawing on Hawthorne 2006*e*) takes *location*, understood as a relation between objects and spacetime points, as its central theoretical primitive. Intuitively, an object is located at a spacetime point if the latter is "inside" the former. The principle can be stated as follows:

Location Plenitude For any property F, there is an object x such that it is metaphysically necessary that for any spacetime point y, x is located at y if and only if y is F.

Note that F can be any property whatsoever—it need not be natural or intrinsic or qualitative. For example, by taking F to be *being a point such that Woody is located at it and there is an odd number of lions*, we get an object such that necessarily, it is located at exactly the same points as Woody if there is an odd number of lions and located nowhere otherwise.[7]

has a consequence which we believe to be false, and which we imagine few friends of plenitude will be willing to embrace: namely, that for any two coincident objects x and y, it is necessary that if x and y both exist they coincide. (Very briefly, and referring the reader to Fairchild 2019 for relevant definitions: suppose for contradiction that a and b coincide at the actual world w_0 but not at some other world w_1. Let E be the set of all "neutral" properties entailed by *either coinciding with a or (coinciding with b and being such that* w_1 *obtains)*, and let A be the set of all other neutral properties instantiated by a at w_0. The ordered pair $\langle E, A \rangle$ is a "nonlocally closed modal profile" at w_0. So by Global Plenitude, there is an object x that has every property in A accidentally at w_0. But two properties in A are *not (coinciding with a and being such that* w_1 *obtains)* and *not (coinciding with b and being such that* w_1 *obtains)*. So each of these properties is one that x possibly lacks: thus, x coincides with a at w_1 and coincides with b at w_1. But since coincidence is symmetric and transitive, this contradicts our assumption that a does not coincide with b at w_1.)

[7] While the idea that objects are "located" at points of *space* is relatively common-sensical, a "location" relation that objects bear to points of space*time* is a theoretical posit; its relation to our ordinary thought about geometrical properties of and relations among ordinary objects and the way in which they change over time is not obvious. The usual setting for this posit is a "B-theoretic" one, where instants of time are taken to be certain three-dimensional slices through spacetime (or to correspond in some natural way to such slices), and geometric predications like 'x is spherical' are taken to predicate relations between objects and instants. But A-theorists, who hold that predicates like 'is spherical' express properties that objects either have or lack simpliciter, also have ways of making sense of the idea that objects are located at spacetime points. One possible A-theoretic approach is a version of the moving spotlight theory (see Deasy 2015) where the location relation between objects and spacetime points is eternal, but where *being present* is a non-eternal property of spacetime points, whose instances always form a three-dimensional spacelike surface: then *being spherical* (simpliciter) will turn on the geometrical relations between the *present* points at which one is located. Alternatively, one might think of the location relation to spacetime as non-eternal: see Dorr unpublished *a* for some options for thinking about this.

Given that there are infinitely many metaphysically possible worlds where there are spacetime points, Location Plenitude entails that for every object *x*, there are infinitely many other objects that are *coincident* with *x* in the sense of being located at exactly the same spacetime points at *x*. Coincident objects must have the same geometric properties and stand in the same geometric relations to other objects: e.g., anything coincident with a table in front of us is table-shaped and in front of us. But certainly not all properties and relations are undiscriminating in this sense, since no two objects have all the same properties.[8] Many ordinary words express discriminating properties: for example, *being a table* and *being a quantity of wood* are discriminating, since some tables coincide with quantities of wood, but no table is a quantity of wood. There are many debatable cases where it is not obvious whether a certain property is undiscriminating, or whether a certain relation is undiscriminating in a particular argument place; in §12.3 we will discuss some hard cases, including colour, mass, parthood, and seeing. In thinking about Location Plenitude and other forms of plenitude, it is tempting but potentially misleading to take for granted that certain of these properties or relations are undiscriminating.[9]

Location Plenitude is a strong claim in several ways.[10] Since it doesn't include any restriction to spacetime points that are "matter-filled" or interesting in any other way, it implies that there are table-shaped objects located in the intergalactic void. This seems acceptable to us, although if physics provided some well-behaved and relevant interpretation of 'matter-filled', it would be perfectly fine for our purposes if one replaced '*y* is *F*' in Location Plenitude with '*y* is matter-filled and *F*' (or, equivalently, if one replaced 'For any property *F*' with 'For any property *F* that necessitates *being matter-filled*').[11] Location Plenitude also includes no restriction that would require the collection of all points at which any given object is located to satisfy any topological well-behavedness condition, e.g. being

[8] Fairchild (2019) calls undiscriminating properties "neutral" properties.

[9] For example, Korman (2015) assumes that plenitude implies that there are 'wide swathes of *highly visible* extraordinary objects' (italics ours).

[10] Location Plenitude is nevertheless weaker in one way than certain other plenitude principles found in the literature (e.g. Yablo forthcoming). Since it says 'an object' rather than 'exactly one object', it is neutral as regards regards whether it ever happens that two objects are *necessarily* located at exactly the same spacetime points. There are several possible reasons for denying the uniqueness claim. Fine (2000) gives a highly debatable putative example of necessary co-location, involving letters written in different languages. A more theoretical motivation arises from certain "hierarchical" plenitude principles, which envision objects as built up in an order reminiscent of the set-theoretic hierarchy, and have been argued for on the basis of premises about *parthood* about which Location Plenitude is silent: see Fine 1999 and Goodman unpublished *c*. A quite different theoretical motivation would appeal to an "identity strength" version of Location Plenitude (see note 31 below) together with the denial of Intensionalism (see §1.4).

[11] If fundamental physics works in a field-theoretic way, it seems unlikely to provide any useful interpretation of 'matter-free'. One might consider points where some chosen fields are zero, but such points will be few and far between even in the intergalactic void, and will not provide a good gloss on the distinction between matter and void that guides folk physics (doubtless in part thanks to the influence of Democritus). In any case, maybe some quite familiar objects, like holes in bagels, are located mostly at matter-free spacetime points.

topologically open or closed; however, one could easily modify Location Plenitude to make it compatible with such constraints by adding a clause of the form '. . . and the collection of all *F* points satisfies condition *C*' at the end.[12] Nor is Location Plenitude restricted to *instantiated* properties: it entails that there are objects—in fact, a vast multitude of objects—that are not located at any points at all. As we are thinking of things, these unlocated objects will include non-concrete possible children of Wittgenstein, the non-concrete possible knife that would have been created if a certain handle had been attached to a certain blade, and so on. Those who deny that there are such contingently non-concrete things might want to restrict the quantifier in Location Plenitude to properties that some spacetime points instantiate; however, given the necessitist commitments built into our basic modal logic, the view that there are no contingently non-concrete things seems hard to sustain.[13]

Location Plenitude characterizes objects in terms of a location relation they bear to spacetime *points*. For various reasons, some might prefer a principle that talks about *regions* rather than points—most obviously, one might hold that there are spacetime regions while denying that there are any spacetime points.[14] We could just replace the word 'point' in Location Plenitude with 'region'; however, it is arguable that the resulting principle will require some further restriction, so as to be compatible with certain necessary laws governing the location relation between objects and regions. For example, if one takes it to be necessary that no object is located at more than one region, one can first replace 'point' with 'region' in

[12] Cartwright (1975) argues that only *regular open* regions can be occupied by material objects. Hudson (2006: ch. 2) defends the competing "liberal" view that every region is possibly occupied by a material object. Uzquiano (2006) responds.

[13] Since our basic modal logic does not also include a commitment to BF, and thus allows for the possibility of *new* things, you might think that we could happily take on board the the claim that there are no contingently non-concrete things by claiming that if Wittgenstein had children they would be new things. But given necessitism it would be extremely hard to deny that there could have *easily* been contingently non-concrete things. For example, if the blade and handle of a certain knife had never been attached, that knife would never have been made and would thus have been non-concrete, although (given necessitism) it would still have been identical to something. Moreover, if the blade and handle had only just failed to be attached, there would have been a good chance of its being concrete, and so its non-concreteness would have been contingent. Thus if one wanted to maintain that there are in fact no contingently non-concrete objects, one would have to think that the actual world is very special in this respect. This raises a pressing security worry concerning any belief that there are no contingently non-concrete objects. And the plasticity posits we defend in this chapter don't seem to help here, since they don't suggest that anything in the sentence 'there are no contingently non-concrete objects' is plastic. Of course, Location Plenitude as stated also implies that there is at least one object that is *necessarily* non-concrete (not located at any points). This seems harmless, but if you don't like it, you can obviously avoid it by replacing 'For any property *F* . . . ' with 'For any possibly instantiated property *F* . . . '.

[14] On views that posit spacetime regions but not points, see Arntzenius and Hawthorne 2005 and Arntzenius 2011 (ch. 4). But there are also live views according to which the fundamental spacetime-constituting entities are neither points nor spacetime regions—see e.g. Earman 1989, Maudlin 2010, 2014, and Dorr 2011. Of course, one need not regard spacetime points as *fundamental* to accept Location Plenitude, or to think that it implies that each material object in the actual world coincides with many others with different modal properties. It is enough if there is some way of "logically constructing" spacetime points, e.g. as certain nested sequences or collections of spacetime regions (as in Whitehead 1920: ch. 4).

Location Plenitude and then replace the final '*y* is *F*' with '*y* is uniquely *F*'.[15] If on the other hand one takes it to be possible for an object to be located at multiple regions, but necessary that if an object is located at a region it is located at all its subregions, one can instead replace '*y* is *F*' with '*y* is a subregion of a region that is *F*'. Some of the apparent choice points here seem merely terminological: for example, if there is one location-theoretic relation *L* that objects can bear to at most one region, we can use it to define another—*bearing L to a subregion of*—which objects can bear to multiple regions, but always bear to the subregions of regions they bear it to (Parsons 2007). But the issues are not *just* terminological: for example, there is a debate about whether there is any location-theoretic relation that an object could bear to multiple spacetime regions without bearing it to any two regions one of which was a subregion of the other.[16] It was partly in order to avoid getting drawn into such debates that we gave pride of place to the version of Location Plenitude with points.[17]

[15] Casati and Varzi (1999: 121) and Parsons (2007) accept this principle, when 'located' is understood as expressing the relation of "exact" location; they take all the other relevant disambiguations of 'located' to be definable from exact location (together with parthood). Parsons adds that exact location can in turn be defined in terms of each of those other location relations (together with parthood). If one identifies regions with nonempty collections of points and defines '*x* is located at *C*' as '*x* is located at all and only the members of *C*', the region-theoretic version of Location Plenitude with 'uniquely' added will become equivalent to the earlier point-theoretic version: for any property *F* of points, the corresponding property of regions is *containing all and only the F points*, and for any property *F* of regions, the corresponding property of points is *belonging to some F region*.

[16] Hudson (2006: ch. 4) takes seriously (without endorsing) the possibility that there is a "location" relation for which such scenarios are possible.

[17] Given an S5 modal logic, the region-theoretic version of Location Plenitude with '*y* is matter-filled and uniquely *F*' is closely related to the view of the 'Plenitude Lover' from Hawthorne 2006*e* (53):

> [L]et a *modal occupation profile* be a function from possible worlds to filled regions of spacetime. The Plenitude Lover says that for any such profile there is an object whose modal pattern of spatiotemporal occupation is correctly described by that profile.

We can understand a modal occupation profile to be a rigid binary relation *F* that each possible world *w* bears to at most one object *x*, such that whenever *Fwz*, *x* is a filled region of spacetime at *x*. *F* "correctly describes the modal pattern of spatiotemporal occupation" of an object *y* just in case for every *w* and *x*, *Fwx* iff at *w*, *y* is located at *x*. Then we can derive the region-theoretic variant of Location Plenitude from Hawthorne's version as follows. Where *X* is any property, let F_X be a rigid binary relation which *w* bears to *y* just in case *w* is a possible world at which *y* is a filled region and *y* is uniquely *X*. BF and ND entail that F_X is functional, since they vindicate the inference from 'at *w* there is a unique thing such that ...' to 'there is a unique thing such that at *w*, ...'. So by Hawthorne's version of plenitude, there is an object *x* such that for any *w* and *y*, *x* is located at *y* at *w* iff $F_X(w, y)$. Since $F_X(w, y)$ just in case *y* is filled and uniquely *X* at *w*, it follows that at each possible world, *x* is located at exactly the filled regions that are uniquely *X*; so by Leibnizian Necessity (which follows from S5: see §1.6), *x* is necessarily located at exactly the filled regions that are uniquely *X*.

In the other direction, we also need the assumption that no world is actual at any other world (which in the framework of §1.6 follows from Rigid Extensionality). Suppose *F* is a rigid functional binary relation such that whenever $F(w, y)$, *w* is a possible world at which *y* is a filled region. Let X_F be the property of being a region to which an actual world bears *F* ($\lambda x.\exists w(\text{Actual}(w) \wedge \text{World}(w) \wedge F(w, x))$). By the region-theoretic version of Location Plenitude, there is an object *x* which necessarily is located at a region iff it is filled and uniquely X_F. Let *w* be any possible world; then (given BF and Rigid Extensionality) at *w*, *w* is the only actual world, so at *w*, $\forall y(\text{Located}(x, y) \leftrightarrow \text{Filled}(y) \wedge F(w, y) \wedge \forall z(F(w, z) \rightarrow y = z))$. But since *F* is rigid and functional, S5 implies that this is true iff $\exists y(F(w, y) \wedge \text{At}(w, \text{Located}(x, y)))$; so *F* correctly describes *x*'s modal pattern of spatiotemporal occupation.

Another plenitude claim with the same structure as Location Plenitude, which would be just as good for our purposes, eschews the language of 'location' and gives centre stage instead to the parthood relation and the property of being a fundamental object:

Fundamental Part Plenitude For any property *F*, there is an object *x* such that it is metaphysically necessary that for any fundamental object *y*, *y* is part of *x* if and only if *y* is *F*.

Fundamental Part Plenitude aspires to be less committal than Location Plenitude as regards what is going on at the fundamental level; it could in principle be applied even to reasoning about unlocated objects like angels and Cartesian souls.[18] Again, there are easy weakenings which would make little difference for most purposes: for example, one could require *F* to be an instantiated or possibly instantiated property.[19]

Also adequate for our purposes would be a weaker kind of plenitude principle, which is similar in general structure to the previous ones, but characterizes the plenitude of material objects by their relations to *one another* rather than to some separate "ground floor" of spacetime points, spacetime regions, or fundamental objects. This principle takes as primitive a relation of *coincidence*, assumed to be necessarily symmetric and transitive:

Coincidence Plenitude For any property *F*, there is an object *x* such that it is metaphysically necessary that for any object *y*, *y* coincides with *x* iff something is *F* and *y* coincides with everything *F*.

The two views are thus equivalent given S5. However, Location Plenitude makes sense when combined with a wide range of S5-denying views. It is also conceptually simpler—a common benefit of formulating things without appealing to the ideology of possible worlds.

[18] If we take all spacetime points to be fundamental objects, and take an object all of whose fundamental parts are spacetime points to be located at exactly those points, Fundamental Part Plenitude will entail Location Plenitude. Conversely, Location Plenitude will entail Fundamental Part Plenitude if we combine it with the (obviously tendentious!) assumptions that all fundamental objects are spacetime points and that an object must have a spacetime point as a part to be located there. Insofar as we don't accept the assumptions required for the two principles to be interchangeable, we might also consider a hybrid principle that replaces 'is located at' with 'has as a part' in Location Plenitude.

[19] Some proponents of a B-theoretic account of time think that properties are never really instantiated temporarily: when we say informally that an object has a property sometimes but not always, what we really mean is that it bears a certain *relation* to some but not all times (Russell 1903: §442). For these philosophers, there is a worry that Fundamental Part Plenitude is too weak to play the role that plenitude principles are supposed to play, since it does not tell us anything about material objects which change their fundamental parts over time. One might argue that this is not a problem, on the grounds that Fundamental Part Plenitude should be understood as a claim about a genuinely binary (and thus permanent) parthood relation, rather than about the more familiar notion of parthood in which things can have parts temporarily (see Sider 2002: ch. 6). But if one is not satisfied with this, one could modify Fundamental Part Plenitude to invoke a time-relative notion of parthood, as follows: for every relation *R*, there is an object *x* such that necessarily, for any fundamental object *y* and any time *t*, *y* is part of *x* at *t* iff *y* bears *R* to *t*.

For example, if *F* is *being a table*, *x* will be such that necessarily if there is a table that coincides with every table, it coincides with that table (and hence with every table), and otherwise it doesn't coincide with anything (which we can equate with its being non-concrete). If one wished, one could weaken Coincidence Plenitude by requiring *F* to be instantiated or possibly instantiated.[20]

For an equivalent formulation, say that a property *F* is *undiscriminating* iff necessarily, an object is *F* if and only if it coincides with something *F*; say that *F* is *unrepeatable* iff necessarily, if *x* and *y* are both *F*, *x* coincides with *y*; and say that an object *x tracks* a property *F* iff necessarily, for any object *y*, *y* coincides with *x* iff *y* is *F*. Then Coincidence Plenitude is equivalent to the claim that every undiscriminating and unrepeatable property is tracked by some object.[21]

If we define '*x* coincides with *y*' as '*x* is located at exactly the same spacetime points as *y*, and *x* is located at at least one spacetime point', Coincidence Plenitude follows from Location Plenitude.[22] Likewise, if we define '*x* coincides with *y*' as '*x* has the same fundamental parts as *y*, and *x* has at least one fundamental part',

[20] Yablo (1987: 310) includes something very similar to Coincidence Plenitude as part of the definition of "full property model"; see also Yablo forthcoming.

[21] Since coincidence is necessarily symmetric and transitive, *coinciding with some and every F thing* is undiscriminating and unrepeatable for any *F*, and necessarily equivalent to *F* if *F* is itself undiscriminating and unrepeatable.

We can also reformulate Coincidence Plenitude in a way that gives it the logical structure of the plenitude principles discussed by authors like Yablo (1987), Leslie (2011), and Fairchild (2019) (see note 5). For the purposes of this note, say that *F* is *essential* to *x* iff necessarily, if *x* coincides with itself, *x* is *F*. We can show that if *F* is undiscriminating, then *x* tracks *F* if and only if the essential undiscriminating properties of *x* are all and only the undiscriminating properties necessitated by *F*. (*Right to left:* Suppose *x* tracks *F* and *F* is undiscriminating. Clearly *F* is essential to *x*, since *x* can coincide with itself only if coincides with something *F* and hence is *F*; and thus all properties *F* necessitates are also essential to *x*. And if *G* is undiscriminating and essential to *x*, then *F* must necessitate *G*, since necessarily, for every *F* object *y*, *y* coincides with *x* (since *x* tracks *F*), so *y* coincides with something *G* (since *x* must coincide with itself to coincide with *y*), so *y* is *G* (since *G* is undiscriminating). *Left to right:* Suppose *x*'s essential undiscriminating properties are exactly those necessitated by *F*. Then since *coinciding with x* is undiscriminating (by the fact that coincidence is necessarily transitive and reflexive), *F* necessitates *coinciding with x*; conversely, since *F* necessitates itself, *F* is essential to *x*, so *coinciding with x* necessitates *coinciding with something F* and hence necessitates *F* itself (since *F* is undiscriminating).)

Using this result, we can give the following (rather less perspicuous) reformulation of Coincidence Plenitude: for every collection *C* of undiscriminating properties, if there is a undiscriminating unrepeatable property *F* such that *C* consists of all and only the undiscriminating properties necessitated by *F*, then there is an object x_C such that *C* contains all and only the essential undiscriminating properties of x_C. And we can likewise reformulate the weakening of Coincidence Plenitude in which *F* is required to be instantiated: for every object *y*, and every collection of *C* of undiscriminating properties of *y*, if there is a undiscriminating unrepeatable property *F* such that *C* consists of all and only the undiscriminating properties necessitated by *F*, there is an object x_C coincident with *y* that has every property in *C* essentially and has no other undiscriminating property of *y* essentially.

[22] Where *F* is any property, let $F^\star$ be *being a point at which some F object is located and such that any two F objects are coincident*. By Location Plenitude, there is an object *x* which is necessarily located at all and only the $F^\star$ points. Since *x* is necessarily not located anywhere (and hence not coincident with anything) unless there exists some *F* object *z* which is coincident with every *F* object, it is necessary that an object *y* coincides with *x* iff it coincides with some and every *F* object. Given the current definition of 'coincides', Location Plenitude is equivalent to the conjunction of Coincidence Plenitude with the following "intra-world" weakening of Location Plenitude (which follows from Location Plenitude by CBF):

Coincidence Plenitude follows from Fundamental Part Plenitude. But for our purposes it does not matter how coincidence is defined, so long as we can take it for granted that it is necessarily symmetric and transitive, and that necessarily, any two coincident objects agree as regards their geometric properties and their geometric relations to other objects.[23]

Note that the inference from any of our plenitude principles to any claim about the *distinctness* of material objects, e.g. that every material object coincides with something distinct from it, depends on Leibniz's Law. If one were willing to jettison Leibniz's Law, one could then consider combining some version of plenitude with the view that it is impossible for distinct objects to coincide. One could even simply replace 'x coincides with y' with '$x = y$ and x is concrete' in Coincidence Plenitude, turning it into a general theory of contingent identity.[24] Indeed, there is a tradition (going back to Carnap 1947) of developing formal systems of quantified modal logic that invalidate Leibniz's Law while validating something like this variant of Coincidence Plenitude.[25] However, while formal sentences similar to

> For any property F, it is metaphysically necessary that there is an object x such that for any spacetime point y, x is located at x if and only if y is F.

For given any property F, Coincidence Plenitude gives us an object x such that necessarily, it coincides with an object y iff y is located at exactly the F points, but "intra-world" Location Plenitude entails that it is necessary that there is such an object y, so we can conclude that x is necessarily such as to coincide with something located at exactly the F points, and hence necessarily located at exactly the F points.

[23] Yablo (1987) unfolds a complex but rather uninformative hierarchy of definitions: coincidence (or "contingent identity") is having all the same "categorical" properties; categorical properties are those which cannot distinguish two objects when one is a "refinement" of the other; x is a "refinement" of y when every "cumulative" property necessary to y is necessary to x; and "cumulative" is undefined.

[24] Note that Coincidence Plenitude implies (given that there are contingent truths) that there are objects which could fail to coincide with themselves. So one cannot simply replace 'x coincides with y' with 'x is identical to y' unless one is giving up not only Leibniz's Law but the necessary reflexivity of identity. We could, however, weaken Coincidence Plenitude to require F to be both necessarily instantiated and unrepeatable.

[25] We are thinking of systems based on the so-called "intensional objects" model theory for first order modal logic described by Cresswell and Hughes (1996: 331) (see also Garson 2001: §1.3, Garson 2006 §13.7 and Hellie, Murray, and Wilson 2021). These models are equipped with a "domain of worlds" W and a "domain of individuals" D, but assignment functions map each individual variable not to a member of D, but to a function from W to D (an "individual concept"), with $x = y$ true at world w on assignment g iff $g(x)(w) = g(y)(w)$. We have the standard clauses for connectives, and clauses for quantifiers that make them range over all functions from W to D. Then for any formulae $C(x, y)$ (playing the role of coincidence) and $F(x)$, (v) (formalizing Coincidence Plenitude) is true at every world in every model where (i)–(iv) are true:

(i) $$\Box\forall x\forall y(C(x,y) \rightarrow C(y,x))$$

(ii) $$\Box\forall x\forall y\forall z(C(x,y) \wedge C(y,z) \rightarrow C(x,z))$$

(iii) $$\Box\forall x\forall y\forall z(C(x,y) \wedge y = z \rightarrow C(x,z))$$

(iv) $$\Box\exists x\neg C(x,x)$$

(v) $$\exists x\Box\forall y(C(y,x) \leftrightarrow \exists z(C(y,z) \wedge F(z) \wedge \forall z'(F(z') \rightarrow C(y,z'))))$$

To see why this argument valid in the model theory, note that for (i)–(iv) to be true at w on g, there must be a function C^* from W to symmetric and transitive but non-reflexive subsets of $D \times D$, such that when w' is accessible from w and g' differs from g at most on x and y, $C(x, y)$ is true at w' on g' iff $\langle g'(x)(w), g'(y)(w)\rangle \in C^*(w')$. (v) can then be witnessed by any individual concept i such that, for any w' accessible from w:

our plenitude principles fall out as a by-product in such systems, they are often interpreted as expressing uncontroversial claims about the unfamiliar category of *individual concepts*, rather than the controversial metaphysical theses about *objects* that they express when taken at face-value.

That said, Leibniz's Law is part of our background logic, so we will take it for granted that plenitude involves a commitment to an abundance of *distinct* coincident objects.

By providing some rigorously stated plenitude principles, this section has attempted to lay to rest any worry to the effect that the plenitudinous vision is too inchoate to be useful for rigorous theory-building, or that its "true spirit" is inconsistent with our other commitments. But while we have suggested that plenitude may help with Tolerance Puzzles, we have not spoken directly to its plausibility. It is to that issue that we now turn.

11.3 The Case for Plenitude

Why should we accept any version of plenitude? The role that plenitude can play in helping us arrive at a stable perspective on Tolerance Puzzles, as we discussed in §11.1 and will elaborate later, is certainly one count in its favour. But there are several other strong and independent motivations for plenitude.[26]

First: plenitude allows that there are many ways in which a linguistic community could talk about the world of material objects without falling into error. Consider,

a. If there is some individual concept h such that $F(z) \wedge \forall z'(Fz' \rightarrow C(z, z'))$ is true at w' on $g[z \mapsto h]$, then $i(w') = h(w')$ for some such h.
b. $\langle i(w), i(w)\rangle \notin C^*(w')$ otherwise.

The "individual concepts" model theory is generally developed in such a way that Leibniz's Law holds for *atomic* predicates, although it does not hold in general (because atomic predicates are interpreted model-theoretically by functions from worlds to sets of n-tuples of members of D, rather than of individual concepts) (see Garson 2006: 233; Hellie, Murray, and Wilson 2021: 4). This odd restriction seems to reflect some mysterious doctrine about the kinds of meanings that are eligible to be expressed by syntactically simple predicates; but as far as we know no such doctrine has ever been explicitly defended. When the quantifiers range unrestrictedly over all individual concepts, this has very strange consequences for the logic of atomic predications. For example, $\forall x(Fx \rightarrow \neg\Diamond Gx)$ can be true at a world in a model only if G is empty at every other world. Also $\forall x(Fx \rightarrow \Diamond Gx)$ can be true in a world in a model only if either $\forall x(Fx \rightarrow Gx)$ is true at that world, or else there is some other world where $\forall xGx$ is true. Thus the following are both valid (where F and G are atomic, and P can be any sentence):

$$\forall x(Fx \rightarrow \neg\Diamond Gx) \rightarrow (P \rightarrow \Box(\neg P \rightarrow \forall x\neg Gx))$$
$$\forall x(Fx \rightarrow \Diamond Gx) \rightarrow (\forall x(Fx \rightarrow Gx) \vee \Diamond\forall xGx)$$

This means that in such a view, there is no prospect for defending even mild quantified essentialist claims like 'All auditoriums are not possibly volcanoes', or mild quantified tolerance claims like 'Every 2m long table could have been 1.9m long'. While Hellie, Murray, and Wilson (2021) offer a treatment of *de re* Tolerance Puzzles which agrees with ours in some important respects, they do not discuss quantified Tolerance Puzzles.

[26] For further discussion of the case for plenitude, see Hawthorne 2006*e* and Fairchild and Hawthorne 2018.

for example, a counterfactual community where people unhesitatingly say things like 'The ship was destroyed' and 'The ship no longer exists' after a ship has sunk to the bottom of the ocean. It's not that they are unaware of the path of the wreckage: rather, they use words like 'wreckage' and 'ship' in a way akin to the way we use the words 'corpse' and 'cat'. Or consider a community where people use 'engine' and 'car' in the same sort of way that we use 'frame' and 'bicycle'—just as we would be inclined to say, 'I made my new bicycle by attaching the wheels of my old bicycle to a new frame' rather than 'I put a new frame in my old bicycle', they are inclined to say, 'I made my new car by putting a new engine inside the body of my old car' where we would say, 'I put a new engine in my old car'. Or consider a community where people use the word 'road' in a way that involves caring about originating matter in the way that we typically do with the word 'table'. They will say, 'If those shipments of tarmac and gravel had never been delivered, then instead of the A34, there would have been a different road following the same route' where we would say 'If those shipments had never been delivered, then the A34 would have been constructed out of entirely different tarmac and gravel.' It is hard to believe that these practices involve their participants in systematic errors where our practice escapes error. When considering variant practices that don't involve issues of ontology, this kind of charity is uncontroversial: who, for example, would be so chauvinistic as to insist that a counterfactual community that applies the verb 'drive' to bicycle-riding are just making a mistake when they say 'I drove to work today' after cycling to work? It seems to us that the same charitable instinct is also in order with regard to the more "ontological" variations of our actual practices.[27]

Even confining our attention to the actual community of English-speakers, one can argue for a rather wide range of modally and/or temporally dissimilar coincident objects just by considering the many respects in which our actual practice is quite flexible. There are a host of detailed questions about the careers of material objects where we can simply go either way: we sometimes talk in a way that takes one answer for granted, and sometimes talk in a way that takes some different answer for granted. Which which way of talking we adopt on a given occasion is often a matter of happenstance; once we have taken a particular presuppositional path, audiences will often go along with it without any internal resistance. For example, consider the question 'Was the house already around when only the foundations had been laid and one wall built?' We wouldn't bat an eye at the speech 'When the foreman died, the house only had one wall' or conversely at the speech 'The foreman never got to see the house, since he died when only one

[27] One might try to buttress this appeal to charity by appealing to epistemological considerations, to the effect that our true beliefs wouldn't count as knowledge unless the relevant counterfactual beliefs were true as well. Given the general lesson that "close" possibilities are more threatening to knowledge than "distant" ones, the best hope for this move involves looking at close counterfactual versions of ourselves, as discussed in the next paragraph. For more on these epistemological considerations see Fairchild and Hawthorne 2018.

wall had been built'. Or consider 'Could a sandcastle have been made of completely different sand?' People seem completely fine with 'If we had made this sandcastle out of the wetter sand from down near the water, it would have lasted a lot longer' (which suggests that the sandcastle is quite tolerant with respect to what sand it is composed of—cf. Strawson's discussion of the Old Bodleian as quoted in §5.3) but also with 'If the sand up here had been any drier, we couldn't have made this sandcastle' (which suggests that the sandcastle is rather intolerant in this respect). While writing this book, we have asked ourselves many questions about which we are inclined to take this flexible attitude. Could this table have had a different top and the same four legs? Could the knife have had a different blade? Would this lectern have been made of ice if the same protons, neutrons, and electrons had been assembled into water molecules but still been arranged according to the same plan by the same artisan? It is hard to believe that our flexibility with regard to these questions involves a systematic proclivity to error. But if none of our moods involves a special disposition to error, there must be a great many objects around with a range of different modal (and temporal) profiles, wide enough to accommodate all of the various perspectives we can adopt.[28]

One can thus get a fair way towards plenitude just by charity to our own community, or by simply reflecting on our own judgements in various ordinary contexts. Mild counterfactual variants of our community take us a good way further, and we can get a very long way if we are disposed to be charitable even towards exotic counterfactual communities.[29] But sooner or later, proponents of plenitude will need to make an abductive leap from case-by-case judgements of multiplicity to some generalization under which those judgements can be subsumed. To our mind, considerations of strength and simplicity speak so strongly in favour of plenitude that such an abductive inference would be well justified even from quite a modest diet of judgements about particular cases. Of course, some

[28] This isn't the only possible way of avoiding error: one could instead think that the different ways of talking involve different interpretations for expressions like 'could have been built of wetter sand', due, most plausibly, to variation in the interpretation of the modal words. Counterpart theory, as developed by Lewis (1986*a*: §4.5), provides a widely discussed vision of such variation might work, although one certainly need not be a counterpart theorist to posit contextual variability in modal operators (Kratzer 1977). It is worth remarking that those like Lewis, who attribute the kinds of flexibility discussed in this paragraph to variation in the modal operators, have tended to resist any case of permanent coincidence among distinct objects, for understandable reasons: it is hard to imagine a satisfying division of labour in which some of the flexibility is handled by appeal to variation as regards which of certain permanently coinciding but distinct objects we are referring to, while other cases are handled by appeal to variation in words like 'could'. Fine (2003) gives a powerful case in favour of certain cases of permanent coincidence, showing how hard it is to explain away the evidence by appeal to context-sensitivity in modal operators (and other relevant predicates). In Chapter 13, we will look at strategies for solving Tolerance Puzzles that appeal to semantic variability in modal operators in place of the plenitude-presupposing semantic plasticity that is appealed to in this chapter.

[29] For example, Hirsch (1976: 361) imagines a community with a word 'incar' which they use in a way that takes for granted the truth of 'An incar got smaller and smaller until it disappeared' in a circumstance where we would say, 'A car drove out of a garage'. See also van Inwagen (1990: 126) on "gollyswoggles", Sosa (1999: 136) on "snowdiscalls", and Hawthorne (2006*c*: vii) on islands.

input beyond these abductive considerations is needed, since a nihilistic ontology that countenances only metaphysical simples is also very simple and systematic. But an ontology that countenanced only objects for which humans have general terms but little else paints a picture of the world that will be altogether offensive to the systematic mind, or at least to those systematic minds that eschew the kind of idealism according to which to be is to be either thought about or spoken about.

Others, no doubt, have different sensibilities. They might be able to make their peace with the thought that, e.g., the Old Bodleian, the QE II, and the A34 are highly tolerant with regard to their originating matter, and not coincident with anything that is only moderately tolerant in that respect, whereas a typical mass-produced table, pork sausage, or mountain is moderately tolerant with respect to its originating matter and not coincident with anything highly tolerant. This kind of position can be thought of as a kind of "particularism" in ontology. Ethical particularists (e.g. Dancy 2014) urge us to eschew the search for general principles and to rest content with moral verdicts informed by vivid engagement with the details of particular cases. Likewise, the "ontological particularist" takes ontology (or whatever we call the branch of philosophy that includes debates about plenitude principles) to be a domain in which explanatory generalizations are not to be had, and abductive considerations of simplicity and strength carry little weight.

While there may be a good case to be made for particularism in art criticism and business ethics, it is utterly out of place in physics and mathematics. And, by our lights, ontology is a lot more like physics and mathematics than like art criticism or business ethics. After all, it is in a natural sense the most general of all the sciences: its objects of study include both those of physics and mathematics but are not exhausted by them—they also include the objects of study of art critics and business ethicists and of practitioners of all other fields of inquiry. A lack of systematicity is thus an important count against any ontological scheme.[30]

Plenitude principles like those of the previous section need not be the end of the line as far as the pursuit of explanatory generalizations is concerned. These principles can be derived from still stronger principles of various kinds, in ways that might be argued to offer a further explanatory advance. One such strengthening is to replace the claims about metaphysically necessary equivalence with claims about *higher-order identity*—a more demanding relation, if Intensionalism is false. For example, one could derive Location Plenitude from the principle that for every property F, there is an object x such that to be an F spacetime point is to be a spacetime point at which x is located. This claim will be rejected by proponents of the restricted Structure-like principles discussed in §7.4, which rule out "interesting" patterns of higher-order identity involving specific objects. But as

[30] Here we are largely recapitulating Fairchild and Hawthorne (2018: 74), who are responding to Korman (2015), who offers the most extended defence of something like the "particularist" outlook.

we discussed in that section, there are plenty of plausible counterexamples to such principles, and for those who accept the counterexamples, it is extremely natural to explain the modal plenitude principles by deriving them from the identity-strength plenitude principles.[31]

An even stronger principle we might consider (following a suggestion we considered at the end of §7.4) is that ordinary material objects are not in the strictest sense *objects* (or individuals) at all, but instead are *properties*. Or in higher-order terms: quantification over material objects is of type $\langle e \rangle$ rather than type e, just as in the idiolect of this book, quantification over "propositions" is of type $\langle\rangle$ rather than type e. On one of the ways in which we suggested fleshing this out, the spacetime points are the only really type-e things, and for a material object to be located at a spacetime point is simply for it to be *instantiated* by that point. According to this hypothesis, for Location Plenitude to be true is just for the following triviality to obtain: for every property F, there is a property G such that it is metaphysically necessary that for any point y, y is G if and only if y is F.[32] While we wish to remain neutral on this proposal, it serves to illustrate the kind of further explanatory progress that plenitude makes possible, but which seems completely off the table for those with a more particularistic bent.

Plenitude is apt to generate incredulous stares. For that reason, it is sometimes regarded as being in conflict with "common sense". Common sense is pictured as a mean between two extreme views that are both incompatible with it: views like plenitude depart from it in the direction of positing too many things, while views like mereological nihilism (according to which there are no composite objects at all) depart from it in the direction of positing too few things.[33] In conversation, some have accused us of being unprincipled in placing great weight on some commonsense beliefs, such as those that support Tolerance premises, while jettisoning the ones that conflict with plenitude. But it seems to us very misleading to lump the dispositions that lead people to reject plenitude together

[31] Identity-strength plenitude principles also suggest a route to a positive answer to the question whether there are distinct but necessarily coincident objects, based on the claim that there are distinct but necessarily coextensive properties (i.e. the denial of Intensionalism). Plausibly, if there are distinct but necessarily coextensive properties, there are cases where *being an F point* and *being a G point* are distinct but necessarily coextensive. Suppose this is the case: then according to identity-strength Location Plenitude, there are objects x and y such that to be a point at which x is located is to be an F point, while to be a point at which y is located is to be a G point. x and y are necessarily colocated. But they are distinct, since the assumption that $x = y$ would imply (by Ref and LL) that to be a point in x is to be a point in y, from which it would follow that to be an F point is to be a G point, contrary to our hypothesis.

By contrast, there is no straightforward route from Intensionalism and even identity-strength Location Plenitude to the view that there *can't* be distinct necessarily coincident objects, since it is unclear how we would get from 'To be a point at which x is located is to be a point at which y is located' to '$x = y$' without relying on Structure-like principles obviously inconsistent with Intensionalism.

[32] The identification also in the same sense "reduces" the question whether there are necessarily coincident objects to the question whether there are necessarily coextensive properties (i.e. whether Intensionalism is true at type $\langle e \rangle$).

[33] For this sort of framing, see Dorr and Rosen 2002 and Korman 2015.

with the ones that anchor the case for Tolerance under the single heading of "common sense". While people who haven't been presented with the arguments for plenitude might indeed default to rejecting various of its consequences when confronted with them for the first time, those dispositions are not very deeply entrenched. Indeed, in some cases, it is quite easy to get people to take back some initial anti-plenitudinous pronouncement. If the table is empty but for a Rubik's cube, someone might offhand say 'one' when asked 'How many cubical objects are on the table?', but when we say, 'What about the little cubes that are parts of the Rubik's cube?' they'll probably think we've caught them out. Someone might initially say 'No' in answer to 'Does anything have three heads and two long beards?', but when we say, 'What about ZZ Top?' they might feel similarly caught out.[34] The pro-Tolerance dispositions, by contrast, are far more integral to our ordinary practices. Even if in a philosophical moment someone was led to claim that, e.g., no table could have been even slightly different in its originating matter, it's a fair bet that after a hearty dinner and a game of backgammon, they'll be back to talking in the usual Tolerance-friendly ways. It is those entrenched beliefs, rather than the offhand judgements that people come up with when prompted in more philosophical or puzzle-like settings, that we regard as good starting points for philosophical inquiry. Of course, a resourceful philosopher who rejects the entrenched beliefs will be able to come up with all sorts of prima facie respectable stories to explain away their apparently inconsistent post-backgammon behaviour. But from our point of view, such philosophers are losing out on a crucial stock of knowledge, without which the prospects for navigating our puzzles are dim.[35]

11.4 Plenitude and Non-Contingency

Even all by itself, plenitude has some power to demystify the puzzles of tolerance. Given plenitude, we don't have to posit any *natural boundaries* in a series of pairwise similar possible worlds starting with the actual world and ending in some

[34] We don't mean to suggest that 'have' obviously has to be interpreted as expressing a relation that ZZ Top bears to its members' beards.

[35] These remarks about plenitude also apply to another of our commitments which some might accuse of "violating common sense", namely necessitism (§1.4). While people coming to the question for the first time can easily be led to espouse contingentism on being presented with the basic pro-contingentist arguments, they can also easily be led to espouse necessitism by going through arguments like 'You couldn't have been distinct from yourself, so you couldn't have been distinct from everything, so it couldn't have been that you were not identical to anything.' And necessitism isn't the sort of doctrine that is hard to stick to when mingling with the folk after backgammon. (Remember that necessitists have good reason to think that there is at least one good interpretation of 'exists'—e.g. concreteness—on which it is contingent whether we exist.) Indeed, necessitists are going to have an easier time mingling than many contingentists, as witnessed by contingentists' continued interest in the project of finding inventive ways of salvaging ordinary claims along the lines of 'There are four possible knives that could be made from these two handles and two blades' (see §1.4).

world where Woody is never created, but some other table is created instead out of different matter. At every step along the series, we lose some objects and gain some objects. Unless Woody is hypertolerant, some step takes us from a world where Woody was made to one where it wasn't. But that doesn't make that step special, since Woody does not enjoy any special natural status that distinguishes it from all the other coincident objects that disappear at other steps. We are free to say that at the actual world, Woody is "at its sweet spot"—at the centre of its zone of tolerance—without thereby having to think of the actual world as "metaphysically special". Other nearby possible worlds also have table-shaped objects that are at their sweet spots, and those objects are no more or less metaphysically important than Woody. These worlds are not any more metaphysically disorderly than the actual world.[36]

On a non-plenitudinous view, by contrast, the hypothesis that Woody is tolerant but not hypertolerant seems to require the existence of one or more inscrutable, magical-seeming natural boundaries in the series, where new objects pop up or old ones disappear. Such boundaries will be objectionable to those who take fundamental physics to be an adequate guide to the natural joints in the space of nearby nomological possibilities. Resistance to positing such boundaries can make the Sorites-style argument from Tolerance to Hypertolerance gripping even to those who have learnt some discipline for resisting the most familiar Sorites arguments for conclusions like 'Everyone is bald' or 'No-one is bald': there is no pressure to think that the step from baldness to non-baldness would have to be a *natural* boundary in the Sorites sequence, since there is no reason to think that baldness is any more natural than other, slightly more or less demanding, hair-distribution properties. Once we have plenitude to work with, parity between the two kinds of Sorites series is restored.[37]

[36] Note that this thought would be blocked if we were considering a Tolerance Puzzle involving a *fundamental* object instead of a table, assuming that there isn't an abundance of coincident fundamental objects whenever there is one. Tolerance Puzzles thus generate interesting pressure in a Hypertolerant (and thus combinatorialist) direction as regards the modal profiles of fundamental objects. And the supervenience-based objections to Hypertolerance from Chapter 6 don't generate any countervailing pressure, since the reasons for regarding an object as fundamental will also be reasons to doubt any supervenience thesis that would make trouble for Hypertolerance. However, there is also less pressure to accept Tolerance since questions about the modal properties of plausibly fundamental objects are pretty remote from ordinary thought—e.g. Tolerance thoughts to the effect that any two spacetime points could be somewhat further apart are neither particularly compelling nor particularly repugnant.

[37] Sider (2002: 9) argues that the kind of intra-world plenitude that he endorses under the name 'four-dimensionalism' can help with the Ship of Theseus puzzle, although he is careful not to suggest that it by itself answers the relevant questions about ship-identity. He operates with a contrast between "metaphysical" and "conceptual" questions, and says that four-dimensionalism resolves the metaphysical puzzle raised by the case: 'the only remaining question is the merely conceptual one of which of these worms counts as a ship'. We are dubious about the metaphysical/conceptual contrast that Sider is invoking here: after all, any question can be reworded as a question about the application of some concept, using a concept-theoretic analogue of semantic ascent. Still, we agree that Sider is on to something: there seems to be some sense in which the debate between two plenitude-lovers who hold different views about the fate of the ship by the end of the story is less "deep" than some other paradigmatic metaphysical debates. This isn't the place for a worked-out account of relative

But as we emphasized in Chapter 3, arguments for Hypertolerance that rely on Sorites-like thoughts about the insignificance of small changes are, though initially tempting, philosophically weak. And the thought that "the actual world isn't special" was already too inchoate to do much heavy lifting as an argument for Non-contingency. As we see things, the most important argument for Non-contingency is the Security Argument of §3.3. Our response to that argument is to reject the Independence premise, by claiming that when we say 'That is tolerant' at nearby worlds where the objects we are *actually* referring to aren't tolerant, we refer not to those things, but to other, tolerant, things. Plenitude isn't strictly speaking *required* for this response; its role is to help us to get an explanatory grip on the many specific ontological claims that are required, by subsuming them under a simple and systematic generalization.

The Independence-denying response doesn't just require the ontological claim that there are appropriately tolerant objects at all the relevant nearby worlds, but the metasemantic claim that when we deploy demonstratives in Tolerance speeches in those worlds, we refer to those tolerant objects rather than to the intolerant objects coincident with them. While this claim does not require one to endorse any more general theory of metasemantics, one might hope to enhance our explanatory grip on the response by subsuming it too under some simple and systematic generalization. It is hard to do quite so well here: metasemantics is a challenging field, where simple generalizations tend to have a short lifespan.[38] However, in the case at hand, one can get a fair amount of explanatory purchase from the following rough-and-ready generalization: *ceteris paribus*, the referents of people's words tend to be such as to make sentences that they presuppose, or implicitly treat at obvious, come out true.[39] But as we have observed, we do in fact have strong dispositions to talk in ways that take tolerance for granted, presupposing that objects are capable of being moderately different from how they actually are in a wide variety of ways in the course of *asserting* various more

depth (a concept that Sider explores further in Sider 2011). But plausibly, part of the contrast is a matter of locality: the dispute will tend not to reverberate into other areas of theoretical inquiry, so that deciding not to care about it would be less of an abdication of philosophical responsibility. Whereas for deniers of plenitude the dispute may be tied up with general principles like 'No object can get from one place to another without following a continuous path', proponents of plenitude will see no prospect for generalizations of that sort statable without invoking some restricting predicate like 'ship' or 'artefact' (see Hawthorne 2006*d*). Another example: plenitude removes the pressure to explain the fact that roads don't have an essence involving originating matter whereas sausages do in terms of some more general principle, statable without using the words 'road' and 'sausage'. That is not to say that plenitude is the only perspective that keeps the relevant disputes relatively "local" or "shallow": if the "particularist" attitude we discussed in §11.3 could be satisfactorily articulated, it might similarly offer some comfort to those who take a "don't-care" attitude towards questions about the modal and temporal robustness of particular kinds of artefacts.

[38] Consider generalizations like 'A demonstrative accompanied by a pointing finger refers to a material object *x* just in case *x* is in line with the finger and there is no other material object in between.' Such generalizations are hopeless given plenitude, and indeed given far weaker thoughts about the variety of material objects: recall the example of the hemisphere and the sphere from note 4.

[39] By contrast imputing falsehood to sentences people *assert* is less of a sign of interpretative failure, since a good interpretation must also avoid imputing uninformativeness.

specific things (e.g. that some particular change would be enough to get the table to fit through a certain door). So given the "charity to presuppositions" maxim, demonstratives should gravitate to tolerant objects. Of course, the maxim is only a *ceteris paribus* generalization. People can, through sensory abnormalities or blunders in reasoning, find themselves not just believing falsehoods, but treating them as obvious common ground. But when fairly natural interpretations that avoid imputing such error are available, one would require some special reason to opt for an imputation of falsehood instead. Even without a fully developed grip on the factors that can make *ceteris* not be *paribus*, we can see that given the extreme similarity between the relevant nearby worlds and the actual world, it is altogether implausible to suppose that an exception to charity is warranted there but not here. And since we know that a wide variety of Tolerance-suggestive claims are true, we know that there is no special reason for the semantic gods to impute error in our own case.[40]

11.5 Plenitude and Common Nouns

So far in this chapter we have been focusing on Tolerance Puzzles stated using demonstratives and proper names. But the distinctive behaviour of demonstratives is neither here nor there when we are dealing with quantificational versions of the puzzles (§2.2). It is to these that we now turn.

The first thing we need to be very clear about is that while (for example) plenitude implies that every table coincides with innumerable other *objects*, it does not imply that every table coincides with many other *tables*. Indeed, it is perfectly compatible with the thesis that no two tables coincide. One might find the combination

[40] Is there some competing vision of metasemantics that would be more hospitable to Tolerance-denial? One might imagine a view that places less of a premium on charity to things taken for granted in ordinary life, instead emphasizing charity to speaker's initial dispositions to respond to philosophical or scientific prompting. For example, people have some disposition to say, 'If the atomic theory of matter is true, then nothing is solid' when they are first exposed to that theory. If one thought it was very important to avoid interpreting dispositions like this one as dispositions to make mistakes, one might be led to a view on which our ordinary practice of applying 'solid' to objects around us is shot through with error. And one could end up saying the same thing about our ordinary Tolerance-friendly practices, if one conjectured that ordinary people would be less confident in conditionals with Tolerance-friendly consequents and, e.g., some form of plenitude in the antecedent.

We take ourselves to know lots of things in conflict with this metasemantic view, such as that tables are solid, so we take ourselves to have decisive evidence against it. But even bracketing all this knowledge, we doubt there is really any stable view here, since people's dispositions when asked to draw out the consequences of views they hadn't thought about before are so shifty and variable that no interpretation could avoid conflict with many of them. Why should people's first halting attempts to do philosophy count for very much at all? One might respond that what really matters are people's dispositions to accept sentences after some kind of idealized rational reflection. Insofar as we have any grip on what idealized rational reflection is supposed to be, it seems plausible that it tends to lead to the utterance of true sentences. But the claim that this is the case is not in competition with our favoured rough-and-ready generalization about metasemantics, since we of course reckon that idealized rational reflection will involve roughly copying the content of this book. (Thanks to Jack Spencer for helpful discussion on this point.)

of plenitude with this thesis mysterious, even if one is willing to grant that some non-tables (such as quantities of wood) coincide with tables. How could our use of the word 'table' be so nuanced as to single out a property possessed by just one of the myriad table-shaped things in the vicinity, given that each is coincident with things with incredibly similar modal profiles? But if there is a mystery here, it has nothing special to do with plenitude. Even leaving *exact* coincidence aside, there remains the fact that wherever there is a table, there are innumerable *approximately* coincident, and otherwise very similar, material objects with slightly different boundaries: starting with any one of the material objects in the vicinity, we can find more by adding or subtracting a few of the atoms around its boundary. Here again one can can make it seem puzzling how our use of the word 'table' could apply to just one member of this vast collection of very similar candidates: this is the basis of the widely discussed "problem of the many".[41] Indeed, the challenge has nothing special to do with reference to *objects*; it can be raised for any vague term. Baldness is just one of many properties which differ from baldness in extremely subtle ways: how can our use of the word 'bald' be so nuanced as to lock on to the property of baldness rather than any of these neighbouring properties? It is quite understandable that someone would find this puzzling, and we don't want to suggest that such puzzlement could only arise from some confusion.[42] But we don't think that the right way of resolving one's puzzlement involves giving up the obvious truth that at most one of the properties that are very similar to baldness is identical to baldness! And whatever techniques we may appeal to in explaining why that trivial claim is not undermined by worries about the putative difficulty of "locking on" to a single property, we can use the same techniques to explain why the claim that just one of the many objects that are very similar to any given table is a table is not undermined by parallel worries about 'table'.

There is thus nothing desperately unstable about combining plenitude with the view that at most one of the objects in any given region is a table. Still, a "many tables" view is an option the plenitude-lover might consider, and indeed is the primary option defended by Leslie (2011) in connection with puzzles of modal variation similar to our Tolerance Puzzles. Leslie focuses on arguments whose premises involve definite descriptions, along the lines of 'The axe before us is tolerant'. Her response is that such arguments fall at the first hurdle, because the

[41] See Geach 1962, Unger 1980, Lewis 1993, McGee and McLaughlin 2000, and Weatherson 2003, 2016.

[42] Nor do we want to take it for granted that we *do* in any sense "lock on" to exactly one of the candidate properties. The tradition of "supervaluationism" (Fine 1975; Lewis 1982; McGee and McLaughlin 1995; Dorr 2003) goes along with the idea that, at least in one important sense, the word 'bald' does not "lock on" to one candidate at the expense of the rest, despite the fact that only one of candidates is the property of being bald. On the "plural signification" way of developing this idea, discussed in Dorr and Hawthorne 2014 (5), in §13.2 below, and in Dorr unpublished b, the word does not even uniquely *refer to* or *express* any one of the candidates (including the property of being bald). Everything we say is intended to be compatible with these views of vagueness.

definite description carries a faulty presupposition of uniqueness. The paradoxes "take us in because of an illusion of singularity" (Leslie 2011: 287).[43]

One glaring worry for the "many tables" approach, which we already discussed in §4.3, is the way it leads to rampant error in ordinary counting speeches: 'I made exactly three tables today', 'There is only one table in my office', and so on. Another central concern is that the approach involves giving up on many compelling quantified Tolerance premises. For example, as we have discussed, there is a reasonably direct route from some fairly ordinary judgements to the conclusion that every table is tolerantly a table (using one of the origin-theoretic closeness relations from §2.4). But on Leslie's view this is false: in fact, every table is coincident with innumerable tables that are living on the edge. And the problem extends to more localized quantified tolerance claims—e.g. claims where the quantifier is restricted to tables in a certain workshop, in a certain room, made by a certain person, and so on—which are even harder to resist without jettisoning large swathes of our ordinary modal discourse. Consider speeches like 'If that shipment hadn't been delayed, every table we made this week would have been veneered in walnut and hence a bit more valuable than it actually is', or 'Since that sander is so hard to handle, each of the tables we made using it could easily have ended up unsaleable thanks to an uneven surface.' On Leslie's view, it looks like such speeches will be systematically false: for example, among the many intolerant tables we made this week, there will be some whose intolerance is of such a sort that they could not have been made at all with the walnut veneer in place of their actual veneer.

In response to both of these worries, one could posit pervasive tacit quantifier domain restriction, claiming that the domains of ordinary quantifiers typically include only tolerant tables and include at most one of any pair of coincident tables.[44] But as we explained in §4.3, we are sceptical about such claims. Unrestricted quantification is not some esoteric achievement that only philosophers can aspire to: it is pretty easy to get non-philosophers to grasp questions involving quantification without any tacit restriction, and by any reasonable diagnostic, a variety of intuitive counting sentences as well as quantified Tolerance premises can be seen to involve unrestricted interpretations of the quantifiers.[45] We conclude

[43] As we noted earlier (note 26 in Chapter 4), Leslie includes a couple of parenthetical remarks that suggest that she is indifferent to the question whether there is co-location of axes and ships or merely of 'axe-like entities' and 'ship-like entities'. But her way of blocking the relevant Tolerance Arguments seems to depend crucially on coincidence of axes and ships, not merely axe-like and ship-like entities. For no matter how many axe-like entities might be before us, so long as only one of them is an *axe*, the definite description 'the axe before us' will not induce any presupposition-failure, and our temptation to accept Tolerance premises involving it cannot be attributed to an "illusion of singularity".

[44] §4.3 also critically discusses another move which would help with the counting worry but would not rescue quantified Tolerance claims, namely the idea that often "we do not count by identity".

[45] As we will discuss in the next chapter, many of the central objections to our view carry over to the domain-restricting approach. The contrast between the two views is any case much diminished if, as Stanley and Szabó (2000) propose, contextual quantifier restriction is implemented compositionally by strengthening the meanings of nouns (or of an expression whose only phonologically realized part is a noun).

that the best option is to stick with a view where, despite the truth of plenitude, ordinary tables do not coincide with, or even largely overlap, any other *tables*. Similar generalizations will be true for the properties expressed by many other common nouns.[46]

Yablo (forthcoming) also develops a plenitude-based approach to Tolerance Puzzles which resembles ours in certain respects. On the view he seems to favour—adapted to be about tables and atoms rather than ships and planks—ordinary tables don't coincide with other tables, and they are tolerant but not hypertolerant with respect to their originating matter. However, the view still does not not accept the quantified Tolerance premise that every table is *tolerantly a table*. Rather, although a given table could have been originally composed by any collection of atoms sufficiently overlapping its actual originating atoms, not all of these collections are such that it could have been *a table* originally composed by them. Indeed, the tablehood of any given table is extremely fragile: even tiny variations in the selection of originating matter would stop it from being a table, although they would not stop it from coinciding with a table. Yablo's picture is that being a table requires being a table-like object that is "at its sweet spot" (as he put it in conversation); at the centre of its "zone of tolerance". On a toy model adapted from his theory of ships, to be a table is to be an object that it necessarily table-shaped if concrete, and composed of some collection of atoms *C* such that it could have been composed of any collection of atoms overlapping *C* by 97 per cent or more, and could not have been composed by any other collection of atoms. Thus replacing even one of the original atoms of a table would take us to a world where it is no longer a table (although it is still concrete), since in that world there are some collections of atoms of which it could not have been composed even though they include more than 97 per cent of the atoms that composed it there.[47]

On this account, the proposition that *every table is tolerant* is thus true in all nearby worlds; indeed, Yablo is sympathetic to the idea that it is necessarily true. So there is no need to appeal to semantic plasticity in the sentence 'Every table

[46] Just to be clear, we mean that *in fact* ordinary tables do not coincide with other tables; in §12.3 we will consider the modal status of this generalization. We say 'ordinary' in order to remain neutral about certain exceptional cases in which it has been argued that some ordinary common noun applies to two coincident objects, such as the letters in Fine 2000 and the road-signs in Johnston (2002). It's hard to do anything exactly similar with tables, but it's not so hard to think of examples where it's arguable that tables temporarily coincide at certain times. The NYU Philosophy Department's seminar room table is composed of five smaller tables: perhaps if four of the smaller tables are going to be destroyed in successive fires, we will finally be left with a situation where the seminar room table coincides with another table that was previously one of its constituents (cf. the forests in Chapter 4).

[47] Yablo misleadingly characterizes an object's distance from the centre of its zone of tolerance in terms of how much replacement is possible, suggesting that once we get to the edge of the zone, all of its *original* parts are irreplaceable. This isn't quite right: on his model, at a world where the *Santa Maria* is made from the first ninety-seven of its actual planks and three other planks, it is still possible to replace one of those first ninety-seven planks by some other plank that is not part of it *there*, but is part of it at the actual world. See note 38 in Chapter 2 for an abstract statement of a "Uniform Tolerance" principle that could be plugged in to a theory that imposes a "sweet spot" requirement on being a table, and some reasons to worry about it.

is tolerant' to explain why we are in no danger of error when we utter it. The proposition that *this* is tolerant (pointing to a table) is true in the actual world but false in some nearby worlds. But here Yablo can appeal, like us, to plasticity in the demonstrative: we are in no danger of error when we say, 'This is tolerant', because the demonstrative in each nearby world refers to the table in that world. (This also explains why there is no danger of error in saying 'This is a table' despite the extreme fragility of the proposition expressed.) Finally, the proposition that every table is *tolerantly a table* is simply false, even in the actual world.[48]

We do not think that Yablo's approach does enough to accommodate ordinary modal judgements. In our ordinary talk about tables, we often presuppose that their *tablehood*, and not merely their concrete existence, is modally pretty robust. Consider 'Since that sander is so hard to handle, there was a good chance that this would be the worst-finished table we made today', or 'If we had had time to apply the veneer, this would have been the most valuable table we ever made', and so on. An error-theoretic account of such judgements is unattractive, and looks particularly unprincipled if one is not willing to be error-theoretic about other ordinary tolerance judgements[49,50]

We conclude that the the best way of deploying the resources of plenitude in connection with typical quantified Tolerance Puzzles is to endorse Tolerance—e.g. by accepting that all tables are tolerantly tables.[51] So unless we are up for

[48] The account that have been describing is one of two accounts that are favourably discussed by Yablo; it is the one that we get if we identify *being a ship* with the property Yablo calls 'being a **SHIP**'. The other option he mentions is to identify *being a ship* with *being a* 𝕊ℍ𝕀ℙ, where a 𝕊ℍ𝕀ℙ is a ship-like object that is possibly a **SHIP**. This delivers something in the vicinity of Leslie's view, since every 𝕊ℍ𝕀ℙ is coincident with immensely many other 𝕊ℍ𝕀ℙs.

[49] Yablo's approach will lead to especially odd consequences if it is generalized to temporal puzzles. Are we to suppose that although a certain painting will still be around after a dab or two, it will no longer be a painting? Or that when the language we currently speak first came to be spoken, it was not yet a language? Of course, Yablo could say something else about some of the temporal puzzles, e.g. following the popular Hypertolerant treatment of the Ship of Theseus. But such a dramatically different treatment of the modal and temporal puzzles leads to some odd surprises. For example, consider possibilities involving things being larger than they actually are, thanks to the addition of extra matter. Presumably Yablo would hold that if extra plastic had been added at the time of the manufacture of some actually thin bucket, that bucket would have been fat and bucket-shaped, but would not have been a bucket. But it would be odd to deny that if we add the extra plastic a little after that time the result would be that the bucket would become fatter while continuing to be a bucket. And it would also be odd to think that it makes all the difference whether the plastic is added during the manufacture or a bit later—this will give a strange kind of power to intuitively irrelevant questions about where to draw the boundary between the time of manufacture and add-ons after the manufacture.

[50] It's a good question what the analogue of Yablo's approach for the chess-to-Twister puzzle would be. Is it that in nearby worlds with slightly different rules, people still play chess but chess is not a game? Or is it that in those worlds, chess is invented but never played?

[51] As we acknowledged in §2.2, there is space for a view on which there probably are counterexamples to quantified Tolerance premises like 'All tables are tolerantly tables' (or even 'All Melltorp tables are tolerantly Melltorp tables'), but such counterexamples are rare enough that the habit of presupposing *specific* tables to be tolerant in various respects still gets to be generally truth-conducive, and perhaps even knowledge-conducive. But note that in the presence of plenitude, this already strange view has even less going for it. According to plenitude, there are many properties T whose extension includes only things coincident with tables and that are tolerantly instances of T: given that such properties are available, it is hard to see why 'table' shouldn't express one of them.

Hypertolerance, we will need to combat the Security Argument for Non-contingency by positing semantic plasticity not only for demonstratives and names but for a wide range of *predicates*, including 'table'.[52] The proposition that every table is tolerantly a table is only contingently true. Indeed, for the reasons we discussed in Chapter 3, the chance of its being true was plausibly quite low, given that each and every episode of table-construction had a nonzero and approximately independent chance of producing a counterexample. Nevertheless, there was no danger of our making a mistake by uttering the sentence 'Every table is tolerantly a table', since at every nearby world *w* where we utter that sentence, we use 'table' to pick out some property *T* such that, at *w*, everything *T* is tolerantly *T*.[53]

Even for those who are comfortable with fine-grained semantic plasticity for demonstratives, the idea that general words like 'table' and 'bucket' are subject to the same kind of plasticity may take some getting used to. Under the influence of Kaplan (1989), many philosophers work with a picture on which demonstratives belong to a special category of 'indexical' words, a category to which most words don't belong. Relative to any language-community, each word is associated with a "character" which is a function that maps things called "contexts" onto entities of the right kind to be referents for that word. Each particular use of a word is "in" a specific context, and refers to whatever the word's character maps that context to. Indexicals have non-constant characters, so that uses of them may vary in reference. By contrast, non-indexicals have constant characters, so that within a single language, every use of a non-indexical refers to the same thing. Within this picture, the semantic plasticity of indexicals is not so surprising, since it is compatible with their having the same *character* at all nearby worlds. By contrast, securing semantic plasticity for a word like 'table' will require one of two moves: either we classify them as indexicals (which might leave the category of non-indexicals rather sparse), or we say that they could very easily have had a different character.

While the bare bones of our approach could be incorporated into the Kaplanian machinery in either way, the already-noted flexibility in our use of many of the relevant common nouns even within the actual world favours the approach that classifies them as indexicals. Recall, for example, that in one mood one can say, 'This sandcastle would have lasted for longer if we had made it from the sand down near the water', while in another, one can say, 'We couldn't have made this sandcastle if this sand had been any drier.' In general, we seem quite willing to

[52] Williamson (1990: 134, 141) likewise suggests that the "indexical component" that helps reduce the indeterminacy in uses of 'that ear-ring', 'that drawing', and 'that language' is present already in the nouns 'ear-ring', 'drawing', and 'language'.

[53] An exactly parallel response is available in the case of the Robustness Argument from Chapter 9. The proposition we actually express by the Robustness premise has a chance—indeed plausibly quite a high chance—of being false; nevertheless, there is no chance of our expressing a false belief by uttering the sentence in the relevant setting.

adjust the interpretation of common nouns in order to fit different presuppositions about the modal profiles of the objects falling under them. Given how many English-speakers there are, it is a fair bet that even over a short period of time in the actual world there is divergence in, say, whether the identity of the artisan is being treated as important for the identity of things being classified as 'tables'. If one insists on a constant character for 'table', one will end up fragmenting the community at the time into a multitude of languages.

However, by contrast with Kaplan's paradigms 'I' and 'today', there is no prospect of a picture where the non-constant character of 'table' corresponds to a conventional rule that one must internalize in order to become a competent user of the word. Like most context-dependent expressions, we will need to assign it a character whose value at a context depends in a very complicated way on the details of the intentions and other psychological states of the agent (and perhaps also other conversational participants) of that context, the goals of the conversation, etc.[54,55]

However it is implemented in the language of semantic values, the picture we are promoting allows that a small difference in a single location—say, a single craftsman's picking a different screw while manufacturing a table—would be enough to make our uses of 'table' refer to a different property. But it is hard to see how this could be an objection it itself, since there is no prospect of a view on which tiny, localized differences can never make a difference to the semantic values of words like 'table'. Uncontroversially, 'table' could have meant something very different—for example, if we had used it in the right way, we would have referred to a property that we actually refer to with 'chair'. But we can get from the actual world to a world where 'table' has that semantic profile by following a series of worlds in which every element differs from its predecessor only by some small,

[54] For some misgivings about the usefulness of the Kaplanian ideology in such cases, see Dorr 2014*b* (§7).

[55] A popular way of implementing context-sensitivity—one that we have no particular objection to—uses the apparatus of *assignment-functions*, and takes context-sensitive expressions to work semantically like variables: their denotation varies from one assignment to another, and just is whatever the assignment maps them to. Most simply, one might simply take a putatively context-sensitive word ('tall', 'ready', 'mistake', 'heap', 'table' . . .), to be a variable, just like '*x*' in a formal language, and leave it at that. However, this doesn't look so plausible since it washes out the semantic differences between different context-sensitive words in the same semantic category (e.g. 'tall' and 'heavy'). One standard way to restore such differences appeals to "presuppositional constraints" which restrict the range of assignment functions relative to which a word has a denotation at all. For example, 'she' is denotationless relative to assignment functions that do not map it to something female. 'Tall' might likewise be denotationless relative to assignment functions that don't map it to a property having to do with height in the right way, and 'table' might be denotationless relative to assignment functions that don't map it to one of some broad family of 'tablehood-like' properties.

Rather than treating the relevant audible word as itself a variable, theorists often instead posit *unpronounced* variables as companions to the audible word, responsible for the assignment-relativity of larger expressions containing it. While there may be a technical case for this when it comes to certain specific kinds of context-sensitive words (like gradable adjectives and relational nouns like 'enemy'), it seems often to be no more than an aesthetic preference. When we say things like '"Table" is context-sensitive', we do not mean to be taking a stand on such questions of deep syntax.

localized difference. At least one of the steps along this path, and likely many steps, has to make a difference to the semantic facts despite being small and localized.

One might object to the fact that the differences that turn out to matter on our view can include differences that are entirely external to the heads of agents. But it's obvious that such factors can matter to the reference of names and demonstratives (William of Ockham 1317–8: II, q. 16, cited in Brower-Toland 2007: 8). And as we have learned from Putnam (1975) and others, external factors are relevant for a wide range of other words, such as 'water'. Is there any principled reason to resist the view that 'table' is like 'water' in the relevant respect? David Chalmers has argued that there is an important category of "Non-Twin-Earthable" words (or utterances of words), whose reference supervenes on head-internal goings-on.[56] Chalmers's candidates for Non-Twin-Earthability include 'and', 'zero', 'philosopher', 'conscious' (Chalmers 2012*a*: 318), 'friend', 'computer' (Chalmers 2005: 473), 'bachelor', 'action', and 'cause' (Chalmers 2012*b*). Our view implies that Non-Twin-Earthability is rarer than Chalmers takes it to be: for example, our argument that 'table' is semantically plastic in ways that render it Twin-Earthable carries smoothly over to 'computer'.[57] However, nothing Chalmers says in support of the specific controversial claims of Non-Twin-Earthability seems to present any particular challenges to our view, which could be implemented within the general framework Chalmers favours.[58]

A more controversial externalist prediction of our view is that developments in the future, such as random choices of planks that will be made tomorrow, can make a difference to the *current* reference of words like 'table'. For these words, meaning for these words not only fails to be in the head, but fails to be in the present plus the past; facts about what property we refer to with 'table' at *t* may have nontrivial objective chances at and after *t*. We will need such "temporal externalism" in

[56] See Chalmers 2012*a* (ch. 7–8) and Chalmers 2012*b*; cf. the closely related notion of "semantic stability" in Bealer (1996). To be more precise: '*E* is Twin-Earthable if there can be a nondeferential utterance of *E* for which there is a possible corresponding utterance by a twin speaker with a different extension' (Chalmers 2012*a*: 317), where in the case of predicates, 'extension' (non-standardly) means 'property expressed', and "twins" are "functional and phenomenal duplicates". The restriction to "nondeferential" utterances is motivated on the grounds that without it, examples in the style of Burge (1979) would make almost every expression count as Twin-Earthable. Yli-Vakkuri and Hawthorne (2018) discuss several challenges in spelling all this out, including some raised by plenitude itself.

[57] Tolerance Puzzles can also be used to motivate Twin-Earthability for some of the other words on Chalmers's list; for example, one could consider a Tolerance Argument based on the premise 'Every action could have been performed somewhat later than it was in fact performed'. Words connected to agency and personhood (including 'friend', 'philosopher', 'bachelor' 'action', and 'experience') raise distinctive issues which we will discuss in Chapter 13. One could try to avoid Twin-Earthability by claiming that the relevant utterances of words like 'computer' count as "deferential" and hence irrelevant to the question whether the words are Twin-Earthable (see previous note). Engaging with this suggestion would require a deeper dive into Chalmers's discussion of deferentiality than we have space for here.

[58] In particular, one can easily implement the semantic plasticity we posit for 'table' by assigning it a two-dimensional semantic value, where the "secondary intension" varies depending on which world is "taken as actual".

order to avoid security worries for future-looking tolerance speeches along the lines of 'Every table that you make tomorrow will be tolerantly a table'. If what we meant by such a speech was fixed by the past and the present, then at the very many nearby worlds that match with respect to the past and present but where intolerant tables are made tomorrow, such speeches would be mistaken. And it would be unfortunate to concede security failure for such speeches, since our ordinary practice displays no tendency to treat future tables any differently from present and past ones.

Temporal externalism is not unprecedented. The 'actually' operator and the name 'Newman-1', stipulated by Kaplan (1968) to refer to the first person to be born in the twenty-second century, also make a good case against the principle that meaning is in general fixed by the present and the past. But unlike those examples, our argument suggests that temporal externalism is very widespread. For reasons given by Brown (2000), previous arguments for widespread temporal externalism (Jackman 1999, 2005; Collins 2006) are not particularly convincing.[59] Our discussion of Tolerance Puzzles puts it on a much firmer footing.

In conclusion, while systematic plasticity in common nouns takes more getting used to than systematic plasticity in demonstratives and names, we see little basis for a split decision. By postulating plasticity in both cases, we can smoothly sustain a Non-contingency-denying treatment of Tolerance Puzzles. When combined with the arguments against the competing treatments of the puzzles that we gave in other chapters, we thus take ourselves to have a very powerful argument for widespread plasticity, and thus for the kind of plenitude such plasticity requires. Of course, we do not expect the argument to persuade everyone. Chapter 13 will address what we see as the two most important challenges to our approach: the worry that the posited form of plasticity will disrupt some large range of speech reports, and the worry that extending our treatment of 'that' and 'table' to 'I' and 'person' will conflict with putatively obvious truths about ourselves. This chapter will also consider some alternative strategies that also rely on plasticity but locate it in different vocabulary. But first, Chapter 12 will develop our approach further, and consider some further choice points which are left open by our basic treatment of Tolerance Puzzles.

[59] In the central contested examples, opponents like Brown postulate meaning-change whereby an initially vague word is later precisified, whereas proponents maintain that the later use-facts make it true that the word already had the precise meaning at the earlier times. The proponents' arguments suggest that meaning-change should be rare, a conclusion we are sceptical of for reasons discussed in Dorr and Hawthorne 2014. The "plural signification" strategy (see §13.2 below) provides an account of the propriety of retrospective speech reports that does not require temporal externalism.

12
Refinements and Choice Points

In the last chapter, we saw how semantic plasticity enabled by plenitude can sustain an approach to Tolerance Puzzles in which Tolerance, Iteration, Persistent Closeness, and Non-hypertolerance are accepted as unequivocally true, while Non-contingency is rejected as unequivocally false. But once plenitude and plasticity are in play, there are some other interesting ways of deploying them. In §12.1, we first explore the prospects for using these resources to shore up a Hypertolerance-embracing treatment of the puzzles, and conclude that the best approach of all is one allows for both hypertolerant and non-hypertolerant contexts. In §12.2 we explore how plenitude and plasticity can be used to resolve the Coincidence Puzzles we introduced in Chapter 4. §12.3 considers some open questions for our view about the sorts of properties liable to be expressed by common nouns like 'table'. §12.4 will briefly consider some questions about other predicates.

12.1 Plenitude, Plasticity, and Hypertolerance

As we saw in Chapter 6, friends of Microphysical Supervenience cannot embrace Hypertolerance across the board as a general strategy for handling Tolerance Puzzles: Microphysical Supervenience rules out Hypertolerance for certain families of extremely fine-grained properties, each of which is fully specific with regard to the microphysical lay of the land. But Microphysical Supervenience is compatible with a "coarse-grained Hypertolerance" picture that endorses Hypertolerance for pretty much any parameter of variation that might come up in ordinary discourse. On this picture, the worlds where a given familiar object is concrete are thoroughly mixed together with worlds where it isn't, so that every world where it is concrete is extremely similar in microphysical respects to some worlds where it isn't. Given that there is no principled basis for saying how worlds get divided up in this way, we should expect a lot of vagueness. Plenitude entails that there are objects with this kind of chaotically discriminating modal profile, thus allowing us to make sense of the posited vagueness in terms of a multiplicity of candidate referents, in the usual way.

Moreover, by combining plenitude with a high level of semantic plasticity, we can answer the main objection that we raised for the coarse-grained hypertolerance view in §6.3. That problem, which provided the basis for Chapter 9's

The Bounds of Possibility: Puzzles of Modal Variation. Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri, Oxford University Press. © Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri 2021.
DOI: 10.1093/oso/9780192846655.003.0013

self-standing Robustness Argument, turned on judgements to the effect that a certain object would be *hard* to get rid of without making some substantial change to some relevant set of underlying facts. The basic worry was that if there are a few worlds that are very similar to actuality in underlying respects but where the facts about the object are different—e.g. worlds where the earring-maker cuts in almost the same place but the actual earrings Lefty and Righty are never created or swap sides—then our robustness judgements in *those* worlds are in error. The appeal to semantic plasticity lets us respond to this Security Argument for Non-contingent Robustness in the same way that we responded to the Security Argument for the non-contingency of Tolerance premises: we can say that at the worlds in question, the relevant names, demonstratives, and common nouns shift their reference in such a way that the the Robustness speeches still come out true. For example, even if the set of worlds that are only microscopically different from actuality contains a few worlds where Lefty is on the right and Righty is on the left, people who say 'this earring' in those worlds while pointing towards Lefty thereby refer not to Lefty, which was unlikely to be on the right, but instead to some object which is *robustly* on the right.[1]

However, as we have seen, plenitude and plasticity also undermine the central argument for Hypertolerance, by offering an answer to the Security Argument for Non-contingency. Moreover, at least when taken out of the blue, the Hypertolerance claims that we are concerned with often elicit initial reactions of incredulity, which strongly suggests that they are false.

Still, even though we don't want to go so far as to embrace the coarse-grained Hypertolerance strategy as a general treatment of the puzzles, we are not inclined to embrace the rejection of Non-contingency as a fully general treatment either. Our favoured approach is a more flexible one, on which certain of the Hypertolerance claims that feature in Tolerance Puzzles are true in some contexts, though not in all. As various earlier examples illustrate (recall the A34, the Old Bodleian, and the sandcastle), it is quite easy to get into a mood where we are willing to assert sentences that logically imply that certain objects could quite readily have been made of completely different originating matter.[2] Likewise, for very many other parameters, it is not that hard to get into a Hypertolerance-friendly frame of mind, given the right kind of contextual priming. For example, if someone unreflectively

[1] In the terminology of Chapter 9, we are proposing to Robustness Arguments by denying Non-contingent Robustness. For reasons given there, Hyperrobustness is untenable in many of the interesting Robustness Arguments, so even those who are willing to embrace a lot of Hypertolerance need to block these arguments in some other way.

[2] Of course this doesn't imply full Hypertolerance with respect to originating matter: perhaps the sandcastle could have been made from the wet sand near the water, but couldn't have been made from certain other equinumerous collections of grains of sand, e.g. some collection that is actually uniformly scattered across the entire beach. But once one has gone as far as the speech allows, it is at least fairly natural to assign 'sandcastle' a denotation that makes Hypertolerance true at least with respect to appropriately sized collections of grains of sand from the beach.

says, 'I have been modifying this painting for many decades, during which it has run the gamut of artistic styles', we may happily accommodate rather than convict the speaker of false presupposition. Since plenitude guarantees the existence of objects that are hypertolerant in the relevant respects, it is natural to think that these shifts in frame of mind do not bring error in their wake, but rather induce a shift in the denotations of relevant expressions (demonstratives, names, common nouns), so that the Hypertolerance-friendly speeches are true.

This need not mean that there is any context in which one can truly and without equivocation affirm the conjunction of many disparate Hypertolerance claims. Whereas 'This sandcastle would have been better if we made it from that wetter sand' is fine, 'This sandcastle would have been better if professional sandcastle-makers had made it last week on Copacabana Beach in the shape of the Tower of London' is a lot harder to make sense of. This suggests a model on which hypertolerance with respect to a dimension of variation that we are focusing on in some context is generally compensated for, in the background, by relative intolerance with respect to other dimensions. Perhaps, for example, the contexts where 'This sandcastle could have been made of completely different sand' is true are contexts in which something like 'This sandcastle couldn't have been made without being the first sandcastle made by us on that beach that day' is also true, despite the fact the latter sentence would be odd to explicitly assert.

Apart from its ability to accommodate various natural hypertolerant frames of mind and associated ways of speaking, the flexible approach also helps with another problem faced a view that takes Non-contingency to be context-insensitively false. As we noted in §3.2, while iterated modal claims are not so common in ordinary practice, if we look hard enough, we can find some pretty ordinary-sounding speeches that suggest that certain Tolerance propositions would have been true in certain counterfactual circumstances, or that they could not easily have been false. Here is one of our examples:

> If we had made the top of that walnut table out of mahogany, then we would have soon realized that it would have looked a lot better if we had made its legs out of mahogany as well.

The flexible contextualist approach can accommodate speeches like this, by claiming that they invoke a context in which we refer to objects that are highly tolerant—perhaps even hypertolerant—along the relevant dimension.[3] There is strong evidence that some kind of contextual shiftiness is in play, since while the

[3] In the example, above, the object being referred to is at least a counterexample to Overlap Essentialism. It is much less clear that it could have been made of *any* atoms chemically matching its actual originating atoms: after all, continuing with '…whereas it would have looked a lot worse if we had made the whole thing out of pine' seems pretty strange.

above speech sounds fine, it does not seem fine to assert it in the same breath as 'We couldn't have made that table if we didn't have any walnut', although the latter sentence is also fine taken by itself.[4]

12.2 Plenitude, Plasticity, and Coincidence Puzzles

We have argued that ordinary tables do not *in fact* coincide with other tables.[5] This leaves it open whether there *could* have been—or could *easily* have been—ordinary tables that coincide with other tables. One might find it objectionable to suppose that the actual world is different from other nearby worlds in being free of table-coincidence, on the grounds that this makes us strangely fortunate to have landed on the truth about the absence of such coincidence. But it should be clear that this security worry can be readily met, once we have embraced the kind of semantic plasticity in words like 'table' that we in any case need for a satisfactory resolution of Tolerance Puzzles. Even if there are nearby worlds where there are coincident ordinary tables, we can say that the property (or properties) that people use 'table' to express at those worlds lack distinct coincident instances at those worlds (in the relevant ordinary cases). Since none of the properties in question are any more natural than the rest, this also sets to rest any impression that the absence of coincident tables in the actual world makes the actual world metaphysically "special" in some objectionable way: where *T* is one of the properties expressed by 'table' in some nearby world *w* where there are coincident tables, the actual world may have coincident *T*s while *w* lacks them. So there is a broader overall symmetry to the relationship between the two worlds, despite the table-theoretic asymmetry.

The view that coincidence of things like tables and buckets could easily have occurred, even though it does not in fact occur, provides one workable solution to the Coincidence Puzzles that were the topic of Chapter 4. As we set up those puzzles, they depend on a "Non-coincidence" premise: for example, that if two given pieces of plastic *A* and *B* had been combined, there would not have been two coincident buckets made out of them. The main motivation we gave for this

[4] A different contextualist strategy for vindicating the ordinary speeches involving embedded modality is to say that the referent of the demonstrative is chosen so as to make Tolerance and Overlap Essentialism express propositions both of which are true at the relevant counterfactual worlds, but at least one of which is false at the actual world. For example, in the above speech, 'this table' might denote an object that is anchored to the quantity of wood that would have been used in the relevant counterfactual circumstance—comprising a mixture of walnut and mahogany—in something like the way in which the referent of 'this table' in a more standard context is anchored to the some actually table-shaped quantity of wood. So the referent of 'this table' in the special context couldn't have been made of wood not overlapping *that* quantity of wood: it is a counterexample to Overlap Essentialism at the actual world, but not at the counterfactual world where we use a mahogany top.

[5] As explained in note 46 to the previous chapter, 'ordinary' is meant inter alia to set to one side cases of "fission" and "fusion" such as the case where the NYU seminar table arguably shrinks to the size of a component table.

premise combined arguments for the absence of coincidence of objects of the relevant kind at the actual world with an appeal to considerations of security and strangeness to preclude a split verdict between the actual world and various nearby worlds. We now have the tools to resist this line of thought.[6]

The plasticity we have found reason to posit in words like 'bucket' (as well as demonstratives like 'this bucket') can also be marshalled in support of a different solution to Coincidence Puzzles. This alternative solution takes the proposition that no two buckets coincide to be necessarily true (or at least true at all close worlds), and thus endorses the Non-coincidence premise. Instead, it puts the blame on the conjunction of the two Robustness premises—e.g. that if *A* and *B* had been combined, they would have composed Flimsy, and that if *A* and *B* had been combined, they would have composed Frail (where Flimsy and Frail are the two buckets actually composed by *A* and *B* respectively). The solution is to appeal to context-shift to allow each of the two Robustness premises to be true in the natural context of its utterance, while denying that they are true together in any context. When we discussed this strategy in §4.4, the challenge we raised was to say which words are shifty and how they shift. With the resources of plenitude at our disposal, it is not hard to tell a story about how this might go. There are (at least) two candidate referents for the name 'Flimsy', both of which are actually made of piece of plastic *A*: if *A* and *B* had been combined, Flimsy_1 would have been made of *A* and *B* whereas Flimsy_2 would not have been made at all. Likewise, there are two candidate referents for 'Frail', both made of *B*: if *A* and *B* had been combined, Frail_1 would not have been made at all, while Frail_2 would have been made of *A* and *B*. And similarly there are two candidate properties expressible by 'bucket': Frail_1 and Flimsy_1 are both necessarily buckets_1 whenever concrete and are necessarily not buckets_2, while Frail_2 and Flimsy_2 are both necessarily buckets_2 whenever concrete and are necessarily not buckets_1. No close worlds contain coincident buckets_1 or coincident buckets_2, although necessarily every bucket_1 coincides with a distinct bucket_2 and vice versa.[7] Given plenitude, there certainly are objects and properties with these profiles, so all that remains is to say why our words 'Flimsy', 'Frail', and 'bucket' pick up on them in variable ways across contexts. But this isn't any more mysterious than the kind of variability we have already made our peace with: when

[6] We can do something similar with temporal Coincidence Puzzles. For example, in the Ship of Theseus puzzle, we can say that at the later time it is natural to use 'ship' to express a property with two instances, one in the harbour and one in the museum, which were coincident earlier; whereas at the time of Theseus's coronation, it is more natural to use 'ship' to express properties with only one instance in the vicinity at that time (though with two instances at the later time). And of course 'ship' at the earlier time may be vague as between a property whose one instance will end up in the museum and a property whose one instance will stay afloat.

[7] It would be implausible to think that the two sets of co-ordinated interpretations of 'Flimsy', 'Frail', and 'bucket' are the only available ones. For example, we can plausibly access an interpretation on which 'If *A* and *B* had been combined, these two buckets would not be buckets but would only be two parts of a larger bucket.' This is particularly natural if the combination of *A* and *B* would have made one an inner layer and the other an outer layer, preserving their actual bucket-like shapes.

we talk in ways that treat certain sentences as obvious, context-sensitive words tend to latch on meanings on which those sentences come out true, if such meanings are available.

So we have two solutions to Coincidence Puzzles available to us, both of which rely on some background plenitudinous ontology and appeal crucially to semantic plasticity (in the first case inter-world, in the second case intra-world). Which should we favour? If forced to pick between we would favour the second strategy, since denying Non-coincidence leads to unpalatable verdicts about counterfactual counting claims such as 'The only way to fit two three-gallon buckets under your sink would be to put one inside the other.' But given the level of intra-world context-sensitivity that we have already found reason to posit, we are more inclined to split the difference, by allowing both for interpretations of 'bucket' where 'There are no coincident buckets' expresses a necessary truth and for interpretations where it expresses a truth that could easily have been false. Interpretations of the first kind are particularly natural when we are dealing with counterfactual counting claims. Interpretations of the second kind come into play when we are making generalizations about the robustness of objects of a given sort, such as 'Every bucket made in that factory would have been a lot more useful if any other slug of plastic had been thrown into the mould along with the one that was in fact used to make the bucket.'

12.3 Three Choice Points Concerning Tablehood

In this section we will consider some choice points concerning the property of tablehood, not settled by the basic idea that in the actual world, every table is tolerantly a table, no table is hypertolerantly a table, and no two ordinary tables coincide. Obviously, similar choices will arise for myriad other common nouns.

First: are there nearby possibilities in which Woody (a typical table) coincides with a table without being a table? Contra Yablo, we maintain that Woody is tolerantly a table: each collection of chemically matching atoms that overlaps Woody's actual originating atoms by 90 per cent is such that Woody could have been a table originally composed by it. But this is compatible with there being other collections of atoms, overlapping Woody's actual originators by less than 90 per cent, such that Woody could have been originally composed of those atoms while being table-shaped, but could not have done so while being a table. One could suppose that as we head out from the actual world by gradually varying which collections of atoms originally compose tables, Woody stops being a table before Woody stops concretely existing altogether. This might go along with the thought that necessarily, an object is a table only if it is tolerant (if not in the 90 per cent way, then at least for some other fraction less than 1). This hypothesis must be sharply distinguished from the hypothesis that necessarily, an object is a table only

if it is *tolerantly a table*. The latter hypothesis is a Necessitated Tolerance premise, and hence leads (given Iteration) to the problematic conclusion that every table is *hyper*tolerantly a table; the former does nothing of the sort.[8]

On reflection we are inclined to the view that Woody's tablehood is more or less as modally robust as Woody's coinciding with a table. When we further probe the everyday modal practices that motivate tolerance judgements in the first place, we find no disposition to acknowledge possibilities where a given table is a non-table coincident with a table. Examining a defective saw, we might naturally argue as follows: 'There was about a 50 per cent chance of this table having a big gouge in its surface. A table with that sort of gouge would be almost worthless. But since this was the only table for which we used the bad saw, there was no significant chance that any other table we made today would be anywhere near worthless. So, there was about a 50 per cent chance of this table being the least valuable table produced today.' Here we are taking it for granted that the table has no chance of being gouged but not a table, which would be a mistake if it had a chance of coinciding with a table without being a table.[9]

Such examples might suggest the stronger generalization that Woody couldn't have *concretely existed* without being a table. But this seems considerably more tendentious. When we consider possibilities where the shape and organization of the relevant quantity of wood is still *somewhat* like its actual shape and organization, but where we still don't want to say that the quantity of wood composes a table, it sometimes seems quite natural to describe them as possibilities where Woody is still around, and is composed of that quantity of wood, but is not a table. For example, we can say things like 'This is a table, but it could instead have been a bench', or 'This is a table, but it could instead have merely been part of a bigger table.'[10] And obviously even more caution is warranted when it comes to generalizations like 'No table could have concretely existed without being a table': even if run-of-the-mill tables aren't counterexamples, it seems far more tendentious when it comes to "borderline" tables.[11]

[8] Some philosophers (e.g. Fine 2005*b*) have argued that certain "kind" properties, perhaps including tablehood, are such that everything that has them has them necessarily: e.g. each person is necessarily a person, and thus would have been a person even if they had never been conceived. This view would of course settle our question in the negative, but we don't see much to recommend it. For one thing, as we point out below, it seems natural enough to say that a certain table could have been a bench rather than a table.

[9] While we are pretty confident that mere permutations of originating matter will not take one to a world where Woody coincides with a table without being a table, we do not maintain that no table, without exception, could have coincided with a table without being a table. There are exotic cases—e.g. some kind of "two tables in one" transformer scenario—where we are not sure what to say.

[10] We are not suggesting that we have to talk like this. We could comfortably describe the very same possibilities by saying things like 'Instead of making this table, I could have made a bench instead' or 'This is a small table, but it could instead have been a much bigger table'. Our semantic picture can easily accommodate such flexibility.

[11] Such generalizations often emerge in the context of theorizing about "sortals" (see Grandy and Freund 2020), a term we would prefer to discard altogether. What And what goes for tables goes for dogs: when we remember that dogs evolved from wolves, it seems extremely tendentious to think that the first dogs were essentially dogs.

The second set of choice points concerns whether objects of a given familiar kind are *uniform* in respect of their ability to tolerate changes of a given sort. For example, nothing we have said rules out the hypothesis that one of the pyramids of Giza could have been anything from 30 per cent to 140 per cent of its actual height but could not have been shorter or taller than that, while another of the pyramids, roughly similar to it as regards its actual height and design, could have been anything from 60 per cent to 200 per cent of its actual height but could not have been shorter or taller than that. Likewise nothing we have said rules out that the rectangular table in this coffee shop could have tolerated losing at most 10 per cent of its length from one end, while it could have lost 30 per cent of its length from the other end, even though there is no interesting difference between the two ends as regards as regards their geometry, composition, or history.[12]

Such unprincipled failures of uniformity seem bizarre. However, we must be cautious, since it is tempting to argue for the relevant uniformity claims by appealing to premises which entail the *impossibility* of the relevant kinds of non-uniformity: e.g. that there couldn't be two categorically similar pyramids with radically different growth potentials, or two categorically similar table-ends which differ as regards how fussy the relevant table is about originating matter at that end. But given our commitments, the necessitated versions of the uniformity claims are hopeless. Since it is not necessary that every table is tolerant, it could easily happen that one table changed its originating matter in such a way as to be far less tolerant in the relevant respect, while some other categorically similar table preserves its actual originating matter, and thus its actual tolerance levels.

A more promising theoretical argument for uniformity appeals to premises about metasemantics. If our linguistic dispositions don't mark any contrast between two pyramids or between the two ends of the table, it might seem that nothing could make the word 'pyramid' or 'table' as used by us express some property whose instances are modally nonuniform. After all, plenitude guarantees that the range of candidate meanings also includes properties whose instances are modally uniform, and these look like better candidates than properties whose instances are modally nonuniform to be expressed by the relevant words. But we also need to be cautious about this kind of argument. The most promising alternative to accepting uniformity is not a view where uniformity fails in some precise but completely inscrutable way, but one on which words like 'pyramid' and 'table' are vague, in such a way that their range of admissible precisifications includes a wide variety of properties whose instances are non-uniformly tolerant in a host of different ways, along with some properties whose instances are uniformly tolerant. The effect of such vagueness would be to make sentences stating uniformity claims and sentences expressing particular ways for uniformity

[12] Note 38 in Chapter 2 suggests a canonical format for "Uniform Tolerance" claims which would rule out surprises along these lines.

to fail neither definitely true nor definitely false. Given the vast range of candidates provided by plenitude, we should already expect large doses of vagueness. While this particular dimension of vagueness is not inevitable, it is not especially egregious as far as general principles of metasemantics are concerned.

That said, we can observe a pretty strong tendency to take various kinds of uniformity claims for granted when considering what kinds of changes objects of various kinds can tolerate. We ask, 'What proportion of a table's matter can we subtract before the table is destroyed?'; hearers readily accept the presupposition that the answer is the same for different tables. Or imagine a machine that makes two paper plates 12 inches in diameter at once by simultaneously punching them, side by side, out of a long strip of cardboard 24 inches wide. Given that the cardboard could easily have been pushed much further into the machine than it actually was, it seems clear that it could easily have happened that two different plates were made instead of the actual ones. But given that the cardboard cannot easily be rotated, it seems clear that it could *not* easily have happened that one of the actual plates was made along with a new plate. This judgement presupposes a kind of uniformity: it would be false if one of the two plates was more tolerant (in the relevant direction) than the other.

If one takes uniformity sentences to be stably true, one will need the relevant words to display a level of semantic plasticity far more fine-grained than what is required merely to secure the stable truth of Tolerance sentences. Suppose that a single circular plate, Plato, was punched out of a large sheet of cardboard, and is uniformly tolerant of shifts in the cardboard, but not hypertolerant with respect to such shifts, so that the set of points on the cardboard at which Plato could have been centred—call it Plato's "zone of tolerance"—makes up a circular region whose centre is the point at which Plato was in fact centred.[13] If the proposition that Plato is at the centre of its zone of tolerance is true at the actual world, then it is false in any possible world where Plato is made centred on any other point in its zone of tolerance. But it would be pretty bizarre to think that we are making no mistake when we say, 'Plato is at the centre of its zone of tolerance', but would have been saying something false if the punching device had been even slightly shifted from its actual position. To get that sentence to be stably true, the reference of 'Plato' in that speech must shift even when the punching device is moved only slightly. Similarly, for an utterance of 'Every plate that was made today was at the centre of its zone of tolerance' to be stably true, 'plate' must be used to express different properties in all those worlds. By contrast, the stable truth of a Tolerance premise like 'Plato could have been made centred on any point within an inch of where it was in fact centred' only requires the reference of 'Plato' (and 'plate') to be different at the close worlds where Plato is within an inch of the edge of its zone of tolerance,

[13] If Overlap Essentialism is true for paper plates, the diameter of Plato's zone of tolerance must be less than twice Plato's own diameter.

which plausibly would require a more than slight shift in the position of the punch. However, if we can learn to live with the amount of semantic plasticity needed to secure the robust truth of Tolerance sentences, the extra dose of shiftiness needed to secure the robust truth of uniformity sentences seems unlikely to introduce any new objections.

We have been talking about whether or not to make room for completely "unprincipled" failures of uniformity, unaccompanied by any articulable explanation of why there should be such failure. But even if we reject non-uniformity of that sort, there remains the possibility of positing some interestingly principled failures of uniformity within a category. We will pursue just one example here. Suppose that the Joneses, living on Elm Street, decide they would rather live on the other end of the street, and hire builders to disassemble the bricks of their house and reassemble them according to the same architectural plan. The builders ask, 'Would you mind if, as we dismantle your house we slowly replace the bricks with new bricks, so we can build another house here?'; the Joneses consent. Meanwhile the Smiths, living on Mulberry Street, decide they would like to upgrade their house by slowly replacing the old bricks with new one; the builders *they* hire ask, 'Would you mind if we take the old bricks and use them to build a new house down the street?'; the Smiths consent. The upshot is that similar brick-moving processes take place on both streets.

We are fine with 'The Joneses' house was moved down the street' and 'The Smiths' house had its bricks upgraded' individually.[14] But can we find a single interpretation on which their conjunction is true? Perhaps: if each family tells us a little about the events without supplying the details, we might go on on this basis to say 'The Smiths and the Joneses have both been living in the same house for quite some time', and it is not so clear that we are making a mistake.[15] There are several ways in which the truth of the conjunction could be explained. One possibility is that 'house' here expresses a property whose instances are "attitude-sensitive" objects: whether a certain process would involve their being dismantled and transported or their undergoing slow replacement of parts depends not just on the physical aspects of that process, but on the mental attitudes of some of the agents involved, such as the owners.[16] This picture preserves an underlying uniformity of modal profile across the category, albeit one that manifests in dramatically

[14] Not that one would never want to impute error in such cases. Salmón (Salmon 1979: 717–18, crediting Kaplan, Kripke, and Wiggins) describes a variant of the Ship of Theseus case where the one replacing the parts is a philosopher hired to move a ship from one museum to another, who does in fact move and reassemble the planks: we judge that the philosopher is just making a mistake when he says, 'I have stolen the ship.'

[15] By contrast, if you say 'Smith is rich and Jones isn't', having picked up the conjuncts from two sources, in a situation where Jones in fact earns more than Smith, it is clear that you have made a mistake.

[16] Johnston (1989, 2010) defends a picture where the attitudes of a person make a difference to what that very person can survive: we will discuss this further in §13.3 below.

different ways due to differences in the circumstances. Another possibility is that 'house' expresses a property had both by an object (the Smiths' house) which would have followed its bricks during any brick-relocation-with-replacement process (no matter how the owners were thinking of it), and by another object (the Jones's house) which would have stayed put with new bricks during any such process. Since we surely don't want to say that any houses would have not been houses if the owners' attitudes had been different, this latter option means giving the attitudinal differences a merely "reference-fixing" role. The property picked out by 'house' just brutely treats different collections of bricks differently—to be a house is to be a "relocation-friendly" object composed of *these* bricks, or an "upgrade-friendly" object composed of *those* bricks, or …—but the differences in owners' attitudes at the actual world explain why we end up talking about that property rather than any of the many other similarly non-uniform properties. We are not sure which way to go here, or whether to hold the line and insist that both or neither of the houses were moved. The question how, if at all, differences in attitude relate to differences in modal profiles in such cases seems like a promising one for further research.

The third set of choice points concerns what to say about the modal profiles of *possible* bearers of the relevant property—e.g. possible tables, possible plates, etc.—that do not concretely exist in the actual world. The demand that tolerance and uniformity speeches express truths places substantial constraints on the modal profiles of the *actual* members of the kind, but leaves a lot of freedom as regards merely possible members. And recall that we cannot demand that tolerance and uniformity speeches express *necessary* truths, on pain of hypertolerance.

To illustrate the variety of options for merely possible members of the kind, let's return to our paper plate punched from a large sheet. Let's suppose (as a sort of proxy for Microphysical Supervenience) that given the background circumstances (including the size and composition of the cardboard, the identity of the artificer, the fact that only one punch is made, etc.), the identity of any plate produced depends entirely on where the punch is centred on the sheet.[17] And let's suppose further that those background circumstances necessitate that exactly one plate is produced.[18] At any two possible worlds where these circumstances obtain and the punch is centred at the same point, the same possible plate is produced. So the relation *x and y are points on the sheet of cardboard such that for some possible plate p, necessarily: if the background circumstances obtain and the punch is centred at x or y, p is made and is a plate* is an equivalence relation. It generates a partition of the cardboard's surface into disjoint regions, each of which is the "production zone"

[17] Cf. Williamson's 'One Earring Per Point' thesis as discussed in §7.2.

[18] One interesting option in response to the below considerations is to drop that; we have already seen some reasons to think that (at least in certain contexts) coincidence of objects like plates might be fairly common at close non-actual worlds.

of some possible plate.[19] If Plato (the plate actually produced) is tolerantly a plate, and uniformly so, we know that one of the cells of this partition is a circular region centred at the point where the punch is actually made. But what are the other cells like? Obviously they can't all be circles of the same size.[20] So, we can't say that all the production zones are alike: some must be smaller or larger or less regular than Plato's.[21] Given that we are not going to be able to say that, several alternative models suggest themselves (see Figure 12.1). On perhaps the most natural picture (b), all the production zones have roughly the same *area*, but many of them are less regular than Plato's. On a more extreme picture (a), every production zone other than Plato's contains just one point: any two points that wouldn't have produced Plato would have produced different plates. So given Iteration, if Plato hadn't been produced, then an intolerant plate would have been produced. (Nevertheless we would have spoken the truth when we said 'A tolerant plate was produced', since

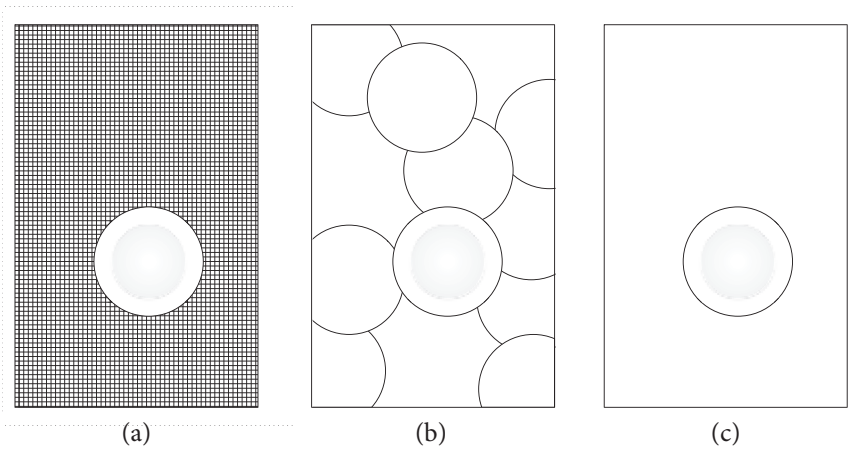

Fig. 12.1 Three models of the plate-stamping case. (a) All production zones except for Plato's are tiny (point-sized). (b) Production zones have roughly equal sizes but differ in shape. (c) Only two production zones, one very large.

[19] This reasoning assumes BF: without it, some points may be such that although a plate could be made centred on them, there is no possible plate that could be made centred on them. It does not depend on Iteration; however, if Iteration fails, the facts about how the sheet is divided into production zones might be contingent. Even though some of the non-concrete possible plates have noncircular production zones, it might still be necessary that every *plate* has a circular production zone.

[20] No topologically connected region can be covered by more than one non-overlapping open or closed discs.

[21] By contrast, if there were only one dimension of variation, as with the cut-angle in Williamson's earring cases, there would be the prospect of having production zones of matching size: for example, Williamson himself suggests that in the world where no earrings are produced, the equivalence relation *x and y are points on the circumference of the disc such that for some possible earring e, e would have been made centred on x if some earring had been made centred on x, and would have been made centred on y if some earring had been made centred on y* partitions the diameter of the disc into four equal quadrants.

we would have expressed a different property by 'plate'.) Or we could go to the other extreme (c) and think that the partition has just two cells, one of which is Plato's production zone and the other of which contains every other point on the surface of the cardboard. So there is just one other possible plate other than Plato which could have been produced given the circumstances. Given Iteration, if this other plate had been produced, it would have been a counterexample to Overlap Essentialism for plates. (Nevertheless we would have spoken the truth when we said 'The plate produced could not have been made of a completely non-overlapping part of the cardboard', since we would have expressed a different property by 'plate'.)

Our default assumption is that the relevant words (e.g. 'plate') are sufficiently vague and context-sensitive that none of the options is determinately ruled out. But it wouldn't harm our general picture if there turn out to be some good arguments for narrowing down the range of admissible resolutions of vagueness, perhaps by eliminating one or both of the extreme options canvassed above.

12.4 Which Properties Are Discriminating?

We have said a lot about the properties expressed by common nouns like 'table'. We will end this chapter by considering some questions about how ordinary predicates other than common nouns behave with respect to the multiplicity of coincident objects generated by plenitude. While these questions are less directly relevant to Tolerance Puzzles, they are still important for proponents of plenitude to think about.

Once we fix on a meaning for 'coincident'—e.g. 'located at the same spacetime points' or 'having the same fundamental parts'—we can define a co-ordinate notion of an *undiscriminating* property: a property is undiscriminating iff necessarily, it is instantiated by both or neither of any two coincident objects. A property that is not undiscriminating might alternatively be "super-discriminating" in the sense that *no* pair of distinct coincident objects could both have it. Or it might be in neither of these categories. Similar questions arise for relations with respect to each of their argument places. We shall look at four tricky cases: colours, mass, seeing (in its direct object argument place), and parthood (in both of its argument places). None of them are super-discriminating: a statue and its coinciding lump typically have the same colour and mass, are both easy to see, and share many parts (such as atoms). But it is also far from obvious that they are undiscriminating.

(i) Colour. Consider a red table in a dark room. According to plenitude, it coincides with something that couldn't be concrete unless in the dark. It also coincides with something that could be concrete in the light, but could only do so while coinciding with a blue thing. Are both of these objects also red? There is a long tradition that ties colours closely to *dispositions*—dispositions to produce

certain kinds of experiences, or to interact in certain ways with light.[22] And having a disposition, in turn, seems to be a modal matter: although there is no uncontroversial analysis of disposition-ascriptions in terms of counterfactuals or chance-ascriptions, it is hard to see how something that couldn't possibly behave in a certain way (and has zero chance of behaving in that way) could nevertheless be disposed to behave in that way.[23] But we should be cautious: if we lean *too* heavily on counterfactual or chance-theoretic criteria for colour, we might be led to the implausible result that any red table in a dark room coincides with a blue object.

If colours are discriminating, but are still shared by (e.g.) wooden tables and their coincident quantities of wood, we face the question whether all of the "table-candidates" coincident with a given table share that table's colour, where a table-candidate is an object that has some property that could easily have been expressed by 'table'. If we decided that some of the table-candidates coincident with a red table are not red, we would need to posit semantic plasticity in 'red' to co-ordinate with the semantic plasticity of 'table', since clearly there are no nearby worlds where 'There is a red table here' expresses a falsehood while the table is still red. While the proposal seems workable, we don't really see any pressure to think that whatever plasticity there may be in 'red' is entangled with the plasticity of 'table' in this way.

(ii) Mass. One might be pushed towards the view that having mass is discriminating by the thought that for an object to have mass, it must "play well with the laws of motion", in a sense that precludes objects with inappropriate modal profiles from having mass (see Hawthorne 2006*d*). This is particularly clear in cases of temporary coincidence: if an object coincident with a planet doesn't move in the ways that would be predicted by the mass of the planet in combination with the gravitational field, one might think that the object doesn't have mass.[24] But even in the case of permanent coincidence, we might hesitate to describe an object as having mass on modal grounds. Could an object be massive if it isn't disposed to move in a straight line when there are no forces acting on it? However, although there clearly is at least one sense of 'mass' in which tables have mass, it is hard to cash out "playing well with the laws" in such a way that they count as

[22] For the view of colours as dispositions (or "powers") to produce visual experiences, see Locke 1689: ch. 7 and Levin 2000. For colours as dispositions to reflect light, see Byrne and Hilbert 1997. For a competing view that takes all such disposition-talk to have a merely reference-fixing role and identifies colours with certain properties having to do with the geometry and atomic composition of an object's outer layer, see Jackson 1996. Since the relevant properties of layers are plausibly undiscriminating, the colours themselves will also turn out to be undiscriminating on this view.

[23] See Manley and Wasserman 2008 for a helpful survey. Note too that it is not at all obvious that relations like *producing* and *reflecting* are undiscriminating (in their subject argument). If objects that produce colour experiences and reflect light can coincide with objects that do not, then that by itself may already generate pressure to say that colours are similarly discriminating, quite apart from considerations to do with dispositions.

[24] The case for discriminatingness is even sharper for *kinetic energy* (relative to a frame), given its close connection to motion. But given the connections between mass and energy it may be difficult to maintain that mass—even rest mass—is undiscriminating if kinetic energy is not.

playing well with the laws, given that (for example) the centre of mass of a table will move discontinuously when a leaf is attached (Hawthorne 2006*d*: 113). This speaks in favour of a relatively undiscriminating treatment. However, there may still be reason to stop short of a completely undiscriminating approach, especially if we are understanding coincidence in location-theoretic terms. For example, one might think that a hole in a bagel coincides with the quantity of cream cheese that fills it, although the quantity has mass and the hole does not.

(iii) Seeing. Seeing is plausibly discriminating (in its second argument) if colours are—although we can see things like panes of glass when things go well, lacking colour is at least a prima facie obstacle to being seen. Relatedly, if with Grice (1961) and Lewis (1980*b*) we hold that seeing an object requires the object to *cause* visual experiences, then seeing will be discriminating if object-causation is.[25] But even if seeing is discriminating, there is reason to think that it isn't *very* discriminating, and in particular that when one sees a table like Woody, one typically sees all of its coincident table-candidates. First, it seems that one couldn't easily have stood right in front of Woody with eyes open while demonstrating and saying 'That's a table' without seeing Woody. (It would be crazy to think that if the sander had been applied a little longer we wouldn't have been able to see Woody!) But second, it also seems that one couldn't easily have demonstratively referred to an object in the relevant way without seeing it. Given our commitment to plasticity in the demonstratives, it follows that for each table-candidate coincident with Woody, there is a nearby world where both Woody and that table-candidate are seen. This doesn't conclusively show that all of the table candidates are seen at the actual world: one could in principle hold that Woody is much easier to see than the other candidates, appealing to plasticity in 'see' to remove any mystery about how this could be so. But we don't see much point to this.[26]

(iv) Parthood. Given the standard philosopher's view that every object is part of itself, the view that parthood is undiscriminating would mean that any two coinciding objects are parts of one another. This is surprising—it seems odd to think that table is part of the quantity of wood—and conflicts with the attractive general principle that parthood is *antisymmetric* (i.e. that mutual parthood implies identity).[27] It also seems doubtful that, e.g., the tabletop is part of the quantity

[25] Note that in the setting of plenitude, one has to be cautious about analysing object-causation in counterfactual terms, since such analyses threaten to multiply the objects which cause any effect in absurd ways: e.g. there will be something coincident with the Sphinx that causes your current visual experience.

[26] One could instead try to motivate the competing view that we see only one thing when we look at Woody by appealing to some premise to the effect that it can't easily happen that someone points in the direction of a thing one sees and uses a perceptual demonstrative without referring to that thing. But such principles seem completely hopeless: for example, we can see someone's head peeking around a corner while using a demonstrative to refer to the person rather than their head.

[27] Authors who maintain that objects like tables and quantities of wood are distinct but parts of one another include Thomson (1998), Hawthorne (2006*d*), and Hovda (2013). Goodman (unpublished *c*) defends antisymmetry, in the context of a view where every table has a coincident portion of matter as a part, but no portion of matter has a table as a part.

of wood. However, it is not obvious that there is any way of combining the view that parthood is discriminating in these ways with the (also popular) view that parthood is a highly *natural* relation (in the sense of Lewis 1983*a*).[28] One might think that the only relevant relations that the tabletop bears to the table but not the quantity of wood involve some messy comparisons of modal profiles that would be very complicated to spell out in precise terms. Assume that the tabletop is part of the table but not the quantity of wood. Now consider a Sorites sequence of coincident objects with gradually varying modal profiles connecting the table to the quantity of wood. It is hard to believe that 'part' draws a non-vague boundary between the objects that have the tabletop as a part and the ones that don't. But if 'part' is vague, that seems to be evidence that parthood is not a highly natural relation.[29] We are not sure that there is any good reason to think that 'part' in its ordinary sense is particularly natural. But if one is wedded to that idea, one may find there is quite a bit of pressure to think that parthood is undiscriminating after all.[30]

These four disparate examples illustrate the wide array of considerations that may be relevant in deciding whether, and to what extent, a given property or relation is discriminating.[31] Although views in the direction of plenitude have been widely discussed, these questions of discriminatingness remain relatively unexplored, and strike us as good topics for future research.

[28] Eddon (2007) defends the claim that parthood is perfectly natural. Bricker (1996) denies that parthood is *perfectly* natural, but is sympathetic to the idea that it has a similar status to *identity* (he puts both in a category of "logical" relations). But identity is plausibly *highly* natural even if not perfectly natural; at least, it is easy to refer to non-vaguely.

[29] In taking these considerations of vagueness to be indicative of a low degree of naturalness for all the relevant discriminating relations, we are implicitly appealing to the idea that highly natural properties are relatively easy to refer to non-vaguely: see Dorr and Hawthorne 2013 §2.10.

[30] On the other hand, one may be able to combine antisymmetry and other intuitive judgments of non-parthood with the view that parthood is highly natural, by adopting a more ambitious plenitude principle in the style of Fine (1999) or Goodman (unpublished *c*), on which each object coincides with a whole hierarchy of other objects which differ as regards their parts, even though many of them necessarily coincide. For example, given any object x, such theories will posit distinct necessarily coincident objects x', x'', … which are not parts of x, such that necessarily, the parts of x' are itself and all the parts of x; the parts of x'' are x'' and all the parts of x'; and so on. Proponents of such views will think that for a given Sorites sequence of modal profiles (all instantiated by coincident objects), there will be many different sequences of objects having these profiles but differing mereologically, so that different choices of sequence will induce different sharp boundaries for 'part'.

[31] Mental properties like *being conscious* and moral properties like *having moral worth* also raise some distinctive issues, which we will be discussing in the next chapter. There are also interesting connections between the examples: for example, if parthood is undiscriminating, it is hard to see how *being partly red* could be discriminating.

13
Alternatives and Challenges

This chapter has two goals. The first is to compare our positive account, as developed in Chapters 11 and 12, with a range of alternatives which also appeal to semantic plasticity in a way that undercuts the Security Argument for Non-contingency, but locate the relevant plasticity in different expressions. The second is to address what we see as the two most important challenges to our view: the threat of rampant error in "cross-world" speech reports, and the uncomfortable surprises that arise when our proposals about things like tables are applied to Tolerance Puzzles concerning people.

13.1 Alternative Sources of Plasticity

Our crucial tool for blocking the Security Argument for Non-contingency is the idea that the sentence expressing Tolerance could easily have been used to say something different than what it in fact says. Although we use it to say something that could easily have been false, we could not easily have used it to say something false. It is natural to ask which subsentential expressions are responsible for this semantic variability in sentences. Our answer is that the relevantly shifty expressions are demonstratives, proper names, and common nouns like 'table'. While this answer doesn't depend on the full-strength version of any of the plenitude theses we surveyed in §11.2, it does require a plethora of coincident objects with suitably different modal profiles. But there are other words in the relevant Tolerance premises to which their semantic plasticity might be traced. Let's look once more at two of our standard Tolerance premises, one using a name and the other using a common noun:

(T1) For every collection of atoms *C* that chemically matches and overlaps the collection of atoms that originally composed Woody by at least 90 per cent, it is possible for Woody to have been originally composed by *C*.

(T2) For every table *x* and every collection *C* of atoms that chemically matches the collection of atoms that originally composed *x* by at least 90 per cent, it is possible for *x* to have been a table originally composed by *C*.

The Bounds of Possibility: Puzzles of Modal Variation. Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri, Oxford University Press. © Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri 2021.
DOI: 10.1093/oso/9780192846655.003.0014

Setting aside really unpromising ideas, like pinning the plasticity on the truth-functional connectives or on the compositional rules whereby the semantic values of sentences are fixed by those of their constituents, there are three alternative sources of plasticity that seem worth considering:

(i) One could blame the modal word 'possible', and the expressions that play the corresponding role in other Tolerance premises (such as 'chance' and the tense operators).

(ii) In the case of quantified Tolerance premises like (T2), one could pin the blame on quantifier words (e.g. 'every' in 'For every table…'), while still (like us) attributing the plasticity of referential Tolerance premises like (T1) to names and demonstratives.[1]

(iii) One could blame the expression 'originally compose' as it occurs in (T1) and (T2). Plausibly, given the connection between 'originally compose' and other mereological notions, this will involve, and can be explained in terms of, plasticity in the word 'part'. More generally, this strategy will locate the relevant dimension of plasticity in predicates that one might think of as characterizing the categorical physical properties of and relations among the relevant objects, e.g. '*n* centimetres tall' in the case of the Great Pyramid puzzle; 'centred *n* inches from this end of the strip of cardboard' in the case of the paper plates, etc.

In principle one could also imagine hybrid views on which the job of explaining the required plasticity at the level of sentences is somehow divided up between several of the relevant kinds of subsentential expressions, e.g. modal operators and quantifiers. But since we haven't thought of any interesting ways for this to work, we will leave it to others to map out these possibilities.

One motivation for exploring these alternative diagnoses of plasticity is the prospect that they might—unlike our own diagnosis—let us block the Security Argument without any commitment to plenitude. But before one gets too excited, it is important to remember that, as we emphasized in §11.2, there are important motivations for plenitude that are independent of its role in helping to solve Tolerance Puzzles. For example, there is the desire to accommodate various ordinary judgements without positing "joints in nature" that just happen to line up with those judgements. And once we have some form of plenitude, it would be rather mysterious if words like 'table' were *not* highly plastic. For given plenitude and the truth of 'No two tables coincide', there is such a vast multiplicity of "candidate" properties that have the right general profile to be referred to by the word 'table' that it seems inevitable that the facts about which of the properties we refer to will

[1] In principle one could attempt to blame the quantifiers even when it comes to premises like (T1), but this seems completely unpromising.

turn out to be highly sensitive to small differences in the underlying facts.[2] So in addition to the specific challenges we will raise for each of (i)–(iii) below, there is to our mind strong theoretical reason of a general kind favouring the diagnosis that pins plasticity on names, demonstratives, and common nouns.

Let's now turn to the three alternatives.

(i) Plasticity in modal operators. Despite the deep differences of principle between the strategy that relies on plasticity in modal operators and the Iteration-denying approach we explored in Chapters 7 and 8, it turns out that many of the objections to Iteration-denial that we raised in those chapters can be retooled as objections to this strategy.[3]

First, in Chapter 7 we pointed out that it is very unpromising to deny Iteration for tense operators like 'sometimes'. Positing plasticity in these operators seems similarly unpromising as a strategy for resolving temporal Tolerance Puzzles. Suppose the painting won't be around anymore in two hours' time, and in one hour's time the painter is going to utter the sentence 'This painting will still be around in one hour's time' while referring to the painting. It seems absurd to suppose that she will be speaking the truth thanks to the fact that she won't mean the same thing by 'in one hour's time'. If this is off the table, those who rely on plasticity in modal operators to resolve modal Tolerance Puzzles will need some other strategy for dealing with temporal Tolerance Puzzles.

Second, in Chapter 7, we saw how deniers of Iteration for metaphysical possibility have the resources for defining a broader modality—ancestral metaphysical possibility—for which Iteration does hold. We noted that if one's strategy for resisting Tolerance Arguments concerning metaphysical modality turns on denying Iteration, one will be pushed towards Hypertolerance for this broader modality. And we pointed out how such Hypertolerance might conflict with some plausible higher-order identities. Similarly, those who posit plasticity in 'metaphysically

[2] See Dorr and Hawthorne 2014 for a careful presentation of this mode of argument from the abundance of candidate meanings to plasticity.

[3] Murray and Wilson (2012) come close to proposing a solution based on plasticity in the modal operators, though (as we discussed in note 14 to Chapter 3) their approach seems more in line with "relativism" than contextualism. By contrast, Hellie, Murray, and Wilson (2021) definitely appeal to Kaplan-style contextualism, both for modal operators and for proper names (which in their models will map to different *individual concepts* relative to different context-parameters). Their favoured strategy for dealing with Tolerance Arguments seems to turn on the context-sensitivity of names rather than that of modal operators. In this respect their view is like ours. However, their models do not include context-sensitivity in *predicates* (like 'table'), so if they wanted to appeal to context-sensitivity to block the case for Non-contingency in *quantified* Tolerance Arguments, they would need to locate it elsewhere. They have several resources that they could draw on here, since in their models, not only modal operators but *variables* give rise to distinctive forms of context-sensitivity.

We have not included the strategy of pinning the context-sensitivity of quantified Tolerance premises on variables rather than common nouns in our list of options in this section because, at least in Hellie, Murray, and Wilson's implementation, such context-sensitivity goes hand in hand with major departures from classical first-order logic. They use an extensively modified version of the "intensional objects" model theory (see note 25 in Chapter 11) which will not validate the quantified Leibniz's Law schema $\forall x \forall y(x = y \rightarrow P[x/z] \rightarrow P[y/z])$ without drastic further stipulations. Moreover, the context-sensitivity they posit in variables will (as in the models of Fara 2008—see note 10 in Chapter 10) invalidate the quantified UI schema $\forall y(\forall x P \rightarrow P[y/x])$.

necessary' seem to have the resources to make sense of a necessity operation that is at least as broad as any of the other operations available to be expressed by 'metaphysically necessary', and are pushed towards a radical, and arguably objectionable, Hypertolerance for this broad modality. If the plasticity of modal operators is implemented in a Kaplanian framework by assigning them non-constant characters, we should be able to simply consider the collection *C* of all properties of propositions that are the content of 'necessary' relative to at least one context, and define our broad necessity as the property of instantiating every member of *C*. Whether or not this property is itself a member of *C*, it should still have the basic logical profile of a modal operation, and it is hard to see how Hypertolerance could fail to be true for it. For example, conjoining the property actually expressed by 'metaphysically necessary' with the property that would have been expressed had Woody been living on the edge gives an interpretation of 'necessary' (and a corresponding interpretation of 'possible') relative to which Woody is considerably more tolerant; the more conjuncts we add, the more radical the modal flexibility.

Third, in Chapter 8, we argued that because of the way in which the expression 'metaphysically possible' was introduced into the philosophical vernacular, if there is a broad possibility operation of the sort just characterized, there is pressure to interpret 'metaphysically possible' as expressing it.[4]

These dialectical parallels do not in any way negate the important theoretical differences between denying Iteration and taking modal operators to be plastic. From the standpoint of this plasticity view, the Iteration-denier is making a mistake analogous to the mistake we would be making if we inferred from the fact that 'This is identical to that' could have expressed a false proposition to the conclusion that it is not necessary that this is identical to that (pointing twice over to the same thing). Likewise, the fact that the word 'possible' could have applied to some impossibilities does not prevent it from being true that everything possibly possible is possible. Nevertheless, the parallels do confirm our instinct that since the paradoxes we are interested in are about buckets, tables, and so on, it is more promising to address them by saying something about the workings of words like 'bucket' and 'table' than by implicitly challenging the breadth of our modal ideology.

(ii) Plasticity in the quantifiers. We have in fact already considered one implementation of a quantifier-blaming strategy in §11.5, in the course of discussing Leslie's (2011) "many axes" proposal. There, we pointed out the proponent of this

[4] The objections to Iteration-denial for 'positive chance' discussed in Chapter 9, by contrast, do not cleanly carry over to the current view. It does seem surprising that there should be a substantial chance of 'chance' being used to denote a different relation between propositions and real numbers. But it is also surprising to say, as we do, that there is a substantial chance of 'table' being used to denote a different property. And it is not obvious that anything in the scientific practice of reasoning about chances would be especially disrupted by the former view.

view can save quantified Tolerance premises by appealing to contextual quantifier domain restriction, using variation in the tacit restrictions in play in different worlds to avoid nearby possibilities of error. This version of the approach admitted the intelligibility of unrestricted quantification, and simply denied the truth of Tolerance on an unrestricted interpretation. We objected that since many ordinary Tolerance speeches have the hallmarks of unrestricted quantification, this view requires either convicting ordinary practice of an implausible level of error or wrongheadedly ignoring communicative intentions. But some theorists have, for other reasons, posited a range of possible quantifier-meanings while being wary of the idea of an unrestricted meaning from which all the rest can be derived by imposing tacit restrictions. For example, Hirsch (2002) suggests that communities pretty like ours could mean different things by the quantifiers, in such a way that in their mouths, different answers to questions like 'Are there temporal parts?' and 'Are there scattered objects?' come out true; moreover, he seems sceptical of the idea that this family of possible meanings contains a maximal element that is at least as broad as all the rest.

There is a large literature on quantifier variance, and it is beyond the scope of this work to engage with it thoroughly.[5] But we will briefly raise a couple of challenges. First, the basis for scepticism about the idea of a maximally broad quantifier-meaning is a bit unclear. In our higher-order setting, ordinary monadic first-order quantifiers are of type $\langle\langle e\rangle\rangle$—they stand for properties of properties of objects. Focusing on candidate meanings for an existential quantifier, we can characterize one such property $\exists_1$ as being "at least as broad as" another such property $\exists_2$ just in case $\exists_2$ necessitates $\exists_1$: necessarily, for any property F, if $\exists_2 F$ then $\exists_1 F$. Within some collection of candidate meanings, a broadest element will be one that is necessitated by all the rest. But if we can make sense of a collection of candidate meanings, we can make sense of the idea of property having *at least one* of the properties of properties in the collection. And as far as its logic is concerned, this property will behave just like a quantifier-meaning of which all the others are restrictions.

Second, even setting aside the question whether there is a broadest quantifier-meaning, the proponent of the view will need to posit some odd and hard-to-explain restrictions on the range of quantifier-meanings we employ in ordinary life. In many ordinary cases, when two people realize that they have been interpreting their quantifiers differently, a natural response is to start using a third quantifier-meaning that is at least as broad as each of the two. For example, if two teachers each report 'Every student passed', it is easy to access a quantifier-meaning that encompasses both groups of students, and indeed we will find it natural to make use of such a meaning when making a speech like 'It's great that every student

[5] We have weighed in elsewhere: see Hawthorne 2009 and Dorr 2005*b*, 2014*a*.

passed' in the presence of both teachers. But the view of Tolerance Puzzles that relies on quantifier plasticity seems to predict pairs of easy-to-access quantifier meanings for which it is incredibly hard to access a third meaning broader than each. Recall for example that we sometimes get in a rather tolerant mood about sandcastles, in which we would be happy to say things like 'Every sandcastle we made today would have been sturdier if we had instead built it the wetter sand from down by the sea', while in other less tolerant moods we will instead say things like 'We couldn't have made any of the sandcastles we made today if this sand had been any drier.' While §11.3 explained this in terms of variation in 'sandcastle', proponents of the quantifier-plasticity strategy will presumably instead invoke variation in the quantifiers. Now suppose we are in a scenario where 'We made exactly three sandcastles today' is true on both interpretations. Then a quantifier-meaning that was broader than each would make the sentence 'We made at least six sandcastles today: three which could been made of entirely different sand, and, coinciding with these, three more which couldn't' come out true. But ordinary practice seems to be completely blind to such a meaning. We leave it as a challenge to those who pin the variation on the quantifiers to explain why such meanings should be so hard to access, given that it is, in other cases, so natural to respond to local conversational variation in quantifier-meanings by ascending to a quantifier meaning that is at least as broad as any of the ones that have been in play.[6]

Third, Hirsch maintains that in the cases he is interested in, the posited variations are by no means confined to the meanings of the quantifiers: "Quantifier variance might thus be said to induce a certain kind of systematic difference of meaning in the word 'touching' and, by the same token, virtually any other general word" (Hirsch 2010: 76). Other authors discussing quantifier-variance have found reasons to agree (Sider 2011: §8.5.1). At least, positing this kind of global variance may help block certain otherwise promising ways of pinning down unique maximally broad meanings for the quantifiers by their logical roles (Dorr 2014*a*). But if the quantifier-variance approach to our puzzles has to go

[6] One option for responding to this challenge would be to say that there just *isn't* any quantifier-meaning that is at least as broad as each of the two meanings in question. But at least if we adopt the standard picture of quantifier-meanings as properties of properties, there is an obvious way to generate such a third meaning, namely by disjoining them (in the case of existential quantifiers) or conjoining them (in the case of universal quantifiers). Provided the input meanings are well-behaved in the sense that theorems of free first-order logic come out true, the disjoined or conjoined meanings will be well-behaved in that way too. As Ted Sider (p.c.) pointed out to us, there are other kinds of well-behavedness which one might take to be required of possible quantifier-meanings that are not automatically preserved by taking conjunctions and disjunctions. For example, the truth of 'If there is a tolerant sandcastle and an intolerant sandcastle, there is a *pair* of sandcastles one of which is tolerant and the other of which is not tolerant' on both input meanings does not guarantee its truth on the conjoined meaning. (Sider made a similar point using plural quantification.) But even if the straightforward conjunctions and disjunctions are ruled out as natural quantifier-meanings on such grounds, we still know of no reason to deny that in the cases at hand there are broader quantifier-meanings that are well behaved in further respects such as preserving closure principles for pairs, collections, sets, pluralities, and so on.

along with the idea that not just the quantifiers but "virtually any other general word" are shifting around at the relevant nearby worlds, the view will involve a strictly more radical thesis of semantic variation than our own proposal. For what it is worth, despite our embrace of plasticity in names and common nouns, we still find it hard to believe that some small change in one carpenter's workshop would suffice for a language-wide semantic revolution.

(iii) Plasticity in "categorical" predicates. Let's turn finally to the strategy that buys security by pleading plasticity for expressions like 'originally compose', 'three feet long', etc.[7] Given the obvious connections between these expressions and mereological and spatial vocabulary, the family of varying words will certainly have to include 'part' and 'located at'. Other words that will plausibly turn out plastic on the approach include myriad ordinary predicates having to do with the physical properties of and physical relations between macroscopic objects: 'solid', 'massive', 'wooden', 'large', 'spherical', 'electrically charged', 'larger than', 'five feet from', 'inside', 'touching', 'kick', 'crush', and so on.[8] Moreover, given that many of the Tolerance and Non-hypertolerance speeches we are trying to save have to do with the circumstances under which things do and don't exist concretely, the family will also need to include 'concrete'.

Let's see how a view like this might handle a simple Tolerance Puzzle. Suppose a three-legged stool, Bill, is made of a top and three legs *A*–*C*, and we are trying to safeguard the following speech: 'This stool could have had any one leg distinct from *A*–*C* as a part along with two of *A*–*C*, but could not have had two legs distinct from *A*–*C* as parts.' Now consider a world $w^\star$ where Bill is made with legs *B*–*D*. At $w^\star$, the corresponding speech substituting *B*–*D* for *A*–*C* had better come out true. On the view in question, 'this stool' still refers to Bill, and the truth of the speech is explained by 'part' standing for some other relation: call it 'parthood*'. At $w^\star$, *B*–*D* are both parts* and parts of Bill. Bill could have had any three legs two of which are among *B*–*D* as parts*, but could not have had two legs not among *B*–*D* as parts*. At any world where *A* and *B* are unused, Bill isn't concrete. But at at least some of those worlds, Bill still has three legs as parts*, and thus is concrete* (where concreteness* is the property expressed by 'concrete' at $w^\star$). For example, there is a world $w^{\star\star}$ where Bill is nonconcrete but concrete*, and has *C*–*E* as parts*.

Note that if we posit plasticity in words like 'concrete' and 'part', it will be very hard not to have it spill over into common nouns like 'stool' as well. *Pace* Fine

[7] Williamson (1990: 141) considers a view like this, on which 'the general terms "drawing" and "language" might pick out determinate classes of objects, even if the constitution relation in which their materials stand to them is undetermined'; he remarks that 'This view may or may not be stable under metaphysical reflection'.

[8] One worry is the possible clash between the posited wholescale shiftiness and the tempting view that some of the relevant predicates—e.g. 'electrically charged'—refer to highly *natural* properties (Lewis 1983*a*) and should therefore be semantically stable. One might keep this concern at bay by thinking that the only highly natural properties in the vicinity are properties of (and relations between) fundamental objects (e.g. particles or spacetime points).

(2005*b*), 'Necessarily every stool is concrete' seems like one of the things that should be stably true; if 'concrete' is shifting around, then 'stool' will need to shift too to safeguard such speeches. This might seem to sacrifice any advantage that the view under consideration might be thought to have over ours. However, as we have seen, the view goes naturally with the idea that names and demonstratives are not nearly as plastic as they are for us: for example, 'this stool' would still refer to Bill if Bill's legs were *B*, *C*, and *D*. And more importantly, the view does not require multiplying objects in anything like the way that our package did. For example, the treatment of the stool example doesn't require that Bill is coincident with any other object.

Trying to further pin down the metasemantic behaviour of names and demonstratives on the view brings to light some further choice points. In the previous example, consider a possible world w^{**} where a stool is made using legs *C*–*E*, and Bill is non-concrete but concrete* (with *C*–*E* as parts*). Call the stool made using *C*–*E* 'Alice'. When we point towards *C*–*E* in w^{**} and say 'this stool', do we refer to Alice or to Bill? There is pressure to think it is Alice: after all, Alice is right there in plain sight in front of us. It would be odd to think that we could quite easily have used 'this stool' to refer to something we can't see! But if we refer to Bill in both the actual world and in w^{*}, but don't refer to Bill in w^{**}, then prima facie we will be landed with a worrying instability in the truth values of sentences about *reference* between actuality and w^{*}. At the actual world we can truly say, 'If a stool had been made with *D* rather than *A* as a part, we would have referred to it when we said "this stool"' (since if a stool had been made from *B*–*D* it would have been Bill, and 'this stool' would have referred to Bill); at w^{*}, by contrast, the corresponding speech about *E* and *B* seems to be false, since 'If a stool had been made with *E* rather than *B* as a part it would have been this stool' is true (bearing in mind that 'made from' at w^{*} denotes a relation that Bill bears to *C*–*E* at w^{**}). This raises a pressing security worry about such speeches using 'refer'. We can, however, try to stabilize things by positing that the co-ordinated plasticity that affects 'part of', 'concrete', and 'stool' also encompasses words like 'refer'. At w^{*}, 'refer' refers to a different relation—reference*—which the relevant uses of 'this stool' stand in to Bill at w^{**} as well as at w^{*}, even though 'this stool' *refers* to Alice at w^{**} and Bill at w^{*}. This plasticity will plausibly then also infect a wide swath of other speech-act and psychological vocabulary.

Plasticity in the very semantic vocabulary that we use in characterizing the phenomenon of plasticity is extremely challenging to think through, especially if it supposed to be co-ordinated with various tranches of non-semantic vocabulary.[9] We see no immediate path here to reducing the view to incoherence. But it is one thing to place some constraints on how the posited plasticity has to work

[9] See Hawthorne 2006*b* for an attempt to navigate some of the difficulties.

in order to combat various Security Arguments, and it is quite another thing to get a real explanatory handle on what the family of "candidate" referents (and candidate referents*, and so on) is for any of the relevant words, and what the co-ordination between the candidates consists in. We leave further pursuit of these pressing challenges for another occasion, hoping that readers will share our sense that this dizzying picture is rather undermotivated given the smoothness of our favoured, plenitude-based package.

13.2 The Challenge from Speech Reports

Suppose that in the actual world, having defined 'tolerant' as in Chapter 1, we utter the sentence 'Every table is tolerantly a table', thereby saying, truly, that every table is tolerantly a table. On the view we are going for, various modest differences in the positioning of the boards in some carpenter's workshop in Idaho would be enough to induce differences as regards what we would have been saying when we made that utterance. If the boards had been shifted in such a way as to produce a table living on the edge, we wouldn't have been saying that every table was tolerantly a table, since we wouldn't have been saying anything false, just as we aren't saying anything false when we make the utterance in the actual world. But we would have been saying *something*—there is a property, call it *being a table**, such that we would have been saying that every table* was tolerantly a table*. And surely in those worlds, the practice of using 'table' to talk about the property of being a table* isn't some rare idiosyncracy that only arises when people start talking about issues of originating matter. To preserve the evenhanded treatment of the worlds required to avoid security worries, we had better say that the practice of using 'table' to denote *being a table** is about as common in the relevant counterfactual world as the practice of using 'table' to denote *being a table* is in the actual world. For example, suppose that in the actual world Alma says, 'There is a round table in my living room', thereby asserting that there is a round table in her living room. In the world at issue, Alma presumably asserts, in uttering that sentence, that there is a round table* in her living room (assuming for simplicity that 'table' is the only word that shifts).

If we take the proposition that there is a round table in Alma's living room to be distinct from the proposition that there is a round table* in Alma's living room, and take it that in uttering the relevant sentence Alma would only assert one proposition, we seem to be driven to the unfortunate conclusion that our ordinary thought and talk about what people say in other possible situations is riven with error. Our ordinary practices seems to presuppose that when a person asserts a proposition by uttering some sentence, then small and distal changes in the world that don't impact whether they utter that *sentence* typically don't impact whether they assert that proposition. For example, we don't feel much of a gap between the

counterfactual direct speech report, 'Whatever happened in Idaho yesterday, Alma would have said "There is a round table in my living room"' and the corresponding counterfactual indirect speech report, 'Whatever happened in Idaho yesterday, Alma would have said that there was a round table in her living room.' Likewise we don't feel much of a gap between 'Alma was likely to say "Dinner is on the the table"' and 'Alma was likely to say that dinner was on the table.' It looks like very few ordinary claims embedding indirect speech reports in counterfactuals and chance-ascriptions will be true, given the high chance of some carpenter or other making one of the relevant choices of planks or cuts, and the obvious risk that the downstream effects of pretty much any ordinary counterfactual antecedent will include such choices.

The problem isn't restricted to speech reports in modal contexts, since we have also been willing to embrace quite extensive variation in the properties expressed by uses of the relevant words within the actual world (even within a short interval of time). For example, we suggested that 'sandcastle' expresses different properties in natural utterances of 'The sandcastles would have been more robust if we made them with the wetter sand' and 'We couldn't have made these sandcastles if this sand had been any drier.' And we suggested that 'house' expresses different properties in natural utterances of 'At that point the house only had one wall' and 'At that point there was no house yet since only one wall had been put up.' This raises a worry about ordinary homophonic indirect reports of 'sandcastle' and 'house' speeches, including even speeches like 'We spent the afternoon making sandcastles' and 'My house has seven windows', which would come out true under all of the relevant interpretations of the count nouns. But we will focus here on the challenge posed by "cross-world" speech reports, since positing cross-world semantic variation is so crucial to our treatment of Tolerance Puzzles.

One way of somewhat restricting the amount of error we have to posit would be to adopt a relatively coarse-grained theory of propositions according to which, for example, the proposition that there is a round table in Alma's living room is identical to the proposition that there is a round table* in Alma's living room, despite the fact that the property of being a table is distinct from (and not even coextensive with) the property of being a table*. After all, as we have been thinking about the properties in question, it is plausibly metaphysically necessary that every table coincides with a table* and vice versa, and hence metaphysically necessary that there is a round table in Alma's living room if and only if there is a round table* in Alma's living room. Given Intensionalism (see §1.4), this suffices for the identity of the propositions. But there are many sentences where substituting tablehood* for tablehood will make a difference to the proposition expressed even on the assumption that Intensionalism is true. For example, the proposition that there is a table in Alma's house that would have arrived earlier if there hadn't been a postal strike isn't even necessarily equivalent to the proposition that there is a

table* in Alma's house that would have arrived earlier if there hadn't been a postal strike. Consider possible worlds where Alma's house contains an intolerant table and a coincident tolerant table* , both constructed very recently from planks some of which were delivered by post. In some of these worlds it will be true that if there hadn't been a postal strike, the table* would have arrived earlier, but the table wouldn't have concretely existed at all. So even in an Intensionalist setting, there will be a threat of error when it comes to counterfactual speech reports like 'Whatever happened in Idaho yesterday, Alma would have said that there was a table in her house that would have arrived earlier if there hadn't been a postal strike.'[10] And even more obviously, Intensionalism will not help to save cross-world reports like 'Whatever happened in Idaho, Alma would have said that this table was ugly', since the relevant shift in the reference of 'this table' will make for an intensional difference. Thus even in an Intensionalist setting, there is still pervasive enough error in speech reporting to generate a serious challenge for the plasticity claims we have been relying on.

The main thing we want to say about this challenge is that very similar challenges will arise no matter what we say about Tolerance Puzzles. Everyone agrees that big enough changes in the facts about language use can make a difference to what people say in uttering a given sentence. At worlds where the use of 'salad' is similar to the actual use of 'soup', people who utter 'Salad is delicious' certainly don't say that salad is delicious. But the property of being salad belongs to an enormous family of properties which draw boundaries in slightly different places (how much greenery? how much dressing? how cooked can it be?…). Given that this family doesn't seem to contain any privileged members that are much more suited to being talked about than the rest, it is hard to resist the conclusion that it is quite typical for even very tiny differences in the underlying facts to make a correspondingly tiny, but real, difference to which properties we express when we utter the word 'salad', and to what we say when we utter the sentence 'Salad is delicious.' This independent motivation for pervasive semantic plasticity already raises a pressing challenge to our ordinary practice of counterfactual speech reporting. Whether or not the supervenience base for facts about what property we express by 'table' extends to goings-on in carpenters' workshops in Idaho, almost everyone will agree that it includes goings-on inside our brains. But pretty much every counterfactual possibility we ever talk about will involve at least some slight difference inside our brains! If only one proposition is asserted at a time, and all the candidates get their turn to be asserted in modal space, then it is entirely obscure how we could be confident that the slight differences in question wouldn't

[10] Modal environments aren't the only ones where the difference between tablehood and tablehood* makes a difference. Other examples include semantic environments ('"Tisch" means table', '"Table" stands for the property of being a table') and attitudinal environments ('She wants a table'; 'She is wondering whether to buy a table').

be such as to make a difference to what we say when uttering sentences like 'Salad is delicious' or 'There is a round table in my living room.'[11]

Elsewhere (Dorr and Hawthorne 2014), we have explored a range of possible reactions to this puzzle. One approach that we discuss involves what we call 'contextual inheritance': the idea that the context-sensitivity of a word as used in an indirect speech report can be resolved in such a way that the word inherits the meaning of the corresponding word used by the target speaker. Such a mechanism can do quite a lot to save straightforward counterfactual speech reports, although it has more trouble with quantified cases. The approach that emerges in the best light is one on which it is typical for people to say an enormous number of different things by a single utterance. This view reconciles the idea that even small differences make some difference to what gets said with the idea that there is in many ordinary cases extensive overlap between the ranges of propositions expressed by an utterance at nearby worlds. It's not obvious that such overlap deals adequately with the problem: a mechanical treatment of cross-world speech reports might expect them to end up expressing a mix of truths and falsehoods. But Dorr and Hawthorne (2014) suggest that systematic error in such reports can be kept at bay by a subtler treatment, on which words tend to take on narrower ranges of meanings when they are used in speech reports. Dorr (unpublished *b*) further develops this picture of "plural signification".

Each of the competing ways of reconciling the plasticity in 'salad' with ordinary cross-world speech reporting practices can be extended straightforwardly to deal with the extra dose of plasticity required by our favoured solution to the puzzles of tolerance. While this additional plasticity perhaps adds a little extra pain, the general shape of the menu of options is so similar that it is hard to think that the problems with speech reporting will turn out to be the Achilles heel of our view.

13.3 People

Almost all of our examples have involved inanimate objects. But of course Tolerance Puzzles also arise for people and animals. For example, it seems clear that each person with a particular DNA sequence could have still been a person while having any of the DNA sequences differing from that one on just one base-pair. The Hypertolerance claim corresponding to this Tolerance premise, which entails that each person could have had any other DNA sequence of the same length as

[11] In Dorr and Hawthorne 2014 we lay out this "metaphysical" motivation for pervasive plasticity in much more detail. We also argue that while a competing "patchwork model" that holds meanings constant across large patches of modal space succeeds in minimizing error in cross-world speech reports, it it is subject to a pressing security objection: there will still be a lot of errors at worlds close to the boundaries of the patches, and these counterfactual errors plausibly infect the epistemic status of speeches far from the boundaries.

the one they actually have, seems quite implausible.[12] But the kind of semantic plasticity required by our strategy for resisting the argument from Tolerance to Hypertolerance seems more disturbing for people than for tables.

Some of the more disconcerting results involve *de re* Tolerance Arguments using the first-person pronoun. To explain why you are not just lucky to be speaking the truth when you say, 'I am tolerant (with respect to my DNA sequence)', you have to think that in nearby worlds where you are not tolerant, you assert something else—something which is true in those worlds—when you assert 'I am tolerant'. In keeping with our general strategy of pinning the shiftiness in such cases on the relevant singular term, this means that when you use 'I' in those nearby worlds to say, 'I am tolerant', you do not refer to yourself. That's strange! You might have thought that the practice of using 'I' as a device of self-reference is extremely entrenched, so that a world would have be drastically different from actuality for someone such as yourself not to refer to themselves by an ordinary, literal use of 'I'.

Another thing that might seem strange is for there to be close possibilities where certain conscious people use the word 'conscious' to express a property other than consciousness—a property they don't have. Consciousness, one might suppose, is the sort of property such that if you have it, it is not easily going to escape your notice, and you are going to want to have a word for it. But there is pressure to posit semantic plasticity in 'conscious' analogous to the pressure we have already considered in the case of 'table', since it seems clear that every conscious thing is tolerantly conscious (e.g. with respect to its DNA sequence), and implausible that any conscious thing is hypertolerantly conscious. This also applies to verbs like 'think' and 'choose', as well as to nouns like 'person' and 'philosopher'.[13]

As with 'table', there are interesting views on which it is just false that every conscious thing is tolerantly conscious, so that there is no special reason to posit plasticity in 'conscious'. On one kind of view, we say that every conscious thing is one of an enormous multitude of conscious things, some but not all of which are tolerantly conscious. The central reasons not to like this view are parallel to those that apply to the analogous view about tables: it conflicts with obvious truths about how many conscious things there are in any given room/bed/bus…, and it conflicts with plausible generalizations in the vicinity of 'Every conscious thing is tolerantly conscious.'[14] On another kind of view (analogous to Yablo's (forthcoming) account

[12] Even if we were to willing embrace such Hypertolerance claims, as explained in §9.3 intuitive Robustness judgements will generate a similar motivation for plasticity.

[13] And note that if we do posit semantic plasticity in these words, that will generate additional pressure to posit close worlds where you don't refer to yourself when you say 'I', given that we don't want there to be close worlds where you say something false when you say 'I am conscious', 'I am a thinking thing', etc.

[14] At one point we were tempted to make an additional epistemological objection to the "many conscious thinkers" view. Prima facie, it is tempting to think that you could never know you were *F* if there were lots of non-*F* thinkers experientially extremely similar to you. For example, it is doubtful that you could know you have hands if most people with experiences are like yours are Boltzmann Brains who lack hands. The similarities between you and the many conscious, experiencing beings who are almost coincident with you according to the "many thinkers" view look in some ways even

of 'ship') there is only one conscious thing in any relevant vicinity, but while that thing may be tolerant, it is not *tolerantly conscious*—indeed, if the view is developed on Yablo's model, its consciousness will be extremely fragile. This view seems particularly unappealing, given that the ordinary modal judgments that support the truth of 'I am tolerant' seem equally forceful when it comes to 'I am tolerantly conscious'.

Given our general perspective, there is a lot of pressure on us to grit our teeth and embrace a picture where a whole host of words like 'I', 'person', 'conscious', 'think' are semantically plastic, and work in such a way as to make 'A typical person/conscious thing/thinker does not overlap any others' literally true.[15] How problematic is this?

In gauging the cost, one should bear in mind that the pressure to posit plasticity in all these words doesn't just come from Tolerance Puzzles. There is independent pressure coming from other kinds of variability in our speech practices involving these words within the actual world. For example: if you have recently had had a pacemaker implanted and you are about to be weighed, you might say, 'I am probably going to be a little heavier now thanks to the pacemaker'; but you might instead decide to say, 'Don't forget to subtract the weight of the pacemaker when you are figuring out how much I weigh.' Likewise, if in the future you had an

more intimate and worrying. But according to the view, each tolerant person is very similar to many overlapping non-tolerant people; so, the tempting principle looks to imply that even the tolerant ones might not be able to know that they are tolerant. However, if we are going to posit semantic plasticity in 'I' in any case, that may help us out here. So long as we are willing to say that none of the thinkers is making a mistake when they say 'I am tolerant', because at least the ones who are not tolerant are not referring to themselves when they say 'I', it is hard to see why the multiplicity would threaten the status of the relevant beliefs as knowledge. (Compare the way in which "many tables" theorist can avoid worries about our failing to express knowledge with speeches of the form 'That table is tolerant' by claiming that demonstratives gravitate to tolerant tables.) On one toy model of such within-world variability in 'I', tolerant thinkers refer to themselves by 'I', while intolerant thinkers refer to one of the tolerant thinkers they overlap with; on another toy model, every thinker refers by 'I' to some *uniformly* tolerant thinker they overlap with, with vagueness in case of ties. Noonan (2010) makes a structurally similar proposal on behalf of the view that each person coincides with a human animal that is not a person, but is a thinker: on his view, while people refer to themselves with 'I', human animals refer not to themselves but to the people with whom they coincide.

[15] Sider (2003) also defends a view that combines a rather abundant ontology of material objects (roughly, the result of deleting the 'necessarily' from Location Plenitude) with the claim that typical conscious things do not overlap any other conscious things. A key part of Sider's defence of this combination is the thesis that consciousness is an *extrinsic* property. Others have seen this thesis as deeply implausible: for example, Merricks (2003*a*) finds it so obvious that "having the phenomenology of consciousness" is intrinsic that he assumes that, in denying that any instantiated intrinsic property suffices for consciousness, Sider must be, weirdly, denying that the phenomenology of consciousness suffices for consciousness. But it is not obvious that we are forced to accept that consciousness is extrinsic, since it is not obvious how the notion of intrinsicness works for properties specified using modal operators. It seems pretty plausible that any *persistent* property (i.e. any property such that necessarily, anything that has it necessarily has it) is ipso facto intrinsic, or at least necessarily coextensive with something intrinsic. This applies even to properties like *being such that necessarily, one is located at all and only those points that are between one and two kilometers from some burning barn*, whose persistence is guaranteed by Iteration. The fact that an object has this property is naturally seen as "just about how things are with that object"; certainly other objects don't seem to be relevant. But if we are happy to count persistent properties like this as intrinsic, there is little pressure to deny the intrinsicness of consciousness. We could for example identify consciousness with the conjunction of some persistent property with some paradigmatically intrinsic property having to do with the physical state inside an object's boundaries, or with a disjunction of such conjunctions.

operation like the one depicted in the film *RoboCop*, where your head and spinal column end up embedded in a hulking metal robot-body, you might in one mood say, 'Now I am mostly made of metal', and in another mood 'All that's left of me is a head and spinal column, but thanks to these devices that are attached to me I am still quite effective as a police officer.' And even closer to home: in one mood we might say, 'Most people are a bit heavier in the evenings just after dinner', and in another 'We only gain weight once food is absorbed into our body.'[16] Given a broadly physicalistic outlook, it is hard to take seriously the idea that there is any important difference between such variations and similar variations in our ways of talking about inanimate objects—for example, the fact that we might at one time say, 'The shoes are a little heavier now thanks to the new insoles' and at another time 'Don't forget to subtract the weight of the insoles when you're weighing the shoes.' In particular, it is hard to take seriously the idea that some of the speeches involving words like 'I' and 'person' betray systematic errors in a way that analogous speeches involving words like 'shoe' do not. Moreover, insofar as our treatment of the examples involving inanimate objects appeals primarily to shiftiness in words like 'shoe', 'table', and 'this' rather than in expressions like 'part of' and 'weighs', it would be strange to think that shiftiness in 'part of' and 'weighs' (rather than in 'person' and 'I') suddenly starts doing all the work when it comes to analogous cases involving people.[17] But positing shiftiness in 'person' and 'I' leads to a model where it is very common, even within the actual world, for people to use 'person' to refer to things that are not people, to use 'I' to refer to objects other than themselves, and so on.[18]

There is thus strong reason, thanks both to Tolerance Arguments and to other aspects of ordinary practice, to think that the whole panoply of agential and

[16] See Hawthorne 2006*b* for similar variations involving slightly different linguistic communities.

[17] Some of the shiftiness in 'heavier' and 'weighs' sentences does look to be due to shiftiness in 'heavier' and 'weighs', rather than in the relevant singular terms and count nouns. For example, 'The car is much heavier now that the whole family is in it' can seem true, but we don't want to say that the people in the car are parts of the car; similarly for 'You are heavier than you were before your coronation thanks to the weight of that crown.' There seems to be a division of labour between shiftiness in words like 'heavier' and shiftiness in the relevant common nouns and singular terms. What would be strange would be for the division to be very different when it comes to 'person' and 'I'.

[18] Johnston (1989, 2010) develops a view that achieves charity to a wide variety of speeches of the form 'I would/would not survive a change of such-and-such sort' *without* denying that 'I' is robustly a device of self-reference, by saying that which sorts of episodes a person would and would not survive are determined by features of that person's psychology such as their self-conception, or their patterns of future-directed concern. For example, people who are in the habit of saying things like 'Teletransportation is a great way of getting from one place to another', and act accordingly, will survive teletransportation. On the other hand, people who say, 'Teletransportation is death' are right that entering a teletransporter would kill *them*, though wrong if they think it would also kill people from the first group. (An analogous view in the spatial case would hold that if you started finding it natural to say things like 'I am entirely constituted by my brain, and the rest of my body is something that surrounds me', you would shrink to the size of a brain.) This sort of view of survival is already very radical when the relevant differences involve deep-seated psychological factors, as on Johnston's proposals. It is even harder to take seriously an extension of the view where the vicissitudes of conversation just prior to the relevant event make the difference between survival and death.

psychological vocabulary is quite shifty.[19] This picture is unsettling: it makes personhood much less *important* in certain ways that one might have thought.

Let's see how this dethronement of people might play out in the moral domain. It is natural to think that questions about the survival of people have a kind of moral significance that questions about the survival of things that are not even conscious lack. Such significance might be partially articulated by an axiological principle along the following lines:

Death Is Bad If there are two possible worlds w_1 and w_2, such that in w_1 a person x dies in the prime of life and is replaced with a new person y, and in w_2 x continues to exist and is at each time similar to whichever of x and y is alive at that time in w_1, and w_1 and w_2 are alike with regard to the lives of all other people, animals, ecosystems etc., then w_2 is *ceteris paribus* considerably worse than w_1.

Consider two worlds w_1 and w_2 such that in w_1 someone dies and is replaced by a new person, whereas in w_2 the original person survives, and everything else is equal in the way required for Death Is Bad to kick in. Suppose, though, that there's a world close to actuality where 'person' is used slightly differently in such a way as to make true 'In w_2 someone dies and is replaced by a new person, whereas in w_1 the original person survives, and everything else is equal.' On this alternative use, 'person' expresses a property whose instances in w_1 and w_2 aren't people, and indeed aren't even conscious in those worlds, though of course 'conscious' as used in the nearby world is true of them. (We might imagine that w_1 and w_2 both involve radical operations in which the amount of continuity in two different respects r_1 and r_2 is close to the boundaries relevant for survival: w_1 features a bit more continuity of type r_1 and a bit less continuity of type r_2; the differences between the two properties expressed by 'person' are due to slightly different weightings of r_1 and r_2.) If we were determined to hold on to person-privileging axiological claims like Death Is Bad, we would be under a lot of pressure to think that 'considerably worse' is also plastic in a corresponding way, so that sentences like Death Is Bad are robustly true. For if we think we *know* Death Is Bad to be true, it is hard to see how such knowledge could be maintained if error was rife at close worlds in this way.[20]

But such plasticity in axiological vocabulary is in tension with the combination of some other ideas. First, we care about the good: when one world is better

[19] Note that this also suggests that the vocabulary in question is all "Twin-Earthable" according to Chalmers's classification (see §11.5): by contrast Chalmers would regard certain expressions like 'conscious' and 'in pain' as paradigms of non-Twin-Earthability.

[20] Death Is Bad might be combined with, but does not require, other kinds of person-privileging axiological claims such as one which would require a significant value-theoretic difference between forms of abortion that involve the death of a person—or of something that would later be a person were it not for the abortion—and forms that do not. Such claims raise analogous issues.

than than another, then (ceteris paribus) we prefer the one that is better. Second, such caring is a virtue: someone whose preferences track goodness in this way is (ceteris paribus) more morally decent than someone who prefers a worse course of history to a better one, or is indifferent between the two. And third, we are not conspicuously lucky to be as morally decent as we are: the actual world isn't like a small island of saints surrounded by vast sea of sinners. The tension should be clear enough. If people at nearby worlds express a different property with 'good', it is completely implausible to suppose that their *preferences* are nevertheless locked onto goodness itself rather than onto that property.

Given these problems, one might well be inclined—as Parfit (1984) was, for closely related reasons—to reject the kind of person-centric axiology encoded in principles like Death Is Bad. One might instead adopt a kind of "moral misanthropy", on which there aren't large value-theoretic contrasts between people and some of the many non-people with which they coincide or largely overlap.[21] On our favoured way of developing this idea, physical continuity brings with it axiological continuity.[22] For example, as we move along continua of cases like those Parfit considers, linking clear cases of survival to clear cases of death-and-replacement, levels of overall moral value vary gradually.[23]

[21] We are referring here only to the coinciding non-people whose modal profiles are very similar to those of people—similar enough to have some of the properties that 'person' could easily have expressed. Plenitude also predicts the existence of a host of objects that are person-like in some respects but whose modal profiles aren't at all like those of people—e.g. an object Juhani* that necessarily coincides with Juhani at any time before which Juhani has never worn a sequinned tuxedo and is nonconcrete at any other time. The role of Juhani* with regard to facts about value seems quite dissimilar from that of any person—if we start with the actual world and gradually modify it until we find a world where Juhani* is only around for thirty years, there is no expectation at all that we are thereby making it worse. Since the use of 'person' would have to be radically different for Juhani* to be in its extension, our argument does nothing to undermine this judgement. In this regard, our argument contrasts with that of Johnston (2017, 2016), who claims that a certain large family of metaphysical views—including but not limited to the "four-dimensionalism" of Lewis (1976)—lead to radically revisionary ethical results. Johnston claims that if there are what he calls "personites"—'shorter-lived very person-like things that extend across part, but not the whole, of a person's life'—they have "moral status". One of his arguments for this turns on the premise that whether something has moral status supervenes on its intrinsic properties (or those of its "physical and mental life"). It is not clear to us where this comes from; but in any case, if we count persistent properties as intrinsic, as suggested in note 15 above, Juhani*'s idiosyncratic modal profile will be an intrinsic property that no person could have. Another of Johnston's arguments turns on the idea whether something has moral status cannot be an "adventitious" matter. But while this premise may work against Lewis, on our view, it is plausibly *necessary* (and thus not adventitious) that Juhani* lacks moral status, though it is true that if Juhani had been murdered after the fashion of Hercules, by being coaxed to don a poisonous sequinned tuxedo, Juhani* would have permanently coincided with something with moral status (namely Juhani).

[22] Qualia-lovers should consider replacing 'physical continuity' with 'physical and phenomenal continuity'. On some but not all conceptions of the extra qualia-theoretic baggage, we will be able to find cases where the qualia-facts are also varying gradually (if at all) across the relevant continua, so including qualia will do little to block the argument against Death Is Bad.

[23] One challenge to this picture involves attitudes like love, respect, and concern which we bear to particular objects. One might think, for example, that genuinely loving a person requires having a non-gradual preferences for histories in which they survive and flourish over histories in which they die, so that the gradual preferences of moral misanthropists are incompatible with love. A fuller defence of the view would require developing a conception of love that is compatible with such gradual preferences.

The view that person-theoretic vocabulary is highly plastic is coherent and principled, and provides a bulwark against relevant Security Arguments. Nevertheless, it is one that takes quite a bit of getting used to. Some philosophers will not be able to get on board, for reasons we have already gestured at: thoughts about the ease of self-reference, or the attention-grabbingness of consciousness, or the axiological centrality of people. If you are in this camp, and you still want insist that you could have been originally composed by a slightly different collection of atoms or have had a slightly different DNA sequence, one important way of invoking semantic plasticity to block the Security Argument for Non-contingency is unavailable to you. There is still the possibility of appealing to semantic plasticity in some other words, for example in 'originally composed by'. But as we saw in §13.1, such views face deep problems of their own; moreover, it is far from clear that they can be developed in such a way as to insulate the person-theoretic vocabulary from plasticity; for example, it will be hard to keep plasticity in 'concrete' from spilling over to many predicates *F* for which 'Necessarily every *F* thing is concrete' seems stably true. The upshot, then, is that person-lovers will find it hard to block the Security Argument for Non-contingency in Tolerance Arguments having to do with people. Given the case for Iteration that we developed in Chapters 7, 8, and 9, that means they will be pushed to accept that people are *hypertolerant* with respect to many different dimensions of variation.

As we discussed in Chapter 5, such an embrace of Hypertolerance can be developed in two different ways. One possibility is to deny Microphysical Supervenience, and another is to accept "coarse-grained" Hypertolerance claims while denying the fine-grained ones that conflict with Microphysical Supervenience, holding that the identities of ordinary objects depend on apparently irrelevant details of the microphysical supervenience base. Against both of these views, §6.3 and §9.3 raised a challenge based on the security of Robustness judgements, to the effect that certain objects had a high chance of having certain relevant properties conditional on certain specifications of the approximate physical situation. The same challenge can be raised in the special case of people. However, Robustness judgements about people are much less firmly entrenched than Robustness judgements about tables, earrings, and the like, and it is not out of the question to embrace a picture where people are extremely non-robust. Traditional Cartesian or Platonic dualists might be happy to claim that at every time until shortly before the embodiment of a person, the chance of that person's ever being embodied was minuscule, even conditional on extremely detailed physical facts, such as the fact that a particular sperm and egg were going to fuse in such a way as to give rise to a single human body. This is surprising, but is not in any obvious way at odds with ordinary beliefs. If person-lovers don't mind going along with traditional dualists on this point, they may be able to stabilize one or other form of hypertolerantism.[24]

[24] Note, however, that the coarse-grained hypertolerance picture seems to require heavy doses of vagueness in order to avert an otherwise pressing worry about the arbitrariness of any specific account

The person-loving point of view thus leads to a position in some ways reminiscent of traditional substance dualism. Its proponents might feel reassured by the thought that they are not are still not going all the way to Cartesianism, since after all, they can still think that people are material objects. However, it is far from clear that there is a satisfying way of staking out an intermediate position of this sort. There is something deeply implausible about the idea that, as one traces a continuous spatial path that starts miles away from a person and ends up in their pineal gland, at some microphysically undistinguished point one will cross a natural boundary from the outside of a person (conscious being, locus of value…) to the inside: a boundary that corresponds to the extension of a property that is robustly referred to across modal space. Assuming that atoms can be part of people in the same sense that they can be parts of tables, it seems preposterous to have a very different attitude to the questions 'Which atoms near the tip of a person's fingernail are inside the boundaries of the person?' and 'Which atoms near the point of a pencil are inside the boundary of the pencil?'. If the answer to the latter turns on the happenstance of semantic plasticity, then the former should surely have the same status.[25] Substance dualists have a reasonable kind of answer to this charge. They can say that the sense in which a point or particle can be "inside" or "outside" a person is quite different from the canonical way in which a point or particle is "inside" or "outside" an inanimate object. Perhaps what is really going on is that people bear a relation of "influence" to spatial points to different degrees, in such a way that when the spatial distance between two points is small, one's level of influence over them is similar. 'Located at' may sometimes have a broad sense in which it is equivalent to the disjunction of location proper (a relation that persons don't bear to any points) and something tied to a threshold of influence. When used in the broad sense, it is subject to semantic plasticity. One could take a similar attitude to words like 'part', 'originally compose', and so on. In that way, the substance dualist can vindicate the claim that the question about the exact boundaries of the table and the question about the exact boundaries of the person are both semantically plastic, although they will have entirely different stories about the sources of plasticity in the two cases. By contrast, if one attempts to combine the view that 'person' and related vocabulary are non-plastic with the claim that people have locations and parts in every sense in which tables do, one is left with the bizarre view that one particular spatial boundary grabs semantic attention even though it looks entirely uninteresting to the physicist's eye.

In addition to being hard to believe in itself, this view also raises sceptical concerns about our knowledge of our boundaries. You don't have to go all *that*

of the dependence works. Insofar as one takes vagueness to involve "multiple candidates", this might in turn make it difficult to use coarse-grained hypertolerance as a basis for resisting plasticity.

[25] By contrast, it doesn't seem preposterous for an anti-physicalist to say the analogous thing about temporal boundaries; certainly the idea that there is a natural temporal boundary in the progression from a sperm and egg to a baby seems quite in the spirit of Cartesian dualism.

far from actuality to reach a world where people say things like 'I am shaped like a brain' rather than 'I am shaped like a body.'[26] For us, such worlds provide no impulse towards scepticism:, since we can say that they and we are both right since their uses of 'I' refer to brain-shaped things. The substance dualists will likewise be unworried, since on their account all that's going on is that the threshold of influence required for 'inside a person' to apply to a point is different, so that again, no-one is making a mistake. But for the non-dualist person-lover, it is very hard to resist the conclusion that one side is making a mistake. They will thus need a far more brazen optimism if they want to stick to ordinary judgements about their boundaries.[27]

So, while our collective first-choice view embraces physicalism, the plasticity of 'person', and moral misanthropy, our second-choice view is the one that embraces substance dualism, the non-plasticity of 'person', and the moral privilege of people. The popular combination of person-loving and physicalism seems much harder to defend.

That concludes our presentation and defence of our favoured way of resolving Tolerance Puzzles, a task that has occupied this and the previous two chapters. As with every serious treatment of these puzzles, ours requires taking on large and controversial commitments that ramify in all sorts of ways elsewhere in philosophy: that's what makes the puzzles so interesting. But to our eyes, the distinctive commitments to plenitude and plasticity that provide the basis for our approach have a systematicity and explanatory power unrivalled by the other approaches. We hope and predict that as their implications for other areas of philosophy are traced out, they will bring further enlightenment in their wake.

[26] Consider also a community whose ancestors were deeply influenced by Clark and Chalmers (1998), where people naturally say things like 'I am a scattered object, because part of my memory system includes my iPhone and diary.'

[27] In principle one could have a view where there are many thinkers in the vicinity but they all are tolerant: perhaps they only differ slightly in their modal profiles, around the very edges. But it is unclear why anyone would want to hold such a view, given its obvious problems doing justice to ordinary uniqueness speeches.

14
Indiscernible Tolerance Arguments

In the final two chapters of this book we will study a special class of Tolerance Arguments, which raise distinctive issues that are not straightforwardly resolved by the previous discussion. These new puzzles involve a narrow family of modalities, which we call *indiscernible* modalities, whose hallmark is that the truth values of qualitative propositions are held fixed, so that any qualitative truths are automatically necessary. This means that in Tolerance Arguments based on indiscernible modalities, there is a distinctive argument for Non-contingency, based simply on the the qualitativeness of the relevant Tolerance premise. This purely metaphysical argument has very little in common with the Security Argument (§3.3) which provides the strongest motivation for Non-contingency in the Tolerance Arguments we have considered up to now. Thus our central resources of plenitude and plasticity, which Chapter 11 used to stabilize the denial of Non-contingency in many other Tolerance Arguments, do not immediately suggest any particular response to these Indiscernible Tolerance Arguments. This chapter will introduce the arguments, and survey options for addressing them while accepting the qualitativeness claims that guarantee the truth of their Non-contingency premises. Chapter 15 will turn to the the option of rejecting the qualitativeness claims. We will not presuppose plenitude, or any of the other positive claims we have made in Chapters 11–13, though we will of course be keeping an eye on ways in which the decision what to say about Indiscernible Tolerance Arguments might interact with the broader dialectic about Tolerance Arguments.

14.1 Qualitativeness

The key concept we need for the new family of arguments is that of *qualitativeness*. This concept has come up in passing in earlier chapters, but hereafter it will take centre stage.

Unfortunately, there is no uncontroversial definition. In the absence of such a definition, one way to introduce the concept is to give a list of examples. Prima facie, the properties, relations, and propositions in the left column in the following list are qualitative, while those on the right are non-qualitative:

The Bounds of Possibility: Puzzles of Modal Variation. Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri, Oxford University Press. © Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri 2021.
DOI: 10.1093/oso/9780192846655.003.0015

being spherical	*being taller than Jean Sibelius*
identity	*being identical to Lasse Virén*
touching	*touching this particular computer*
being a city-dweller	*being a New Yorker*
owning	*being preferred to by Jean Sibelius*
that no computer is spherical	*that Jean Sibelius owns a spherical computer*
that most people own a computer	*that not everyone is a New Yorker*

However, lists of examples are somewhat tendentious, since as we will see, the puzzles we will be discussing in this chapter can be used to motivate some surprising denials of qualitativeness.

Another potentially helpful way into the concept of qualitativeness connects it to the concept of *symmetry*: the only way it could happen that two distinct objects have exactly the same qualitative properties would be for the universe to be perfectly symmetric, under a symmetry that maps one of those two objects to the other.[1]

While there is no definition of qualitativeness that is both uncontroversial and potentially illuminating to someone who found the notion unclear, there are some that have one of these two properties. One fairly uncontroversial but not so illuminating definition, which we will consider further in Chapter 15, characterizes qualitativeness (for a given type) in terms of a relation of *aboutness*, which entities of that type can bear to collections of objects: a qualitative entity is one that is about the *empty* collection of objects. (By contrast, e.g., *being taller than Jean Sibelius* is about the singleton collection of Jean Sibelius and not about the empty collection.) This seems unlikely to help anyone who found qualitativeness unclear, since they will likely find the relevant notion of aboutness equally unclear. Several potentially more illuminating definitions of qualitativeness have been proposed, but they require controversial commitments of various sorts. Some employ other difficult notions, like fundamentality and naturalness.[2] Others employ only familiar logical vocabulary, but are defensible only in the presence of certain debatable premises

[1] This gloss requires somewhat broadening the familiar geometric conception of a symmetry: if two co-located particles have exactly the same qualitative properties, a function that maps one to the other could count as a "symmetry of the universe" even though there are no symmetries geometrically speaking.

[2] Most influentially, Lewis (1986*a*: 60) takes qualitative properties to be those whose distribution supervenes on the pattern of *perfectly natural* properties and relations. (See Dorr 2019: §4.3 for discussion.) A related approach takes the qualitative entities to be those which are *definable from* some special (e.g. fundamental or perfectly natural) properties and relations, where definability is not spelled out in terms of supervenience. For example, Adams (1979), in his classic exposition of the notion of a qualitative property (which Adams calls a 'suchness'), suggests taking the qualitative properties to be those definable from "basic suchnesses". None of these theories are straightforwardly expressible in our higher-order framework as it stands, since they seem to require a way of collecting together entities from infinitely many types (unless we are willing to take for granted that the special properties and relations only inhabit some finite set of types).

about fineness of grain.[3] Our discussion of qualitativeness in what follows is compatible with, but does not require, any of these definitions.

Despite the lack of an uncontroversial definition and the wide range of debatable cases, we take the notion of qualitativeness to be well understood. Moreover, we will be assuming some basic principles about qualitativeness, which are guided by the idea *whatever is definable from qualitative ingredients is itself qualitative*. For example: the conjunction of two qualitative properties is qualitative; the converse of a qualitative relation is qualitative; the result of predicating a qualitative operation of a qualitative proposition is qualitative; the proposition that every member of a collection of qualitative propositions is true is qualitative. In our formal language, qualitativeness can be expressed by a family of predicates Qual_σ of type $\langle\sigma\rangle$ for any σ, and the general pattern behind these examples can be captured by what we will call the "Basic Theory of Qualitativeness", given by the four axioms in Figure 14.1.

Qualitative Closure $\forall x_1 \ldots \forall x_n(\text{Qual}(x_1) \wedge \cdots \wedge \text{Qual}(x_n) \rightarrow \text{Qual}(a))$, where $n \geq 0$ and a is some complex term containing no constants and no free variables other than $x_1, \ldots, x_n$.

Qualitative Constants $\text{Qual}(c)$ where c is $\neg$, $\wedge$, $\vee$, $\forall_\sigma$, $\exists_\sigma$, $=_\sigma$, Rigid_σ, or Qual_σ (for any type σ).

Qualitative Collections $\forall F(\forall x(Fx \rightarrow \text{Qual}(x)) \rightarrow \exists C(\text{R}C \wedge \text{Qual}(C) \wedge \forall x(Fx \leftrightarrow Cx)))$.

Persistent Qualitiveness $\forall x(\text{Qual}(x) \rightarrow \Box\,\text{Qual}(x))$ (where $\Box$ is metaphysical necessity).

Fig. 14.1 The Basic Theory of Qualitativeness.

[3] One candidate definition is considered by Khamara 1988, Dorr (2016*b*), and Goodman (unpublished *d*): proposition p is qualitative iff there is no property X and object y such that $p = Xy$; more generally, a relation Y of type $\langle\sigma_1, \ldots, \sigma_n\rangle$ is qualitative iff it cannot be formed by plugging an object into an argument place of some other relation, i.e. iff there is no object x and Y' of type $\langle e, \sigma_1, \ldots, \sigma_n\rangle$ such that $Y = \lambda z_1 \ldots z_n.Y'(x, z_1, \ldots, z_n)$. Note that if Booleanism is true (see note 22 in Chapter 1), then *nothing* is qualitative according to this definition. For example, there are no qualitative propositions, since according to Booleanism $p = (\lambda x.p \wedge (Fx \vee \neg Fx))(x)$ for any proposition p, property F, and object x; we can do something similar in any other type. Likewise, if the principle of *vacuous Beta-conversion* (Chapter 1, note 13) is true, nothing is qualitative on this definition, since for any p and x, $p = (\lambda y.p)(x)$. And the definition would also be untenable in a very fine-grained theory where failures of non-vacuous Beta-conversion are rampant: a proponent of such a theory might think that for no proposition p, property F, and object x is it true that $\neg p = Fx$, but surely it would be absurd to think that the negation of every proposition is qualitative.

Adams (1979) adopts an interesting hybrid approach: he adopts something like the Khamara definition as an account of being a *basic* suchness, but then defines qualitativeness (*being a suchness*) as being "definable from" the basic suchnesses.

Qualitative Closure captures the basic idea that things definable in qualitative terms are themselves qualitative. Qualitative Constants adds that among the qualitative resources we get to draw on in defining new qualitative things are all our usual logical constants as well as qualitativeness itself. Qualitative Collections says roughly that any collection of qualitative entities is itself qualitative.[4] Persistent Qualitiveness says that everything qualitative is metaphysically necessarily qualitative. All four principles, and their (metaphysical) necessitations, seem like natural conditions of adequacy for any candidate definition of qualitativeness.

Since we allow for meaningful attributions of qualitativeness to entities of any type, we allow for meaningful attributions of qualitativeness to *objects*. This is somewhat unfamiliar, and might seem unintelligible. But if we don't initially understand a notion of qualitativeness for objects, we can introduce one by stipulating that for x to be a qualitative object is for x's haecceity—$\lambda y.y = x$, the property of being identical to x—to be a qualitative property. It is not obvious that there *are* any qualitative objects in this sense. But it is also not obvious that there aren't. For example, if one takes *being greater than* to be a qualitative relation, and thinks that to be identical to God is to be greater than anything else, one will take God to have a qualitative haecceity, and thus to be a qualitative object according to our proposed definition (see Russell 2018: §2). Likewise if one takes *coinciding with* to be a qualitative relation and *being an iron sphere* to be a qualitative property, and thinks that there is an object x such that to be identical to x is to be something that, necessarily, coincides with every iron sphere if there is an iron sphere and any two iron spheres coincide, and otherwise doesn't coincide with anything.[5] Our discussion will be independent of the question whether there are or could be qualitative objects.

Persistent Qualitiveness is somewhat less integral to the topic than the other three axioms, since it concerns the interaction of qualitativeness with modality rather than the pure theory of qualitativeness. Still, it is widely taken for granted by metaphysicians theorizing about qualitativeness, and it is hard to think of a compelling reason for denying it: what else would have to be true for redness, say, to be non-qualitative?[6] It might be given up in the context of certain particularly uncompromising implementations of a "combinatorialist" account of the range

[4] This claim—i.e. $\forall C(\mathsf{R}C \rightarrow \forall x(Cx \rightarrow \mathrm{Qual}(x)) \rightarrow \mathrm{Qual}(C))$—entails Qualitative Collections given Rigid Comprehension, and follows from Qualitative Collections on the assumption of Rigid Extensionality (see §1.5). We choose the weaker formulation for the same reason that we prefer not to presuppose Rigid Extensionality.

[5] The existence of such an x follows from the "identity strength" version of Location Plenitude discussed in §11.3.

[6] One might wonder whether denying ND would pose a problem for Persistent Qualitiveness. But even if there are possibilities where something that is in fact qualitative becomes identical to something that is in fact non-qualitative, it seems far more plausible to suppose that the identification would involve the non-qualitative thing becoming qualitative rather than the opposite. Certainly if one thought that the proposition that John is Cian was non-qualitative and possibly identical to the qualitative proposition that everything is self-identical (as an ND-denying Classicist might think), one should take the former to be possibly qualitative rather than taking the latter to be possibly non-qualitative.

of metaphysical possibilities.[7] But for our purposes it wouldn't matter much if Persistent Qualitiveness were rejected on these grounds, since we could respond by replacing metaphysical possibility in the relevant claims with a narrower notion of possibility on which true attributions of qualitativeness (to entities of whichever type we are concerned with) are "held fixed" and thus automatically counted as necessary.[8]

There is one further principle about qualitativeness which might seem like a natural addition to our list, namely the converse of (the $n \geq 1$ case of) Qualitative Closure.[9] According to this principle, if a can be defined entirely in terms of one or more ingredients $x_1, \ldots, x_n$, and a is qualitative, all of $x_1, \ldots, x_n$ must be qualitative: non-qualitativeness is hereditary. But this is far more tendentious than the above principles, and we will not assume it. Given Intensionalism (see §1.4), this principle would have the undesired consequence that *everything* is qualitative. Recall that according to Intensionalism, there is only one impossible proposition $\bot$. $\bot$ is qualitative, since it can be expressed by a closed sentence involving only logical constants, namely $\forall pp$. But the hereditariness principle entails that whenever $q \wedge \neg q$ is qualitative, q is qualitative; given Intensionalism, this will entail that every proposition is qualitative.[10] We can do something parallel for properties and relations of any type. Since we don't want to presuppose the truth or falsehood of Intensionalism, we will not take a stand on whether non-qualitativeness is hereditary.[11]

14.2 Indiscernible Tolerance Arguments

Using the notion of qualitativeness, we can define a kind of modality in which all qualitative truths are held fixed. Given some starting modality $\Diamond$—e.g.

[7] For example, Bacon's (2020) "Logical Combinatorialism" implies that it is *broadly* possible for any *fundamental* property to instantiate any instantiated "pure" property of properties. This would conflict with Persistent Qualitiveness if one held that (i) some fundamental properties are qualitative; (ii) qualitativeness is pure; and (iii) metaphysical necessity is the broadest necessity. Bacon (2018*a*) himself denies (iii), but we defended it in Chapter 8. (i) looks very hard to deny on the assumption that there are any fundamental properties—indeed it is plausibly necessary that all fundamental properties are qualitative. And (ii) is at least natural given that on Bacon's conception other logical constants are pure, though as Goodman (unpublished *b*) shows, Logical Combinatorialism implies some surprising denials of purity.

[8] Where σ is the type in question, p is possible in the narrower sense iff p is metaphysically compossible with the collection of all true propositions q such that for some x of type σ, $q = \text{Qual}\, x$. It's an interesting question whether we could define a modality that holds fixed qualitativeness facts for all types simultaneously, but this would not be needed for our purposes.

[9] This is equivalent to the principle $\forall x(\text{Qual}(a) \rightarrow \text{Qual}(x))$, where x is free in a. Note that the $n = 0$ case of Qualitative Closure, which is not a conditional at all, just amounts to the claim that all *combinators*—closed terms like $\lambda p.p$ with no constants—express qualitative entities.

[10] This reasoning does not require the full strength of Intensionalism or even of Classicism (see Chapter 8). Just from Booleanism (see note 22 in Chapter 8), we can derive that if any proposition is qualitative all are.

[11] The Khamara theory of qualitativeness (see note 3) accepts the hereditariness principle and thus requires giving up Intensionalism to avoid trivialization.

metaphysical or nomic possibility—say that *p* is *indiscernibly ◇-possible* just in case for every collection *C* of true qualitative propositions, ◇(*p* and every proposition in *C* is true). Likewise, *p* is *indiscernibly □-necessary* just in case there is a collection *C* of true qualitative propositions such that □(if every proposition in *C* is true, then *p*).

There are some general theoretical views on which indiscernible *metaphysical* possibility is a quite uninteresting status. Here is one such view, which we encountered in Chapter 5:

Weak Anti-haecceitism Every truth is metaphysically necessitated by a qualitative truth.

In terms of our new terminology, Weak Anti-haecceitism is equivalent to the claim that every truth is indiscernibly metaphysically necessary. As we discussed in §5.1, Weak Anti-haecceitism doesn't seem very plausible to us. It seems to require either an absurd hyperessentialism, according to which nothing could be concrete while having any qualitative properties other than those it in fact instantiates, or a weirdly conspiratorial form of moderate essentialism, according to which the range of possible qualitative profiles for an object is nontrivial, but somehow turns out never to include any of the qualitative profiles actually instantiated by other objects. We will proceed on the assumption that Weak Anti-haecceitism is false, and indeed that there is a large space of metaphysically possible worlds that agree with actuality as regards all qualitative propositions.

Assuming that indiscernible modalities aren't uninteresting because of Weak Anti-haecceitism, they are interesting for our purposes because of the availability of a distinctive kind of argument for the Non-contingency premise in certain Tolerance Arguments involving them. Recall that in our abstract schema for Tolerance Arguments, Non-contingency says that Tolerance is necessary if true, where Tolerance is the proposition that for every object *x* and properties *F* and *G*, if *x* is *K* and *F*, and *G* is close to *F*, then it is possible that *x* is *K* and *G*. If 'possible' and 'necessary' here stand for some indiscernible possibility and necessity, then Non-contingency follows from the premise that *Tolerance is a qualitative proposition*: the signature feature of indiscernible modality is that every true qualitative proposition is indiscernibly necessary. And given Qualitative Closure and Qualitative Constants, the claim that Tolerance is a qualitative proposition can in turn be derived from the following claims: (i) *K* is a qualitative property of objects; (ii) closeness is a qualitative relation between properties; (iii) indiscernible possibility is a qualitative operation. For Tolerance is expressed by a closed sentence whose only nonlogical constants express *K*, closeness, and indiscernible possibility. Moreover, given that indiscernible possibility was itself defined in terms of some starting modality (e.g. metaphysical possibility) together with qualitativeness (which by Qualitative

Constants is itself qualitative), we can replace (iii) with the claim that possibility in the sense of that starting modality is a qualitative operation.

This is strikingly different from the arguments for Non-contingency premises considered in Chapter 3. It is purely metaphysical in character: there is no mention of belief or knowledge. Nor does it rely in any way on dubious Soritical thoughts. But of course it is only available in a very restricted range of Tolerance Arguments, namely those based on indiscernible modalities. And it remains to be seen whether any such argument constitutes a Tolerance *Puzzle*, i.e. is such that its remaining premises are prima facie plausible (or at least not obviously false), while its Hyper-tolerance conclusion is prima facie implausible (or at least not obviously true).

Our best shot in trying to set up such a puzzle will be to take *K* to be some moderately specific "kind" property like *being a table*. If we instead chose something like *thing identical to Woody* or *table produced in this workshop* or *Ikea table*, the argument for qualitativeness would be a non-starter. Meanwhile, if we chose something like *table-shaped object*, the Tolerance premise isn't likely to be very plausible, at least to those who are open to views like Plenitude on which table-shaped objects vary widely in their modal properties. For our closeness relation, it seems best to choose something like our usual "origin-closeness" relation—some plausibly qualitative relation among non-qualitative properties. That will give a Tolerance premise involving indiscernible possibility a fighting chance, since there is no obvious reason (setting Weak Anti-haecceitism aside) why modifying an object's originating atoms should require modifying any qualitative fact. By contrast, if we used something like similarity of *shape*, such a Tolerance premise won't look plausible, since changing a table's shape would in typical cases require modifying the qualitative facts.

With these desiderata in view, let's consider a Tolerance Argument based on an indiscernible modality, with *K* as 'table' and closeness as "origin-closeness": the relation that *F* bears to *G* when *F* is *being originally composed of C* and *G* is *being composed of D* for some collections of atoms *C* and *D* that "chemically match"—contain equally many atoms with each atomic number—and also overlap by at least 90 per cent. In order to make Persistent Closeness unproblematic, we had better focus not on indiscernible *metaphysical* possibility but on indiscernible *atomic* possibility. Atomic possibility, as defined in §2.4, is metaphysical compossibility with the collection comprising all truths attributing atomic numbers to individual atoms and attributing distinctness to pairs of individual atoms. *Indiscernible* atomic possibility thus amounts to metaphysical compossibility with the larger collection of truths that also includes all the qualitative truths. When we speak of indiscernible possibility and necessity in what follows, we will have in mind indiscernible atomic possibility and necessity.

The Tolerance premise in this argument says, intuitively, that we can make any modest variation in the original material composition of any table while leaving the qualitative facts unchanged and merely varying which atoms play which

qualitative roles. Spelling out the definition of 'origin-close' and 'indiscernibly atomically possible', we can state it explicitly as follows:

> For any collections of atoms *C* and *D* overlapping by at least 90 per cent and containing equally many atoms with each atomic number, any table *x*, and any collection *P* of true propositions each of which is either (a) qualitative, or (b) of the form *a is an atom with atomic number n*, or (c) of the form *a and a′ are distinct atoms*: if *x* is originally composed by *C*, then it is metaphysically possible that every member of *P* is true and *x* is a table originally composed by *D*.

Notice that this sentence is couched entirely in general terms: we do not need to single out any particular tables, atoms, collections of atoms, or anything like that in order to express it. So prima facie it seems plausible that it expresses a qualitative proposition, and thus automatically indiscernibly necessary if true. More carefully, using Qualitative Closure and Qualitative Constants, we can derive the qualitativeness of this Tolerance premise, and hence the truth of Non-contingency, from the premise that each of the following is qualitative:

being an atom
being an atomic number
being originally composed by
being metaphysically possible
being a table

Given constants expressing these properties and relations, it is a straightforward exercise to write down a formal sentence expressing our Tolerance premise, using standard techniques for spelling out the relevant cardinality comparisons in higher-order terms.[12]

We do not claim that it is *obvious* that all of the above properties and relations are qualitative. In Chapter 15 we will be looking in more detail at views that hold that at least one of them is not qualitative, focusing on views that deny the qualitativeness of *being a table*. But they certainly *seem* qualitative. Without some special argument, one would not be tempted to classify any of them along with properties like *being a New Yorker, being identical to Jean Sibelius*, etc. If the previous section's attempts to explain the notion of qualitativeness were at all successful, the idea of denying the qualitativeness of any of these properties or relations should seem pretty surprising.

Let's now turn to the remaining premises of our Tolerance Argument and its conclusion. Iteration for indiscernible possibility follows from Iteration from atomic possibility, which as we saw in §2.4, follows in turn from Iteration for

[12] For example, '*C* and *D* contain equally many atoms with each atomic number' will mean $\forall F(\text{AtomicNumber}(F) \rightarrow \exists R(\forall xyx'y'(Rxy \rightarrow (Rxy' \rightarrow y = y') \wedge (Rx'y \rightarrow x = x'))) \wedge \forall x((Fx \wedge Cx) \rightarrow \exists y(Rxy) \wedge (Fx \wedge Dx) \rightarrow \exists y(Ryx)))$.

metaphysical possibility, a thesis which we defended at length in Chapters 7 and 8.[13] Since everything atomically necessary is indiscernibly necessary, and Persistent Closeness is true for atomic necessity (see §2.4), it is also true for indiscernible necessity. And finally, since indiscernible possibility is stronger than atomic possibility which is in turn stronger than metaphysical possibility, Non-hypertolerance follows from the metaphysical-possibility version of Non-hypertolerance, which follows from the claim that there is some table originally composed of at least ten atoms, and some collection of atoms chemically matching those atoms, such that it is metaphysically impossible for that table to have been originally composed by that collection.[14] Although this isn't beyond controversy, as we discussed in Chapter 5, it is certainly far being obviously false.

But what about our Tolerance premise—is there even a prima facie case for its truth? Certainly its connection with any ordinary modal judgements is far more remote than that of the Tolerance premises we have been focusing on in earlier chapters. Claims about indiscernible possibility don't come up much outside of philosophy, so some sort of philosophical argument would be needed to bridge the gap from a Tolerance premise involving some broader modality (such as metaphysical or atomic possibility) to anything involving an indiscernible modality. And we have already seen one principled reason for resisting such arguments, namely Weak Anti-haecceitism, according to which indiscernible possibility coincides with truth. This rules out the Tolerance premise so long as there is some table and some collection of atoms that chemically matches the atoms that originally composed that table, and overlaps them by at least 90 per cent but less than 100 per cent.

[13] For suppose it is indiscernibly possibly indiscernibly possible that p. That is: for every collection Q of qualitative truths, it is atomically possible that every member of Q is true and every collection Q' of qualitative truths is such that it is atomically possible that every member of Q' is true and p. Let Q be any collection of qualitative truths. Then by the qualitativeness of qualitativeness (which is part of Qualitative Constants) together with Qualitative Collections, the proposition that Q is a collection of qualitative truths—call it q—is itself a qualitative truth. Let Q^+ be a collection that contains that proposition along with every proposition in Q; by Qualitative Collections, Q^+ is also a collection of qualitative truths. So instantiating our assumption with Q^+, we have that it is atomically possible that every member of Q^+ is true and every collection Q' of qualitative truths is such that it is atomically possible that every member of Q' is true and p. But (by Persistence: see §1.5) it is atomically necessary that q is in Q^+, and thus it is atomically necessary that if every member of Q^+ is true, Q is a collection of qualitative truths. So we can instantiate Q' to Q to derive that it is atomically possible that it is atomically possible that every member of Q is true and p. So by Iteration for atomic possibility, it is atomically possible that every member of Q is true and p. Since this holds for every collection Q of qualitative truths, we can conclude that p is indiscernibly possible.

Note that this reasoning does *not* depend on the (prima facie plausible) Persistent Qualitativeness principle, according to which everything qualitative is metaphysically necessarily qualitative: whether or not this is true, the qualitativeness of qualitativeness guarantees that its analogue for indiscernible modality is true, which is all we need.

[14] Any two chemically matching collections containing at least ten atoms can be connected by a short finite chain in which neighbouring collections overlap by at least 90 per cent. By contrast, a collection containing nine or fewer atoms cannot overlap any collection other than itself by at least 90 per cent. Hence Hypertolerance in this argument is vacuously true of any table originally composed by nine atoms or fewer.

But even if we set Weak Anti-haecceitism aside, the Tolerance premise is still so bold that there might seem to be little cost in giving it up. Let C be Woody's originating collection of atoms, and let a be one of the carbon atoms in C. Even supposing that Weak Anti-haecceitism is false, is it really plausible that every other carbon atom x in the universe, even one in a distant galaxy, is such that it is indiscernibly possible for Woody to be a table composed by the collection comprising x together with all of C except for a? Leave Woody out of it: is it even plausibly indiscernibly possible for each such collection to compose *any* table?[15]

There is a view that would motivate this claim, namely that qualitative roles can be freely distributed over atoms, at least so long as one preserves each atom's atomic number. We can spell this idea out as follows:

Atom Permutability Whenever π is a permutation of atoms that preserves atomic numbers, and P is a collection of propositions each of the form Fa, where a is an atom and F is a qualitative property instantiated by the atom to which π maps a, it is metaphysically possible for every member of P to be true.[16]

(Here a "permutation of atoms that preserves atomic numbers" is just a binary relation π such that whenever πab, a and b are atoms with the same atomic number, and if $\pi a'b'$, $a = a'$ iff $b = b'$.) But Atom Permutability is very strong. It would be false if, for example, some atoms of carbon-12 couldn't have been atoms of carbon-14 (i.e. have had eight rather than six neutrons, while still having six protons, i.e. atomic number 6). It would be false (assuming the universe is non-symmetric) if some pairs of oxygen atoms which originate in the same supernova couldn't have originated in distinct supernovae. And it would be false if some

[15] As we discussed in Chapter 2, similar worries arise even for the much weaker Tolerance premises involving metaphysical and atomic possibility: there, we focused on collections of atoms with members drawn from distant galaxies. We suggested controlling for such concerns by adding piecemeal restrictions along the lines of 'not too scattered' and 'in this factory'; but none of the restrictions we suggested there are both plausibly qualitative and such as to eliminate worries about collections that couldn't play the qualitative role of a table's actual originating collection.

[16] Atom Permutability is a little weaker than the intuitive thought it aims to capture, because of the way it treats symmetric worlds. Consider a mirror-symmetric world with four atoms, one red and one green on each side. Atom Permutability says that it is indiscernibly possible for each atom to be either red or green, and thus to instantiate either of the two instantiated qualitative roles. But it is compatible with the claim that the pair of atoms on each side is such that it would be indiscernibly impossible for them to be on opposite sides. We could strengthen Atom Permutability to rule out this claim by also allowing P to contain propositions of the form Rab where a and b are atoms and R is a qualitative relation that holds between the atom to which π maps a and the atom to which π maps b. But even the stronger form does not fully capture the intuitive idea when it comes to more exotic kinds of symmetric worlds in which there are triples of objects that differ qualitatively although none of their constituent pairs do. (It is unclear that this is possible for *atoms*, but it might plausibly hold for some triples of points in a featureless Euclidean space where any pair of distinct points is qualitatively just like any other.) There are further possible strengthenings of Atom Permutability which imply not only that any pair of distinct atoms could have played the role of any other such pair, but that the same is true for all n-tuples and even for transfinite sequences. We suppress the details, since stating them requires some work to encode quantification over n-tuples and transfinite sequences in higher-order logic.

atoms that are formed from other atoms by nuclear fission and fusion couldn't have been such that those other atoms were instead formed from them. A analogue of Atom Permutability for arbitrary permutations of *fundamental* objects is, we think, rather attractive, as part of some broader "combinatorialist" vision.[17] But real-world atoms are not fundamental objects: from the metaphysical standpoint, they are a lot more like tables, reading groups, and water waves than like the atoms of Democritus or the minima of Epicurus. And when we bear this in mind, it starts to look like the collection of atoms that could play the exact qualitative role of a given atom *a* might be quite a lot smaller than the collection of all atoms with the same atomic number as *a*.

These reflections look like bad news for the Tolerance premise based on 'table', indiscernible atomic possibility, and origin-closeness. But we can bypass such worries by restricting the Tolerance premise to those collections of atoms that *could* play the exact qualitative role actually played by a given table's originating collection. To do this, we can replace the relation of *chemical match* (containing equally many atoms with each atomic number) with *modal match*, defined as follows:

> *C* modally matches *D* := *C* and *D* are finite, and for any qualitative property *Q*, it is indiscernibly atomically possible that *C* has *Q* iff it is indiscernibly atomically possible that *D* has *Q*.

Our new closeness relation, call it "strong origin-closeness", holds between *F* and *G* iff *F* is *being originally composed of C* and *G* is *being originally composed of D* for some modally matching collections of atoms *C* and *D* that overlap by at least 90 per cent. In what follows we will focus on the Tolerance Argument based on this closeness relation, with indiscernible possibility and *K* as 'table'. For future reference, Figure 14.2 spells this argument out in full.

Given that the definition of atomic possibility includes holding fixed atomic numbers, modal match guarantees chemical match. The new versions of Tolerance and Hypertolerance thus follow from the old versions, since if a collection *C* contains *n* atoms with some given atomic number, it is atomically necessary and a fortiori indiscernibly necessary that it does so. How much stronger modal match is than chemical match depends on how much room there is within the space of indiscernibly possible worlds for atoms to take on the qualitative roles actually played by other atoms (with the same atomic number). If Weak Anti-haecceitism is true, so that indiscernible possibility coincides with truth, then *no* collection of atoms modally matches any distinct collection, unless the universe is perfectly symmetric. In that case Indiscernible Tolerance and Indiscernible Hypertolerance are both vacuously true. That's all right: we already knew that Weak Anti-haecceitism would

[17] See Dorr and Hawthorne 2013 and Bacon 2020.

Indiscernible Tolerance Every table originally composed by some collection of atoms is such that it is indiscernibly possible for it to be a table originally composed by any collection that at least 90 per cent overlaps and modally matches that collection.

Indiscernible Non-contingency If Indiscernible Tolerance is true, it is indiscernibly necessary.

Indiscernible Iteration Whatever is indiscernibly possibly indiscernibly possible is indiscernibly possible.

Indiscernible Persistent Closeness If two collections of atoms overlap by at least 90 per cent and modally match, it is indiscernibly necessary that they do so.

Indiscernible Hypertolerance Every table originally composed by some collection of atoms is such that it is indiscernibly possible for it to be a table originally composed by any collection that can be connected to that collection by any finite chain of collections in which neighbouring elements overlap by at least 90 per cent and modally match.

Fig. 14.2 Indiscernible Table Argument.

render Tolerance Arguments based on indiscernible modalities uninteresting in some way or other, and replacing chemical match with modal match just means that rather than Tolerance being false for uninteresting reasons, Hypertolerance is true for uninteresting reasons. But on the assumption that Weak Anti-haecceitism is false, and that atoms don't have such super-demanding essences that the atomic-possibility analogue of Weak Anti-haecceitism is true, Indiscernible Hypertolerance seems hard to live with. For even if atoms aren't as interchangeable as Atom Permutability makes out, they still seem *pretty* interchangeable. A collection of atoms produced in usual ways (e.g. by fusion in supernovae) will plausibly modally match *many* other collections, including both collections that largely overlap it and collections that don't overlap it at all. If so, we can expect that many ordinary tables are such that their originating collection can be connected to some entirely non-overlapping collection by a chain of modally matching collections in which adjacent elements overlap by at least 90 per cent. Any such table must either be a counterexample to Indiscernible Hypertolerance or a counterexample to the following principle, which we discussed in Chapter 5:

Overlap Essentialism For every table x and collection of atoms C: if C originally composed x, then necessarily, x is not originally composed by any collection of atoms with no members in common with C.

And while we found some reasons in §5.3 to suspect that *some* tables are counterexamples to Overlap Essentialism, it would be very surprising if considerations having to do with qualitativeness forced us to conclude that *all* tables (or all tables made of ordinary, supernova-generated atoms) are counterexamples.

Let's consider how the shift from chemical match to modal match affects the remaining premises of the Indiscernible Table Argument. Indiscernible Iteration is unaffected. Indiscernible Non-contingency is different from the version using chemical match; but the qualitativeness-based argument for its truth is the same as before, since modal match was defined entirely in terms of logical constants plus 'indiscernibly atomically possible' (which is itself defined in terms of logical constants plus 'atom' and 'atomic number'). Indiscernible Persistent Closeness, however, raises a new issue. If Iteration or the 5 axiom fails for atomic possibility, so that there are contingently impossible or contingently possible propositions, then it might be that some collections of atoms that modally match are such that it is atomically possible for them not to modally match. But the need to rely on Iteration is not a problem since we need it as a premise in any case. And while the 5 axiom remains potentially controversial even when Iteration is granted, so long as we accept Iteration, we should be able to bracket any doubts about the 5 axiom by a further narrowing of the operative modality. One natural strategy is to mimic the definition of atomic possibility by holding fixed facts about which pairs of collections modally match in the same way that atomic possibility holds fixed the atomic numbers of atoms.[18] In the rest of the chapter we will assume that atomic possibility obeys $\mathsf{H}_{\mathsf{S5R}}$, leaving it as an exercise for those who doubt that this is true on our current definition of 'atomic possibility' to adapt the argument to fit.

Given $\mathsf{H}_{\mathsf{S5R}}$ for atomic possibility, there are three remaining strategies for dealing with the Indiscernible Tolerance Argument: we could deny Indiscernible Tolerance; we could accept Indiscernible Hypertolerance; or we could deny Indiscernible Non-contingency. The next chapter will develop the third of these options (our favourite). The remaining two sections of this chapter will discuss the first two options. As we we shall see, there is plenty to learn from them as well.

14.3 Rejecting Indiscernible Tolerance

Assuming we are not willing to give up our standard Tolerance premises (involving metaphysical possibility, atomic possibility, positive objective chance, and so

[18] A different option is to replace metaphysical necessity in the definition with the corresponding operator $\Box_5$ as defined in Appendix D, which is guaranteed to obey the 5 axiom if metaphysical necessity obeys Iteration. Once we have Iteration, there is good reason to think that the space of metaphysically possible worlds that don't differ from actuality as regards what is metaphysically possible is large and varied, and hard to see a principled reason for a package that accepts Tolerance when stated in terms of atomic possibility but denies it for the narrower version defined in terms of $\Box_5$.

on), the idea of rejecting Indiscernible Tolerance might initially seem completely unpromising. If it is possible for a certain collection of atoms *C* to play the exact qualitative role actually played by Woody's originating atoms, the possible worlds where they do that might seem to be clearly the best candidates to be worlds where they compose Woody, so that if it is possible for them to compose Woody at all, it is possible for them to do so while playing that role.

The "best candidates" thought that drives this argument against Indiscernible Tolerance is interesting enough to be worth stating more precisely. The idea is that there would be something weirdly *disorderly* about Woody's modal profile if it could have been originally composed by a certain collection *C* but only by *C*'s playing some qualitative role other than that of Woody's actual originating collection. So let's call it:

Orderlininess If *x* is a table originally composed by a collection of atoms *C*, and *Q* is the collection of all qualitative properties of *C*, and it is atomically possible for *x* to be a table originally composed by *D*, and it is atomically possible for *D* to instantiate all properties in *Q*, then it is atomically possible for *x* to be a table originally composed by *D* while *D* instantiates all properties in *Q*.

By combining Orderlininess with the Tolerance premise of our ordinary Tolerance argument involving atomic possibility—let's call this 'Atomic Tolerance'—we can derive Indiscernible Tolerance. For suppose *x* is a table originally composed of *C*, and *D* is a modally matching collection that overlaps *C* by at least 90 per cent. Let *Q* collect all qualitative properties of *C*, and let *p* be any true qualitative proposition. By Atomic Tolerance, *x* could (in the sense of atomic possibility) have been originally composed by *D*. Since *C* and *D* modally match, *D* could have instantiated every property in *Q*. So by Orderlininess, *D* could have instantiated every property in *Q* while originally composing *x*. But instantiating every property in *Q* entails *being such that p*; hence *D* could have originally composed *x* while *p* was true. In other words, it is indiscernibly possible for *D* to have originally composed *x*.

We were initially gripped by this case for Indiscernible Tolerance, but subsequently realized that matters are more delicate. Insofar as we are convinced by the qualitativeness-based motivation for Indiscernible Non-contingency, we know that Tolerance premises involving indiscernible modality are liable to blow up in our faces in a way that more familiar kinds of Tolerance premises are not. And indeed, someone could easily get into a mood where Orderliness and Indiscernible Tolerance would both seem like complete non-starters. Suppose that it is atomically possible for Woody to have been originally composed by a collection of atoms *C* just like its actual originating atoms except that one atom *a* is replaced by some other atom *b* that isn't actually part of Woody, where *a* and *b* are modally

interchangeable in the sense that they could have swapped qualitative roles while every other atom played the same qualitative role. In addition, suppose for simplicity that Woody is a "uniformly tolerant table" (see note 38 in Chapter 2)—for some percentage *n*, strictly between 0 per cent and 100 per cent, the collections of atoms which chemically match its originating atoms and could have composed Woody while Woody was a table are all and only those that overlap Woody's actual originating atoms by at least *n*%. By Iteration, Woody could not have been a uniformly tolerant table with threshold *n*% without having its exact actual original composition. Insofar as there is a good case that Indiscernible Tolerance is a qualitative proposition, there is an equally good case that *being part of a uniformly tolerant table with threshold n*% is a qualitative property. So, one might reason that in order to swap *all* the qualitative properties of *a* and *b*, one would have to go to an indiscernibly possible world where *b* has the property *being part of a uniformly tolerant table with threshold n*%, which means that the table that *b* is part of at the world in question cannot be Woody, but must be some other possible table. On this vision, then, even though it is possible for Woody to be composed by *C*, and possible for *a* and *b* to switch qualitative roles, it is not possible for both things to happen together.

This particular argument rests on the somewhat tendentious premise that Woody is uniformly tolerant; but the underlying idea is more general. Unless Woody is *hypertolerant*, just about any way of swapping atoms around that moves some atoms across Woody's boundary can be expected to bring Woody slightly "closer to the edge" in some respect; and under this chapter's assumptions about qualitativeness, how close one is to the edge in any given respect is a qualitative matter.

So was the initial appeal of Orderlininess just a mirage? When one first encounters Indiscernible Tolerance, it is natural to think something along the following lines: 'Surely Woody could still exist in a world where two atoms are switched and the complete pattern of shapes, colours, masses, charges, distances and so forth was the same; but if things were like that, then all the qualitative facts would be the same.' In thinking like this, one is implicitly appealing to some kind of supervenience claim, to the effect that all qualitative truths are metaphysically necessitated by the the truths about the pattern of instantiation of some collection of "basic" qualitative properties and relations. So it is worth investigating whether a case for Orderlininess could be made based on such a supervenience claim. While there are many things that might work, as usual we will focus on a version that gives *physics* a special role:

Qualitative Microphysical Supervenience Every possibly true qualitative proposition is metaphysically necessitated by some possibly true qualitative microphysical proposition.

Intuitively, the "qualitative microphysical propositions" should be those specifying the pattern of distribution of certain basic microphysical properties and relations; see Chapter 6 for some options for making sense of this.[19]

Given Qualitative Microphysical Supervenience, we can derive claims of indiscernible possibility from claims about "micro-indiscernible" possibility, a modality defined just like indiscernible possibility but where we hold fixed only *microphysical* qualitative truths rather than all qualitative truths. So, in particular, Indiscernible Tolerance follows from the combination of Qualitative Microphysical Supervenience and the following principle:

Micro-Indiscernible Tolerance Every table originally composed by some collection of atoms is such that it is *micro*-indiscernibly possible for it to be a table originally composed by any collection that at least 90 per cent overlaps and modally matches that collection.

And for the same reason that Indiscernible Tolerance follows from the combination of Atomic Tolerance and Orderlininess, Micro-Indiscernible Tolerance follows from the combination of Atomic Tolerance and a version of Orderlininess restricted to microphysical properties:

Micro-Orderlininess If x is a table originally composed by a collection of atoms C, and Q is the collection of all qualitative microphysical properties of C, and it is atomically possible for x to be a table originally composed by D, and it is atomically possible for D to instantiate all properties in Q, then it is atomically possible for x to be a table originally composed by D while D instantiates all properties in Q.

[19] Qualitative Microphysical Supervenience is intimately related to the Microphysical Supervenience principle discussed in Chapter 6, which results from it when both occurrences of 'qualitative' are deleted. Nevertheless the two principles are logically independent. If one were led to deny Microphysical Supervenience by an embrace of some fine-grained Hypertolerance claims, one might hold on to Qualitative Microphysical Supervenience, holding that whenever multiple possible worlds agree on all microphysical propositions they also agree on all qualitative propositions, and disagree merely as regards the identities of the (non-microphysical) objects playing certain qualitative roles. Conversely, a proponent of Microphysical Supervenience might be led to deny Qualitative Microphysical Supervenience on the basis of the argument that it is possible for Woody to be a table in a world where the microphysical qualitative truths are just the same but one of Woody's atoms swaps *microphysical* qualitative roles with some other atom, but it is impossible for this to happen while Woody remains uniformly tolerant, so that the possibility in question differs from actuality as regards the qualitative proposition that *every table is uniformly tolerant* despite agreeing on all microphysical qualitative propositions. However, the two supervenience principles do form a natural package. And there are stronger claims which entail both: most obviously, the thesis that *every proposition is microphysical*, which (as we discussed in Chapter 7) is arguably the best expression of the physicalistic impulse that led us to embrace supervenience claims in the first place.

Micro-Orderlininess is weaker than Orderlininess, and less likely to seem like a dialectical non-starter.

Both Qualitative Microphysical Supervenience and Micro-Orderlininess can certainly be resisted. However, many of the objections to Qualitative Microphysical Supervenience are neither here nor there in the current context, since they suggest substitute principles which could do the same work; in this regard they are similar to various objections to Microphysical Supervenience which we set aside in §6.4. For example, if one thought that qualitative facts about the pattern of distribution of phenomenal qualia were a counterexample to Qualitative Microphysical Supervenience, one could throw those into the supervenience base alongside the qualitative microphysical facts.[20] The kind of failure of Qualitative Microphysical Supervenience that would be required to block the argument is radical: it challenges not just the claimed special status of the qualitative facts of *physics*, but the more general picture where there are some fundamental properties and relations whose pattern of distribution suffices to pin down all the qualitative truths.[21] If Indiscernible Tolerance Puzzles led us to give up this picture, that would indeed be a dramatic confirmation of their importance! But in our view this would be a heavy cost.

But what can be said in favour of Micro-Orderlininess? Is there anything really so bad about a package that accepts Atomic Tolerance, but claims that in order to get Woody to be composed by any other collection of atoms, we have to make some modification to the overall pattern of microphysical facts rather than just permuting the atoms into different microphysical qualitative roles? The suggestion naturally raises a worry about arbitrariness: it is not obvious that there is any way of telling a *principled* story about how the microphysical qualitative facts would need to be different for Woody to be composed by any given collection of atoms.[22] However, when we bear in mind how pervasive vagueness is in this area, it is not so clear that there is anything wrong with the a lack of such a principled story. If 'table' is vague, we might think that some or all of the ways of making it precise will involve making many arbitrary choices that do not correspond to any identifiable trends in our use of 'table'. If so, Micro-Orderlininess may be false on at least some

[20] If one took there to be uninstantiated, fundamental, non-microphysical properties and relations whose failure to be instantiated is not settled by the qualitative microphysical facts, one could replace metaphysical possibility with "inner sphere possibility" (Lewis 1983*a*): metaphysical compossibility with the uninstantiatedness of all uninstantiated fundamental properties and relations.

[21] Goodman (unpublished *e*) argues for giving up Qualitative Microphysical Supervenience as a way out of a certain Tolerance Puzzle. In effect, his argument is that while there are cases where Micro-Indiscernible Tolerance is true, Indiscernible Tolerance must be false since it leads via a Tolerance Argument to a false Hypertolerance conclusion.

[22] However, we will later consider one such story, based on a strong form of origin essentialism, which will makes for failures of Micro-Orderlininess of a perfectly principled sort.

precisifications, and perhaps on all. The argument from Micro-Orderlininess is thus far far from being a decisive case for Indiscernible Tolerance.[23]

So the option of blocking the Indiscernible Tolerance Argument by rejecting Indiscernible Tolerance is starting to look more defensible. But there is a different way of generating potential trouble for this strategy, based on "robustness" judgements to the effect that it would have been hard for our actual tables *not* to be tables without the relevant underlying microphysical facts being significantly different (cf. §6.3 and §9.3). Recall that denying Indiscernible Tolerance just means that *some* of the relevant permutations of atoms disrupt the identity of the actual tables. But on further investigation, it emerges that if one wants one's rejection of Indiscernible Tolerance to help with the Indiscernible Table Argument, one will be pushed towards the more radical claim that every table is fragile with regard to *almost all* permutations: intuitively speaking, the tables of the actual world are quite rare within the space of all indiscernibly possible worlds. As we will see, that radical claim in turn generates a potential worry about robustness. But first, let's see why it is hard for deniers of Indiscernible Tolerance to resist it.

The pressure comes not from the denial of Indiscernible Tolerance on its own, but from the package required to motivate that denial via the Indiscernible Table Argument. This package not only includes Indiscernible Non-contingency, and Indiscernible Non-hypertolerance, but also accepts the qualitativeness of *being a table*, *originally composing*, etc. (which provides the interesting argument for Indiscernible Non-contingency), and moreover accepts various natural strengthenings of Indiscernible Non-hypertolerance (which by itself is very weak), such as Overlap Essentialism. To see why this package makes actual tables rare in the space of indiscernibly possible worlds, let's consider a simple example. Suppose that the universe consists of a large number of gold atoms $a_1, \ldots, a_n$. There is just one table in the universe, Goldy, which is composed by atoms $a_1, \ldots, a_m$ (where $m \ll n$). There are no qualitative symmetries, so each atom a_i instantiates a maximally specific qualitative property Q_i that distinguishes it from all other

[23] Even if you didn't mind blocking the Indiscernible Table Argument by giving up Indiscernible Tolerance and Micro-Orderlininess, there is a different qualitativeness-based argument for Hypertolerance that proponents of Qualitative Microphysical Supervenience might find worrying. Even if one doesn't buy the Security Argument that Atomic Tolerance is atomically necessary if true, one might still be gripped by the following "compatibility thesis": if Atomic Tolerance is true, it is compatible with every atomically possible, microphysical, qualitative proposition. The thought is that for any micropattern, there is *some* way of assigning specific atoms to qualitative roles that is compatible with all tables being tolerantly tables: one could find this quite gripping even after conceding that some *other* ways of assigning specific atoms to the same qualitative roles lead to there being tables that aren't tolerantly tables. But given Qualitative Microphysical Supervenience, any *qualitative* proposition that's compatible with every atomically possible microphysical qualitative proposition must be atomically necessary. And given the qualitativeness of tablehood (etc.), Atomic Tolerance is qualitative. So the compatibility thesis implies that Atomic Tolerance is atomically necessary if true, thus delivering the corresponding Hypertolerance claim by our standard Tolerance Argument. We won't further explore the merits of the compatibility thesis here. It is not clear to us that its initial appeal will survive critical scrutiny, whether or not tablehood is qualitative.

atoms. Suppose moreover that Atom Permutability is true, so that every one-to-one function F from $a_1, \ldots, a_n$ to $Q_1, \ldots, Q_n$ corresponds to an indiscernibly possible proposition p_F. There are vastly many ($n!$, to be precise) such propositions. No two of them are consistent, but they all necessitate every qualitative truth. How many of them are compatible with Goldy's being a table? We can begin to probe this by considering a subclass of n of these propositions:

$$
\begin{aligned}
p_1 &\coloneqq Q_1a_1 \wedge Q_2a_2 \wedge \cdots \wedge Q_{n-1}a_{n-1} \wedge Q_na_n \\
p_2 &\coloneqq Q_1a_2 \wedge Q_2a_3 \wedge \cdots \wedge Q_{n-1}a_n \wedge Q_na_1 \\
&\vdots \\
p_k &\coloneqq Q_1a_k \wedge \cdots \wedge Q_{n+1-k}a_n \wedge Q_{n+2-k}a_1 \wedge \cdots \wedge Q_na_{k-1} \\
&\vdots \\
p_n &\coloneqq Q_1a_n \wedge Q_2a_1 \wedge \cdots \wedge Q_{n-1}a_{n-2} \wedge Q_na_{n-1}
\end{aligned}
$$

Each possibility comes from its predecessor by shunting each atom one step along in the sequence of qualitative roles. p_1 is true and the rest are false. Since we are assuming that *being a table*, *being an atom*, and *originally composing* are qualitative, $Q_1, \ldots, Q_m$ all entail *being among the atoms that originally compose the one and only table*. Suppose we thought Goldy could still be a table while p_2 was true. Then we would also be forced to say the same as regards p_3. For given our assumptions about qualitativeness, the following is a true qualitative proposition:

> There are atoms $x_1, \ldots, x_n$ such that $Q_1x_1 \wedge \cdots \wedge Q_nx_n$, and for every table y, it is possible that (y is a table and $Q_1x_2 \wedge Q_2x_3 \wedge \cdots \wedge Q_{n-1}x_n \wedge Q_nx_1$).

It is thus necessitated by p_2, since p_2 necessitates every qualitative truth. So if it is possible for Goldy to be a table while p_2 is true, it is possibly possible for Goldy to be a table while p_3 is true. By Iteration it follows that this is possible. Continuing in this way, our assumption that p_2 is compatible with Goldy's tablehood leads to the implausible conclusion that each p_i is compatible with Goldy's tablehood. This conclusion entails that Goldy is a counterexample to Overlap Essentialism, since p_{m+1}, for example, entails that the only table there is is composed of $a_{m+1}, \ldots, a_{2m}$ which do not overlap Goldy's actual atoms $a_1, \ldots, a_m$ at all. So Goldy's being a table can't after all be compatible with p_2! By similar reasoning it can't be compatible with p_3, since if it were, it would also be compatible with $p_5, p_7, \ldots$, again making Goldy a counterexample to Overlap Essentialism. In fact, of our n propositions $p_1, \ldots, p_n$, only one—namely p_1, the true one—is compatible with Goldy's tablehood. And we can do the same thing starting with any p_F that is compatible with Goldy's tablehood: each such p_F belongs to an "orbit" of n other propositions which we get by shunting the atoms around in the order of $Q_1, \ldots, Q_n$; and if in fact Goldy isn't a counterexample to Overlap Essentialism, at

most one of the n propositions in the given orbit can be compatible with Goldy's tablehood. So, in our big family of propositions p_F, at most $1/n$ of them are compatible with Goldy's tablehood.[24] Given that n is very large, this substantiates our claim that the package we are looking at requires Goldy to be "rare" across the space of indiscernibly possible worlds.

The assumption that there is only one table is of course false, and the assumption of Atom Permutability is tendentious. But similar results can be got on more realistic and safer assumptions.[25] The general issue is that since small permutations can typically be composed with themselves many times to yield big permutations, the proposition that tables are tolerant with respect to a small permutation will imply (given the qualitativeness of tablehood) that tables are also tolerant with respect to all the big permutations that can be built up by composing that permutation with itself, a conclusion that is no more plausible than Indiscernible Hypertolerance. The lesson is that escaping the puzzle by denying Indiscernible Tolerance requires embracing extreme level of indiscernible *intolerance*. Almost any permutation of atoms that maps an atom that is part of some table x to an atom that is not part of x is such that x couldn't still be a table in an indiscernibly possible world where atoms swap roles in accordance with that permutation.

While this conclusion is not objectionable in itself (or at least, not significantly more objectionable than the bare denial of Indiscernible Tolerance), it is worrisome because there is a risk that it will conflict with many ordinary "robustness" judgements. We generally assume that a given table couldn't easily have failed to be created without there being substantial differences as regards which bits of wood are assembled, which plans are used, which artificers are involved, etc. There was a high *objective chance* of our making Woody conditional on our making a table from roughly the wood we actually used for Woody, according to roughly the same plan, etc. The worry is that these many other tables that are made in indiscernibly possible worlds differing from actuality by minor permutations of atoms will have to compete for room with Woody, so that Woody cannot after all be robust in the way we want.

Let's crudely think of the robustness of Woody as amounting to Woody's being present at a high proportion of 'close' worlds. Of course, since most close worlds aren't indiscernibly possible, Woody's rarity in the indiscernibly possible worlds doesn't immediately imply that Woody is rare in the close worlds. The problem for robustness arises on the assumption that there will be a substantial overlap between

[24] Note that this reasoning requires that $m \leq n/3$. If we had $m > n/2$ we could say that Goldy's tablehood is completely tolerant, compatible with all of the p_i, without giving up Overlap Essentialism. And even if, say, $n = 27$ and $m = 10$, Overlap Essentialism does not rule out the hypothesis that Goldy's tablehood is compatible with p_1, p_{10}, and p_{19} and not with any of the other p_i.

[25] With multiple tables, the above reasoning will go through in the same way except that instead of an upper bound of $1/n$ we will have t/n, where t is the number of tables having exactly the same number of atoms as Goldy.

the worlds close to the actual world and the worlds close to nonactual indiscernibly possible worlds derived from actuality by small permutations of atoms. Suppose most worlds close to actuality are also close to another indiscernibly possible world w', where Woody's actual qualitative role is played by a different table Woody′. If Woody is robust at the actual world, Woody′ must be robust at w' (since robustness is presumably qualitative). Thus, barring a high chance of there being coincident tables, it can't be that Woody is a table at most worlds close to actuality and Woody′ is a table at most worlds close to w'.

This robustness problem could be entirely avoided by denying that any world is both close to actuality and close to some nonactual indiscernibly possible world.[26] Is such a thought defensible? It would certainly be hopeless if one were thinking of 'closeness' in terms of some intuitive notion of similarity. Intuitively, a world just like actuality except that one atom near Woody's surface has swapped places with another atom a few nanometres away from it is *extremely* similar to actuality—more similar to actuality than any world where, e.g., we shift the saw over by a millimetre while cutting the planks for Woody's top. Since some of the latter worlds are certainly "close" in the sense at issue for the robustness judgements, a similarity-driven conception of closeness will force the small-permutation worlds to be close as well.

But it is just a mistake to think of closeness in terms of an intuitive notion of similarity that works like that. The role of the notion is to represent the range of worlds that matters for the purposes of ordinary, practical modal judgements. And there is a strong case that this range is always restricted to worlds that *match actuality with regard to the course of history up to some time*. This has been most discussed in the case of counterfactuals: when we are thinking about what would have happened if things had gone differently in some respect at some time, we tacitly hold fixed a very wide range of facts about what happened at earlier times.[27] It is also true for ascriptions of chance: we standardly take facts about history up to t to have chance 1 at t (see the discussion of "History Fixity" in Chapter 9). And there is a reasonable case that ordinary talk about luck, danger, safety, what could easily have happened, and so on works in a similar way: if there has never been a positive chance that p, it couldn't easily have been that p. But under certain assumptions about physics, this necessary condition for closeness will preclude any non-actual indiscernibly possible world from being close to actuality. Suppose that we live in a Democratean world of eternal particles moving around in a fixed background space (never coinciding and never being arranged in a perfectly symmetric way). Then any indiscernibly possible world that differs microphysically from the actual world must differ as regards which eternal particles play which qualitative roles,

[26] Or in chance-theoretic terms: whenever w is indiscernibly possible but non-actual, some proposition that in fact has chance 1 has chance 0 at w.

[27] See Lewis 1979 and Dorr 2016*a* (251).

which means it must be in a different state at every time. So if closeness requires perfect match on some initial stretch of history, Democrateanism implies that no world is close both to the actual world and to some other indiscernibly possible world.[28]

The claim that the history of every close world perfectly matches that of the actual world for some initial segment becomes much harder to defend if the laws of nature are *deterministic*. For with deterministic laws, there won't *be* any worlds with perfectly matching initial segments of history where the actual laws of nature are true. There is some temptation to think that such determinism means that our ordinary "easy possibility" judgements are radically error-ridden: nothing could easily have happened that didn't happen; no one who survived was ever in danger of dying; no one ever had a good chance of winning but ended up losing.... But this temptation must be resisted. Determinism is a live physical hypothesis, and our ordinary judgements are not taking a stand on rejecting it.[29] So given determinism, we will have to choose between saying that the laws of nature could easily have been false (would have been false if we had done otherwise; recently had a high chance of being false; ...) and saying that certain truths about the distant past could easily have been false (would have been false if we had done otherwise; recently had a high chance of being false; ...). Neither option is comfortable. But the good news for the latter option (see Dorr 2016*a*: §4) is that even if the laws don't let us hold fixed the *completely specific* truth about the state of the world over any interval of time, they will, if they are anything like the deterministic laws posited in live physical hypotheses, allow us to hold fixed the *approximate* truth about the state of the world over any given finite span of history. 'Approximate' can be made as demanding as we please, so long as it allows some nonzero amount of divergence in the value of every continuous dynamical parameter (e.g. the distances between particles). Our practice of "holding history fixed" in evaluating counterfactuals, chance claims, easy possibility claims, etc. provides good reason for thinking that the only worlds relevant for such judgements are worlds whose state matches actuality extremely closely for some long stretch of history. Given the

[28] The picture is a little more complicated if new particles are sometimes created: then there will be room for indiscernibly possible worlds to match in their history up to some time, differing only as regards the identities of particles created subsequent to that time. But this doesn't matter, since the reasons we gave for thinking that our actual tables are rare in the space of indiscernibly possible worlds do not count against the view that there is a time t, much earlier than the creation of Woody, such that Woody is a table in all indiscernibly possible worlds that differ only with regard to the identities of the particles created after t. This would be less plausible if Woody were a table been made out of freshly created particles; but our ordinary robustness judgements don't really apply to tables made by such extraordinary processes.

[29] We are being a bit high-handed here, given the extensive literature arguing for the incompatibility of determinism with the truth of ordinary claims of the form 'x could have done otherwise' (see, e.g., van Inwagen 1986). If the incompatibilists are right, that is good news for the strategy of denying Indiscernible Tolerance.

availability of this option, we think it better to hold fixed the laws and allow the past to vary.[30]

But this does not automatically disrupt the case that no world is both close to actuality and close to some other indiscernibly possible world. Suppose for example that we understand 'approximate match' in such a way that for w' to approximately match w at some time t, the distance between any two particles at t at w' must differ from the distance between those particles at w by less than some tiny d, considerably smaller than the smallest separation between any two particles at the actual world. Even disregarding any match-conditions involving other dynamical parameters, this demanding standard as regards inter-particle distances guarantees that no world approximately matching actuality at some time will also approximately match any other indiscernibly possible world at that time. And it is completely compatible with the deterministic character of the laws that all "close" worlds (in the sense relevant to everyday modal judgements) approximately match actuality during some long chunk of history, by some such demanding standard.

Thus, even given determinism, the conclusion that actual tables are rare in the space of indiscernibly possible worlds need not generate a clash between ordinary robustness judgements and the denial of Indiscernible Tolerance. Perhaps the most natural way of fleshing out this picture would involve a kind of origin-essentialism entertained favourably by Kripke, on which objects like tables are extremely demanding with regard to the state of the universe at times sufficiently far before their their creation.[31] The absence of the actual tables from almost all non-actual indiscernibly possible worlds will in that case follow from a more general principle according to which they are absent from all worlds whose early history doesn't match (or at least approximately match) that of the actual world.[32] This form of origin-essentialism also suggests a different line of response to our earlier argument based on Orderlininess. Origin-essentialism provides a principled (and non-arbitrary) answer to the question why making Woody out of C while C instantiates qualitative role Q should be impossible, if it is possible to

[30] See Loewer 2001, Eagle 2011, and Dorr 2016*a*.

[31] 'Ordinarily when we ask intuitively whether something might have happened to a given object, we ask whether the universe could have gone on as it actually did up to a certain time, but diverge in its history from that point forward so that the vicissitudes of that object would have been different from that time forth. *Perhaps* this feature should be erected into a general principle about essence. Note that the time in which the divergence from actual history occurs may be sometime before the object itself is actually created' (Kripke 1972: n. 57).

[32] This certainly isn't the only principled way to combine robustness with rarity at indiscernibly possible worlds. A quite different thought would be to deny Qualitative Microphysical Supervenience: then we could attribute Woody's absence from almost all nonactual indiscernibly possible worlds to the fact that any change in Woody's composition will make a difference to the qualitative facts about the exact tolerance level of the table playing a certain role. On this picture, there are many close *micro*-indiscernibly possible worlds where Woody is around, but such worlds are not indiscernibly possible *simpliciter*, since they differ as regards such non-microphysical qualitative facts.

make Woody out of *C*, possible to make Woody out of a collection instantiating *Q*, and possible for *C* to instantiate *Q*. The answer is that the really crucial time for Woody—the time when it really hangs in the balance whether Woody or some other table will be made—is not the time of its creation, but some much earlier time. For Woody to be made, every atom, including those in *C*, must behave approximately as it actually does in the distant past. Permuting *C* into qualitative role *Q* disrupts this, whereas getting Woody to be composed by *C* need not disrupt it, since this can be achieved by changes which only (non-microscopically) affect the more recent, less crucial portion of history.

Thus, so long as we are thinking of non-actual but indiscernibly possible worlds as differing from actuality by permuting some countable collection of fundamental particles, the strategy of blocking the Indiscernible Table Argument by denying Indiscernible Tolerance seems workable. Its consequences are somewhat counter-intuitive, but not in conflict with the central range of "easy possibility" judgements motivating Tolerance. Unfortunately, it is unclear whether this good news can carry over to other pictures of how indiscernibly possible worlds differ from actuality at the fundamental physical level. For example, let's consider a General Relativity-inspired picture on which the only fundamental objects are spacetime points. On this picture, indiscernible possibilities with the same spacetime points will correspond to permutations of the spacetime points. Perhaps not all such permutations correspond to indiscernible possibilities. For example, it is often suggested that the *topological* properties of spacetime regions are essential to them. In that case, it will, for example, be metaphysically impossible for two spacetime points to swap qualitative roles while the roles of all other spacetime points remain the same, since that would require a topologically open region containing the first point but not the second to possibly play the qualitative role actually played by the region derived from it by removing the first point and adding the second, a role that entails not being open. However, unless one holds that *no* nontrivial permutations correspond to indiscernible possibilities (either because Weak Anti-haecceitism is true, or because the actual spacetime points have ultra-demanding essences), it seems inevitable that the group of permutations that do correspond to indiscernible possibilities will be *continuous*. That is: for any nonzero spatiotemporal distance, it will contain nontrivial permutations so small that they map every spacetime point to one within that distance of it. This need create no difficulty on the assumption of indeterminism: there, as before, we could require perfect match in the distant past as a condition for the creation of actual tables. But matters are more difficult in the deterministic setting, since the earlier strategy of requiring "approximate match" will no longer work. If our test for approximate match allows any leeway at all, there will be indiscernibly possible worlds that pass it, because the permutation of spacetime points to which they correspond is so tiny. And since we can build up arbitrarily big permutations by composing arbitrarily small ones, there is no prospect (given Indiscernible

Iteration and the qualitativeness of 'table') of saying that tables are indiscernibly tolerant with respect to all sufficiently small permutations, but not indiscernibly hypertolerant.

Note that the worry here is *not* that the spacetime picture supports something like Anti-haecceitism, as many philosophers of physics seem to think.[33] If they were right, that would of course pose an entirely different problem for the denial of Indiscernible Tolerance. We don't think they are right, but entering into this debate is beyond the scope of this book. Our claim is that theories like GR support the view that *if* Weak Anti-haecceitism is false, then there are indiscernible possibilities *arbitrarily* similar to actuality (on any physically natural way of measuring similarity), thereby generating a tension between ordinary robustness claims and the kind of fragility required for Indiscernible Tolerance to fail.

To sum up: while the strategy of responding to the Indiscernible Table Argument by rejecting Indiscernible Tolerance has turned out to be much more defensible than it initially seemed, its prospects have turned to be interestingly tied up with questions about the character of the underlying physics. On an old-fashioned picture of the world as consisting of fundamental particles in a fixed background space, the strategy looks pretty promising. But with a spacetime-theoretic picture, it becomes far more challenging to reconcile ordinary robustness judgments with the claim that our actual tables are rare within the space of indiscernible possibilities.

14.4 Indiscernible Hypertolerance and Microphysical Supervenience

Let's turn next to the option of dealing with the Indiscernible Table Argument by embracing Indiscernible Hypertolerance. One not-so-interesting way of developing this involves embracing Weak Anti-haecceitism (or at least its slightly weaker atomic-possibility analogue). In that case atoms have such super-demanding essences as to preclude there being any modally matching distinct collections of atoms, making Indiscernible Hypertolerance vacuously true. But let's continue to

[33] Views reminiscent of anti-haecceitism (or its restriction to spacetime points or regions) have been widely defended as a response to the "hole argument" (Earman and Norton 1987): examples include Maidens (1992: 135) ("there is no such thing as the transworld identity of points"); Butterfield (1989: 23) ("any point is a part of just one possible word"); Brighouse (1994: 122) ("we individuate spacetime points by qualitative similarity"); Hoefer (1996) ("points . . . do not have primitive identity"); Pooley (2005: 9) ("there is no real contest: anti-haecceitism is the clear winner"). But there is less consensus than meets the eye. Following Lewis (1986*a*: §4.4), some of these authors, including Butterfield and Brighouse, express the anti-haecceitism-like views they accept using the terminology of possible worlds, and embrace Lewis's suggestion that the inference from this world-theoretic "anti-haecceitism" to the modal thesis *we* call 'Weak Anti-haecceitism' should be blocked by adopting a counterpart-theoretic semantics. (Lewis calls this 'cheap haecceitism'.) For arguments that those who take spacetime points or regions to be "substances" should *not* accept anti-haecceitism, see Belot and Earman 2001.

put this unattractive option to one side, and assume that there is plenty of modal match—enough to make Indiscernible Hypertolerance incompatible with Overlap Essentialism.

In earlier chapters, we have been willing to take Hypertolerant solutions to various Tolerance Puzzles quite seriously. Our central argument in Chapter 6 against Hypertolerance as a general strategy for dealing with Tolerance Puzzles turned on the conflict between Microphysical Supervenience and certain very fine-grained Hypertolerance claims, involving families of properties which can only be instantiated at worlds where a given maximally specific microphysical proposition ("microstory") is true. Given Microphysical Supervenience, any such "microspecific" property can only be instantiated at one possible world. So if two microspecific properties could be instantiated by distinct objects, they are not both possibilities for any one object.

The Indiscernible Tolerance Arguments we have been considering so far in this chapter are not fine-grained in this way. But it is worth investigating whether they can be modified so that Hypertolerance conflicts with Microphysical Supervenience, without substantially diminishing the plausibility of the other premises. To generate the conflict, the field of the closeness relation should consist of microspecific properties, at least two of which are incompatible but incorporate the same microstory. And for the Tolerance premise regarding indiscernible possibility not to be obviously false, the closeness relation had better be such that whenever one property is close to another, they agree as regards which *qualitative* microphysical propositions they incorporate, differing only as regards which microphysical things play which qualitative microphysical roles.

Arguments fitting this bill can be constructed under the far-fetched, but plausibly metaphysically possible, supposition that the universe is *qualitatively symmetric*.[34] For example, we can suppose that the universe is, throughout its history, symmetric under rotations of 180° around some axis. To keep things simple, let's also suppose that the universe consists just of a finite number n of simple, Democratean atoms, and that these atoms are modally interchangeable, so that any two of them could swap qualitative roles while all the rest continue to play the

[34] The possibility of a qualitatively symmetric world is not beyond question: it has recently been questioned by Goodman (unpublished *a*) on the basis of the (often disputed) Conditional Excluded Middle (CEM) axiom (see note 9 to Chapter 7). However, the considerations that motivate Goodman's denial of the possibility of symmetry also put some pressure on Microphysical Supervenience. Moreover, if one accepted Microphysical Supervenience while denying the possibility of symmetry, one would have good reason to doubt that atoms are modally interchangeable in the way we are supposing. For even Goodman does not deny the possibility of *micro*-symmetry. For any two atoms a and b, we can consider a full microphysical specification m of a micro-symmetrical world where they are the only two atoms and they have the same micro-qualitative properties. Goodman thinks that such worlds won't be fully qualitatively symmetric (because of supposedly qualitative properties like *being the atom that would have been concrete if only one of the two was concrete*). But given Microphysical Supervenience, m is incompatible with these further qualitative properties being distributed in the opposite way. Thus the soundness of Goodman's argument against the possibility of symmetry would in several ways change the dialectic of this section.

same roles. The considerations that make Indiscernible Tolerance plausible suggest that tables in such a world would be tolerant with respect to all such swaps. That is:

Swap Tolerance For any table x, any n distinct atoms $a_1, \ldots, a_n$, and any qualitative relation Q such that $Qa_1 \ldots a_n$, it is indiscernibly possible that $Qa_2a_1a_3 \ldots a_n$ while x is a table composed by the same collection of atoms modulo the interchange of a_1 and a_2.[35]

For short: *every table is swap-tolerantly a table*. This claim does not follow from Indiscernible Tolerance, which only requires that for every collection of atoms 90 per cent overlapping a table's actual originating atoms there is *at least one* way of permuting roles among atoms compatible with the table's being composed by that collection. This does not rule out that there are many other small permutations—even mere swaps—which would get rid of some table, or require it to jump in some weird way. (For example, where x is a table composed by $a_1 \ldots a_k$, and a_{k+1} is an external atom, Indiscernible Tolerance requires that there is some indiscernibly possible world where x is composed by $a_2 \ldots a_{k+1}$, but is compatible with the hypothesis that the only such worlds are, e.g., one where a_1 swaps roles with a_{k+1} while a_2 swaps roles with a_3, whereas x couldn't be concrete at all if a_1 merely swapped roles with a_{k+1}.) Nevertheless, it is hard to see why anyone would accept Indiscernible Tolerance but not Swap Tolerance. In particular, insofar as considerations of *robustness* motivate Indiscernible Tolerance (as we explored in the previous section), they also motivate Swap Tolerance.

Swap Tolerance can serve as the Tolerance premise in the Tolerance Argument where K is 'table', the modality is indiscernible possibility, and closeness is the following relation among micro-specific properties:

> F is swap-close to G := For some distinct atoms $a_1, \ldots, a_n$, some qualitative n-ary relation Q, and two collections of atoms C_1 and C_2 that are identical modulo the interchange of a_1 and a_2: F is *being originally composed by* C_1 *and such that* $Qa_1 \ldots a_n$ and G is *being originally composed by* C_2 *and such that* $Qa_2a_1a_3 \ldots a_n$.

Since every permutation of our finite collection of atoms can be built up by composing some finite sequence of swaps, the conclusion of this Tolerance Argument is equivalent to the following claim:

[35] We can take n here to be a fixed parameter equal to the number of atoms. If we want to treat it as a variable, we can do so by finding some way of coding arbitrary finite sequences of objects in higher order terms and taking R to range over binary relations between objects and such sequence-codes. This is not so hard: for example, we can take sequence-codes to be rigid binary relations between Church numbers (of type $\langle\langle\langle\rangle\rangle, \langle\rangle\rangle$, say) and objects. Similar coding tricks can be used to define a notion of swap-tolerance that works as it intuitively should even if there are infinitely many atoms.

Swap Hypertolerance For any table x composed by a collection of atoms C, any n distinct atoms $a_1, \ldots, a_n$, any qualitative relation Q such that $Qa_1 \ldots a_n$, and any permutation π mapping $a_1, \ldots, a_n$ to $b_1, \ldots, b_n$: it is indiscernibly possible that $Qb_1 \ldots b_n$ while x is a table composed by the the collection of all atoms b such that π maps some member of C to b.

For short: every table is *swap-hypertolerantly a table*.

But given even Weak Microphysical Supervenience, Swap Hypertolerance cannot be true on the assumption that the world is qualitatively symmetric under rotations of 180°, and that there are tables which are located away from the axis of symmetry. For under such circumstances, Swap Hypertolerance tells us that the tables could have been composed of certain other atoms, namely the symmetry partners of the ones that in fact compose them, while the microphysical facts were undisturbed. But Weak Microphysical Supervenience tells us that nothing could be different without a disturbance of the microphysical facts. In more detail: consider the permutation π that maps every atom to its symmetry partner. Let x be a table, let C be the collection of its originating atoms, and let $D \neq C$ be the collection of their symmetry partners. Then where $a_1, \ldots, a_n$ are all the atoms, $b_1, \ldots, b_n$ are their respective symmetry partners, and Q is some maximally specific qualitative relation, $Qb_1 \ldots b_n$ is necessarily equivalent to $Qa_1 \ldots a_n$, and thus given Weak Microphysical Supervenience necessitates every truth, including the truth that x is not originally composed by D. But Swap Hypertolerance entails that it is atomically possible that $Qb_1 \ldots b_n$ and x is originally composed by D: contradiction.

The Non-contingency premise of the Tolerance Argument whose conclusion is Swap Hypertolerance can be supported, as before, by appeal to certain premises about qualitativeness: the only new nonlogical primitive needed by our new Tolerance premise is the property of being a microphysical proposition, whose claim to qualitativeness seems strong. And Persistent Closeness is guaranteed in the same way as before. So if Iteration-denial is off the table, the supposition that Swap Tolerance is true in a symmetric Democratean universe (with some off-centre tables) is incompatible with the supposition of Weak Microphysical Supervenience. Iteration and the qualitativeness premises thus force a surprising conclusion: necessarily, if a table is one of two qualitatively indiscernible tables in a world of finitely many modally interchangeable Democratean atoms where Weak Microphysical Supervenience is true, then it is *not* swap-tolerantly a table.[36]

[36] There are in principle two ways in which a table x could fail to be swap-tolerantly a table. First, there could be two atoms a and b such that it is impossible for x to be table while a and b swap roles. Or, second, it could that there is no such pair of atoms, but there are two atoms a and b such that it is not possible for a and b to swap roles while x has the same originating atoms modulo the substitution of a and b: the swap could force x to "jump" in some unexpected way. Note that in a world where no two tables overlap, the only way for a table x to be a counterexample of the second kind would be for it to be the that if a and b had swapped, x would have been composed by some collection of atoms with at most one atom in common with x's actual originating atoms. For given our qualitativeness premises, *being*

The result that tables in symmetric universes can't be swap-tolerantly tables is surprising. But given that the universe is not actually symmetric, the result is compatible with the actual truth of Indiscernible Hypertolerance, and indeed with the stronger claim that necessarily, so long as the universe is not symmetric, all tables are swap-hypertolerantly tables. It is certainly jarring to think that tables at symmetric worlds are so dramatically more fragile than tables at non-symmetric worlds. But this does not strike us as an overwhelming cost. Symmetric worlds are pretty exotic, and it would not be so surprising to think that we are apt to fall into error when we extend our ordinary ways of reasoning about the modal properties of ordinary objects to such worlds.

But even setting symmetric worlds aside, Swap Hypertolerance still has some unpalatable consequences under the assumption of Weak Microphysical Supervenience. Assume, for simplicity, that no collection of atoms originally composes two distinct tables. Then Swap Hypertolerance and Weak Microphysical Supervenience jointly imply that when a table has a qualitative property, it is indiscernibly impossible for it to lack that qualitative property.[37] This is strange, and sits oddly with the motivations for Indiscernible Tolerance. Following Adams (1979), suppose that the universe is *nearly* symmetric: there are two tables, Castor and Pollux, on either side of the axis of near-symmetry, but Castor has a scratch whereas Pollux doesn't have a scratch. It seems plausible that Castor could have been scratch-free while Pollux was scratched, and indeed that Castor could have had all the qualitative properties that Pollux in fact has and vice versa. Moreover, it seems implausible that this would require getting rid of any of the actual atoms

such that no distinct but overlapping collection composes a table is a qualitative property of collections of atoms; so if x's actual originating collection C has this property, the collection C' that comes from C by swapping b for a must have it in a world where a and b swap roles. So there is a world like that where x is a table not composed by the relevant collection C', it must be composed there by a collection that doesn't overlap C' at all, and hence overlaps C by at most one atom.

[37] Suppose that $a_1, \ldots, a_n$ are all the atoms; Q is a maximally specific qualitative relation such that $Qa_1 \ldots a_n$; π is a permutation mapping $a_1, \ldots, a_n$ to $b_1, \ldots, b_n$; x is the only table originally composed by $a_1, \ldots, a_m$; and x has qualitative property F. Then by Swap Hypertolerance, it is possible that $Qb_1 \ldots b_n$, and every table composed by $b_1 \ldots b_m$ is F, and x is the only table originally composed by $b_1 \ldots b_m$. But if Weak Microphysical Supervenience is true, Q entails *being* $x_1 \ldots x_n$ *such that every truth is necessitated by the truth that* $Qx_1 \ldots x_n$. So it can't *also* be possible that $Qb_1 \ldots b_n$ and x is not F. Since this holds for any permutation π, we can conclude that x's being F is indiscernibly necessary.

Even without the assumption that no collection originally composes two tables, we get a very similar consequence: when a table is originally composed by a collection of atoms having a certain qualitative property, it is indiscernibly impossible for it to be originally composed by a collection of atoms lacking that property. In principle, one could allow for qualitative differences that don't involve qualitative changes in its originating collection. For example, if a single collection of atoms first composes a Chippendale-style table Chippy, whose splintered remains are later assembled into a modernist table Splinter, Swap Hypertolerance and Weak Microphysical Supervenience do not rule out the possibility that certain ways of permuting the atoms lead to indiscernibly possible worlds where Chippy plays the qualitative role Splinter actually plays and vice versa. However, this leeway is uninteresting, since there are no principled grounds for thinking that every table is swap-hypertolerantly a table that do not extend to similar claims like 'Every modernist table is swap-hypertolerantly a modernist table.'

or introducing any new ones. But these natural modal thoughts are incompatible with the combination of Weak Microphysical Supervenience and Swap Hypertolerance.[38]

The thesis that it is indiscernibly impossible for tables to play different qualitative roles than they in fact play is also a consequence of Weak Anti-haecceitism. But while we think this is an implausible consequence of Weak Anti-haecceitism, the current package is even stranger. We are assuming that that it is indiscernibly possible for *atoms* to play different qualitative roles: for example, it is indiscernibly possible for the atoms that in fact composed Castor to compose an unblemished table while those that in fact composed Pollux compose a scratched table. What seems really strange is that this modal freedom in the atoms should not be accompanied by any analogous modal freedom in tables.

We saw earlier that those drawn to Indiscernible Hypertolerance are forced to make exceptions for symmetric worlds. The further oddities we have just pointed out provide strong motivation for them to make more exceptions, allowing for failures of Swap Hypertolerance (and thus also of Swap Tolerance) also in certain other situations where distinct tables play sufficiently similar, though not identical, qualitative roles. If we switch from tables to some other kind like *grains of sand*, there may even be pressure to reject the analogue of Swap Hypertolerance at the actual world: even though the actual world is not even approximately symmetric, it still seems quite plausible that there are certain pairs of grains of sand which could have swapped qualitative roles.

So to our minds, the strategy of accepting Indiscernible Hypertolerance (and the Indiscernible Tolerance Argument that motivates it) is looking rather unattractive, at least when combined with a picture where modal match is plentiful. It can't be accepted as a necessary truth because of symmetric worlds, and any way of trying to draw the line between the cases where Swap Hypertolerance holds and those where Swap Tolerance fails seems uncomfortable. And there are additional problems which arise on any way of drawing the line. For as we will now argue, there are convincing reasons to think that objects can't *cross* the line: when something is swap-tolerantly a table, it's indiscernibly impossible for it not to be, and vice versa. For example, if the line is drawn in such a way that Swap Hypertolerance holds everywhere except for symmetric worlds, this means that it's atomically impossible for a tables in a symmetric world to be created in a non-symmetric world, and vice versa.

To see how this consequence arises, the following definition will be helpful:

[38] As Adams (1979) points out, the idea that near-symmetries could have been broken in the opposite way is especially compelling when we consider worlds that start out in a perfectly symmetric state, and become non-symmetric only later. For example, if tables Castor and Pollux started out symmetric and Castor subsequently acquired a defect thanks to some spontaneous probabilistic process, it's especially strange to deny that it could have been that Pollux acquired a matching defect while Castor remained unblemished.

> x is *permutation-indifferent* := for any $n+1$-ary qualitative relation Q, any distinct atoms $a_1, \ldots, a_n$, and any permutation π that maps $a_1, \ldots, a_n$ to $b_1, \ldots, b_n$: if it is atomically possible that $Qxa_1 \ldots a_n$, it is atomically possible that $Qxb_1 \ldots b_n$.

Informally, an object is permutation-indifferent iff its modal profile is entirely indiscriminate with respect to (the actual) atoms. Note that given Microphysical Supervenience, an object in a symmetric world cannot be permutation-indifferent unless it is itself symmetric, in the sense that whenever it bears a qualitative relation to some atoms it bears the same relation to their symmetry partners. And given our assumption that H_{S5} holds for atomic possibility (see §14.2), permutation-indifference must be an atomically non-contingent property: anything permutation-indifferent is atomically necessarily permutation-indifferent, and that anything not permutation-indifferent is atomically necessarily not permutation-indifferent. For being permutation-indifferent is just a matter of one's range of (atomic) possibilities fitting a certain pattern, and under H_{S5}, what is possible for a thing is a non-contingent matter.[39] This means in particular that any object that could fail to be itself symmetric at some symmetric atomically possible world is not only *possibly*, but (atomically) *necessarily*, not permutation-indifferent: it will bear qualitative modal relations like *being [not] possibly composed of* $y_1 \ldots y_n$ *while* $Qy_1 \ldots y_n$ to some atoms but not others in every atomically possible world.

The definition of permutation indifference makes no special mention of tablehood; but for tables, there is little daylight between being permutation indifferent and being swap-tolerantly a table. In fact our qualitativeness premises entail that every permutation-indifferent table is swap-hypertolerantly a table: just consider the relation *being* $x, a_1, \ldots, a_n$ *such that* x *is a table originally composed by* $a_1, \ldots a_k$ *and* $Qa_1 \ldots a_n$, for a given qualitative Q. And we also get something *close* to the converse of this: any table that is swap-hypertolerantly a table *and does not have exactly the same originating atoms as any other table* is permutation-indifferent.[40]

[39] Where $\Diamond$ is atomic possibility, 'x is permutation-indifferent' means:

(i) $\quad \forall Ra_1 \ldots a_n b_1 \ldots b_n \pi((\mathrm{Qual}\, R \wedge \mathsf{R}\pi \wedge \bigwedge_i \pi a_i b_i \wedge \Diamond Rxa_1 \ldots a_n) \rightarrow \Diamond Rxb_1 \ldots b_n)$

Given S5 and $\Box\forall R(\mathrm{Qual}\, R \rightarrow \Box\, \mathrm{Qual}\, R)$, we have $\Diamond\, \mathrm{Qual}\, R \rightarrow \mathrm{Qual}\, R$. Also, since $\Box(\mathsf{R}R \wedge \bigwedge_i \pi a_i b_i \rightarrow \Box(\mathsf{R}R \wedge \bigwedge_i \pi a_i b_i))$ (by Persistent Rigidity and Persistence), we have $\Diamond(\mathsf{R}R \wedge \bigwedge_i \pi a_i b_i) \rightarrow (\mathsf{R}R \wedge \bigwedge_i \pi a_i b_i)$. And by the 4 and 5 axioms, $\Diamond\Diamond Rxa_1 \ldots a_n \rightarrow \Diamond Rxa_1 \ldots a_n$ and $\Diamond Rxb_1 \ldots b_n \rightarrow \Box\Diamond Rxb_1 \ldots b_n$. We can thus strengthen (i) to

(ii) $\quad \forall Ra_1 \ldots a_n b_1 \ldots b_n \pi((\Diamond\, \mathrm{Qual}\, R \wedge \Diamond(\mathsf{R}\pi \wedge \bigwedge_i \pi a_i b_i) \wedge \Diamond\Diamond Rxa_1 \ldots a_n) \rightarrow \Box\Diamond Rxb_1 \ldots b_n)$

which by K implies

(iii) $\quad \forall Ra_1 \ldots a_n b_1 \ldots b_n \pi\Box((\mathrm{Qual}\, R \wedge \mathsf{R}\pi \wedge \bigwedge_i \pi a_i b_i \wedge \Diamond Rxa_1 \ldots a_n) \rightarrow \Diamond Rxb_1 \ldots b_n)$

Finally we can move the box to the front using BF to derive the necessitation of (i).

[40] For suppose that $a_1, \ldots, a_n$ are all the atoms there are; x is the only table originally composed by $a_1, \ldots, a_m$, and is swap-hypertolerantly a table; R is a qualitative $n+1$-ary relation such that it's atomically possible that $Rxa_1 \ldots a_n$; and π is a permutation mapping $a_1, \ldots, a_n$ to $b_1, \ldots, b_n$. Let T be

In principle, one could think that matters are different when it comes to unusual tables which share their exact originating atoms with other tables (e.g. made earlier or later); but such exceptionalism is both unpromising and uninteresting. So from now on we will assume that necessarily any table is permutation-indifferent if and only if it is swap-hypertolerantly a table.

Given the non-contingency of permutation-indifference, this assumption implies that whatever might be the necessary and sufficient conditions for a table to be swap-hypertolerantly a table, it is necessary that if those conditions obtain, then every table is such that it is impossible for that table to be a table unless those conditions obtain, and vice versa. For example, suppose we go for the most conservative response to the symmetry argument according to which necessarily, any table is swap-hypertolerantly a table unless it has a symmetry partner. Then we will have to say that any table with a symmetry partner is such that it couldn't still be a table without having a symmetry partner. This is very strange! Consider a rotationally symmetric world with two tables, Castor and Pollux. One might have thought that the symmetry could have been broken in various ways, e.g. by Castor having a scratch not corresponding to any scratch on Pollux, or by Castor's being destroyed earlier than Pollux, or even by Castor's never being made in the first place. But the conservative view under consideration prohibits this, since it implies that necessarily any table in an asymmetric universe is permutation-indifferent. Since neither Castor nor Pollux is permutation-indifferent, neither of them could have been permutation-indifferent (with the same atoms), and thus neither could have a table in an asymmetric universe (with the same atoms).[41]

One might try to avoid this strangeness by drawing the line less conservatively, crafting some notion of 'partner' that demands less (perhaps much less) than perfect qualitative indiscernibility, and claiming that necessarily, tables that lack partners are swap-hypertolerantly tables, while tables that have partners aren't even swap-tolerantly tables. But that just moves the strangeness to a different and even less promising place, since we will now have to say that tables that only barely fail to have partners couldn't have had partners, and that tables that only barely have partners couldn't have lacked them.

the relation *being atoms* $y_1, \ldots, y_n$ *such that for every table z, if z is originally composed by* $y_1, \ldots, y_m$, *then* $Rzy_1 \ldots y_n$. *T* is qualitative by the the qualitativeness premises, and instantiated by $a_1, \ldots, a_n$. Since *x* is swap-hypertolerantly a table, it is atomically possible that $Tb_1 \ldots b_n$ and *x* is a table composed of $b_1, \ldots, b_m$. It follows that it is atomically possible that: *x* is a table composed of $b_1, \ldots, b_m$ and for every table *z* composed of $b_1, \ldots, b_m$, it is atomically possible that $Rzb_1 \ldots b_n$. Hence by Iteration it is atomically possible that $Rzb_1 \ldots b_n$.

[41] This argument is a version of a very influential argument in Adams 1979 (§V). Adams's Castor and Pollux are not tables but "globes", like those in Black 1952 (but perhaps populated). Adams is arguing against Weak Anti-haecceitism from the possibility of qualitative symmetry (which he has already argued for). But his mode of argument supports the stronger conclusion that there are possible worlds that agree on all qualitative propositions, but differ as regards which qualitative properties are instantiated by which *globes*.

Throughout this section we have been reasoning under the simplifying assumption that the world is built from a finite stock of modally interchangeable Democratean atoms. But much of the argumentation would still go through using other conceptions of the microphysical ground floor, and many weakenings or adaptations of the permutability assumption. So long as Weak Anti-haecceitism and its atomic-possibility analogue are false, there will be *some* group of permutations of the microphysical objects that correspond to indiscernible possibilities. Insofar as there is a good case for Indiscernible Tolerance, there will be a good case for the claim that all tables are tolerantly tables with respect to any "sufficiently small" permutations in this group. But given the qualitativeness premises and Iteration, this claim will entail that all tables are also tolerantly tables with respect to any permutation that can be built up by composing some finite sequence of "sufficiently small" permutations. This might well mean (depending on whether the universe is in some appropriate sense infinite or finite) that they are tolerantly tables with respect to *all* permutations that correspond to indiscernible possibilities. One could stake one's hopes on the idea that the universe will turn out to be infinite, so that tables in the actual world can still be claimed to be tolerant with respect to "finite-sized" permutations, without being tolerant with respect to the kinds of "global" permutations that make for trouble in symmetric worlds, given Microphysical Supervenience. But this still doesn't address our other worry, about the implausibility of ruling out certain plausible possibilities of role-swapping among tables (or among grains of sand). More generally, we are not inclined to make our views about the modal behaviour of tables a hostage to empirical fortune in this particular way.

We have been looking in this section at an approach to Indiscernible Tolerance Arguments that accepts Indiscernible Tolerance, Iteration, and the qualitativeness premises, and thus accepts Indiscernible Hypertolerance, but does not accept anything like Anti-haecceitism (which makes Indiscernible Hypertolerance vacuously true). We have shown that unless it is combined with the denial of Microphysical Supervenience, the package has some bizarre and unpalatable consequences. Meanwhile, for reasons explained in the previous section, the only remotely promising motivations for Indiscernible Tolerance already turn on something in the vicinity of Microphysical Supervenience. Thus even if one were willing to jettison Microphysical Supervenience, the hypertolerant package still seems undermotivated.[42] If we were forced to accept the qualitativeness premises, then assuming the underlying physics was favourable, we would be inclined to go for the strategy of denying Indiscernible Tolerance, avoiding the robustness problems by relying on some super-demanding essentialism with regard to the distant past,

[42] The package will also have to reckon with a version of the challenge from robustness, but we think the concerns already raised are damaging enough as to not require going into this as well.

as suggested in §14.3. If the underlying physics was less favourable, we would have some hard thinking to do.

But in fact, the qualitativeness premises—and specifically, the premise that *tablehood* is qualitative—look quite shaky, especially in the light of the picture developed in Chapter 11 on which the word 'table' is semantically highly plastic. In the next chapter we will develop our favoured response to the Indiscernible Tolerance Arguments, which blocks the case for Indiscernible Non-contingency by denying the qualitativeness of properties like tablehood.

15
Non-Qualitativeness and Aboutness

In the previous chapter, we saw that certain prima facie plausible assumptions about qualitativeness suffice for the truth of Non-contingency in certain Tolerance Arguments involving indiscernible modalities. In particular, given the background theory of qualitativeness we laid out in §14.1, the plausible assumption that all of the following are qualitative suffices for the truth of the Non-contingency premise in the Indiscernible Table Argument (see Figure 14.2):

being an atom
being an atomic number
being originally composed by
being metaphysically possible
being a table

If Indiscernible Non-contingency is true, and the denial of Iteration for metaphysical necessity is off the table, we then have a difficult choice between denying Indiscernible Tolerance and accepting Indiscernible Hypertolerance. As we saw in the last chapter, the second option is hard to maintain for those like us who embrace Microphysical Supervenience. The first option emerged as more defensible, but was to a considerable extent hostage to empirical fortune.

In the present chapter we present our preferred response to the puzzle, which is to accept both Indiscernible Tolerance and Indiscernible Non-hypertolerance while denying Indiscernible Non-contingency. This requires denying the qualitativeness of at least one of the five properties and relations listed above. The most plausible culprit, and the one we will focus on in this chapter, is *being a table*.

15.1 The Non-Qualitativeness of Tablehood

The idea that *being a table* is a non-qualitative property might initially seem bizarre. But we need to be careful. It really would be bizarre to suppose that *being a table-shaped object* is non-qualitative.[1] But on any view that allows that tables sometimes coincide with table-shaped objects that are not tables, there is a very

[1] Here we are imagining that 'table-shaped' is spelled out in geometrical terms. It would not be bizarre if 'table-shaped' is taken to mean something like 'similar in shape to a paradigm table' where what counts as a paradigm table might vary from world to world.

The Bounds of Possibility: Puzzles of Modal Variation. Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri, Oxford University Press. © Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri 2021.
DOI: 10.1093/oso/9780192846655.003.0016

significant difference between the two properties. And while many of the table-shaped objects coincident with a given table will be excluded from tablehood on qualitative grounds—for example, some of them will be *possibly spherical*, whereas typical tables are plausibly not possibly spherical—the need to disqualify all but one of any a given collection of coincident table-shaped objects makes it natural to expect that non-qualitative factors will also play a role.[2]

By denying the qualitativeness of tablehood, we can smoothly combine acceptance of Indiscernible Tolerance with rejection of Indiscernible Hypertolerance. Our picture will be that every table is still a table in all the "close" indiscernibly possible worlds—those derived from the actual world by some suitably "small" permutation of the microphysical objects. Moreover the originating atoms of a table in each such world are whichever atoms there play the qualitative roles of its actual originating atoms.[3] But while the tablehood of each table is in this way robust, some of its qualitative properties are far more fragile: for example, the qualitative property of being uniformly tolerant (or "at one's sweet spot") will be disrupted by any permutation that does not preserve a table's actual originating atoms.[4] At an indiscernibly possible world where a given table still exists but doesn't play its actual qualitative role, *something* must of course play that role (since the proposition that something plays the role is qualitative, hence indiscernibly necessary). But typically, the object that plays the role at other indiscernibly possible worlds will be something that is not a table, either at that world or at the actual world.[5]

In fact, the plenitude and plasticity package that emerged in Chapter 11 already makes it quite natural to think that tablehood is non-qualitative. We avoided worries about the security of our belief that every table is tolerantly a table by claiming that 'table' is highly semantically plastic, so that a change as small as moving one saw in one workshop would make a difference to what is expressed by the relevant uses of 'table'. Consider all of the other properties expressed by uses of 'table' in nearby worlds that differ slightly with respect to the choices of originating matter for tables. The crucial difference between those worlds and the actual world seems to lie in the haecceitistic facts as regards which portions of matter make up which tables. So it seems eminently plausible that the difference in the properties expressed is also haecceitistic in character. The property expressed in those worlds

[2] A view like Yablo's (see Chapter 11, §11.5), where being a table requires being modally "at one's sweet spot", would make the qualitativeness of tablehood much more plausible: after all, *being at one's sweet spot* seems qualitative. But as we have already said, we regard the failure of tables to be tolerantly tables (and the failure of people to be tolerantly people) as too high a price to pay.

[3] Thus, in the Democratean setting of §14.4, we would accept Swap Tolerance.

[4] In the Democratean setting, this means that tables are not "permutation-indifferent" even with respect to permutations that merely swap two atoms.

[5] Thus not only tablehood itself, but all of the more specific properties that ordinary tables have that necessitate tablehood must be non-qualitative.

privileges the collections that constitute tables in those worlds in the same way that the property we express privileges the actual table-constituting collections.

We shouldn't overstate the case: the semantic plasticity of 'table' doesn't *force* one to think that the properties it typically expresses are haecceitistic ones. The pressure to invoke semantic plasticity by claiming that 'table' expresses different properties at two worlds arises when there is reason to think that the instances of the property expressed in one world still instantiate the property at the other world, but differ in modal respects (e.g. by being intolerant). But as we saw in §14.3, it is coherent to suppose that while tables are tolerant, they are indiscernibly intolerant. On that picture, worlds that qualitatively match actuality but differ by some permutation of atoms that disrupts the facts about which atoms compose tables generally do not contain the same tables as the actual world; so there is no pressure to think that 'table' expresses different properties at such indiscernibly possible worlds. It is still true that moving one saw in one workshop would be enough to stop uses of 'table' from referring to tablehood, but it is false that the crucial difference between the actual world and the world where the saw is moved is haecceitistic. Rather, what drives the semantic shift on that picture is some detailed features of the qualitative situation—e.g. qualitative facts about the later histories of the atoms that played certain highly specific roles in the early history of the universe—and the properties expressed are qualitative properties that are sensitive to those detailed qualitative differences. Nevertheless, having bought into plenitude and plasticity, we are not much inclined to go for this esoteric picture.

Myriad other properties could be substituted for tablehood in the Indiscernible Table Argument without substantially affecting the prima facie plausibility of any premise, or the prima facie implausibility of the conclusion. The argument for the non-qualitativeness of tablehood from Indiscernible Tolerance and Indiscernible Non-hypertolerance will thus carry over to all these other properties. As we discussed in Chapter 13, this category includes properties like *being conscious*, *being in pain*, *being a thinking thing*, and the like. And while we don't have much sympathy with philosophers who claim to find it obvious that tablehood is qualitative, it is hard to be similarly dismissive of philosophers who hold up consciousness or thought, or more specific consciousness- or thought-entailing properties like *being in pain* and *knowing that one is self-identical*, as paradigms of qualitativeness.[6] However, the sensibility that regards the qualitativeness of such properties as non-negotiable is likely to go hand in hand with the view that rejects

[6] It's an interesting question whether, having accepted the non-qualitativeness of properties like *being in pain*, one should also accept the non-qualitativeness of properties of *events* like *being a pain* (or *being an experience of red*, etc.). It seems odd to think of events as "composed" by atoms, but one might be able to find some other relation which could play a similar role in an Indiscernible Tolerance Argument, e.g. a relation *occurring at* between events and times or spacetime points. Since *being in pain* is plausibly equivalent to *being the subject of a pain event*, if one wanted to preserve the qualitativeness of *being a pain event* while denying the qualitativeness of *being in pain*, one would need to deny the qualitativeness of the *being a subject of* relation.

out of hand the idea that the kind of semantic plasticity we have posited in words like 'table' extends to words like 'conscious' and 'think'. As we argued in §13.3, the most stable elaborations of this perspective involve a localized embrace of Hypertolerance in the special case of thinkers, conscious beings, and so on, taking us surprisingly far in the direction of Cartesianism. If on the other hand one is comfortable with the plasticity of 'conscious' and 'think', it is doubtful that there is any solid basis for insisting on the qualitativeness of the properties they express.

15.2 From Non-Qualitativeness to Plenitude

We have suggested that the package of plenitude and plasticity, which we advocated as a strategy for dealing with other Tolerance Arguments, also helps to undermine Indiscernible Non-contingency by undermining the prima facie case for the qualitativeness of tablehood. In this section we will argue in the other direction, showing that the denial of Indiscernible Non-contingency generates pressure to go pretty far in the direction of plenitude. The upshot is an entirely new kind of argument for plenitude.

The following intuitive thought will serve as our springboard is that: since properties like *being indiscernibly tolerant*, *not being hypertolerant*, and *being table-shaped* are qualitative, if our actual tables are indiscernibly tolerant but not hypertolerant, then the proposition that there are table-shaped objects that are indiscernibly tolerant but not hypertolerant will be indiscernibly necessary (since it is qualitative and true). From this intuitive thought and mundane facts about actual tables, we can start to generate the kind of abundance of coincident objects that plenitude-lovers embrace, and that opponents of plenitude abhor. Coincident with any table-shaped object that is "at its sweet spot", there will be many others that are not at their sweet spots, and which vary as to which collections would put them at their sweet spots.

To flesh this out, consider a very simplified case: suppose that there are just twelve Democratean atoms $a_1, \ldots, a_{12}$ arranged as in Figure 15.1, and that atoms $a_1, \ldots, a_6$ count as composing a very tiny spoon, which we shall name Nano.[7] $a_1, \ldots, a_{12}$ instantiate twelve distinct qualitative roles $Q_1, \ldots, Q_{12}$, respectively. For any non-repeating sequence $\pi = \langle \pi_1, \ldots, \pi_{12} \rangle$ consisting of the numbers from 1 to 12 in some order, let p_π be the proposition that a_1 has Q_{π_1} and …and a_{12} has $Q_{\pi_{12}}$. Let's assume that the atoms are modally interchangeable, so that each p_π is indiscernibly possible. And finally, let's also assume that Nano is uniformly tolerant with threshold 2/3: for each π, if $\pi_1, \ldots, \pi_6$ include at least four of the numbers 1–6, then it is indiscernibly possible for Nano to be spoon-shaped while p_π is true,

[7] We found designing a tiny spoon easier than designing a similarly tiny table; the claim that every spoon is indiscernibly tolerant but not indiscernibly hypertolerant is just as plausible as its analogue for tables.

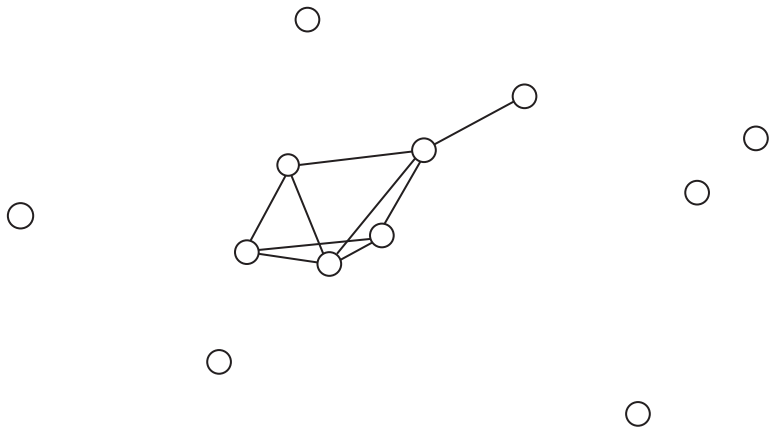

Fig. 15.1 A nanospoon and its environment.

and otherwise, it is indiscernibly impossible for Nano to be concrete while p_π is true. Thus, at least as regards indiscernibly possible worlds, Nano needs at least four of $a_1, \ldots, a_6$ as parts, but any four of them will do. It doesn't care how they are arranged inside it, or how the remainder are arranged outside it.

For any object x and collection of atoms Y, say that Y is a "sweet spot collection" for x just in case for any collection Z with the same number of members as Y, it is indiscernibly possible for x to be composed by Z if and only if Z has at least four members in common with Y. So our stipulation about Nano can be rephrased as the claim that the collection whose members are $a_1, \ldots, a_6$ is a sweet spot collection for Nano. Note that an object can have at most one sweet spot collection, since no two collections can agree as regards which equinumerous collections overlap them by at least four. Also, given our standing assumption that the logic of atomic possibility (and hence also of indiscernible possibility) is H_{55} (see §14.2), an object has its sweet spot collection indiscernibly necessarily.

When F is any property and x is an object, say that F is "sweet for x" just in case F is instantiated by all and only the members of x's sweet spot collection. So in the world of our example, the property of being an atom that is part of Nano is sweet for Nano, and so is the disjunction $Q_1 \vee \cdots \vee Q_6$. Both of these properties are only contingently sweet for Nano, since it is only contingently true that their instances are $a_1, \ldots, a_6$.

A little combinatorics shows that there are 262 different six-membered collections of instantiated qualitative roles that contain at least four of $Q_1, \ldots, Q_6$.[8] Each corresponds to a distinct qualitative property, the disjunction of its members. Each of these qualitative properties is indiscernibly possibly such that its instances are $a_1, \ldots, a_6$ in a world where at least four of $a_1, \ldots, a_6$ instantiate properties among $Q_1, \ldots, Q_6$. Hence, each is indiscernibly possibly sweet for Nano while Nano is

[8] One collection that contains all of $Q_1, \ldots, Q_6$; $36 = \binom{6}{5}\binom{6}{1}$ that contain five of $Q_1, \ldots, Q_6$ and one of $Q_7, \ldots, Q_{12}$; and $225 = \binom{6}{4}\binom{6}{2}$ that contain four of $Q_1, \ldots, Q_6$ and two of $Q_7, \ldots, Q_{12}$.

spoon-shaped. So each is such that there it is indiscernibly possible for there to be a spoon-shaped object for which it is sweet. But the properties are all qualitative, and *being spoon-shaped* is qualitative, and *being sweet for* is a qualitative relation. So for each of the properties, the proposition that there is a spoon-shaped object for which it is sweet is a qualitative proposition, and thus can be indiscernibly possible only by being true. So, for each of our 262 properties, there is a spoon-shaped object for which it is sweet. Moreover, since no object has more than one sweet spot collection, any two properties both of which are sweet for the same object must be coextensive, and none of our 262 properties are coextensive. So we can conclude that there are *in fact* at least 262 distinct spoon-shaped objects. Given how the atoms are arranged, this means there are at least 262 objects that coincide with Nano. And once we turn from tiny spoons to macroscopically sized objects, these numbers will become astronomical.

So, we have an argument for a signature consequence of plenitude—vast hordes of material objects coinciding with some given object—whose only premises are that the object in question is indiscernibly tolerant (with respect to its original composition); that it is not hypertolerant; and—let's not forget—that Iteration holds for metaphysical (and hence also indiscernible) possibility. This is interestingly different from, and complementary to, the arguments for plenitude that we considered in Chapter 11. It does not rely on any premise to the effect that people who talked in certain ways would not be making a mistake, or on any thought about the implausibility of positing "natural joints" among the properties in the vicinity of a given modal profile.

Admittedly, the multiplicity of coinciding objects that we can argue for by these means falls a long way short of what we get from plenitude: all of the objects in our multiplicity have modal profiles of the same "shape", differing only with regard to the position of the actual world relative to that shape. In other words, they have exactly the same *necessary* qualitative properties. So, we will still need a step of abductive generalization if we want an argument that gets us all the way to some plenitude principle of the sort discussed in § 13.2. But even before taking this step, we already have a view that will seem repugnant to anyone who finds plenitude repugnant. Moreover, even without the full strength of plenitude, the view already generates a sufficient abundance of material objects to provide the array of available interpretations for demonstratives and words like 'table' required to get our plasticity-based response to the Security Argument off the ground.

15.3 Aboutness

The thesis that tablehood is haecceitistic (non-qualitative) leads naturally to further questions. Is it "just a little bit haecceitistic", like the property of a metre from a particular elementary particle a_1. Or is it "massively haecceitistic", like the property

of being x_1 metres from a_1, and x_2 metres from a_2, and ... and $x_{10^{100}}$ metres from $a_{10^{100}}$, for 10^{100} distinct elementary particles $a_1 \ldots a_{10^{100}}$? To sharpen this kind of question, we shall use a somewhat unfamiliar notion of *aboutness*. In this section we will introduce this ideology with some care, both because it can (as we will see in subsequent sections) be used to raise some interesting challenges for our view, and because we want to advertise its utility as a tool for refining metaphysical debates involving qualitativeness.[9]

We will understand aboutness for entities of a given type as a relation between entities of that type and *collections of objects* (i.e. rigid properties of type $\langle e \rangle$). Intuitively, x is about collection C just in case we can define x entirely in terms of terms of qualitative ingredients together with objects belonging to C. For example: the property of *rating Lasse Virén above Sebastian Coe* is about the two-membered collection whose members are Virén and Coe, since it can be reached by plugging Lasse Virén and Sebastian Coe into two of the three open arguments of the three-place qualitative relation *x rates y above z*. It is also about any larger collection that includes them. More generally, a basic property of aboutness as we conceive of it is monotonicity: when an entity is about a collection of objects, it is also about any larger collection that includes that collection. If we can define the entity using only resources from the smaller collection, then of course we can also define it using only resources from the larger one. The qualitative entities of a given type are those that are about the *empty* collection of objects (and hence also about every other collection of objects).[10]

Our basic theory about aboutness is given in Figure 15.2. The first axiom captures the monotonicity principle just introduced. The next three are generalizations of the corresponding axioms from the Basic Theory of Qualitativeness (Figure 14.1). These three axioms work together with the crucial new axiom, Object Aboutness, to let us deduce a wide range of aboutness claims. For example, we can deduce that that for any object x belonging to some collection C, x's haecceity—i.e. $\lambda y.y = x$, the property of being identical to x—is about C. For x is about C by Object Aboutness; and identity is about C by Constant Aboutness; and by Aboutness Closure, $\lambda y.Ryx$ is about C whenever R and x are both about C. The final axiom, Persistent Aboutness, says that relations of aboutness obtain of necessity, naturally generalizing Persistent Qualitiveness.[11]

[9] Our thinking about this notion has been shaped by conversations with Jeremy Goodman, whose influence on the discussion that follows is pervasive.

[10] Given how we are talking about aboutness it would be natural also to add $\mathrm{About}(x, C) \rightarrow \mathsf{R}C$ as an axiom. But in fact such an axiom does no real work, since Aboutness Monotonicity already implies that About is extensional in its second argument. Our focus on the case where the second argument is a collection just serves to make it vivid that all that matters in that argument is which things have the property, rather the nature of the property itself.

[11] Note that even if we allow for $\mathrm{About}(x, Y)$ to be true when Y is a non-rigid property of objects, as suggested in note 10, Persistent Aboutness must be restricted to the case where the second argument is rigid (i.e. is a collection). Persistent Aboutness is potentially more controversial than the other axioms

Aboutness Monotonicity $\forall x(Cx \rightarrow Dx) \rightarrow \text{About}(x, C) \rightarrow \text{About}(x, D)$.

Aboutness Closure $\text{About}(x_1, C) \rightarrow \cdots \rightarrow \text{About}(x_n, C) \rightarrow \text{About}(a, C)$, where $x_1 \ldots x_n$ are variables of any types, and a is some complex term containing no constants and no free variables other than $x_1, \ldots, x_n$ (where $n \geq 0$).

Constant Aboutness $\text{About}(c, C)$ where c is any of $\neg$, $\wedge$, $\vee$, $\forall_\sigma$, $\exists_\sigma$, $=_\sigma$, Rigid_σ, or About_σ, for any type σ.

Collection Aboutness $(\forall x(Fx \rightarrow \text{About}(x, C)) \rightarrow \exists D(\mathsf{R}D \wedge \text{About}(D, C) \wedge \forall x(Fx \leftrightarrow Dx)))$.

Object Aboutness $Cx \rightarrow \text{About}(x, C)$ (where x is a variable of type e).

Persistent Aboutness $\mathsf{R}C \rightarrow \text{About}(x, C) \rightarrow \Box\,\text{About}(x, C)$ (where $\Box$ is metaphysical necessity).

Fig. 15.2 The Basic Theory of Aboutness.

Just like the notion of a qualitative object, the notion of an *object* being about a collection of objects (which is crucial to Object Aboutness) is somewhat unfamiliar. But if we wanted to, we could just *define* 'object x is about C' either as 'x belongs to C' or as 'x's haecceity is about C' (thus upgrading the conditional we just proved to a biconditional).[12]

There is another principle about aboutness which one might be tempted to add to the list, but which is much more tendentious than those in Figure 15.2, namely that aboutness is *closed under intersection*: if x is about C and x is about D, and E contains all and only the objects that belong to both C and D, then x is about E. Certain plausible higher-order identities that featured in Chapter 7 would, if true, refute that principle. Consider for example:

(1) To be in the Tri-State Area is to be in New York, New Jersey, or Connecticut.

Assuming that *being in* is qualitative, Object Aboutness, Aboutness Closure, and Constant Aboutness together entail that the property of being in the Tri-State Area is both about the singleton collection containing only the Tri-State Area

for reasons discussed in §14.1; we could avoid the need to rely on it by focusing on a narrower modality stipulated to hold fixed aboutness facts for the relevant type.

[12] Note that the converse of Object Aboutness—i.e. the principle that no object *not* belonging to a collection is about that collection—will be controversial given that biconditional. For example, nothing in the theory rules out the possibility that some objects' haecceities are qualitative, so that both they and the objects whose haecceities they are will be about the empty collection.

and about the three-membered collection containing New York, New Jersey, and Connecticut. The intersection principle would then entail that it is about the empty collection, i.e. qualitative. But unless one thinks that every property whatsoever is qualitative, it is completely implausible that being in the Tri-State Area is qualitative.

The plural character of our notion of aboutness takes some getting used to. It is more common for metaphysicians to theorize using a notion of aboutness, or "involvement", understood as as a binary relation between entities of a given type and *single* objects.[13] Starting with this singular notion of aboutness, one might be tempted to define a plural notion of aboutness as follows: x is about collection C just in case every object that x is (singularly) about belongs to C. This definition automatically validates Aboutness Monotonicity, but also validates the tendentious claim that plural aboutness is closed under intersection, and for that reason makes Aboutness Closure problematic. For example, if one accepts (1) and thinks that none of *being in New York*, *being in New Jersey*, and *being in Connecticut* is singularly about the Tri-State Area, one will have a counterexample to Aboutness Closure: each of the three properties is about the collection comprising all objects distinct from the Tri-State Area, but their disjunction is not. Since Aboutness Closure is central to our conception of plural aboutness, we will not assume the tempting definition of plural in terms of singular aboutness, and will not explore the singular notion further here.

It is debatable whether aboutness draws distinctions between metaphysically necessarily equivalent properties, relations, and propositions. One might think it does, if one thinks that (1) is false although the corresponding metaphysically necessitated biconditional is true: the property of being in the Tri-state Area is necessarily equivalent to the property of being in New York, New Jersey, or Connecticut, but only the latter is about the collection of New York, New Jersey, and Connecticut. But for our current purposes it will be convenient to be able to bypass such hyperintensional issues, for the same reason that our argument for the non-qualitativeness of tablehood extends naturally to the thesis that tablehood is not even necessarily equivalent to a qualitative property. So, let's say that a property, proposition, or relation is *intensionally about* a collection of objects just in case it is necessarily equivalent to one that is about that collection, and that an object is intensionally about a collection of objects just in case its haecceity is necessarily equivalent to a property that is about that collection. (Of course, if Intensionalism is true, intensional aboutness just is aboutness.)

[13] Singular aboutness might in turn be defined in a way that generalizes the Khamara (1988) definition of qualitativeness (see note 3 in Chapter 14): Y of type $\langle \sigma_1, \ldots, \sigma_n \rangle$ is about object x iff for some Y' of type $\langle e, \sigma_1, \ldots, \sigma_n \rangle$, $Y = \lambda z_1 \ldots z_n . Y'(x, z_1, \ldots, z_n)$.

Note that in the presence of plenitude, intensional aboutness behaves in some surprising ways. For example, for every proposition p that entails that there is a concrete object, there is an object x that is necessarily concrete iff p. So every such p is about some one-membered collection. But so long as one doesn't make the mistake of assuming that intensional aboutness is closed under intersection, nothing disastrous follows from this.

From now on intensional aboutness will be our main concern, so we will just call it 'aboutness' and reserve 'strong aboutness' for the previous notion. And we will bear down on the focal question 'How is tablehood non-qualitative?' by asking which collections tablehood is about.

15.4 What Is Tablehood About?

We can probe the question what tablehood is about using aboutness-theoretic analogues of the Indiscernible Tolerance Argument. For any collection of objects C, and any given variety of possibility, we can introduce a corresponding notion of "C-indiscernible possibility" that stands to *being about C* as our earlier notion of indiscernible possibility stands to qualitativeness. To be C-indiscernibly possible is to be compatible with every collection of true propositions about C. Equivalently: to be true at some possible world that agrees with actuality not just as regards the pattern of qualitative properties and relations, but as regards the particular positions in that pattern occupied by each of the members of C. Just as qualitative propositions are automatically indiscernibly necessary when true, propositions about C are automatically C-indiscernibly necessary when true.

Generalizing our argument for the non-qualitativeness of tablehood, we can use Tolerance Arguments based on C-indiscernible modalities to argue for conclusions to the effect that tablehood is not about a given collection C. As in § 14.2, we will work with the closeness relation defined in terms of "modal match", and use atomic possibility as the base notion in terms of which C-indiscernible possibility is defined. Bundling the Tolerance and Non-hypertolerance premises together, the form of the argument will be as follows:

C-**moderation** Every table is C-indiscernibly tolerantly a table, but not every table is C-indiscernibly hypertolerantly a table.

C-**non-aboutness** Tablehood is not about C.

As background premises we assume Iteration (for metaphysical possibility, hence for atomic possibility, and hence also for C-indiscernible possibility); Persistent Closeness; and that all the relevant properties and relations other than tablehood

and C (e.g. *originally composing*, *being an atom* and *being metaphysically possible*) are qualitative.[14] Given the last assumption, tablehood's being about C would, by Aboutness Closure, entail that the proposition *that every table is C-indiscernibly tolerantly a table* is about C, since this proposition can be built up from tablehood and C together with qualitative ingredients. If so, this proposition is C-indiscernibly necessary if true. But then, given Iteration and Persistent Closeness, we would have everything we need for a Tolerance Argument whose conclusion is that every table is C-indiscernibly hypertolerantly a table, contradicting C-moderation.

We have already, in effect, given an argument of this form for the case where C is the empty collection $\varnothing$, so that 'about C' means 'qualitative'. If our case for $\varnothing$-moderation is accepted, it can plausibly be extended to certain non-empty collections. For example, consider the collection X of all the xenon atoms in the universe. It is by no means obvious that tablehood is not about X: after all, if tablehood were qualitative, it would be about every collection. But the argument that tablehood is not qualitative extends naturally to an argument that it is not about X, since any reason for thinking that every table is indiscernibly tolerantly a table will plausibly extend to the claim that every table is X-indiscernibly tolerantly a table. At least, this is the case on the assumption that no xenon atoms are parts of tables, or integral to the processes from which tables originate. Given this, if one is prepared to grant that for every table x and collection of atoms C that modally matches and 90 per cent overlaps x's originating atoms there is a qualitatively indiscernible world where x originates in C, it would be quite implausible to maintain that for certain such x and C, the only such worlds fail to be X-indiscernibly possible, i.e. involve some variation in the qualitative roles played by some xenon atoms. The question what the members of X are up to in a given indiscernibly possible world just seems entirely irrelevant to the question what the actual tables are doing there.[15]

Arguments of this form don't get us very far: for example, in the case where some members of C are originating parts of tables, it is unclear why anyone would

[14] Thinking about the elementary constituents of larger atoms might suggest that *being an atom* fails to be qualitative for the same sort of reason that *being a table* fails to be qualitative; but one could run the whole argument using elementary particles rather than atoms.

[15] You could resist this, holding that the xenon atoms are relevant to the tables, perhaps in some inscrutably vague way, or perhaps because tables are subject to a strong form of origin essentialism of the sort encountered in §14.3, on which no actual table could have been produced unless some early segment of the history of the world were the same in all respects, including the distribution of individual xenon atoms. But if you were willing to play either of those cards, then you should have already been suspicious of the argument that every table is *indiscernibly* tolerant, which can be resisted in similar ways, as we discussed in Chapter 14. Note also that if one replaces Indiscernible Tolerance with Swap Tolerance (see §14.4), the generalization to the analogous claim about X-indiscernible possibility (which could do similar work in an argument that tablehood is not about X) is automatic if no members of X are originating parts of any table.

accept *C*-moderation. But we can get a lot more mileage out of our argument-form if we embed the premise under a further possibility operator:

Possible *C*-moderation It is metaphysically possible that (every table is *C*-indiscernibly tolerantly a table, but not every table is *C*-indiscernibly hypertolerantly a table).

The argument from Possible *C*-moderation to *C*-non-aboutness is valid given the metaphysically necessary truth of the background premises—Iteration, Persistent Closeness, and the general principles about aboutness from §15.3, crucially including Persistent Aboutness.[16]

The argument from Possible *C*-moderation to *C*-non-aboutness is a quite powerful tool. Consider for example the collection *A* of all atoms that are ever parts of tables. It is not at all plausible that *in fact* every table is *A*-indiscernibly tolerantly a table: indeed, one might naturally think that the facts about which atoms originate which tables don't vary *at all* over the space of *A*-indiscernibly possible worlds. But this thought does not apply in non-actual metaphysical possibilities where none of the members of *A* or their constituent particles exist concretely. Assuming that the proposition that all tables are indiscernibly tolerantly tables is in fact true, it is plausible that it could still be true in a world like that. And if all tables were indiscernibly tolerantly tables at such a world, it is plausible that they would also be *A*-indiscernibly tolerantly tables. As we have said, it is implausible that (at the actual world) certain indiscernibly possible changes to the original composition of certain tables can only be made if we simultaneously permute some xenon atoms into the roles of other xenon atoms. For similar reasons, it is implausible that at the worlds in question (where all members of *A* and their constituent particles are non-concrete), some such indiscernibly possible changes can only be made if we simultaneously permute some members of *A* into other qualitative roles.[17] Meanwhile, assuming it is actually true that some tables are not indiscernibly hypertolerantly tables—e.g. because some of them are not counterexamples to Overlap Essentialism—it is plausible that this could still have been the case in a world where all the atoms that are actually parts of tables are non-concrete, but which is still like the actual world in containing many tables, unused

[16] Suppose tablehood is about *C*. Then by Persistent Aboutness, it necessarily about *C*, so *C*-indiscernible Non-contingency is necessarily true. Given that Iteration and Persistent Closeness are also both necessarily true, it follows by a Tolerance Argument that it's necessary that if every table is *C*-indiscernibly tolerantly table, every table is *C*-indiscernibly hypertolerantly a table.

[17] It's important that we get rid of the elementary particles that actually compose the members of *A* along with the atoms themselves. If those particles were still around, then it is plausible that despite being non-concrete, the members of *A* would have interestingly different qualitative profiles involving properties like *being an x such that are some elementary particles that are less than one metre apart and possibly compose x*, which would be disturbed by some of the relevant permutations.

table-parts, and so on. But any table that was not indiscernibly hypertolerantly a table would a fortiori not be *A*-indiscernibly hypertolerantly a table.

We thus have a powerful strategy for strengthening the conclusion that tablehood is non-qualitative, i.e. not about the empty collection, to the conclusion that it is also not about various other collections. In many cases (such as that of *A*) this stronger conclusion seems as defensible as the original conclusion that tablehood is non-qualitative. But there is at least one collection for which this stronger conclusion will be problematic, namely the universal collection *U*, containing all objects whatsoever. In the next section we will turn to that collection, which as we will see, poses an interesting challenge to the argumentative strategy of this section.

15.5 New Objects

The thesis that everything (of any given type) is about *U*, the universal collection of all objects, seems quite hard to resist. A property not about *U* would, in terms of our intuitive gloss, be one that couldn't be "built" or "defined" by starting with all the qualitative properties and relations and then plugging in objects. Even though this isn't ruled out by the Basic Theory of Aboutness, it is hard, prima facie, to get any grip on how there could be such properties.

If every proposition is about *U*, then no false proposition is *U*-indiscernibly possible. A fortiori, no table originally composed by a collection of atoms that modally matches and 90 per cent overlaps some other collection is *U*-indiscernibly tolerantly a table. But it is not obvious that it is *necessary* that every proposition is about *U*. And so it is not obvious that it is *impossible* for a table like that to be *U*-indiscernibly tolerantly a table. Of course, given Persistent Aboutness, if every proposition is about *U*, then every proposition is necessarily about *U*. But unless we accept the Barcan Formula (BF: $\forall x \Box Fx \rightarrow \Box \forall x Fx$ or $\Diamond \exists x Fx \rightarrow \exists x \Diamond Fx$), there is no straightforward route from 'every proposition is necessarily about *U*' to 'necessarily every proposition is about *U*'. And BF is controversial. Many philosophers have rejected BF for first-order quantification on the grounds that that there could be *new objects*, distinct from all the objects there actually are, and thus not belonging to *U*. This requires BF to fail, since every object is such that it necessarily belongs to *U*. But if it is possible for there to be objects not in *U*, it is plausibly also possible for there to be objects that aren't even *about U*. For example, while a new knife made of an actually existing blade and handle would still be about *U* if it was about that blade and handle, it is hard to see how a fundamental particle or spacetime point could be about *U* without belonging to *U*. And if there could be *objects* that were not about *U*, then plausibly there could also be entities of other types not were not about *U*, such as the haecceities instantiated by such new objects.

The most widely discussed views that reject BF and accept the possibility of new objects are *contingentist* ones, that also accept the possible non-being of the objects there in fact are. And the most widely discussed view that accepts necessitism is that of Williamson (2013*b*), who also accepts BF. But although necessitism precludes old objects from ceasing to be, it does not (on its own) preclude new objects from coming to be. Necessitism is a theorem of our basic modal logic, H_{KA} but BF is not. BF *is* (as Prior showed: see note 46 in Chapter 1) a theorem of the stronger modal logic H_{S5}, and in fact all that needed for this derivation beyond H_{KA} is the B axiom ($\forall p(p \rightarrow \Box\Diamond p)$) and its necessitation. In §8.2 we also presented a different, and harder-to-resist, argument for BF from certain premises about the interaction of metaphysical necessity with the 'actually' operator. On the other hand, some philosophers (including at least one of the authors) are rather tempted to give up BF in favour of a view on which whatever objects there are, there could be more—perhaps even a higher cardinality of objects. So, even given our necessitist starting point, it is worth carefully exploring views where there could be objects not in *U*. (Moreover, much of the discussion that follows would carry over to a contingentist setting.)

But once one embraces the view that there could be *thoroughly* new objects—objects not even *about U*—along with ∅-moderation (i.e. the view that tables are (in fact) indiscernibly tolerantly tables but not indiscernibly hypertolerantly tables), a puzzle arises. At a rough first pass: if there could be thoroughly new objects, then presumably there could have been thoroughly new tables made of thoroughly new atoms. And if ∅-moderation is in fact true, it is natural to think that ∅-moderation could still be true if there were such thoroughly new tables. But since swapping around new matter would seem to make no difference to the facts about *U*, there is little room between ∅-moderation and *U*-moderation when it comes to such universes. But as we have seen, assuming that tablehood is about *U*, *U*-moderation cannot hold in any possible circumstances.[18]

Let's go through this more carefully. The possibilities we have in mind are ones where there are new microphysical objects that are independent of *U* in a sense that goes slightly beyond merely not being about *U*. For any collection *C*, say that *C* is *U*-independent iff every indiscernibly metaphysically possible proposition that is about *C* is also *U*-indiscernibly metaphysically possible. Let an *alien-matter universe* be one composed entirely of a *U*-independent collection of new microphysical objects. In other words, an alien-matter universe is one composed of microphysical objects which could permute into any instantiated

[18] Don't confuse the claim that table*hood* is about *U* with the claim that necessarily every *table* is about *U*. The former claim does not imply the latter, any more than the true claim that being table-shaped is about the empty collection (i.e. is qualitative) implies that necessarily all table-shaped things are about the empty collection. Prima facie, if one takes tables in the actual world to be about certain collections of atoms including those that originally compose them, one would expect tables made of thoroughly new atoms to be about some collections of those atoms, and hence not about *U*.

qualitative roles that are possible for them without any difference in the truths about U. In an alien-matter universe, every actual *microphysical* object must be non-concrete, but it bears emphasis that we do not require all members of U to be non-concrete: indeed on our favoured plenitudinous picture, U includes objects that are necessarily concrete if anything is. But given Microphysical Supervenience, members of U that are concrete in an alien-matter universe can't be sensitive to facts about which new objects play which roles.[19] The possibility of an alien-matter universe is not strictly guaranteed by the possibility of new objects, or even by the possibility that every microphysical object is thoroughly new.[20] But if one countenances the possibility of thoroughly new objects, it is entirely natural to countenance the possibility of an alien-matter universe.

If there could be an alien-matter universe at all, there could be one with tables in it. Given that in fact Ø-moderation is true, we would like to think that it could still be true in an alien-matter universe. On the other hand, if tablehood is about U, U-moderation is necessarily false. But any table that isn't indiscernibly hypertolerantly a table would facto not be be U-indiscernibly hypertolerantly a table. And it is hard to see how the tables in an alien-matter universe could be indiscernibly tolerantly tables without being U-indiscernibly tolerantly tables. To make sense of this, one would have to tell a story on which some moderate differences in the composition of new tables out of new atoms can only be achieved by changing the qualitative pattern of U. But this would be very strange: after all, given the definition of an alien-matter universe, all the possible permutations of roles among the new microphysical objects can be implemented without changing the U-facts at all, and it seems quite odd to suppose that permuting the atoms while keeping the same tables around is in some case only possible if one does change the U-facts. Indeed, it is natural to think that tables in an alien-matter universe would, or at least could, be *about* the new microphysical objects, just as actual tables are plausibly about the actual microphysical objects; but the only way such a

[19] Microphysical Supervenience implies that every proposition is necessarily equivalent to a microphysical proposition. Given the way we are understanding "microphysical proposition", it should be straightforward that every microphysical proposition is about the collection M of all microphysical objects, so that given Microphysical Supervenience, every proposition about U is also (intensionally) about M. Thus, so long as indiscernibly metaphysically possible propositions about the new microphysical objects can all be made true without disrupting the boring non-concreteness of the members of M, they can also be made true without changing the truth value of any proposition about U.

[20] Being thoroughly new requires only that one's haecceity is not necessarily *equivalent* to any proposition about U, but is compatible with there being many necessary connections that fall short of equivalence. For example, one could imagine a universe composed of two thoroughly new electrons x_1 and x_2 playing two qualitative roles Q_1 and Q_2 such that for certain actual electrons a_1 and a_2, x_1 is possibly co-concrete with a_1 and not a_2, while x_2 is possibly co-concrete with a_2 and not a_1. Then even if if it indiscernibly metaphysically possible for x_1 to play role Q_2 and x_2 to play role Q_1, this will not be U-indiscernibly metaphysically possible, since it is incompatible with the truth that a_1 is possibly co-concrete with something Q_1. But this is all compatible with the assumption that it is possible for x_1 and x_2 both to be non-concrete and for there to be two different new electrons y_1 and y_2 that bear exactly the same differential relationship to a_1 and a_2. If so, x_1 and x_2 will still be thoroughly new, although the collection comprising the two of them is not U-independent.

table could be indiscernibly tolerantly a table would be for it to be *U*-indiscernibly tolerantly a table.[21]

We can condense the foregoing discussion into the following argument for Possible *U*-moderation, and hence for the claim that tablehood is not about *U*:

Alien Matter There could be an alien-matter universe.

Alien Compatibility If there could be an alien-matter universe, then there could be an alien-matter universe in which every table is indiscernibly tolerantly a table, and not every table is indiscernibly hypertolerantly a table.

Irrelevance of *U* Necessarily, if a table in an alien-matter universe is indiscernibly tolerantly a table, it is *U*-indiscernibly tolerantly a table.

Possible *U*-moderation Possibly, every table is *U*-indiscernibly tolerantly a table, and not every table is *U*-indiscernibly hypertolerantly a table.

As explained in the previous section, for any collection of objects, Possible *C*-Moderation implies that tablehood is not about *C*. Possible *U*-moderation thus implies that tablehood is not about *U*, which conflicts with the natural assumption that every property is about *U*.

In this remainder of this chapter we will discuss three reactions to this argument. First, one could accept the conclusion and deny that every property is about *U*. Second, one could deny Alien Matter, most plausibly as part of a package that denies that possibility of new objects altogether. Third, one could deny Alien Compatibility. To anticipate: we are disinclined towards the first option, but open to both the second and third. (We won't reopen the question of Irrelevance of *U*, since we haven't found a promising treatment of the puzzle based on the denial of this premise.)

First: while we started this section with an impressionistic case that every property is about *U*, we didn't really scrutinize that thesis, and in fact it is not obviously true. There is a reasonably popular picture, defended by Plantinga (1974), Bealer (1993), and others, that accepts BF for higher order quantification but rejects it for first-order quantification: there can be new objects, but there can't be new properties, relations, and propositions. On this view, certain properties, such as those that would be the haecceities of new objects if there were any, seem prima facie not to be about *U*. Given Persistent Aboutness, these would-be haecceities cannot be qualitative, since they wouldn't be qualitative if there were objects whose haecceities they were. And they also cannot be "built" from qualitative ingredients

[21] For all of the relevant collections of atoms (i.e. those modally match their originating collection and overlap by at it by at least 90 per cent), the proposition that the table was originally composed by that collection is about the new microphysical objects, and thus *U*-indiscernibly possibly true if indiscernibly possibly true.

together with the objects there *in fact* are. So, one might think that considering properties like tablehood provides a surprising new argument for a view in the Plantingan spirit.

The package that endorses higher-order BF but not first-order BF involves a prima facie disunity that some may find implausible. But even setting such theoretical considerations aside, the package is in tension with the plenitude principles that we have endorsed. Let's focus here on one of these principles, Coincidence Plenitude. Recall from §11.2 that this is the claim that *every undiscriminating, unrepeatable property is tracked by some object*, where:

> object *x tracks* property *F* := necessarily, all and only the *F* things coincide with *x*.
>
> *F* is *undiscriminating* := necessarily, an object is *F* if and only if it coincides with something *F*.
>
> *F* is *unrepeatable* := necessarily, any two *F* things coincide.

The tension can be made vivid by considering the combination of Coincidence Plenitude with the following further plausible thesis, which seems like a natural (though not inevitable) way of filling out a Plenitudinous vision:

No Tight Connection There couldn't be two distinct, possibly-concrete, tightly connected objects.

where

> *x* and *y* are *tightly connected* := necessarily, for any *z*, *z* coincides with *x* iff *z* coincides with *y*.

No Tight Connection implies that no property could be tracked by more than one possibly-concrete object. Using this, we can derive a contradiction from the assumption that there could be a concrete object not in *U*. For if there could be such an object, there could be a property that was the haecceity of such an object, so by higher-order BF there *is* a property *F* that *could* be the haecceity of such an object.[22] Now consider the property *F** of *coinciding with something with F as its haecceity*. *F** is undiscriminating and unrepeatable (since coincidence is necessarily transitive and symmetric, and no two objects can share a haecceity).

[22] It is natural to think that any property that could be a haecceity but isn't in fact haecceity would be uninstantiated, but in fact in the setting of the weak modal logic required for BF-denial to be consistent, there is no logical obstacle to the existence of instantiated properties that are not haecceities but could be haecceities—for example, if it is possible for John and Cian to be identical, then (at least given Booleanism) it is possible for the property of being either John or Cian to be a haecceity.

So by Coincidence Plenitude, there is an object x that tracks F^*. By Iteration, x necessarily tracks F^*. But necessarily, if there were a concrete object not in U with F as its haecceity, it would also track F^*, and thus be tightly connected to x, and thus be identical to x by No Tight Connection, and thus be in U (since x is necessarily in U): contradiction.

The combination of Coincidence Plenitude, No Tight Connection, and higher-order BF thus implies that it is not possible for there to be an *alien-matter universe* as defined above. This undermines the argument for the surprising claim that tablehood is not about U.

One might contemplate escaping this result by weakening Coincidence Plenitude to the claim that every *instantiated* undiscriminating and unrepeatable property is tracked by some object. This restriction does not run afoul in any obvious way of our general case for plenitude, and might seem especially tempting in a setting where the domain of properties includes uninstantiated would-be haecceities. But in fact the argument above still goes through with the restricted version of Coincidence Plenitude. The restricted principle implies that there is at least one "worldbound" object, Picky, which is actually concrete and couldn't have been concrete without every actually true proposition being true. So, instead of F^* in the above argument, consider the property F^{**} of *either coinciding with Picky or with something whose haecceity is F*. F^{**} is not only undiscriminating and unrepeatable but instantiated, and so even the weaker version of Coincidence Plenitude serves up an object x that tracks it. But necessarily, if there were a concrete object y not in U with F as its haecceity, F would be necessarily instantiated (by y), so Picky would be necessarily non-concrete (since necessarily Picky is concrete only if F is uninstantiated), so F^{**} would be necessarily equivalent to *coinciding with y*, so y would be tightly connected to x, and thus by No Tight Connection y would be identical to x, and hence y would be in U.

No Tight Connection plays an important role in both of the above arguments. What happens if we give it up? Going back to full Coincidence Plenitude, we still get the conclusion that for any property F that could be the haecceity of something not in U, there is an object x (in U) such that necessarily, any object with F as its haecceity is tightly connected to x. Without No Tight Connection, it doesn't follow that x *itself* could have F as its haecceity. But the weaker conclusion already comes very close to ruling out the possibility of an alien-matter universe. In particular, it rules out the possibility of an alien-matter universe containing a new microphysical object y which instantiates some undiscriminating qualitative property Q, but is such that it is indiscernibly metaphysically possible for it to lack Q. For given our definition of "alien-matter universe", if it is indiscernibly metaphysically possible for y to lack Q, it must also be U-indiscernibly metaphysically possible. But it isn't U-indiscernibly metaphysically possible, since any new microphysical object y will have a "shadow" x in U, such that necessarily x is Q iff y is Q. (One can derive a similar result using the restricted version of Coincidence Plenitude

by judicious reliance on Picky.) In principle one could reconcile this conclusion with an affirmation of Alien Matter by insisting on bizarre constraints on the modal profiles of alien matter. But since Alien Compatibility is only tenable on the assumption that it could be indiscernibly possible for alien atoms that are parts of tables to swap places with alien atoms that aren't, we will in any case block the combination of Alien Matter and Alien Compatibility, thus again undermining the argument that tablehood is not about *U*.[23]

Given our sympathies towards plenitude, we are thus disinclined to the option of denying that tablehood is about *U* as part of a package that accepts higher-order BF while rejecting first-order BF. Of course, the thesis that tablehood is not about *U* does not logically require a commitment to higher-order BF, but we don't see a promising way to develop that thesis in the absence of such a commitment.

Let's turn to the second option, that of denying Alien Matter. While we have just seen some unexpected ways in which this option might be combined with the denial of first-order BF, the most natural way to develop the option involves a full embrace of BF, and thus of the thesis that necessarily every object is in *U*. This view could be secured by strengthening our basic modal logic for metaphysical necessity to H_{S5}, a package which we find pretty appealing.[24] However, while adopting H_{S5} will certainly undercut the case for Possible *U*-indiscernible Moderation, it may not set the underlying worry fully to rest. Many necessitists are comfortable with a use of the word 'exist', not equivalent to 'be identical to something', where we get to say things like 'Only a few of the possible knives that could have been made of these blades and handles exist' and the like. Following Williamson, we have been saying 'be concrete' instead of using 'exist' in this way, to avoid the risk that someone might think we meant 'exist' in the sense of 'be identical to something'. But whereas the contrast between the concrete and the non-concrete only makes sense for *objects*, one might naturally think that the contrast between the existent and the nonexistent also makes sense in other types. The nonexistent properties might, for example, include the haecceities of the nonexistent possible children of Wittgenstein—or at least, of nonexistent possible particles and spacetime points.[25]

[23] Recall that 'indiscernibly hypertolerantly a table' is defined in terms of *modal match* (see §14.2): thus in an non-symmetric alien world where it was indiscernibly impossible for atoms to vary at all in their qualitative undiscriminating properties, and where no atom coincides with any other, no collection of atoms would modally match any other, and thus any table would vacuously count as indiscernibly hypertolerant.

[24] One can also consistently accept BF without B or ND: see Bacon 2018*a*.

[25] Although Williamson is disciplined in avoiding 'exist', he seems to feel the pull to have a notion that works like this. In Williamson 2013*b* (314) he introduces the notion of 'chunkiness', suggesting that it might be cashed out as 'being grounded in the concrete'. This sounds like it might make sense in higher types, and exclude at least some uninstantiated haecceities. Another group of theorists who should be in a position to make sense of a discriminating notion of existence are those contingentists who take themselves to be in a position to "simulate" necessitism by stipulatively introducing new meanings for the quantifiers that make necessitism come out true. In §1.4 we invited those in this position to apply their translation-scheme to everything we say in this book. We hope they have been remembering to do so! But since they also have the original meanings of the quantifiers hanging around, they will be able

If one can make sense of a notion of existence that works like this, it would be very natural to suppose that every *existent* property (or entity of any type) is about the collection *E* of all existent objects. If we can also convince ourselves that tablehood is existent, it will follow that tablehood is about *E*, and hence that Possible *E*-indiscernible Moderation is false.[26] We will then need to find some way of resisting the analogue for *E* of the above argument for Possible *U*-moderation, with 'alien-matter universe' redefined using *E* rather than *U*. And even if one had made one's peace with the denial that there could be an alien-matter universe in the old sense, it is quite natural to think that this remains possible in the new sense.

However, given Coincidence Plenitude and No Tight Connection (see above) and the assumption that no object being concrete without existing, we can show that necessarily every object is (intensionally) about *E*. Consider some concrete object Picky that couldn't be concrete if anything were otherwise. Then any non-concrete object *x* is about the collection comprising Picky together with an object *y* that necessarily coincides with whichever of *x* and Picky is concrete: *being identical to x* is necessarily equivalent to *being something such that necessarily, one coincides with it if and only if one coincides with y and not with Picky*. And even without No Tight Connection, we can still conclude that necessarily every object *x* has all of its undiscriminating qualitative properties with *E*-indiscernible necessity, which would undermine the plausibility of the *E*-analogue of Alien Matter. It is thus far from clear that in the setting of H_{S5} with plenitude one can generate a compelling existence-theoretic variant of our puzzle.[27] And indeed, the combination of H_{S5} and plenitude is one of our favourite responses to the puzzle.

Finally, we come to our third option, where we accept both that tablehood is about *U* and that there could be an alien-matter universe, but deny Alien Compatibility. This response to the argument will be plain sailing for those who have already convinced themselves that *in the actual world* either some tables are

to use them to introduce a discriminating notion of existence: to exist is to be identical to *something*, in the original sense of 'something'. And this works in every type: if one previously accepted higher-order contingentism (understood as a claim using the original quantifier-meanings), simulating necessitism will require introducing new meanings for higher-order as well as first-order quantifiers.

[26] How might one argue that tablehood is existent? It's not obvious. The mere fact that tablehood is *instantiated* is not enough: as we are imagining the view being developed, the negation of the haecceity of any nonexistent object will be an instantiated but nonexistent property. And the fact that we manage to *refer* to tablehood is not obviously decisive, since there are examples in the literature that suggest that reference to nonexistent objects (e.g. 'Noman' from Salmon 1987) is possible. Arguments that such reference is impossible because the referents can't be "singled out in thought" tend, among other things, to neglect the phenomenon of vagueness: see also Bacon 2013.

[27] On the other hand, with the contrast between existent and nonexistent properties at one's disposal, there are fairly natural ways of weakening Coincidence Plenitude while still preserving its general spirit (and its capacity to resolve Tolerance Puzzles). For example, one might try restricting the quantifier to properties that are not only undiscriminating and unrepeatable, but *could not be instantiated without being existent*. This weakened version of Coincidence Plenitude does not automatically generate concrete "shadow" objects that track properties defined by reference to nonexistent objects, and might perhaps allow for a revival of the existence-theoretic version of the puzzle. We will not explore the matter further here.

not indiscernibly tolerantly tables or all tables are indiscernibly hypertolerantly tables (perhaps on the grounds that tablehood is qualitative). But for those like us who accept ∅-moderation (and thus give up the qualitativeness of tablehood), there is more of a challenge: if ∅-moderation is in fact true, then prima facie one would have thought it could still be true in an alien-matter universe. Isn't it odd to think that we just got lucky to get the actual atoms, which allow for the construction of moderately tolerant tables, rather than some alien collection of atoms which would not allowed for this?

But anyone who has been sympathetic with our overall strategy for solving Tolerance Puzzles will have learned, by now, not to be too shocked by the idea that the actual world is in certain ways special as regards the modal behaviour of tables. The hallmark of our view is that certain universal generalizations about the modal behaviour of things like tables which one might offhand take to be metaphysically necessary are in fact only contingently true, and even rather fragile. But we are not just lucky to speak the truth when we endorse these generalizations, since the relevant *sentences* are stably such as to express truths. In this connection we focused on Tolerance generalizations, but we also extended the same treatment to Robustness generalizations, and suggested that in some contexts, Non-coincidence generalizations have the same status (§12.2). Against this background, it wouldn't be especially costly to think, for example, that any table made out of entirely new atoms would be indiscernibly intolerant, although we would under such circumstances have expressed a different property by 'table', so that sentences like 'Every table is indiscernibly tolerant' would still have been true in our mouths.

There are at least two natural ways of developing the strategy of denying Alien Compatibility. The first holds that tables in alien-matter universes would be surprisingly intolerant; the second holds that such tables would be surprisingly hypertolerant.[28] To be more specific, we can imagine the proponent of the first approach as holding on to the necessary truth of Overlap Essentialism, thus ruling out the existence of tables that are indiscernibly tolerantly tables in any alien-matter universe with enough matter that anything that was indiscernibly

[28] These are not the only ways of developing the strategy. One could instead think that the way tablehood works in an alien-matter universe is something like the way Yablo thinks tablehood actually works, so that its instances are indiscernibly tolerant as far as their concrete existence is concerned, but are not indiscernibly tolerantly *tables*. We are not particularly more averse to this approach than to those discussed in the main text.

One could alternatively think that that there is rampant coincidence of tables in alien-matter worlds, so that there would be (e.g.) a mixture of indiscernibly intolerant and indiscernibly tolerant tables, with each of the former coinciding with one of the latter. This way of developing the view strikes us as less promising: it would be quite strange to think that there couldn't be an alien-matter universe containing exactly one table.

Least plausibly of all, one could deny that there could be any tables at all in an alien-matter universe (while still allowing that there could be such a universe with matter arranged tablewise). As far as we can see this has nothing going for it.

hypertolerant would be a counterexample to Overlap Essentialism. Meanwhile, the proponent of the second approach will think that counterexamples to Overlap Essentialism would be pervasive in any alien-matter universe where the new matter is sufficiently abundant.

The second approach (of embracing hypertolerance for tables in alien-matter universes) is untenable in full generality. As we saw in §14.4, Microphysical Supervenience plus the qualitativeness of tablehood imposes some limits on the kinds of indiscernible tolerance claims that could be true in *symmetric* universes, since Microphysical Supervenience rules out saying that one of two qualitatively indiscernible tables in a symmetric universe could have been originally composed of the atoms that in fact composed the other table without any change in the qualitative relations among microphysical objects. The claim that tablehood is about *U* imposes corresponding limits when it comes to symmetric alien-matter universes where Microphysical Supervenience is true, since the symmetries of those universes will plausibly preserve not only all qualitative properties and relations, but all properties and relations that are about *U*. Thus, as regards such possibilities, we have to make our peace with the view that tables would be surprisingly indiscernibly intolerant—e.g., that the premise Swap Tolerance from §14.4 would be false.

This falls short of establishing that tables in alien-matter universes would in general be indiscernibly intolerant. One could make a special exception for symmetric universes, maintaining that tables in non-symmetric alien-matter universes would still be hypertolerant. Such special pleading, while a little inelegant, does not strike us as a decisive cost.

The intolerant and hypertolerant ways of fleshing out the denial of Alien Compatibility both take some getting used to. Each has interesting features, but none of them seem to have the makings of a decisive argument for one over the other. If no such arguments are forthcoming, it would be unsurprising if the right thing to say (on the assumption that alien-matter worlds are possible) was that the predicate 'table' is vague, with different precisifications corresponding to different views about what modal surprises are in store at alien-matter worlds. Our general picture is that such vagueness will arise by default where competing desiderata leave reference unsettled, and ordinary practice gives no clear priority to some of the desiderata over others.

Both views involve drawing some unexpected contrast between the modal properties that tables in alien-matter universes would have and the modal properties tables actually have. Someone might find such contrasts so unsettling as to be inclined to draw some sweeping metaphysical conclusion not involving the word 'table': e.g. that there couldn't after all have been new objects; that not all properties are about *U*; that Microphysical Supervenience is false (Goodman unpublished *e*); or even that Iteration fails for metaphysical necessity. While we are open to some of these conclusions, e.g., that there couldn't have been new objects, we place little

store in arguments based on the repugnance of the relevant contrasts. After all, we have already had to learn to live with the fact that ordinary Tolerance premises are highly contingent, failing even in some quite nearby possible circumstances. Once this contrast between the actual world and others has been absorbed, the special new kinds of contrast that might be required if there could be an alien-matter universe should not seem like such a bitter pill to swallow.

This concludes our treatment of Tolerance arguments based on indiscernible modality. Our detailed discussion has not been motivated by the worry that there is a deep problem for our view in the vicinity, but rather by a desire to show how the related ideologies of qualitativeness and aboutness can be used to raise important, difficult, and in some cases unfamiliar, metaphysical questions. As with the more familiar Tolerance arguments treated in earlier chapters, the arguments relating to qualitativeness and aboutness turn out to open up a rich and intricate landscape of views worthy of serious engagement. Whether or not readers embrace the solutions we have argued for—based on plenitude, plasticity, and the non-qualitativeness of properties like tablehood—we hope that the logical and conceptual tools we have provided will help them to navigate that landscape for themselves.

APPENDIX A

Modal Logics with and without Necessitation

The results of this appendix concern the modal logic H_{KA} presented in §1.4, which extends H_0 with the following three new axioms (closing under MP and GEN):

$\Box^*$**N**	$\Box^*\forall\vec{v}\Box P$, where $\vdash_{\mathsf{H}_0} P$.
$\Box^*$**K**	$\Box^*\forall p\forall q(\Box(p \to q) \to (\Box p \to \Box q))$.
$\Box^*$**Duality**	$\Box^*\forall p(\Diamond\neg p \leftrightarrow \Box\neg p)$.

Here $\vec{v}$ stands for any zero or more variables. $\Box^*$ is the ancestral necessity operator corresponding to $\Box$, defined as *having every finite iteration of* $\Box$:

$$\Box^* := \lambda p.\forall X(\mathbf{F}_\Box X \to Xp)$$
$$\mathbf{F}_\Box := \lambda Y.\forall Z(Z(\lambda p.p) \to \forall X(ZX \to Z(\lambda p.X\Box p)) \to ZY)$$

The first result is that H_{KA} contains the logic H_{K}, defined to be the smallest extension of H_0 that contains K and Duality and is closed under MP, GEN, and the necessitation rule NEC. The nontrivial task here is to show that the work of NEC in generating theorems in H_{K} is fully subsumed by the claims of ancestral necessity characteristic of H_{KA}. The second cluster of results concerns certain stronger logics H_{K4}, H_{S4}, and H_{S5} derived from H_{K} by adding certain additional axioms and closing under MP, GEN, and NEC; we show that these can also be characterized as the results of adding certain similar axioms to H_{KA} and closing under MP and GEN. This means that H_{KA} itself lies between H_{K} and H_{K4}.

We will begin by collecting some consequences in H_0 of the above definitions of $\mathbf{F}_\Box$ and $\Box^*$.

Lemma A1. (a) $\vdash_{\mathsf{H}_0} \mathbf{F}_\Box(\lambda p.p)$
(b) $\vdash_{\mathsf{H}_0} \mathbf{F}_\Box X \to \mathbf{F}_\Box(\lambda p.X\Box p)$
(c) $\vdash_{\mathsf{H}_0} \Box^* P \to P$
(d) $\vdash_{\mathsf{H}_0} \Box^* P \to \Box^*\Box P$
(e) $\vdash_{\mathsf{H}_0} \Box^* P \to \Box^*\Box^n P$ for all n
(f) $\vdash_{\mathsf{H}_0} \Box^* P \to \Box^n P$ for all n

Proof: (a) and (b) are immediate from the definition of $\mathbf{F}_\Box$. (c) follows from (a) by UI and Eβ. (d) follows from (b) via $\Box^* P \to (\mathbf{F}_\Box X \to (\lambda p.X\Box p)P)$. (e) follows from (d) by metalinguistic induction on n, and (f) follows from (c) and (e). □

The next lemma will be our standard resource for proving generalizations $\forall X(\mathbf{F}_\Box X \to P)$ about the finite iterations of necessity.

Lemma A2 (Induction)**.**

$$\vdash_{\mathsf{H}_0} P[(\lambda p.p)/X] \to \forall X(\mathbf{F}_\Box X \to P \to P[(\lambda p.X\Box p)/X]) \to \forall X(\mathbf{F}_\Box X \to P)$$

Proof: Let $Z \coloneqq \lambda X.\mathbf{F}_\Box X \wedge P$. Then $\vdash_{\mathsf{H}_0} P[(\lambda p.p)/X] \rightarrow Z(\lambda p.p)$ by Lemma A1a and Eβ. Also $\vdash_{\mathsf{H}_0} \forall X(\mathbf{F}_\Box X \rightarrow P \rightarrow P[(\lambda p.X\Box p)/X]) \rightarrow \forall X(ZX \rightarrow Z(\lambda p.X\Box p))$ by Lemma A1b, Eβ, UI, GEN, and UD. Given the definition of $\mathbf{F}_\Box$, the result follows by Eβ, UI, GEN, and UD. □

Turning now to H_{KA}, we begin with something quite unsurprising. For every theorem P of H_0, $\Box^*\Box P$ is an instance of $\Box^*$N; but we cannot immediately infer $\Box^* P$ from this (since we do not have the converse of Lemma A1d). But it is still a theorem:

Lemma A3. $\vdash_{\mathsf{H}_{\mathsf{KA}}} \Box^* P$ whenever $\vdash_{\mathsf{H}_0} P$.

Proof: By Induction, it suffices to show that whenever $\vdash_{\mathsf{H}_0} P$, (i) $\vdash_{\mathsf{H}_{\mathsf{KA}}} (\lambda p.p)P$, which is obvious by Eβ; and (ii)$\vdash_{\mathsf{H}_{\mathsf{KA}}} \mathbf{F}_\Box X \rightarrow XP \rightarrow (\lambda p.X\Box p)P$, which follows by Eβ from the fact that $\vdash_{\mathsf{H}_{\mathsf{KA}}} \mathbf{F}_\Box X \rightarrow X\Box P$, which is true since $\Box^*\Box P$ is an instance of $\Box^*$N. □

(Lemma A3 does a lot of the work of the more complicated $\Box^*$N, and one might have hoped to recover the latter from the former. After all, Lemma A3 already gives us $\Box^*\forall\vec{v}P$ and $\forall\vec{v}\Box^* P$ when $\vdash_{\mathsf{H}_0} P$, and given Lemma A1d, these imply $\Box^*\Box\forall\vec{v}P$ and $\forall\vec{v}\Box^*\Box P$. But there is no way with just Lemma A3 to get the quantifier into the intermediate position it occupies in $\Box^*$N.)

Lemma A4 (Distribution)**.** For any P and Q, $\vdash_{\mathsf{H}_{\mathsf{KA}}} \mathbf{F}_\Box X \rightarrow X(P \rightarrow Q) \rightarrow XP \rightarrow XQ$ and $\vdash_{\mathsf{H}_{\mathsf{KA}}} \Box^*(P \rightarrow Q) \rightarrow \Box^* P \rightarrow \Box^* Q$

Proof: For any operator O, let $K(O)$ be the open formula $O(p \rightarrow q) \rightarrow Op \rightarrow Oq$. We will prove that $\vdash_{\mathsf{H}_{\mathsf{KA}}} \mathbf{F}_\Box X \rightarrow K(X)$; the first part of the result follows by instantiating p and q to P and Q, and the second part is an immediate consequence of the first part. By Induction, it suffices to show

(i) $\vdash_{\mathsf{H}_{\mathsf{KA}}} \forall p\forall q K(\lambda p.p)$

(ii) $\vdash_{\mathsf{H}_{\mathsf{KA}}} \mathbf{F}_\Box X \rightarrow \forall p\forall q K(X) \rightarrow \forall p\forall q K(\lambda p.X\Box p)$

(i) is straightforward since $\forall p\forall q K(\lambda p.p)$ is a theorem of H_0 by UI, Eβ, and GEN. To derive (ii), we use the following three H_0-theorems (each of which is derived from a UI-instance by Eβ):

$$\vdash_{\mathsf{H}_0} \forall p\forall q K(X) \rightarrow X(\forall p\forall q K(\Box) \rightarrow K(\Box)) \rightarrow X\forall p\forall q K(\Box) \rightarrow XK(\Box)$$
$$\vdash_{\mathsf{H}_0} \forall p\forall q K(X) \rightarrow XK(\Box) \rightarrow X\Box(p \rightarrow q) \rightarrow X(\Box p \rightarrow \Box q)$$
$$\vdash_{\mathsf{H}_0} \forall p\forall q K(X) \rightarrow X(\Box p \rightarrow \Box q) \rightarrow X\Box p \rightarrow X\Box q$$

Combining these three yields

$$\vdash_{\mathsf{H}_0} \forall p\forall q K(X) \rightarrow X(\forall p\forall q K(\Box) \rightarrow K(\Box)) \rightarrow X\forall p\forall q K(\Box) \rightarrow X\Box(p \rightarrow q) \rightarrow X\Box p \rightarrow X\Box q$$

by PC. But since $\vdash_{\mathsf{H}_0} \forall p\forall q K(\Box) \rightarrow K(\Box)$, by Lemma A3 we have $\vdash_{\mathsf{H}_{\mathsf{KA}}} \mathbf{F}_\Box X \rightarrow X(\forall p\forall q K(\Box) \rightarrow K(\Box))$. And by $\Box^*$K, $\vdash_{\mathsf{H}_{\mathsf{KA}}} \mathbf{F}_\Box X \rightarrow X\forall p\forall q K(\Box)$. Thus in H_{KA} we can simplify the previous conditional to

$$\vdash_{\mathsf{H}_{\mathsf{KA}}} \mathbf{F}_\Box X \rightarrow \forall p\forall q K(X) \rightarrow X\Box(p \rightarrow q) \rightarrow X\Box p \rightarrow X\Box q$$

By Eβ this implies $\vdash_{\mathsf{H}_{\mathsf{KA}}} \mathbf{F}_\Box X \rightarrow \forall p\forall q K(X) \rightarrow K(\lambda p.X\Box p)$, from which an application of GEN and UD yields (ii). □

These two lemmas give us the closure of each finite iteration of $\Box$, and thus of $\Box^*$, under arguments valid in H_0:

Lemma A5 (Closure)**.** If $\vdash_{\mathsf{H}_0} P_1 \to \cdots \to P_n \to Q$, then $\vdash_{\mathsf{H}_{\mathsf{KA}}} \mathbf{F}_\Box X \to XP_1 \to \cdots \to XP_n \to XQ$ and hence $\vdash_{\mathsf{H}_{\mathsf{KA}}} \Box^* P_1 \to \cdots \to \Box^* P_n \to \Box^* Q$.

Proof: By metalinguistic induction on n, using Lemma A3 for the base clause and Distribution for the induction step. □

The last preliminary results we will need—and the key place we will need the extra strength of $\Box^*$N over Lemma A3—are in the vicinity of CBF:

Lemma A6. (a) $\vdash_{\mathsf{H}_{\mathsf{KA}}} \mathbf{F}_\Box X \to X\Box\forall\vec{\nu}P \to X\forall\vec{\nu}\Box P$

(b) $\vdash_{\mathsf{H}_{\mathsf{KA}}} \Box^*\Box\forall\vec{\nu}P \to \Box^*\forall\vec{\nu}\Box P$

(c) $\vdash_{\mathsf{H}_{\mathsf{KA}}} \Box^*\forall\vec{\nu}P \to \Box^*\forall\vec{\nu}\Box P$

Proof: For (a), note first that by UI,

$$\vdash_{\mathsf{H}_0} \forall p\forall q(\Box(p \to q) \to \Box p \to \Box q) \to \Box(\forall\vec{\nu}P \to P) \to \Box\forall\vec{\nu}P \to \Box P$$

and hence by GEN and UD,

$$\vdash_{\mathsf{H}_0} \forall p\forall q(\Box(p \to q) \to \Box p \to \Box q) \to \forall\vec{\nu}\Box(\forall\vec{\nu}P \to P) \to \Box\forall\vec{\nu}P \to \forall\vec{\nu}\Box P$$

By Closure, this gives us

$$\vdash_{\mathsf{H}_{\mathsf{KA}}} \mathbf{F}_\Box X \to X\forall p\forall q(\Box(p \to q) \to \Box p \to \Box q) \to X\forall\vec{\nu}\Box(\forall\vec{\nu}P \to P) \to X\Box\forall\vec{\nu}P \to X\forall\vec{\nu}\Box P$$

But since we have $\vdash_{\mathsf{H}_{\mathsf{KA}}} \mathbf{F}_\Box X \to X\forall p\forall q(\Box(p \to q) \to \Box p \to \Box q)$ by $\Box^*$K and $\vdash_{\mathsf{H}_{\mathsf{KA}}} \mathbf{F}_\Box X \to X\forall\vec{\nu}\Box(\forall\vec{\nu}P \to P)$ by $\Box^*$N, this simplifies to $\vdash_{\mathsf{H}_{\mathsf{KA}}} \mathbf{F}_\Box X \to X\Box\forall\vec{\nu}P \to X\forall\vec{\nu}\Box P$.

Part (b) follows immediately from part (a), and part (c) follows from (b) together with Lemma A1c. □

We now have all we need to establish our first main result:

Proposition A7. $\mathsf{H}_\mathsf{K} \subseteq \mathsf{H}_{\mathsf{KA}}$.

Proof: Let T be the set of all formulae P such that $\vdash_{\mathsf{H}_{\mathsf{KA}}} \Box^*\forall\vec{\nu}P$ for all $\vec{\nu}$. By construction, T is closed under GEN, and by Lemma A6c, it is closed under NEC. Since $\forall\vec{\nu}P$ is a theorem of H_0 whenever P is, T includes H_0 by Lemma A3. By Closure and the fact that $\vdash_{\mathsf{H}_0} \forall\vec{\nu}(P \to Q) \to \forall\vec{\nu}P \to \forall\vec{\nu}Q$, T is closed under MP. Finally, since $\forall p\forall q(\Box(p \to q) \to \Box p \to \Box q)$ belongs to T by $\Box^*$K and $\vdash_{\mathsf{H}_0} \forall p\forall q(\Box(p \to q) \to \Box p \to \Box q) \to \forall\vec{\nu}(\Box(p \to q) \to \Box p \to \Box q)$, $\Box(p \to q) \to \Box p \to \Box q$ belongs to T by Closure.

Given that H_K is defined to be the smallest extension of H_0 containing $\Box(p \to q) \to \Box p \to \Box q$ and closed under MP, GEN, and NEC, it follows that $\mathsf{H}_\mathsf{K} \subseteq T$. But since $\vdash_{\mathsf{H}_{\mathsf{KA}}} \Box^*\forall\vec{\nu}P \to \forall\vec{\nu}P$ by Lemma A1c and $\vdash_{\mathsf{H}_{\mathsf{KA}}} \forall\vec{\nu}P \to P$ by UI, $T \subseteq \mathsf{H}_{\mathsf{KA}}$. Hence $\mathsf{H}_\mathsf{K} \subseteq \mathsf{H}_{\mathsf{KA}}$. □

This result carries over to various extensions of H_{KA}. Where A is any formula or set of formulae, let $\mathsf{H}_\mathsf{K} + A +$ NEC be the smallest extension of H_K containing A and closed under MP, GEN and NEC, and let $\mathsf{H}_{\mathsf{KA}} + A$ be the smallest extension of H_{KA} containing A and closed under MP and GEN.

Proposition A8. If A is a set of closed sentences and $\Box^*P$ is in B whenever P is in A, then $\mathsf{H_K} + A + \mathrm{NEC} \subseteq \mathsf{H_{KA}} + B$.

Proof: As in the proof of Proposition A7, but with $\mathsf{H_{KA}} + B$ instead of $\mathsf{H_{KA}}$. We need to show one other thing, namely that T includes every sentence P in A; since $\Box^*P$ is in B for each such P, and $\vdash_{\mathsf{H}_0} P \rightarrow \forall\vec{\nu}P$, $P \in T$ follows by Closure. □

Our second set of results involves certain extensions of $\mathsf{H_{KA}}$ and $\mathsf{H_K}$ that include the 4 axiom, which we take as a closed sentence:

4 $\qquad \forall p(\Box p \rightarrow \Box\Box p)$

We will show that in the presence of 4 and $\Box 4$, the extra strength that $\mathsf{H_{KA}}$ adds over $\mathsf{H_K}$ disappears. Here is the key fact:

Lemma A9. $\vdash_{\mathsf{H}_0} 4 \rightarrow \forall p((p \wedge \Box p) \leftrightarrow \Box^* p)$

Proof: By Lemmas A1c and A1f, we have $\vdash_{\mathsf{H}_0} \Box^* p \rightarrow p$ and $\vdash_{\mathsf{H}_0} \Box^* p \rightarrow \Box p$, so we need only prove the version where $\leftrightarrow$ is replaced with $\rightarrow$. This follows in turn from the claim that $\vdash_{\mathsf{H}_0} 4 \rightarrow \forall X(\mathbf{F}_\Box X \rightarrow (MX \vee NX))$, where we define $MX :\!= \forall q(q \rightarrow Xq)$ and $NX :\!= \forall q(\Box q \rightarrow Xq)$. By Induction, we can establish this by showing that

(i) $\quad \vdash_{\mathsf{H}_0} 4 \rightarrow (M(\lambda p.p) \vee N(\lambda p.p))$, and

(ii) $\quad \vdash_{\mathsf{H}_0} 4 \rightarrow \mathbf{F}_\Box X \rightarrow (MX \vee NX) \rightarrow (M(\lambda p.X\Box p) \vee N(\lambda p.X\Box p))$

(i) is trivial since Eβ guarantees $\vdash_{\mathsf{H}_0} N(\lambda p.p)$. (ii) follows from two simpler facts: first, $\vdash_{\mathsf{H}_0} MX \rightarrow N(\lambda p.X\Box p)$, which is true since $\vdash_{\mathsf{H}_0} MX \rightarrow \Box q \rightarrow X\Box q$, and second, $\vdash_{\mathsf{H}_0} 4 \rightarrow NX \rightarrow N(\lambda p.X\Box p)$, which is also true since by UI $\vdash_{\mathsf{H}_0} 4 \rightarrow NX \rightarrow \Box q \rightarrow X\Box q$ and so $\vdash_{\mathsf{H}_0} 4 \rightarrow NX \rightarrow N(\lambda xp.X\Box p)$. □

We can use this to simplify the axiomatization of theories that include 4 as well as $\mathsf{H_{KA}}$. Let $\mathsf{H_{K-}}$ be the very weak fragment of $\mathsf{H_K}$ axiomatized by adding: K, $\Box$K, Duality, $\Box$Duality, and all instances of the following two schemas to H_0 (and closing under MP and GEN):

N $\quad \Box P$ whenever P is a theorem of H_0

□N $\quad \Box\forall\vec{\nu}\Box P$ whenever P is a theorem of H_0

Adding the 4 axiom to this weak theory (and closing under MP and GEN) gives a theory that includes $\mathsf{H_{KA}}$:

Lemma A10. $\mathsf{H_{KA}} \subseteq \mathsf{H_{K-}} + 4$

Proof: By Lemma A9, every instance $\Box^*\forall\vec{\nu}\Box P$ of $\Box^*$N follows in H_0 from 4, the $\Box$N-instance $\Box\forall\vec{\nu}\Box P$, and $\forall\vec{\nu}\Box P$ which follows by GEN from the N-instance $\Box P$. Likewise, $\Box^*$K follows from 4, K, and $\Box K$, and $\Box^*$Duality follows from 4, Duality, and $\Box$Duality. □

Given these facts, we can readily give NEC-free characterizations of $\mathsf{H_{K4}}$ and a wide range of stronger theories. For any set of closed sentences A, let $\Box A$ be $\{\Box P : P \in A\}$ and $\Box^* A$ be $\{\Box^* P : P \in A\}$.

Proposition A11. If $4 \in A$, then $\mathsf{H_K} + A + \text{NEC} = \mathsf{H_{KA}} + \Box^* A = \mathsf{H_{KA}} + A + \Box A = \mathsf{H_{K-}} + A + \Box A = \mathsf{H_K} + A + \Box A$.

Proof: (i) $\mathsf{H_K} + A + \text{NEC} \subseteq \mathsf{H_{KA}} + \Box^* A$ by Proposition A8.

(ii) $\mathsf{H_{KA}} + \Box^* A \subseteq \mathsf{H_{KA}} + A + \Box A$ since $\vdash_{\mathsf{H_0}} 4 \to (P \wedge \Box P) \to \Box^* P$, by Lemma A9.

(iii) $\mathsf{H_{KA}} + A + \Box A \subseteq \mathsf{H_{K-}} + A + \Box A$ since $\mathsf{H_{KA}} \subseteq \mathsf{H_{K-}} + 4$ by Lemma A10.

(iv) $\mathsf{H_{K-}} + A + \Box A \subseteq \mathsf{H_K} + A + \Box A$ since every instance of N, $\Box$N, $\Box$K, and $\Box$Duality follows from K and Duality by NEC and GEN.

(v) $\mathsf{H_K} + A + \Box A \subseteq \mathsf{H_K} + A + \text{NEC}$ since each member of $\Box A$ follows by NEC from a member of A. □

Taking $A = \{4\}$, this gives us several axiomatizations of the theory $\mathsf{H_{K4}}$, which we define as $\mathsf{H_K} + 4 + \text{NEC}$ (i.e. the smallest set of formulae including $\mathsf{H_0}$, K, Duality, and 4, and closed under NEC as well as MP and GEN):

Corollary A12. $\mathsf{H_{K4}} = \mathsf{H_{KA}} + \Box^* 4 = \mathsf{H_{KA}} + 4 + \Box 4 = \mathsf{H_{K-}} + 4 + \Box 4 = \mathsf{H_K} + 4 + \Box 4$.

Two other systems that play an important role in some of our discussions are $\mathsf{H_{S4}} := \mathsf{H_K} + 4{+}\text{T}{+}\text{ NEC}$ and $\mathsf{H_{S5}} := \mathsf{H_K} + 4{+}\text{T}{+}\text{B}{+}\text{ NEC}$, where

T	$\forall p(\Box p \to p)$
B	$\forall p(p \to \Box\Diamond p)$

We can use Proposition A11 to give NEC-free characterizations of these theories, using the T axiom to omit some redundancies.

Corollary A13. $\mathsf{H_{S4}} = \mathsf{H_{KA}} + \Box^* 4 + \Box^* \text{T} = \mathsf{H_{KA}} + \Box 4 + \text{T} + \Box\,\text{T} = \mathsf{H_{K-}} + \Box 4 + \text{T} + \Box\,\text{T} = \mathsf{H_K} + \Box 4 + \text{T} + \Box\,\text{T}$.

Corollary A14. $\mathsf{H_{S5}} = \mathsf{H_{KA}} + \Box^* 4 + \Box^* \text{T} + \Box^* \text{B} = \mathsf{H_{KA}} + \Box 4 + \text{T} + \Box\,\text{T} + \Box\,\text{B} = \mathsf{H_{K-}} + \Box 4 + \text{T} + \Box\,\text{T} + \Box\,\text{B} = \mathsf{H_K} + \Box 4 + \text{T} + \Box\,\text{T} + \Box\,\text{B}$.

Finally, it's worth noting that these theories are closed not only under NEC but under "NEC*", i.e. the rule $P/\Box^* P$.

Proposition A15. If a theory T contains $\mathsf{H_0}$ and 4 and is closed under MP, GEN, and NEC, then it is also closed under NEC*, i.e. the rule $P/\Box^* P$.

Proof: Suppose $\vdash_T P$; then $\vdash_T \Box P$ by NEC and $\vdash_T \Box P \to \Box^* P$ by Lemma A9, so $\vdash_T \Box^* P$ by MP. □

APPENDIX B

Rigidity and Ancestral Iteration

In the absence of Iteration, the logic $\mathsf{H_{KA}}$ discussed in §1.4 and Appendix A is interestingly stronger than $\mathsf{H_K}$. But it is still weaker than one might wish, in some striking ways. For example, although for any tautology T we can prove $\Box^*\mathsf{T}$, as far as we can tell, there is no way to prove $\Box^*\Box^*\mathsf{T}$, or even $\Box\Box^*\mathsf{T}$.[1] These claims seem very plausible, quite independently of the status of Iteration (for $\Box$). In response, one might explore a whole hierarchy of possible enrichments of the axioms of $\mathsf{H_{KA}}$ that would achieve some desirable extra strength. But for us no such exploration is needed, since in the context of the rigidity axioms we introduced and motivated in §1.5, the situation becomes much simpler. In the logic $\mathsf{H_{KR}}$ that adds the rigidity axioms to $\mathsf{H_{KA}}$, we can in fact prove that $\Box^*$ conforms to $\mathsf{H_{S4R}}$, the result of adding the rigidity axioms to the theory $\mathsf{H_{S4}}$ discussed in Appendix A. It follows from this that $\mathsf{H_{KR}}$ is closed both under NEC and under the "NEC*" rule, $P/\Box^*P$.

The fact that the rigidity axioms can play this role should not be so surprising. In §7.3 we gave an informal argument that Iteration holds for $\Box^*$ based on thinking of $\Box^*$ as an infinite conjunction of all of the finite iterations of $\Box$. But the intuition behind that argument depends essentially on applying ordinary modes of reasoning about *plurals* to the finite iterations of $\Box$, in the way that the rigidity axioms are designed to license. Without rigidity (or Iteration for $\Box$), we would have no guarantee that there couldn't be new finite iterations of $\Box$, in addition to all the ones there in fact are.

We will start with three results which do not require the rigidity axioms. The first says that the property of *being a finite iteration of necessity* is ancestrally necessary to everything that has it:

Lemma B1. $\vdash_{\mathsf{H_{KA}}} \mathbf{F}_\Box X \rightarrow \Box^*\mathbf{F}_\Box X$

Proof: By Induction, it suffices to show that (i) $\vdash_{\mathsf{H_{KA}}} \Box^*\mathbf{F}_\Box(\lambda p.p)$, and (ii) $\vdash_{\mathsf{H_{KA}}} \mathbf{F}_\Box X \rightarrow \Box^*\mathbf{F}_\Box X \rightarrow \Box^*\mathbf{F}_\Box(\lambda p.X\Box p)$. For (i), note that $\vdash_{\mathsf{H_0}} \mathbf{F}_\Box(\lambda p.p)$ by Lemma A1a, so $\Box^*\mathbf{F}_\Box(\lambda p.p)$ is an instance of $\Box^*$N. For (ii), note that $\vdash_{\mathsf{H_0}} \mathbf{F}_\Box X \rightarrow \mathbf{F}_\Box(\lambda p.X\Box p)$ by Lemma A1b, so by Closure, $\vdash_{\mathsf{H_{KA}}} \Box^*\mathbf{F}_\Box Y \rightarrow \Box^*\mathbf{F}_\Box(\lambda p.Y\Box p)$. □

The second result says that when X is a finite iteration of necessity, the operators $\lambda p.X\Box p$ and $\lambda p.\Box Xp$ that result by adding an extra $\Box$ before or after X are ancestrally necessarily coextensive:

Lemma B2. $\vdash_{\mathsf{H_{KA}}} \mathbf{F}_\Box X \rightarrow \Box^*\forall q(\Box Xq \leftrightarrow X\Box q)$

Proof: By Induction, it suffices to show that (i) $\vdash_{\mathsf{H_{KA}}} \Box^*\forall q(\Box(\lambda p.p)q \leftrightarrow (\lambda p.p)\Box q)$, and (ii) $\vdash_{\mathsf{H_{KA}}} \mathbf{F}_\Box X \rightarrow \Box^*\forall q(\Box Xq \leftrightarrow X\Box q) \rightarrow \Box^*\forall q(\Box(\lambda p.X\Box p)q \leftrightarrow (\lambda p.X\Box p)(\Box q))$. For (i), note that $\Box^*\forall q\Box((\lambda p.p)q \leftrightarrow q)$ is an instance of $\Box^*$N. So by $\Box^*$K and Closure, $\vdash_{\mathsf{H_{KA}}} \Box^*\forall q(\Box(\lambda p.p)q \leftrightarrow \Box q)$, which by Closure implies $\vdash_{\mathsf{H_{KA}}} \Box^*\forall q(\Box(\lambda p.p)q \leftrightarrow (\lambda p.p)\Box q)$. For (ii), note that $\Box^*\forall q\Box((\lambda p.X\Box p)q \leftrightarrow X\Box q)$ is an instance of $\Box^*$N. So by $\Box^*$K and

[1] Showing that these sentences are not theorems of $\mathsf{H_{KA}}$ would require model theory beyond the scope of this book.

Closure, $\vdash_{\mathsf{H_{KA}}} \Box^*\forall q(\Box(\lambda p.X\Box p)q \leftrightarrow \Box X\Box q)$. But since $\vdash_{\mathsf{H_0}} \forall q(\Box(\lambda p.X\Box p)q \leftrightarrow \Box X\Box q) \rightarrow \forall q(\Box Xq \leftrightarrow X\Box q) \rightarrow \forall q(\Box(\lambda p.X\Box p)q \leftrightarrow (\lambda p.X\Box p)(\Box q))$, we can apply Closure to get $\vdash_{\mathsf{H_{KA}}} \Box^*\forall q(\Box Xq \leftrightarrow X\Box q) \rightarrow \Box^*\forall q(\Box(\lambda p.X\Box p)q \leftrightarrow (\lambda p.X\Box p)(\Box q))$. □

The third result says that finite iterations of necessity can freely be added under universal quantifiers inside the scope of $\Box^*$:

Lemma B3. $\vdash_{\mathsf{H_{KA}}} \Box^*\forall\vec{v}P \rightarrow \mathbf{F}_\Box X \rightarrow \Box^*\forall\vec{v}XP$

Proof: By Induction, it suffices to show (i) $\vdash_{\mathsf{H_{KA}}} \Box^*\forall\vec{v}P \rightarrow \Box^*\forall\vec{v}(\lambda p.p)P$, and (ii) $\vdash_{\mathsf{H_{KA}}} \Box^*\forall\vec{v}P \rightarrow \mathbf{F}_\Box X \rightarrow \Box^*\forall\vec{v}Xp \rightarrow \Box^*\forall\vec{v}(\lambda p.X\Box p)p$. (i) is immediate from Closure. For (ii): suppose $\mathbf{F}_\Box X$, and $\Box^*\forall\vec{v}XP$; then by Lemma A1d, we have $\Box^*\Box\forall\vec{v}Xp$. So by Lemma A6b, $\Box^*\forall\vec{v}\Box Xp$. By Lemma B2 and Closure, this gives us $\Box^*\forall\vec{v}X\Box p$, and hence also $\Box^*\forall\vec{v}(\lambda p.X\Box p)p$. □

Next we turn our attention to the modal logic $\mathsf{H_{KR}}$ that adds the following four rigidity axioms to $\mathsf{H_{KA}}$.

$\Box^*$**Rigid Comprehension**	$\Box^*\forall Y\exists C(\mathsf{R}C \wedge \forall\vec{z}(C\vec{z} \leftrightarrow Y\vec{z}))$
$\Box^*$**Persistent Rigidity**	$\Box^*\forall C(\mathsf{R}C \rightarrow \Box\mathsf{R}C)$
$\Box^*$**Persistence**	$\Box^*\forall C\forall\vec{z}(\mathsf{R}C \rightarrow C\vec{z} \rightarrow \Box C\vec{z})$
$\Box^*$**Inextensibility**	$\Box^*\forall C\forall Y(\mathsf{R}C \rightarrow (\forall\vec{z}(C\vec{z} \rightarrow \Box Y\vec{z}) \rightarrow \Box\forall\vec{z}(X\vec{z} \rightarrow Y\vec{z})))$.

The crucial consequences of these axioms which we need for our result are given by the next three lemmas, which concern $\Box^*$-analogues of Persistent Rigidity, Persistence, and Inextensibility.

Lemma B4 (Persistent* Rigidity)**.** $\vdash_{\mathsf{H_{KR}}} \mathsf{R}C \rightarrow \Box^*\mathsf{R}C$

Proof: By Induction, it suffices to show (i) $\vdash_{\mathsf{H_{KR}}} \mathsf{R}C \rightarrow (\lambda p.p)\mathsf{R}C$ and (ii) $\vdash_{\mathsf{H_{KR}}} \mathsf{R}C \rightarrow \mathbf{F}_\Box X \rightarrow X\mathsf{R}C \rightarrow (\lambda p.X\Box p)\mathsf{R}C$. (i) is immediate from Eβ. For (ii), we have $\vdash_{\mathsf{H_{KR}}} \mathbf{F}_\Box X \rightarrow X\forall C(\mathsf{R}C \rightarrow \Box\mathsf{R}C)$ by $\Box^*$Persistent Rigidity, which implies $\mathbf{F}_\Box X \rightarrow X\mathsf{R}C \rightarrow X\Box\mathsf{R}C$ by Closure. □

Lemma B5 (Persistence*)**.** $\vdash_{\mathsf{H_{KR}}} \mathsf{R}C \rightarrow C\vec{z} \rightarrow \Box^*C\vec{z}$

Proof: By Induction, it suffices to show that (i) $\vdash_{\mathsf{H_{KR}}} \mathsf{R}C \rightarrow \forall\vec{z}(C\vec{z} \rightarrow (\lambda p.p)C\vec{z})$ and (ii) $\vdash_{\mathsf{H_{KR}}} \mathsf{R}C \rightarrow \mathbf{F}_\Box X \rightarrow \forall\vec{z}(C\vec{z} \rightarrow XC\vec{z}) \rightarrow C\vec{z} \rightarrow X\Box C\vec{z}$. (i) is immediate from Eβ. For (ii), suppose that $\mathsf{R}C$ and $\mathbf{F}_\Box X$. Then by Persistent* Rigidity, $X\mathsf{R}C$. But by Persistence, $X\forall C\forall\vec{z}(\mathsf{R}C \rightarrow C\vec{z} \rightarrow \Box C\vec{z})$; so by Distribution, $XC\vec{z} \rightarrow X\Box C\vec{z}$. So by UI and PC, $\forall\vec{z}(C\vec{z} \rightarrow XC\vec{z}) \rightarrow C\vec{z} \rightarrow X\Box C\vec{z}$. □

Lemma B6 (Inextensibility*)**.** $\vdash_{\mathsf{H_{KR}}} \mathsf{R}C \rightarrow \forall\vec{z}(C\vec{z} \rightarrow \Box^*Y\vec{z}) \rightarrow \Box^*\forall\vec{z}(C\vec{z} \rightarrow Y\vec{z})$

Proof: We prove the following claim, from which the result follows immediately.

$$\vdash_{\mathsf{H_{KR}}} \mathsf{R}C \rightarrow \mathbf{F}_\Box X \rightarrow \forall Y(\forall\vec{z}(C\vec{z} \rightarrow XY\vec{z}) \rightarrow X\forall\vec{z}(C\vec{z} \rightarrow Y\vec{z}))$$

By Induction it suffices to show the following:

(i) $\vdash \mathsf{R}C \to \mathbf{F}_\Box X \to \forall Y(\forall\vec{z}(C\vec{z} \to (\lambda p.p)Y\vec{z}) \to (\lambda p.p)\forall\vec{z}(C\vec{z} \to Y\vec{z}))$

(ii) $\vdash \mathsf{R}C \to \mathbf{F}_\Box X \to \forall Y(\forall\vec{z}(C\vec{z} \to XY\vec{z}) \to X\forall\vec{z}(C\vec{z} \to Y\vec{z})) \to \forall Y(\forall\vec{z}(C\vec{z} \to X\Box Y\vec{z}) \to X\Box\forall\vec{z}(C\vec{z} \to Y\vec{z}))$

(i) is immediate from Eβ. For (ii), suppose $\mathsf{R}C$, $\mathbf{F}_\Box X$, and $\forall Y(\forall\vec{z}(C\vec{z} \to XY\vec{z}) \to X\forall\vec{z}(C\vec{z} \to Y\vec{z}))$. Instantiating Y with $\lambda p.\Box Yp$ gives

$$\forall\vec{z}(C\vec{z} \to X(\lambda p.\Box Yp)\vec{z}) \to X\forall\vec{z}(C\vec{z} \to (\lambda p.\Box Yp)\vec{z})$$

Given the assumption that $\mathbf{F}_\Box X$, we can simplify this by using Closure to apply Eβ inside X:

$$\forall\vec{z}(C\vec{z} \to X\Box Y\vec{z}) \to X\forall\vec{z}(C\vec{z} \to \Box Y\vec{z})$$

But by Persistent* Rigidity and the assumption that $\mathsf{R}C$, $X\mathsf{R}C$, and hence by Inextensibility and Distribution,

$$X\forall\vec{z}(C\vec{z} \to \Box Y\vec{z}) \to X\Box\forall\vec{z}(C\vec{z} \to Y\vec{z}).$$

Combining the last two lines gives us our desired conclusion:

$$\forall\vec{z}(C\vec{z} \to X\Box Y\vec{z}) \to X\Box\forall\vec{z}(C\vec{z} \to Y\vec{z}). \qquad \Box$$

Putting these facts together with Lemma B1, we can show that $\mathbf{F}_\Box$ is ancestrally necessarily coextensive with something rigid:

Lemma B7. $\vdash_{\mathsf{H_{KR}}} \exists C(\mathsf{R}C \wedge \Box^*\forall X(CX \leftrightarrow \mathbf{F}_\Box X))$.

Proof: By Rigid Comprehension, there is something rigid coextensive with $\mathbf{F}_\Box$; call it C. By Lemmas A1a and A1b, $C(\lambda p.p)$ and $\forall Y(CY \to C(\lambda p.Y\Box p))$; so by Persistence*, $\Box^*C(\lambda p.p)$ and $\forall Y(CY \to \Box^*C(\lambda p.Y\Box p))$. By Inextensibility*, the latter implies $\Box^*\forall Y(CY \to C(\lambda p.Y\Box p))$. But by the definition of $\mathbf{F}_\Box$ we know that

$$\vdash_{\mathsf{H_0}} C(\lambda p.p) \wedge \forall Y(CY \to C(\lambda p.Y\Box p)) \to \forall Y(\mathbf{F}_\Box Y \to CY).$$

So we can apply Closure to derive $\Box^*\forall Y(\mathbf{F}_\Box Y \to CY)$.

For the other direction, we have by assumption that $\forall Y(CY \to \mathbf{F}_\Box Y)$, so by Lemma B1, $\forall Y(CY \to \Box^*\mathbf{F}_\Box Y)$, so by Inextensibility*, $\Box^*\forall Y(CY \to \mathbf{F}_\Box Y)$. $\Box$

It follows from this that $\mathbf{F}_\Box$ itself behaves as if it were rigid—in particular, it satisfies an analogue of Inextensibility:

Lemma B8. $\vdash_{\mathsf{H_{KR}}} \forall Y(\mathbf{F}_\Box Y \to \Box^*P) \to \Box^*\forall Y(\mathbf{F}_\Box Y \to P)$

Proof: Fix a rigid C such that $\Box^*\forall X(CX \leftrightarrow \mathbf{F}_\Box X)$: such a C exists by Lemma B7. Assume $\forall Y(\mathbf{F}_\Box Y \to \Box^*P)$; then $\forall Y(CY \to \Box^*P)$, so $\Box^*\forall Y(CY \to P)$ by Inextensibility*. Combining this with $\Box^*\forall Y(\mathbf{F}_\Box Y \to CY)$, Closure yields $\Box^*\forall Y(\mathbf{F}_\Box Y \to P)$. $\Box$

Finally, by combining this result with the CBF-style results we proved earlier in this appendix (before introducing the rigidity axioms), we have:

Lemma B9. $\vdash_{\mathsf{H_{KR}}} \Box^*\forall\vec{v}P \to \Box^*\forall\vec{v}\Box^*P$ for any P.

Proof: By Lemma B3, $\vdash_{\mathsf{H_{KA}}} \Box^*\forall\vec{v}P \rightarrow \forall X(\mathbf{F}_\Box X \rightarrow \Box^*\forall\vec{v}XP)$. Hence by Lemma B8, $\vdash_T \Box^*\forall\vec{v}P \rightarrow \Box^*\forall X(\mathbf{F}_\Box X \rightarrow \forall\vec{v}XP)$, so by Closure, $\vdash_T \Box^*\forall\vec{v}P \rightarrow \Box^*\forall\vec{v}\forall X(\mathbf{F}_\Box X \rightarrow XP)$ and hence $\vdash_T \Box^*\forall\vec{v}P \rightarrow \Box^*\forall\vec{v}\Box^*P$. □

Setting $\vec{v}$ as the empty sequence, P as p, and applying GEN gives us Iteration for $\Box^*$:

Corollary B10. $\vdash_{\mathsf{H_{KR}}} \forall p(\Box^*p \rightarrow \Box^*\Box^*p)$.

Lemma B11. $\mathsf{H_{KR}}$ is closed under the rule $P/\Box^*\forall\vec{v}P$.

Proof: Let T be the set of all formulae P such that $\vdash_{\mathsf{H_{KR}}} \Box^*\forall\vec{v}P$ for all sequences of variables $\vec{v}$; to prove the result we show that T includes every theorem of $\mathsf{H_{KR}}$. For the same reason as in the proof of Proposition A7, T is closed under GEN and MP, and includes $\mathsf{H_0}$. To show that T includes every instance of $\Box^*$N, note that each such instance is of the form $\Box^*\forall\vec{u}\Box Q$ where Q is a theorem of $\mathsf{H_0}$. Then $\Box^*\forall\vec{v}\forall\vec{u}\Box Q$ is also an instance of $\Box^*$N, so $\vdash_{\mathsf{H_{KR}}} \Box^*\forall\vec{v}\Box^*\forall\vec{u}\Box Q$ by Lemma B9, and hence $\Box^*\forall\vec{u}\Box Q \in T$. To show that T includes $\Box^*$K, $\Box^*$Duality, and each of the rigidity axioms, note that these are all of the form $\Box^*Q$ where Q is closed. Thus $\vdash_{\mathsf{H_0}} Q \rightarrow \forall\vec{v}Q$, so by Closure $\vdash_{\mathsf{H_{KA}}} \Box^*Q \rightarrow \Box^*\forall\vec{v}Q$. Hence $\vdash_{\mathsf{H_{KR}}} \Box^*\forall\vec{v}Q$, and thus by Lemma B9, $\vdash_{\mathsf{H_{KR}}} \Box^*\forall\vec{v}\Box^*Q$, i.e. $\Box^*Q \in T$. □

We now have all the ingredients we need to show that $\Box^*$ "obeys S4" in $\mathsf{H_{KR}}$. More precisely: whenever P is a theorem of $\mathsf{H_{S4R}}$ not containing the primitive $\Diamond$ constant, $P[\Box^*/\Box]$—the result of substituting $\Box^*$ for every occurrence of $\Box$—is a theorem of $\mathsf{H_{KR}}$. Here $\mathsf{H_{S4R}}$ is the result of adding the rigidity axioms to $\mathsf{H_{S4}}$. By Proposition A11 and Corollary A13, its restriction to the $\Diamond$-free language can be characterized as the smallest set of $\Diamond$-free formulae that (i) contains every theorem of $\mathsf{H_0}$; (ii) contains $K := \forall p\forall q(\Box(p \rightarrow q) \rightarrow \Box p \rightarrow \Box q)$; (iii) contains $4 := \forall p(\Box p \rightarrow \Box\Box p)$; (iv) contains $T := \forall p(\Box p \rightarrow p)$; (v) contains every instance of Rigid Comprehension, Persistent Rigidity, Persistence or Inextensibility; (vi) is closed under MP; (vii) is closed under GEN, and (viii) is closed under NEC.[2]

Proposition B12. If $\vdash_{\mathsf{H_{S4R}}} P$ then $\vdash_{\mathsf{H_{KR}}} P[\Box^*/\Box]$ (where P doesn't contain $\Diamond$).

Proof: Let T be the set of all $\Diamond$-free P such that $\vdash_{\mathsf{H_{KR}}} P[\Box^*/\Box]$. So what we want to show is that T contains the $\Diamond$-free part of $\mathsf{H_{S4R}}$, for which it suffices to show that T has the eight properties just listed.

(i) is immediate from the fact that $\vdash \mathsf{H_0}$ treats $\Box$ no differently from a variable of the same type.

(ii) $K[\Box^*/\Box]$ is $\forall p\forall q(\Box^*(p \rightarrow q) \rightarrow \Box^*p \rightarrow \Box^*q)$, which is a theorem of $\mathsf{H_{KA}}$ by Distribution.

(iii) $4[\Box^*/\Box]$ is $\forall p(\Box^*p \rightarrow \Box^*\Box^*p)$ which is a theorem of $\mathsf{H_{KR}}$ by Corollary B10.

(iv) $T[\Box^*/\Box]$ is $\forall\, p(\Box^*p \rightarrow p)$, which is a theorem of $\mathsf{H_0}$ by Lemma A1c.

(v) All instances of Rigid Comprehension are in T since they are unchanged by the substitution of $\Box^*$ for $\Box$. The results of substituting $\Box^*$ for $\Box$ in instances of Persistent Rigidity, Persistence, and Inextensibility are in T by Lemmas B4–B6 (Persistent* Rigidity, Persistence*, and Inextensibility*).

[2] We don't need to mention Duality, since it is obviously conservative over the $\Diamond$-free fragment. The un-necessitated versions of the rigidity axioms are sufficient, since $\mathsf{H_{S4R}}$ is closed under NEC* by Proposition A15.

(vi) T is closed under MP since H_{KR} is and substitution commutes with $\rightarrow$.

(vii) T is closed under GEN since H_{KR} is and substitution commutes with $\forall$.

(viii) T is closed under NEC since by Lemma B11, if $\vdash_{\mathsf{H}_{\mathsf{KR}}} P[\Box^*/\Box]$, then $\vdash_{\mathsf{H}_{\mathsf{KR}}} \Box^*(P[\Box^*/\Box])$, and $\Box^*(P[\Box^*/\Box]) = (\Box P)[\Box^*/\Box]$. □

APPENDIX C

Consequences of the Rigidity Axioms

This appendix will prove some further facts about the theory H_{KR}, axiomatized by the rigidity axioms from Figure 1.5, the modal axioms from Figure 1.4, and the nonmodal axioms from Figure 1.3. First, we explore a candidate definition of the primitive rigidity predicates in terms of persistence and inextensibility, and show that this definition would be adequate for our purposes. Second, we explore a claim about rigidity that is not encoded in our axioms, namely that if something lacks a rigid property, it necessarily lacks it. In the presence of the other axioms, this turns out to be equivalent to the necessity of distinctness; moreover both it and Inextensibility are redundant in the presence of the B axiom. Third, we prove that the Boolean Completeness and Actuality principles discussed in §1.5 are theorems of H_{KR}, and show that under certain circumstances they could take over the role of the rigidity axioms. Fourth, we discuss the MaxCon and Atomicity principles which came up in §1.6: while neither is a theorem of H_{KR}, they can be shown to be equivalent in that logic; moreover, both become theorems if we add the characteristic principles of S5 to H_{KR}.

First: consider the following candidate definition of the rigidity predicates R_σ:

$$\begin{aligned}\mathsf{R}' &:= \lambda X.\ \text{Persistent}\ X \wedge \text{Inextensible}\ X\\ \text{Persistent} &:= \lambda X.\Box^*\forall\vec{z}(X\vec{z} \rightarrow \Box X\vec{z})\\ \text{Inextensible} &:= \lambda X.\Box^*\forall Y(\forall\vec{z}(X\vec{z} \rightarrow \Box Y\vec{z}) \rightarrow \Box\forall\vec{z}(X\vec{z} \rightarrow Y\vec{z}))\end{aligned}$$

For any formula P, let P' be $P[\mathsf{R}'/\mathsf{R}]$, i.e. the result of substituting R'_σ for each primitive predicate R_σ in P. Let $\mathsf{H}_{\mathsf{KR}'}$ be H_{KA} + $\Box^*$Rigid Comprehension′ + $\Box^*$Persistent Rigidity′. Then we have:

Proposition C1. (a) If $\vdash_{\mathsf{H}_{\mathsf{KR}}} P$ then $\vdash_{\mathsf{H}_{\mathsf{KR}'}} P'$, and (b) $\mathsf{H}_{\mathsf{KR}'} \subseteq \mathsf{H}_{\mathsf{KR}}$.

Proof: For (a), it suffices to show that $\Box^*$Persistence′ and $\Box^*$Inextensibility′ are both theorems of H_{KA}, and hence also of $\mathsf{H}_{\mathsf{KR}'}$. This follows from the factivity of $\Box^*$ (Lemma A1c) and the closure of H_{KA} under $\Box^*$-necessitation (Lemma B11). This means that the result of applying $'$ uniformly to any derivation in H_{KR} is a derivation in $\mathsf{H}_{\mathsf{KR}'}$.

For (b), we need only note further that $\Box^*$Rigid Comprehension′ and $\Box^*$Persistent Rigidity′ are also theorems of H_{KR}. The latter fact is a consequence of the fact that $\vdash_{\mathsf{H}_{\mathsf{KR}}} \Box^*\forall p(\Box^* p \rightarrow \Box^*\Box^* p)$, which follows from Corollary B10 and Lemma B11. For the former fact, note first that by $\Box^*$Persistence, $\Box^*$Inextensibility, and Closure, we have $\vdash_{\mathsf{H}_{\mathsf{KR}}} \Box^*(\mathsf{R}C \rightarrow \mathsf{R}'C)$; but by Persistent* Rigidity (Lemma B4) and Lemma B11, $\vdash_{\mathsf{H}_{\mathsf{KR}}} \Box^*\forall C(\mathsf{R}C \rightarrow \Box^*\mathsf{R}C)$. Combining this with Rigid Comprehension yields $\Box^*$Rigid Comprehension′. □

This means that $\mathsf{H}_{\mathsf{KR}'}$ includes all the theorems of H_{KR} in which none of the primitive predicates R_σ occur. Thus, while proponents of H_{KR} are not forced to accept the identity of R'_σ with R_σ, they have no quarrel with someone who stipulatively defines R as R' and accepts every theorem of H_{KR} under this definition.

We can also note an even simpler axiomatization of $\mathsf{H}_{\mathsf{KR}'}$:

Proposition C2. $\mathsf{H}_{\mathsf{KR}'} = \mathsf{H}_{\mathsf{KA}} + \Box^*$Rigid Comprehension$' + \Box^*\forall p(\Box^* p \rightarrow \Box^*\Box^* p)$.

Proof: By Corollary B10 and Lemma B11, $\Box^*\forall p(\Box^* p \rightarrow \Box^*\Box^* p)$ is a theorem of $\mathsf{H}_{\mathsf{KR}'}$, since it is R-free and a theorem of H_{KR}. Conversely, $\Box^*$Persistent Rigidity$'$ follows in H_{KR} from $\Box^*\forall p(\Box^* p \rightarrow \Box^*\Box^* p)$, since $\mathsf{R}'C$ is β-equivalent to a conjunction of formulae each beginning with $\Box^*$. □

Since $\Box^*\forall p(\Box^* p \rightarrow \Box^*\Box^* p)$ is a theorem of H_{S4}, it follows that $\mathsf{H}_{\mathsf{KR}'} \subseteq \mathsf{H}_{\mathsf{S4}} +$ □Rigid Comprehension$'$.

Second: consider the following axiom, which might strike one as a natural addition to H_{KR}:

Negative Persistence $\quad \forall C(\mathsf{R}C \rightarrow \forall \vec{x}(\neg C\vec{x} \rightarrow \Box\neg C\vec{x}))$

This turns out to be equivalent in H_{KR} to the necessity of distinctness:

Proposition C3. For any types $\sigma_1, \ldots, \sigma_n$, $\vdash_{\mathsf{H}_{\mathsf{KR}}} \mathrm{NP}_{\sigma_1,\ldots,\sigma_n} \leftrightarrow \mathrm{ND}_{\sigma_1}$, where

$$\mathrm{ND}_{\sigma_1} := \forall x \forall y(x \neq_{\sigma_1} y \rightarrow \Box(x \neq_{\sigma_1} y))$$
$$\mathrm{NP}_{\sigma_1,\ldots,\sigma_n} := \forall C(\mathsf{R}_{\sigma_1,\ldots,\sigma_n} C \rightarrow \forall \vec{x}(\neg C\vec{x} \rightarrow \Box\neg C\vec{x}))$$

Proof, left to right: Suppose $x \neq_{\sigma_1} y$. Choose any $x_2, \ldots, x_n$ of types $\sigma_2 \ldots, \sigma_n$. By Rigid Comprehension, there is a rigid C such that $\forall \vec{y}(Cz_1 \ldots z_n \leftrightarrow (z_1 = x \wedge z_2 = x_2 \wedge \cdots \wedge z_n = x_n))$. $\Box Cxx_2 \ldots x_n$ by Persistence, whereas $\Box\neg Cyx_2 \ldots x_n$ by $\mathrm{NP}_{\sigma_1,\ldots,\sigma_n}$. Since $\vdash_{\mathsf{H}_0} Cxx_2 \ldots x_n \rightarrow \neg Cyx_2 \ldots x_n \rightarrow x \neq y$, we can infer $\Box x \neq y$ by Closure. *Right to left:* Suppose C is rigid and $\neg Cx_1 \ldots x_n$. By ND, $\forall \vec{z}(C\vec{z} \rightarrow \Box z_1 \neq x_1 \vee \cdots \vee \Box z_n \neq x_n)$, and hence $\forall \vec{z}(C\vec{z} \rightarrow \Box(z_1 \neq x_1 \vee \cdots \vee z_n \neq x_n))$, by Closure. So by Inextensibility, $\Box\forall \vec{z}(C\vec{z} \rightarrow (z_1 \neq x_1 \vee \cdots \vee z_n \neq x_n))$. But of course $\Box(x_1 = x_1 \wedge \cdots x_n = x_n)$ by $\Box^*$N, so by another application of Closure, $\Box\neg Cx_1 \ldots x_n$. □

One might wonder whether adding $\Box^*$Negative Persistence as an extra axiom would make $\Box^*$Inextensibility redundant. The answer is no; however, if we add the ancestral necessity of the Barcan Formula along $\Box^*$Negative Persistence, $\Box^*$Inextensibility does become derivable. This follows from the following theorem:

Proposition C4. $\vdash_{\mathsf{H}_{\mathsf{KA}}} \mathrm{BF}_{\sigma_1,\ldots,\sigma_n} \rightarrow \mathrm{NP}_{\sigma_1,\ldots,\sigma_n} \rightarrow \mathrm{Inextensibility}_{\sigma_1,\ldots,\sigma_n}$, where $\mathrm{BF}_{\sigma_1,\ldots,\sigma_n} := \forall F^{\langle\sigma_1,\ldots,\sigma_n\rangle}(\forall \vec{x}\Box F\vec{x} \rightarrow \Box\forall \vec{x}F\vec{x})$.

Proof: Suppose $\mathrm{BF}_{\vec{\sigma}}$, $\mathrm{NP}_{\vec{\sigma}}$, $\mathsf{R}C$ and $\forall \vec{z}(C\vec{z} \rightarrow \Box Y\vec{z})$. By NP we have $\forall \vec{z}(\neg C\vec{z} \rightarrow \Box\neg C\vec{z})$. Putting the last two facts together we have $\forall \vec{z}(\Box Y\vec{z} \vee \Box\neg C\vec{z})$, which by Closure implies $\forall \vec{z}\Box(C\vec{z} \rightarrow Y\vec{z})$, and hence $\Box\forall \vec{z}(C\vec{z} \rightarrow Y\vec{z})$ by $\mathrm{BF}_{\vec{\sigma}}$. □

Moreover, the following theorem shows that if we add the ancestral necessity of the B axiom ($\Box^*\forall p(\Diamond\Box p \rightarrow p)$) as well as $\Box^*$Persistent Rigidity and $\Box^*$Persistence, we will get $\Box^*$Negative Persistence for free:

Proposition C5. $\vdash_{\mathsf{H}_{\mathsf{KA}}} \mathrm{B} \rightarrow \mathrm{Persistent\ Rigidity}_{\vec{\sigma}} \rightarrow \mathrm{Persistence}_{\vec{\sigma}} \rightarrow \mathrm{NP}_{\vec{\sigma}}$.

Proof: Suppose C is rigid and $\neg C\vec{x}$. By Persistent Rigidity, $\Box\mathsf{R}C$, so by Persistence $\Box(C\vec{x} \rightarrow \Box C\vec{x})$. By K and Duality this implies $\Diamond C\vec{x} \rightarrow \Diamond\Box C\vec{x}$, and hence $\Diamond C\vec{x} \rightarrow C\vec{x}$ by B, and thus $\Box\neg C\vec{x}$. □

Since the B axiom also implies BF, the upshot of the last two results is $\Box^*$Inextensibility is redundant once we have the ancestral necessity of B.

Third: consider the following claims, which were stated without proof in §1.5:

Boolean Completeness
$\forall F \exists X(\forall Y(FY \rightarrow \Box\forall\vec{z}(X\vec{z} \rightarrow Y\vec{z})) \wedge \forall Z(\forall Y(FY \rightarrow \Box\forall\vec{z}(Z\vec{z} \rightarrow Y\vec{z})) \rightarrow \Box\forall\vec{z}(Z\vec{z} \rightarrow X\vec{z})))$

Actuality $\exists p(p \wedge \forall q(q \rightarrow \Box(p \rightarrow q)))$

We will show that both are theorems of $\mathsf{H_{KR}}$.

Proposition C6. $\vdash_{\mathsf{H_{KR}}}$Boolean Completeness.

Proof: We consider the case of propositions; the generalization to properties and relations is parallel. Fix F. By Rigid Comprehension, there is a rigid property coextensive with F; call it C, and let p be $\forall q(Cq \rightarrow q)$. Since by Persistence every F proposition is necessarily C, p necessitates every F proposition. Suppose that r also necessitates every F proposition. Then *being true if r is* $(\lambda q.r \rightarrow q)$ is a property that every C proposition has necessarily. So by Inextensibility, it is necessitated by C: $\Box\forall q(Cq \rightarrow (r \rightarrow q))$. By Closure, this implies $\Box(r \rightarrow \forall q(Cq \rightarrow q))$, i.e. that r necessitates our p. □

Proposition C7. $\vdash_{\mathsf{H_{KR}}}$Actuality.

Proof: By Rigid Comprehension, there is a rigid C such that $\forall p(Cp \leftrightarrow p)$. We claim that the proposition that every instance of C is true—$\forall p(Cp \rightarrow p)$—witnesses Actuality. It is true, since C is coextensive with truth. And suppose q is true: then Cq, so by Persistence $\Box Cq$, and hence $\Box(\forall p(Cp \rightarrow p) \rightarrow q)$ by Closure. □

Since the primitive rigidity predicates do not occur in Boolean Completeness or Actuality, there is obviously no prospect of using either or both of them to derive any of the rigidity axioms. But one might still hope to derive Rigid Comprehension′, which only involves the non-primitive rigidity predicate R', from one or both of them. It turns out in the setting of $\mathsf{H_{S5}}$, Boolean Completeness and Actuality both imply, and are in fact equivalent to, Rigid Comprehension′. So if we are happy with the definition of 'rigid' as 'persistent and inextensible', either one of these axioms will suffice to axiomatize the theory of rigidity given $\mathsf{H_{S5}}$.

Proposition C8. $\vdash_{\mathsf{H_{S5}}}$Boolean Completeness ↔ Rigid Comprehension′

Proof, right to left: we showed above that $\vdash_{\mathsf{H_{S4}}}$ Rigid Comprehension′ $\rightarrow P'$ whenever $\vdash_{\mathsf{H_{KR}}} P$; given Proposition C6, the fact that $\mathsf{H_{S4}} \subseteq \mathsf{H_{S5}}$, and the fact that Boolean Completeness′ is just Boolean Completeness, this implies that $\vdash_{\mathsf{H_{S5}}}$ Rigid Comprehension′ → Boolean Completeness.

Left to right: let F be some arbitrary property (the case of relations is parallel), and let $F^=$ be the property of being the haecceity of something F: $\lambda X.\exists y(Fy \wedge X = (\lambda z.z = y))$. By Boolean Completeness $F^=$ has a least upper bound; call it C_F. We will show that C_F is coextensive with F, persistent, and inextensible. To do so, we first note that $\forall y(Fy \rightarrow \Box C_F y)$, since whenever y is F its haecceity $(\lambda z.z = y)$ is $F^=$ and thus necessitates C_F. Also $\forall y(\neg Fy \rightarrow \Box\neg C_F y)$, since

when y is not F, each F thing is necessarily distinct from y by ND, meaning that $\lambda z.z \neq y$ is necessitated by each $F^{=}$ property, and hence by their least upper bound C_F. These facts together imply that F is coextensive with C_F. They also imply that $\forall y(\Box\neg C_F y \vee \Box C_F y)$. This implies the persistence of C_F: $\forall y(\Box\neg C_F y \vee \Box\Box C_F y)$ by Iteration, hence $\forall y\Box(\neg C_F y \vee \Box C_F y)$ by Closure, hence $\Box\forall y(\neg C_F y \vee \Box C_F y)$ by BF, i.e. $\Box\forall y(C_F y \rightarrow \Box C_F y)$. By parallel reasoning we also get the negative persistence of C_F: $\Box\forall y(\neg C_F y \rightarrow \Box\neg C_F y)$. Using Closure, this trivially implies $\forall X\Box(\forall y(\neg C_F y \vee \Box Xy) \rightarrow \forall y(\Box C_F y \vee \Box Xy))$ and hence $\forall X\Box(\forall y(C_F y \rightarrow \Box Xy) \rightarrow \forall y\Box(C_F y \rightarrow \Box Xy))$. A final appeal to Iteration and BF yields $\Box\forall X\Box(\forall y(C_F y \rightarrow \Box Xy) \rightarrow \Box\forall y(C_F y \rightarrow \Box Xy))$ □

Proposition C9. $\vdash_{\mathsf{H}_{\mathsf{S5}}}$ Actuality ↔ Rigid Comprehension′.[1]

Proof: The right-to-left direction is straightforward given Proposition C7, for the same reason as in Proposition C8. For the left-to-right direction, suppose w witnesses Actuality and X is any property. Let C be $\lambda z.\Box(w \rightarrow Xz)$. C is coextensive with X since w necessitates all and only the truths. Iteration implies immediately that C is persistent. The 5 axiom similarly implies that C is negatively persistent, i.e. $\Box\forall z\Diamond(w \wedge \neg Xz) \rightarrow \Box\Diamond(w \wedge \neg z)$. But as we saw in the final stage of the previous proof, negative persistence suffices for inextensibility given Iteration and BF. □

Fourth and finally: consider the following principles from §1.6:

MaxCon $(\mathsf{R}C \wedge \Diamond\forall p(Cp \rightarrow p)) \rightarrow$
$\exists C'(\mathsf{R}C' \wedge \forall p(Cp \rightarrow C'p) \wedge \Diamond\forall p(C'p \rightarrow p) \wedge \forall q(C'q \vee C'\neg q))$

Atomicity $\Diamond p \rightarrow \exists p'(\Box(p' \rightarrow p) \wedge \Diamond p' \wedge \forall q(\Box(p' \rightarrow q) \vee \Box(p' \rightarrow \neg q)))$

These claims are equivalent in H_{KR}:

Proposition C10. $\vdash_{\mathsf{H}_{\mathsf{KR}}}$ MaxCon ↔ Atomicity.

Proof, left to right: Suppose MaxCon and $\Diamond p$. By Rigid Comprehension, there is a singleton collection C such that $\forall q(Cq \leftrightarrow q = p)$. By the necessity of identity, $\forall q(Cq \rightarrow \Box q = p)$, so by Inextensibility, $\Box\forall q(Cq \rightarrow q = p)$; since $\Diamond p$ this gives us $\Diamond(p \wedge \forall q(Cq \rightarrow q = p))$ and hence $\Diamond\forall q(Cq \rightarrow q)$. So we can detach the consequent of MaxCon, to conclude that there is a "maximal consistent" collection C' extending C. Let p' be $\forall q(C'q \rightarrow q)$. Since $C'p$, $\Box C'p$ by Persistence, hence $\Box(p' \rightarrow p)$. And since by Persistence p' necessitates every member of C', it necessitates each proposition or its negation.

Right to left: suppose Rigid $C \wedge \Diamond\forall p(Cp \rightarrow p)$. Let r be $\forall q(Cq \rightarrow q)$; by Atomicity there is a possibly-true proposition p' that necessitates r and necessitates every proposition or its negation. By Rigid Comprehension, there is a rigid C' such that $\forall p(C'p \leftrightarrow \Box(p' \rightarrow p))$. By Persistence $\forall p(Cp \rightarrow \Box Cp)$, so we have $\forall p(Cp \rightarrow \Box(p' \rightarrow p))$ and hence $\forall p(Cp \rightarrow C'p)$. Since $\forall p(C'p \rightarrow \Box(p' \rightarrow p))$, by Inextensibility we have $\Box\forall p(C'p \rightarrow (p' \rightarrow p))$, and hence $\Box(p' \rightarrow \forall p(C'p \rightarrow p))$; given $\Diamond p'$ this implies $\Diamond\forall p(C'p \rightarrow p)$. And finally, since p' necessitates each proposition or its negation, C' contains each proposition or its negation. □

[1] This is a strengthening of Theorem 11.5 in Gallin 1975, which shows that the necessitation of Actuality (Gallin's 'At') is equivalent in higher-order S5 to the necessitation of Rigid Comprehension′ (equivalent to Gallin's E^{σ}). Our result is slightly stronger, since even in H_{S5}, Actuality and Rigid Comprehension′ do not imply their own necessitations.

Neither MaxCon nor Atomicity is a theorem of H_{KR} (for proof see Bacon and Dorr forthcoming). Both do, however, follow if we add the further principles of H_{S5}. Indeed, Atomicity is equivalent in H_{S5} to the necessary truth of Actuality.

Proposition C11. $\vdash_{\mathsf{H}_{\mathsf{S5}}}$ Atomicity $\leftrightarrow \Box$ Actuality.

Proof, left to right: Let an *atom* be a possible proposition that settles every question:

$$\text{Atom } w \coloneqq \Diamond w \wedge \forall q(\Box(w \rightarrow q) \vee \Box(w \rightarrow \neg q))$$

It's easy to see that in H_{S5}, any atom is necessarily an atom. For if w is an atom, $\Box\Diamond w$ by 5; also $\forall q\Box(\Box(w \rightarrow q) \vee \Box(w \rightarrow \neg q))$ by Iteration and K, and hence $\Box\forall q(\Box(w \rightarrow q) \vee \Box(w \rightarrow \neg q))$ by BF. But by the T axiom, Actuality is equivalent to the claim that there is a *true* atom. So if every atom is necessarily an atom, every atom is compatible with its being a true atom, and thus with *there being* a true atom, i.e. with Actuality. Since an atom necessitates every proposition with which it is compatible, it follows that every atom necessitates Actuality. But given Atomicity, this implies that Actuality is necessarily true, since if it were possibly false, its falsehood would have to be compatible with some atom by Atomicity.

Right to left: $\Box$Actuality implies that every possible proposition is compatible with there being a true atom:

$$\Diamond p \rightarrow \Diamond(p \wedge \exists w(w \wedge \forall q(q \rightarrow \Box(w \rightarrow q))))$$

and hence by Closure,

$$\Diamond p \rightarrow \Diamond\exists w\forall q(\Box(w \rightarrow p) \wedge w \wedge (\Box(w \rightarrow q) \vee \Box(w \rightarrow \neg q))).$$

By BF and CBF, this yields

$$\Diamond p \rightarrow \exists w\forall q\Diamond(\Box(w \rightarrow p) \wedge w \wedge (\Box(w \rightarrow q) \vee \Box(w \rightarrow \neg q)))$$

from which we can use K to derive

$$\Diamond p \rightarrow \exists w(\Diamond\Box(w \rightarrow p) \wedge \Diamond w \wedge \forall q(\Diamond\Box(w \rightarrow q) \vee \Diamond\Box(w \rightarrow \neg q))).$$

By 5 we can replace each $\Diamond\Box$ with $\Box$, which yields Atomicity. $\Box$

Given Proposition C9, it follows that H_{S5} + $\Box$Rigid Comprehension′ = H_{S5}+Atomicity.

APPENDIX D

Narrower Modalities in Higher-Order S4

Appendix B showed how given an operator $\Box$ obeying the relatively weak higher order logic H_{KR}, one can define a perhaps "broader" operator $\Box^*$ guaranteed to obey the stronger logic $\mathsf{H}_{\mathsf{S4R}}$. This appendix will work in the opposite direction, showing that if we start with an operator $\Box$ obeying H_{S4}, we can naturally define a perhaps *narrower* operator $\Box_5$ obeying H_{S5}, as well as an intermediate operator $\Box_{\neq}$ guaranteed to obey ND but perhaps not all of H_{S5}. $\Box_{\neq}$ and $\Box_5$ are, intuitively, maximally strong weakenings of $\Box$ that make true the relevant additional logical principles.[1]

The definitions of $\Box_{\neq}$ and $\Box_5$ make sense for any starting $\Box$; and so long as $\Box$ obeys H_{S4}, the operations so defined will be, in various ways that we will lay out, naturally distinguished relative to $\Box$. Of course, the metaphysical interest of this will depend on the metaphysical interest of the starting $\Box$: garbage in, garbage out. One particularly interesting setting is that of Classicism, where as we saw in §8.2, the operator $\Box_{\mathsf{T}} := (\lambda p.p = \mathsf{T})$ obeys H_{S4} and is moreover logically distinguished in ways that give it an excellent claim on the label 'the broadest necessity'. More generally, as discussed in §8.4, $\Box$ will enjoy a similar kind of logical naturalness so long as it is "identity-based" in the sense that we can prove $\Box = \lambda p.Ap = Bp$ for some A, B where $\vdash_{\mathsf{H}_0} (Ap = Bp) \rightarrow p$. For these conditions on $\Box$ mean that case we will automatically have $\forall p(\Box p \rightarrow Xp)$ for any other operator X that obeys H_{K}.[2] In what follows we will not assume anything about $\Box$ beyond H_{S4} (except for the very end when we will bring in $\mathsf{H}_{\mathsf{S4R}}$), but for convenience we will refer to $\Box$ as "broad necessity"; we will pronounce $\Box\forall\vec{x}(F\vec{x} \rightarrow G\vec{x})$ as 'F entails G' and $\Box\forall\vec{x}(F\vec{x} \leftrightarrow G\vec{x})$ as 'F is broadly equivalent to G'. Throughout this appendix, $\vdash$ means $\vdash_{\mathsf{H}_{\mathsf{S4}}}$.

We will begin by defining an array of properties of operations. (The variable X will always have type $\langle\langle\rangle\rangle$.)

$$
\begin{aligned}
\mathrm{N} &:= \lambda X.\forall p(\Box p \rightarrow Xp)\\
\mathrm{K} &:= \lambda X.\forall p\forall q(X(p \rightarrow q) \rightarrow Xp \rightarrow Xq)\\
4 &:= \lambda X.\forall p(Xp \rightarrow XXp)\\
\mathrm{T} &:= \lambda X.\forall p(Xp \rightarrow p)\\
\mathrm{B} &:= \lambda X.\forall p(p \rightarrow X\neg X\neg p)\\
\mathrm{J} &:= \lambda X.\forall p(p \rightarrow X\Diamond p)\\
\mathrm{I}_\sigma &:= \lambda X.\forall y^\sigma\forall z^\sigma(y \neq z \rightarrow X(y \neq z))\\
\mathrm{F}_\sigma &:= \lambda X.\forall Y^{\langle\sigma\rangle}(X\forall zYz \rightarrow \forall zXYz)
\end{aligned}
$$

Combinations abbreviate conjunctions, e.g. $\mathrm{NK} := \lambda X.\mathrm{N}X \wedge \mathrm{K}X$. An initial $\Box$ applies to the entire conjunction, e.g. $\Box\mathrm{NK} := \lambda X.\Box(\mathrm{N}X \wedge \mathrm{K}X)$.

[1] If $\Box$ obeys $\mathsf{H}_{\mathsf{S4R}}$, $\Box_{\neq}$ and $\Box_5$ will too; however, the rigidity axioms are not important to the main results of this appendix.

[2] Indeed, we only need the weak logic $\mathsf{H}_{\mathsf{K}-}$ discussed in Appendix A.

K, 4, T, B, I_σ, and F_σ are familiar. (The last two correspond to ND and BF for a given type σ—we would have used those familiar labels if the letters N and B weren't already in use.) N is self-explanatory; in the Classicist setting, where $\Box$ is $\Box_\top$, it is equivalent to $\lambda X.X\top$. The only unfamiliar property is J, which turns out to be intimately related to the necessity of distinctness. When combined with N and K, it entails each I_σ:

Proposition D1. $\vdash \mathrm{NKJ}X \to \mathrm{I}_\sigma X$ for every σ.

Proof: By definition of J, $\vdash \mathrm{J}X \to y \neq z \to X\Diamond(y \neq z)$. By the necessity of identity (for $\Box$), we have $\vdash \Box(\Diamond(y \neq z) \to y \neq z)$, hence $\vdash \mathrm{N}X \to X(\Diamond(y \neq z) \to y \neq z)$, and thus $\vdash \mathrm{NK}X \to X\Diamond(y \neq z) \to X(y \neq z)$. Combining these we get $\vdash \mathrm{NKJ}X \to y \neq z \to X(y \neq z)$. $\Box$

Moreover, if we go beyond H_{S4} by assuming that $\Box$ is "identity-based" in the sense explained above, we can turn this around and show that J is in fact *equivalent* to $I_{\langle\rangle}$ (the necessity of propositional distinctness) in the presence of N and K:

Proposition D2. $\vdash (\Box = (\lambda p.Ap = Bp)) \to \mathrm{NKI}_{\langle\rangle}X \to \mathrm{J}X$.

Proof: Suppose $\Box = (\lambda p.Ap = Bp)$. Since $\vdash p \to \neg\Box\neg p$, $p \to A\neg p \neq B\neg p$. Hence $\forall p(\mathrm{I}_{\langle\rangle}X \to p \to X(A\neg p \neq B\neg p))$. But we also have $\Box(A\neg p \neq B\neg p \to \Diamond p)$ and hence $\mathrm{N}X \to X(A\neg p \neq B\neg p \to \Diamond p)$; also $\vdash \mathrm{K}X \to X(A\neg p \neq B\neg p \to \Diamond p) \to X(A\neg p \neq B\neg p) \to X\Diamond p$. Putting these together, we have $\vdash \mathrm{NKI}_{\langle\rangle}X \to \forall p(p \to X\Diamond p)$. $\Box$

The properties on the above list and their necessitations are the sorts of things that might with more or less plausibility be included in accounts of "being a necessity operation" in the sense relevant to claims like 'metaphysical necessity is the broadest necessity operation'. In §8.2, we considered six candidate definitions of 'necessity operation' in the setting of Classicism, four of which were N, $\Box$N, NK, and $\Box$NK.[3] We tentatively suggested that $\Box$NK ("Nec_6") was the best of these candidates; we noted that $\Box\mathrm{NK}X$ is a sufficient (though not necessary) condition for $P[X/\Box]$ to be true whenever $\vdash_{\mathsf{H}_{\mathsf{KA}}} P$, and that $\Box$ is itself $\Box$NK and entails every other $\Box$NK operation. But the special structural role of $\Box$ which undergirds its claim to the label 'metaphysical necessity' goes well beyond its distinctive position in the class of $\Box$NK operators. That structural role is summed up by the following three facts:

(A) $\Box$ materially implies every N operation: $\vdash \forall X(\mathrm{N}X \to \forall p(\Box p \to Xp))$.

(B) $\Box$ *entails* every $\Box$N operation: $\vdash \forall X(\Box\mathrm{N}X \to \Box\forall p(\Box p \to Xp))$.

(C) $\Box$ is itself $\Box$NK4T.

(They are trivial to prove: (A) just unpacks the definition of N; (B) follows from (A) by necessitation and the K axiom, while (C) is immediate from the K, 4, and T axioms of H_{S4}.) Given (A) and (C), it follows that for any property C of operations which is entailed by $\Box$NK4T and entails N, $\Box$ is extensionally minimal among the C operations, and indeed broadly necessarily equivalent to *having every C operation*. And given (B), if C also entails $\Box$N, $\Box$ also *entails* all the C operations.

Later in §8.2, we said that if ND fails for $\Box_\top$, Williamson's argument for ND from the logic of 'actually' (Williamson 1996; see §4.2) might reasonably be regarded as a good reason to deny that metaphysical necessity is identical to or coextensive with $\Box_\top$. We suggested that

[3] The remaining two were $\lambda X.X^*\mathrm{N}X$ and $\lambda X.X^*\mathrm{NK}X$, where X^* is the ancestral of X, defined like $\Box^*$ (see Appendix A).

those who accept this argument should not give up on the slogan 'metaphysical necessity is the broadest necessity', but should refine their understanding of what 'necessity operation' means in the slogan, so as to include "respecting ND". Let's assume that $\Box$ is identity-based, so that we can work with J rather than than the various I_σ properties. Then perhaps the most plausible candidate for the refined sense of 'necessity operation' is $\Box$NKJ. This might be deemed too strong; but at the very least, any "ND-respecting" necessity operation should be NKJ. Happily, we don't need to spend a lot of time looking at the different possible ways of precisifying the refinement, since as we will show, there is a particular operation $\Box_{\neq}$ that occupies a distinctive structural position within all of the candidate classes. Its role is summed up by the following facts:

($A_{\neq}$) $\Box_{\neq}$ materially implies every NKJ operation.

($B_{\neq}$) $\Box_{\neq}$ entails every $\Box$NKJ operation.

($C_{\neq}$) $\Box_{\neq}$ is itself $\Box$NK4TJ. (By Proposition D1 this implies that it is also $\Box I_\sigma$ for every type σ.)

(($A_{\neq}$) and ($C_{\neq}$) will be proved below after we have introduced the definition of $\Box_{\neq}$; ($B_{\neq}$) follows immediately from ($A_{\neq}$) by necessitation and the K axiom.) ($A_{\neq}$) and ($C_{\neq}$) together imply that for any property of operations *C*, so long as *C* is entailed by $\Box$NK4TJ and entails NKJ, $\Box_{\neq}$ is extensionally minimal among the *C* operations, and indeed broadly necessarily equivalent to *having every C operation*. ($B_{\neq}$) implies that if *C* also entails $\Box$NKJ, then $\Box_{\neq}$ *entails* all the *C* operations. Thus, to the extent that one is convinced by the case for ND, this should make $\Box_{\neq}$ a better candidate than $\Box$ to be metaphysical necessity (if the two are distinct).

§8.2 went on to observe that there are arguments for BF somewhat similar to Williamson's argument for ND, using some other principles about the interaction between metaphysical necessity and 'actually'. These arguments might suggest that "respecting BF" belongs alongside "respecting ND" in an account of the relevant sense of "necessity operation". As with ND, we don't need to spend a lot of time deciding what exactly "respecting BF" should mean, because there is a particular operation $\Box_5$ whose claim to be identical to or at least coextensive with metaphysical necessity will emerge no matter what we decide. Its structural role can be summed up by the following results:

(A_5) $\Box_5$ materially implies every $NKJF_{\langle\rangle}$ operation.

(B_5) $\Box_5$ entails every $\Box NKJF_{\langle\rangle}$ operation.

(C_5) $\Box_5$ is itself $\Box$NK4TB. (As we will show, this implies that it is also $\Box JI_\sigma F_\sigma$, for every type σ.)

As with $\Box_{\neq}$, (A_5) and (C_5) jointly imply that $\Box_5$ is broadly necessarily equivalent to *having every C operation* whenever *C* is entailed by $\Box$NK4TB and entails $NKJF_{\langle\rangle}$. (B_5), which follows from ($A_{\neq}$) by necessitation and the K axiom, implies that $\Box_5$ furthermore *entails* every *C* operation if *C* entails $\Box NKJF_{\langle\rangle}$. And ($C_5$) moreover means that the logic of $\Box_5$ is nice in several important further respects. $\Box$NK4TB is naturally glossed as 'S5 necessity operation', given that whenever $\vdash_{H_{S5}} P$, $\vdash_{H_{S4}} \Box NK4TBX \rightarrow P[X/\Box]$.[4] Then ($C_5$) and ($B_5$)

[4] This definition of 'S5 operator' is suggested in Dorr 2016*b* (70 and n. 62); (B_5) and (C_5) establish the conjecture there that "*being mapped to a truth by every S5 operator* . . . will turn out to itself be and S5 operator".

imply that $\Box_5$ is an S5 necessity operation that entails every other, and is broadly equivalent to *having every S5 necessity operation*.

There is one further way form of "logical well-behavedness" that is plausible for metaphysical necessity, and which is not guaranteed even by its being $\Box$NK4TB, namely the truth of the rigidity axioms for metaphysical necessity. Even if we assume the rigidity axioms for $\Box$—i.e. strengthen the background logic to $\mathsf{H}_{\mathsf{S4R}}$—the claim that an operation X is $\Box$NK4TB does not guarantee that it obeys the rigidity axioms. There is no issue about Rigid Comprehension, which does not contain $\Box$. And the broad necessity of Persistent Rigidity and Persistence automatically guarantee that any N operation will obey these principles, and any $\Box$N operations will broadly necessarily obey them. But Inextensibility is not automatic: for example, it is plausible that *having chance 1* is $\Box$NK4TB, but not plausible that it obeys Inextensibility (since the collection of all propositions of the form *the dart doesn't land on this point* doesn't have chance 1 of containing only truths even though each of its members has chance 1 of being true). Still, we will show in $\mathsf{H}_{\mathsf{S4R}}$ that $\Box_{\neq}$ and $\Box_5$ both do obey Inextensibility; thus all theorems of $\mathsf{H}_{\mathsf{S4R}}$ hold when $\Box$ is replaced with $\Box_{\neq}$ and $\Box_5$, and all theorems of $\mathsf{H}_{\mathsf{S5R}}$ hold for $\Box_5$.

We haven't yet defined $\Box_{\neq}$ and $\Box_5$, and nothing of philosophical significance turns on the choice of definition: facts ($\mathrm{A}_{\neq}$), ($\mathrm{C}_{\neq}$), (A_5), and (C_5) imply that in each case there is a whole array of possible definitions of the form $\lambda p.\forall X(CX \rightarrow Xp)$ which will be provably broadly equivalent. So we are free to choose whatever definitions will make the proofs go most smoothly. We have found it most convenient to pick C to be $\Box$NKJ in the case of $\Box_{\neq}$ and $\Box\mathrm{NKJF}_{\langle\rangle}$ in the case of $\Box_5$. So here (finally) are the official definitions:

$$\Box_{\neq} := \lambda p.\forall X(\Box\mathrm{NKJ}X \rightarrow Xp)$$
$$\Box_5 := \lambda p.\forall X(\Box\mathrm{NKJF}_{\langle\rangle}X \rightarrow Xp)$$

In what follows, we will first lay out some relations between the properties B, J, I_σ, and F_σ, thus making good on the parenthetical promise in (C_5). Next we will prove a general result that implies ($\mathrm{A}_{\neq}$) and (A_5), before turning to the proof of (C_5) and ($\mathrm{C}_{\neq}$). Finally, we will introduce some alternative definitions of $\Box_{\neq}$ and $\Box_5$ which make it easy to show (in $\mathsf{H}_{\mathsf{S4R}}$) that Iteration holds for them.

By a result of Prior's (1957), the B axiom and its necessitation jointly imply BF in H_{KA} (see note 47 in Chapter 1). Corresponding to this, we have the following fact:

Proposition D3. $\vdash \mathrm{NKB}X \wedge X\mathrm{NKB}X \rightarrow \mathrm{F}_\sigma X$.

Proof:

1.	$\vdash \Box(\neg X\neg\neg Yz \rightarrow \neg\forall z(Xz))$	UI, PC, NEC	
2.	$\vdash \mathrm{N}X \rightarrow X(\neg X\neg\neg Yz \rightarrow \neg\forall z(XYz))$	1, Df. N	
3.	$\vdash \mathrm{NK}X \rightarrow X\neg X\neg\neg Yz \rightarrow X\neg\forall z(XYz)$	2, Df. K	
4.	$\vdash \mathrm{NKB}X \rightarrow \neg Yz \rightarrow X\neg\forall z(XYz)$	3, Df. B	
5.	$\vdash \mathrm{NKB}X \rightarrow \neg X\neg\forall z(XYz) \rightarrow \forall zYz$	4, PC, GEN	
6.	$\vdash \Box(\mathrm{NKB}X \rightarrow \neg X\neg\forall z(XYz) \rightarrow \forall zYz)$	5, NEC	
7.	$\vdash \mathrm{N}X \rightarrow X(\mathrm{NKB}X \rightarrow \neg X\neg\forall z(XYz) \rightarrow \forall zYz)$	6, Df. N	
8.	$\vdash \mathrm{NK}X \rightarrow X\mathrm{NKB}X \rightarrow X\neg X\neg\forall z(XYz) \rightarrow X\forall zYz$	7, Df. K	
9.	$\vdash \mathrm{NKB}X \rightarrow X\mathrm{NKB}X \rightarrow \forall z(XYz) \rightarrow X\forall zYz$	8, Df. B	□

Similarly, corresponding to the Priorian proof of ND from the B axiom in H_{KA} (see note 46 in Chapter 1), we have the following fact:

Proposition D4. $\vdash \mathrm{NKB}X \wedge X\mathrm{N}X \to \mathrm{J}X$.

Proof:

1.	$\vdash \Box(\mathrm{N}X \to \neg X \neg p \to \Diamond p)$	Df. N
2.	$\vdash \mathrm{N}X \to X(\mathrm{N}X \to \neg X \neg p \to \Diamond p)$	1, Df. N
3.	$\vdash \mathrm{NK}X \to X\mathrm{N}X \to X(\neg X \neg p \to \Diamond p)$	2, Df. K
4.	$\vdash \mathrm{NK}X \to X\mathrm{N}X \to X\neg X \neg p \to X\Diamond p$	3, Df. K
5.	$\vdash \mathrm{NKB}X \to X\mathrm{N}X \to p \to X\Diamond p$	4, Df. B

□

By Proposition D1 this implies:

Corollary D5. $\vdash \mathrm{NKB}X \wedge X\mathrm{N}X \to \mathrm{I}_\sigma X$.

Also, since $\Box\mathrm{NKB}X$ entails $\Box X\mathrm{NKB}X$, Propositions D3 and D4 together imply:

Corollary D6. $\vdash \Box\mathrm{NKB}X \to \Box\mathrm{NKJF}_\sigma X$.

One other fact about the I_σ properties is worth noting, namely that in the presence of NK, all the different I_σ properties except for I_e are equivalent (I_e may be had without the rest):

Proposition D7. $\vdash \mathrm{NKI}_{\langle\sigma_1,\ldots,\sigma_n\rangle}X \to \mathrm{I}_{\sigma_0}X$ for any types $\sigma_0 \ldots \sigma_n$.

Proof: Let y, z be of type σ_0 and $u_1, \ldots, u_n$ of types $\sigma_1, \ldots, \sigma_n$. Since $\vdash y \neq z \leftrightarrow (\lambda u_1 \ldots u_n.y = z) \neq (\lambda u_1 \ldots u_n.y = y)$, we have both

$$\vdash \mathrm{I}_{\langle\sigma_1,\ldots,\sigma_n\rangle}X \to y \neq z \to X((\lambda u_1 \ldots u_n.y = z) \neq (\lambda u_1 \ldots u_n.y = y))$$

and

$$\vdash \mathrm{NK}X \to X((\lambda u_1 \ldots u_n.y = z) \neq (\lambda u_1 \ldots u_n.y = y)) \to X(y \neq z)$$

Putting these together gives: $\vdash \mathrm{NKI}_{\langle\sigma_1,\ldots,\sigma_n\rangle}X \to y \neq z \to X(y \neq z)$. □

The next result is the important one for establishing (A_5) and $(\mathrm{A}_{\neq})$:

Proposition D8. For any conjunction C of N, K, J, I_σ, F_σ: every C operation is coextensive with one that is $\Box C$, i.e. $\vdash \forall X(CX \to \exists X'(\Box CX' \wedge \forall p(Xp \leftrightarrow X'p)))$

Proof: Suppose C is one of those conjunctions, and X is an C operation. Let $X' := \lambda p.(CX \to Xp)$. X' is obviously coextensive with X. Moreover we can show in each case that $\Box CX'$. If N is a conjunct of C, then $\Box\mathrm{N}X'$, i.e., $\Box\forall p(\Box p \to CX \to Xp)$, by definition of N. Similarly if K is a conjunct we have $\Box\forall p\forall q((CX \to X(p \to q)) \to (CX \to Xp) \to (CX \to Xq))$. If J is a conjunct we have $\Box\forall p(p \to CX \to X\Diamond p)$. If I_σ is a conjunct we have $\Box\forall y\forall z(y \neq z \to CX \to X(y \neq z))$. And if F_σ is a conjunct we have $\Box\forall F^{\langle\sigma\rangle}((CX \to X\forall zYz) \to \forall z(CX \to XYz))$.[5] □

[5] The reasoning also extends to the case where C has NK4 as a conjunct (although not where C includes 4 without NK). Looking at the definitions of N, K, and 4 we can see that

$$\Box\forall p(\mathrm{N}X \to X(Xp \to (CX \to Xp)))$$
$$\Box\forall p(\mathrm{K}X \to X(Xp \to (CX \to Xp)) \to XXp \to X(CX \to Xp))$$
$$\Box\forall p(4X \to Xp \to XXp)$$

When C entails NK4, we can combine these to get $\Box\forall p(CX \to Xp \to X(CX \to Xp))$, which is equivalent to $\Box\forall p((CX \to Xp) \to CX \to X(CX \to Xp))$, or $\Box 4X'$. We can also extend the list to include T by changing the definition of X' to $\lambda p.(CX \to Xp) \wedge (CX \to p)$; the reasoning for the other properties still goes through with this more complicated definition. Proposition D8 does not extend to the case where C is B or NKB. Something a bit more complicated is true, though: if $\mathrm{NKB}X$ and $X\mathrm{NKB}X$, then $\Box\mathrm{NKB}X'$ for some X' coextensive with X.

Given this, (A_5) and ($A_{\neq}$) are straightforward:

Theorem D9. $\vdash \forall X(\mathrm{NKJ}X \to \forall p(\Box_{\neq} p \to Xp))$ and $\vdash \forall X(\mathrm{NKJF}_{\langle\rangle}X \to \forall p(\Box_5 p \to Xp))$.

Proof: Suppose NKJX and $\Box_{\neq} p$. By Proposition D8, there is some X' coextensive with X such that $\Box\mathrm{NKJ}X'$. By definition of $\Box_{\neq} p$ we must have $X'p$; so Xp by the coextensiveness of X and X'. Similarly for $\Box_5$. □

Facts (B_5) and ($B_{\neq}$) follow immediately by necessitation:

Corollary D10. $\vdash \forall X(\Box\mathrm{NKJ}X \to \Box\forall p(\Box_{\neq} p \to Xp))$ and $\vdash \forall X(\Box\mathrm{NKJF}_{\langle\rangle}X \to \Box\forall p(\Box_5 p \to Xp))$.

To establish (C_5) and ($C_{\neq}$), we will need to consider a couple of ways of making new operations out of old ones. For any operation X, define

$$X^2 := \lambda p.X(Xp)$$
$$\widehat{X} := \lambda p.\exists q(q \wedge \Box(\neg X\neg q \to p))$$

While the significance of X^2 is self-explanatory, $\widehat{X}$ is less familiar. The following fact helps to illuminate its relation to X:

Proposition D11. $\vdash p \to \widehat{X}\neg X\neg p$ and $\vdash \Box\mathrm{NK}X \to \neg\widehat{X}\neg Xp \to p$.

Proof: Trivially, $\vdash p \to (p \wedge \Box(\neg X\neg p \to \neg X\neg p))$, and hence by $\mathrm{EG}_{\mathrm{sub}}$, $\vdash p \to \exists q(q \wedge \Box(\neg X\neg q \to \neg X\neg p))$, i.e. $\vdash p \to \widehat{X}\neg X\neg p$. For the second part, $\Box\mathrm{NK}X$ implies $\neg p \to (\neg p \wedge \Box(\neg X\neg\neg p \to \neg Xp))$, and hence by $\mathrm{EG}_{\mathrm{sub}}$, $\neg p \to \exists q(q \wedge \Box(\neg X\neg q \to \neg Xp))$, i.e. $\neg p \to \widehat{X}\neg Xp$. □

The structure of Proposition D11 is suggestively analogous to the mixing axioms $p \to \mathsf{G}\neg\mathsf{H}\neg p$ and $\neg\mathsf{G}\neg\mathsf{H}p \to p$ in tense logic; it suggests that one can think of $\widehat{X}$ as a kind of "left inverse" of X. This naturally suggests the question under what circumstances we can swap X and $\widehat{X}$ in Proposition D11. As it turns out, $\mathrm{NKJF}_{\langle\rangle}X$ is a sufficient condition for this:

Proposition D12. $\vdash \mathrm{NKJF}_{\langle\rangle}X \to p \to X\neg\widehat{X}\neg p$ and $\vdash \mathrm{NKJF}_{\langle\rangle}X \to \neg X\neg\widehat{X}p \to p$

Proof: For the first part:

1.	$\vdash p \to \forall q(X\neg q \vee (\neg X\neg q \wedge p))$	PC, GEN
2.	$\vdash \mathrm{J}X \to p \to \forall q(X\neg q \vee X\Diamond(\neg X\neg q \wedge p))$	1, Df. J
3.	$\vdash \mathrm{NK}X \to \forall q(X\neg q \to X(q \to \Diamond(\neg X\neg q \wedge p)))$	Df. NK
4.	$\vdash \mathrm{NK}X \to \forall q(X\Diamond(\neg X\neg q \wedge p) \to X(q \to \Diamond(\neg X\neg q \wedge p)))$	Df. NK
5.	$\vdash \mathrm{NKJ}X \to p \to \forall qX(q \to \Diamond(\neg X\neg q \wedge p))$	2, 3, 4
6.	$\vdash \mathrm{NKJF}_{\langle\rangle}X \to p \to X\forall q(q \to \Diamond(\neg X\neg q \wedge p))$	5, Df. $\mathrm{F}_{\langle\rangle}$
7.	$\vdash \mathrm{NKJF}_{\langle\rangle}X \to p \to X\neg\exists q(q \wedge \Box(\neg X\neg q \to \neg p))$	6, Df. NK
8.	$\vdash \mathrm{NKJF}_{\langle\rangle}X \to p \to X\neg\widehat{X}\neg p$	7, Df. $\widehat{X}$

The proof for the second part is the same, substituting $\neg p$ for p and p for $\neg p$ throughout. □

Unsurprisingly, various modalized properties of X are inherited by X^2:

Proposition D13. (i) If $\Box \mathrm{N}X$ then $\Box \mathrm{N}X^2$; (ii) if $\Box \mathrm{NK}X$ then $\Box \mathrm{NK}X^2$; (iii) if $\Box \mathrm{NKJ}X$ then $\Box \mathrm{NKJ}X^2$; (iv) if $\Box \mathrm{NKF}_\sigma X$ then $\Box \mathrm{NKF}_\sigma X^2$

Proof: For (i): Suppose $\Box \mathrm{N}X$. Then $\Box \forall p(\Box p \to Xp)$, so $\Box \forall p \Box(\Box p \to Xp)$ (by 4 and CBF), so $\Box \forall p(\Box p \to \Box Xp)$, so $\Box \forall p(\Box p \to XXp)$ (by NX).

(ii): Suppose $\Box \mathrm{NK}X$ Then by $\Box \mathrm{K}X$, $\Box \forall p \forall q(\Box(X(p \to q) \to (Xp \to Xq)))$, so $\Box \forall p \forall q(X(X(p \to q) \to (Xp \to Xq)))$ since $\Box \mathrm{N}X$, so $\Box \forall p \forall q(XX(p \to q) \to X(Xp \to Xq))$ and hence $\Box \forall p \forall q(XX(p \to q) \to (XXp \to XXq))$ given $\Box \mathrm{K}X$.

(iii): Suppose $\Box \mathrm{NKJ}X$. Since $\Box \mathrm{J}X$, $\Box \forall p \Box(p \to X \Diamond p)$, so $\Box \forall p(Xp \to XX \Diamond p)$ since $\Box \mathrm{NK}X$. Instantiating p with $\Diamond p$, we get $\Box \forall p(X \Diamond p \to XX \Diamond \Diamond p)$. Using $\Box \mathrm{J}X$ again, this becomes $\Box \forall p(p \to XX \Diamond \Diamond p)$, and finally by S4 and $\mathrm{NK}X^2$ (see (i) and (ii)), we have $\Box \forall p(p \to XX \Diamond p)$, i.e. $\Box \mathrm{J}X^2$.

(iv): Suppose $\Box \mathrm{NKF}_\sigma X$. We have $\Box \forall Y \Box(\forall z XYz \to X \forall z Yz)$, so $\Box \forall Y(X \forall z Xyz \to XX \forall z Yz)$ since $\Box \mathrm{NK}X$. But also $\Box \forall Y(\forall z XXyz \to X \forall z XYz)$ since $\Box \mathrm{F}_\sigma X$. Hence, $\Box \forall Y(\forall z XXyz \to XX \forall z Yz)$, i.e. $\Box \mathrm{F}_\sigma X$. □

We can prove some facts of a similar sort about $\widehat{X}$.

Proposition D14. $\vdash \Box \mathrm{NK}X \to \Box \mathrm{NKJ}\widehat{X}$.

Proof: Since H_{S4} is closed under necessitation, it suffices to show $\mathrm{NKJ}\widehat{X}$ on the assumption that $\Box \mathrm{NK}X$.

$\mathrm{N}\widehat{X}$ is just $\forall p(\Box p \to \exists q(q \wedge \Box(\neg X \neg q \to p)))$ which is trivial.

For $\mathrm{K}\widehat{X}$, suppose that $\widehat{X}(p \to r) \wedge \widehat{X}p$, i.e. that for some q and q',

$$q \wedge \Box(\neg X \neg q \to (p \to r)) \wedge q' \wedge \Box(\neg X \neg q' \to p)$$

It follows that

$$q \wedge q' \wedge \Box((\neg X \neg q \wedge \neg X \neg q') \to r)$$

And since $\Box \mathrm{NK}X$, we have $\Box(\neg X \neg(q \wedge q') \to (\neg X \neg q \wedge \neg X \neg q'))$. So we get

$$q \wedge q' \wedge \Box(\neg X \neg(q \wedge q') \to r)$$

which entails $\exists q(q \wedge \Box(\neg X \neg q \to r))$, i.e. $\widehat{X}r$.

For J, note that since $\Box \mathrm{NK}X$, we have $\Box(\neg X \neg p \to \Diamond p)$ for any p; hence $p \to (p \wedge \Box(\neg X \neg p \to \Diamond p))$, so by EG, $p \to \exists q(q \wedge \Box(\neg X \neg q \to \Diamond p))$, i.e. $p \to \widehat{X} \Diamond p$. □

Proposition D15. $\vdash \Box \mathrm{NKF}_{\langle\rangle} X \to \Box \mathrm{NKJF}_\sigma \widehat{X}$ for every σ.

Proof: Given Proposition D14 it suffices to prove $\Box \mathrm{NKJF}_{\langle\rangle} X \to \mathrm{F}_\sigma \widehat{X}$. Assume $\Box \mathrm{NKF}_{\langle\rangle} X$. Let y be a variable of type σ. Since $\Box \mathrm{NK}X$, $\Box(\neg X \neg \forall z \widehat{X} Yz \to \forall z \neg X \neg \widehat{X} Yz)$. By Proposition D12, this yields $\Box(\neg X \neg \forall z \widehat{X} Yz \to \forall z Yz)$. Since $\Box \mathrm{NK}\widehat{X}$ (by Proposition D14), we can infer that $\Box(\widehat{X} \neg X \neg \forall z \widehat{X} Yz \to \widehat{X} \forall z Yz)$. But by Proposition D11, $\Box(\forall z \widehat{X} Yz \to \widehat{X} \neg X \neg \forall z \widehat{X} Yz)$; so $\Box(\forall z \widehat{X} Yz \to \widehat{X} \forall z Yz)$. □

Now we are ready to prove facts $(\mathrm{C}_{\neq})$ and (C_5).

Theorem D16. $\vdash \Box \mathrm{NKTJ4}\Box_{\neq}$ and $\vdash \Box \mathrm{NKTB4}\Box_5$.

Proof: Since H_{S4} is closed under necessitation, it suffices to prove the un-necessitated versions.

For N: $\vdash \Box \forall p(\Box p \to \forall X(CX \to Xp))$ whenever C entails N, as both $\Box \mathrm{NKJ}$ and $\Box \mathrm{NKJF}_{\langle\rangle}$ do.

For K: Whenever C entails K, as both $\Box$NKJ and $\Box$NKJF$_{\langle\rangle}$ do, we have $\vdash \forall X(CX \to X(p \to q)) \to \forall X(CX \to Xp) \to \forall X(CX \to Xq)$.

For T: Since $\vdash \Box\text{NKJF}_{\langle\rangle}(\lambda q.q)$, $\vdash \Box\forall p(\Box_{\neq} p \to (\lambda q.q)p)$ and $\vdash \Box\forall p(\Box_5 p \to (\lambda q.q)p)$.

For J: $\vdash \forall p(p \to \forall X(CX \to X\Diamond p))$ whenever C entails J, as $\Box$NKJ does.

For B:

1. $\vdash \forall X(\Box\text{NKJF}_{\langle\rangle}X \to \Box\text{NKJF}_{\langle\rangle}\widehat{X})$	Proposition D15
2. $\vdash \forall X(\Box\text{NKJF}_{\langle\rangle}X \to X\Box\text{NKJF}_{\langle\rangle}\widehat{X})$	1, Df. N
3. $\vdash p \to \forall X(\Box\text{NKJF}_{\langle\rangle}X \to X\neg\widehat{X}\neg p)$	Proposition D12
4. $\vdash p \to \forall X(\Box\text{NKJF}_{\langle\rangle}X \to X(\Box\text{NKJF}_{\langle\rangle}\widehat{X} \wedge \neg\widehat{X}\neg p))$	2, 3, Df. NK
5. $\vdash p \to \forall X(\Box\text{NKJF}_{\langle\rangle}X \to X(\exists Y(\Box\text{NKJF}_{\langle\rangle}Y \wedge \neg Y\neg p)))$	4, Df. NK
6. $\vdash p \to \Box_5\neg\Box_5\neg p$	5, Df. $\Box_5$

For 4: We have already proved $\Box$NKJ$\Box_{\neq}$, so by Proposition D13 we also have $\Box$NKJ$\Box_{\neq}{}^2$, and thus $\Box\forall p(\Box_{\neq} p \to \Box_{\neq}{}^2 p)$ by fact (B$_{\neq}$). Likewise, we have proved $\Box$NKB$\Box_5$, which implies $\Box\text{NKJF}_{\langle\rangle}\Box_5$ by Corollary D6, so $\Box\text{NKJF}_{\langle\rangle}\Box_5{}^2$ by Proposition D13, and thus $\Box\forall p(\Box_5 p \to \Box_5{}^2 p)$ by fact (B$_5$). □

There is one more thing we need to check, namely that when we strengthen the background logic of $\Box$ from H$_{\text{S4}}$ to H$_{\text{S4R}}$, we also get H$_{\text{S4R}}$ for $\Box_{\neq}$ and $\Box_5$ (and thus H$_{\text{S5R}}$ for $\Box_5$). Happily, showing this will also provide an occasion for some further exploration of the interesting $\widehat{\ }$ operation introduced above. First, this operation can be used to give an alternative characterization of $\Box_{\neq}$:[6]

Proposition D17. $\vdash \Box\forall p(\Box_{\neq} p \leftrightarrow \widehat{\Box} p)$

Proof, left to right: $\Box$NKJ$\widehat{\Box}$ by Proposition D14, so $\Box_{\neq}$ entails $\widehat{\Box}$ by fact (B$_{\neq}$).

Right-to-left:

1.	$\vdash q \to \forall X(\text{J}X \to X\Diamond q)$	Df. J
2.	$\vdash \Box(\Diamond q \to p) \to \forall X(\text{NK}X \to (X\Diamond q \to Xp))$	Df. NK
3.	$\vdash \exists q(q \wedge \Box(\Diamond q \to p)) \to \forall X(\text{NKJ}X \to X(\Diamond q \to p))$	1, 2
4.	$\vdash \widehat{\Box} p \to \Box_{\neq} p$	3, Df. $\widehat{\Box}$, Df. $\Box_{\neq}$.

□

Once we move to H$_{\text{S4R}}$, we can establish a companion fact about $\Box_5$, namely that it is broadly equivalent to $\widehat{\widehat{\Box}}$. (We leave it as an open question whether this can be shown already in H$_{\text{S4}}$.) In fact, showing this only requires the following consequence of the rigidity axioms (see Proposition C7):

$\Box$Actuality $\quad \Box\exists p\, \text{W}_@\, p$

where

$$\text{W}_@ := \lambda p.p \wedge \forall q(q \to \Box(p \to q))$$

This provides a helpful alternative characterization of $\widehat{\ }$ when applied to $\Box$NK operations (such as $\Box$ and $\widehat{\Box}$).

[6] Thanks to Jeremy Goodman, from whom we learned about $\Box_{\neq}$ via this definition of it.

Proposition D18. $\vdash_{\mathsf{H}_{\mathsf{S4R}}} \Box\mathrm{NK}X \to \Box\forall p(\widehat{X}p \leftrightarrow \forall w(\mathrm{W}_@\, w \to \Box(\neg X\neg w \to p)))$.

Proof: The claim follows by necessitation, K, and the definition of $\widehat{\ }$ from

$$\vdash_{\mathsf{H}_{\mathsf{S4R}}} \Box\mathrm{NK}X \to \forall p(\exists q(q \wedge \Box(\neg X\neg q \to p)) \leftrightarrow \forall w(\mathrm{W}_@\, w \to \Box(\neg X\neg w \to p))).$$

For the left-to-right direction, suppose $\mathrm{NK}X$, q, $\Box(\neg X\neg q \to p)$, and $\mathrm{W}_@\, w$. Then $\Box(w \to q)$, so by $\Box\mathrm{NK}X$, $\Box(\neg X\neg w \to \neg X\neg q)$ and hence $\Box(\neg X\neg w \to p)$. For the right-to-left direction, note that since $\vdash_{\mathsf{H}_{\mathsf{S4R}}} \exists w(\mathrm{W}_@\, w)$ and $\vdash \mathrm{W}_@\, w \to w$, $\vdash_{\mathsf{H}_{\mathsf{S4R}}} \forall w(\mathrm{W}_@\, w \to \Box(\neg X\neg w \to p)) \to \exists q(q \wedge \Box(\neg X\neg q \to p))$. □

One payoff is a strengthened version of Proposition D15:

Proposition D19. $\vdash_{\mathsf{H}_{\mathsf{S4R}}} \Box\mathrm{NKJ}X \to \Box\mathrm{F}_\sigma\widehat{X}$.

Proof: By Proposition D18, it suffices to show that

$$\vdash \Box\mathrm{NKJ}X \to \forall w(\mathrm{W}_@\, w \to \forall z\Box(\neg X\neg w \to Yz) \to \Box(\neg X\neg w \to \forall z Yz))$$

We will give a proof from assumptions: assume (a) $\mathrm{W}_@\, w$; (b) $\Box\mathrm{N}X$; (c) $\Box\mathrm{K}X$; (d) $\Box\mathrm{J}X$; (e) $\forall z\Box(\neg X\neg w \to Yz)$. Note that given (b) and (c), (d) is equivalent to the dual (d′): $\Box\forall p(\neg X\neg\Box p \to p)$ which will be more useful here.

1.	$\Box(w \to \forall z\Box(\neg X\neg w \to Yz))$	(a), (e)
2.	$\Box X(w \to \forall z\Box(\neg X\neg w \to Yz))$	(b), 1
3.	$\Box(\neg X\neg w \to \neg X\neg\forall z\Box(\neg X\neg w \to Yz))$	(c), 2
4.	$\Box(\neg X\neg w \to \forall z\neg X\neg\Box(\neg X\neg w \to Yz))$	(b), (c), 3
5.	$\Box(\neg X\neg w \to \forall z(\neg X\neg w \to Yz))$	(d′), 4
6.	$\Box(\neg X\neg w \to \forall z Yz)$	8, H_K

□

Combining Propositions D19 and D14, we have

Corollary D20. $\vdash_{\mathsf{H}_{\mathsf{S4R}}} \Box\mathrm{NK}X \to \Box\mathrm{NKJF}_\sigma\widehat{\widehat{X}}$.

Setting $X = \Box$ in this provides the advertised alternative characterization of $\Box_5$:

Proposition D21. $\vdash_{\mathsf{H}_{\mathsf{S4R}}} \Box\forall p(\Box_5 p \leftrightarrow \widehat{\widehat{\Box}}p)$

Proof, left to right: immediate from Corollary D20 and fact (B_5).

Right to left:

1. $\vdash q \to \forall p(\Box\neg p \vee (\Diamond p \wedge q))$	H_{S4}
2. $\vdash q \to \mathrm{NJ}X \to \forall p(X\neg p \vee X\Diamond(\Diamond p \wedge q))$	1, Df. $\mathrm{NB}_\Diamond$
3. $\vdash q \to \mathrm{NKJ}X \to \forall pX(p \to \Diamond(\Diamond p \wedge q))$	2, Df. NK
4. $\vdash q \to \mathrm{NKJF}_{\langle\rangle}X \to X\forall p(p \to \Diamond(\Diamond p \wedge q))$	3, Df. $\mathrm{F}_{\langle\rangle}$
5. $\vdash q \to \mathrm{NKJF}_{\langle\rangle}X \to X\widehat{\Diamond}q$	Df. $\widehat{\Box}$
6. $\vdash (q \wedge \Box(\widehat{\Diamond}q \to p)) \to (\mathrm{NKJF}_{\langle\rangle}X \to Xp)$	5, Df. NK
7. $\vdash \exists q(q \wedge \Box(\widehat{\Diamond}q \to p)) \to \forall X(\mathrm{NKJF}_{\langle\rangle}X \to Xp)$	6, GEN, ED, UD
8. $\vdash \widehat{\widehat{\Box}}p \to \Box_5 p$	7, Df. $\widehat{\widehat{\Box}}$, Df. $\Box_5$

□

These nice results make it natural to wonder about $\widehat{\widehat{\Box}}$—does it play some further interesting role? In $\mathsf{H}_{\mathsf{S4R}}$ at least the answer is no, because of the following fact:

Proposition D22. $\vdash \Box\mathrm{NKB}X \rightarrow \Box\forall p(Xp \leftrightarrow \widehat{X}p)$.

Proof: If $\Box\mathrm{NKB}X$ then $\Box\mathrm{NKJF}_{\langle\rangle}X$ (by Corollary D6). Hence, $\Box\forall p(Xp \rightarrow \widehat{X}\neg X\neg Xp)$ by D12; but $\Box\forall p(\neg X\neg Xp \rightarrow p)$ since $\mathrm{NKB}X$, so $\Box\forall p(\widehat{X}\neg X\neg Xp \rightarrow \widehat{X}p)$ since $\mathrm{NK}\widehat{X}$ (by Proposition D14); hence $\Box\forall p(Xp \rightarrow \widehat{X}p)$. Conversely, $\Box\forall p(\widehat{X}p \rightarrow X\neg X\neg\widehat{X}p)$ since $\Box BX$, but $\Box\forall p(\neg X\neg\widehat{X}p \rightarrow p)$ by Proposition D12, so $\Box\forall p(X\neg X\neg\widehat{X}p \rightarrow Xp)$ since $\mathrm{NK}X$; hence $\Box\forall p(\widehat{X}p \rightarrow Xp)$. □

But by Corollary D20 and Theorem D16, we know that $\vdash_{\mathsf{H}_{\mathsf{S4R}}} \Box\mathrm{NKB}\widehat{\widehat{\Box}}$ (in fact, $\Box\mathrm{NKTB4}\widehat{\widehat{\Box}}$), so we have:

Corollary D23. $\vdash_{\mathsf{H}_{\mathsf{S4R}}} \Box\forall p(\widehat{\widehat{\Box}}p \leftrightarrow \widehat{\widehat{\Box}}p)$.

That fact is specific to $\Box$: for an arbitrary X, there is no guarantee that $\widehat{\widehat{X}}$ is $\Box\mathrm{NKB}$, and so no guarantee that it coincides with $\widehat{\widehat{X}}$. But so long as X is $\Box\mathrm{NK}$, we get nothing new (up to broadly necessary equivalence) after the first *three* applications of $\widehat{\ }$, thanks to the following fact:

Proposition D24. $\vdash \Box\mathrm{NKJF}_{\langle\rangle}X \rightarrow \Box\forall p(Xp \leftrightarrow \widehat{\widehat{X}}p)$

Proof: Suppose $\Box\mathrm{NKJF}_{\langle\rangle}X$. Then by Proposition D12, we have both $\Box(\widehat{\widehat{X}}p \rightarrow X\neg\widehat{X}\neg\widehat{\widehat{X}}p)$ and $\Box(\neg\widehat{X}\neg\widehat{\widehat{X}}p \rightarrow p)$; given that $\Box\mathrm{NK}X$, the latter implies $\Box(X\neg\widehat{X}\neg\widehat{\widehat{X}}p \rightarrow Xp)$, so we can conclude that $\Box(\widehat{\widehat{X}}p \rightarrow Xp)$. For the other direction, we only need Proposition D11, which gives us both $\Box(Xp \rightarrow \widehat{\widehat{X}}\neg\widehat{X}\neg Xp)$ and $\Box(\neg\widehat{X}\neg Xp \rightarrow p)$; from the latter and $\Box\mathrm{NK}\widehat{\widehat{X}}$, we get $\Box(\widehat{\widehat{X}}\neg\widehat{X}\neg Xp \rightarrow \widehat{\widehat{X}}p)$; and hence $\Box(Xp \rightarrow \widehat{\widehat{X}}p)$. □

Given Corollary D20, one condition sufficient in $\mathsf{H}_{\mathsf{S4R}}$ for $\Box\mathrm{NKJF}_{\langle\rangle}X$ is for $X = \widehat{\widehat{Y}}$ for some Y such that $\mathrm{NK}Y$. So for any such Y, four applications of $\widehat{\ }$ amounts to the same thing as two.

Using our alternative characterizations of $\Box_{\neq}$ and $\Box_5$, it is easy to show that Inextensibility is true when $\Box$ is replaced by $\Box_{\neq}$ or $\Box_5$. Persistent Rigidity and Persistence are immediate from the fact that $\Box_{\neq}$ and $\Box_5$ are entailed by $\Box$. What remains is Inextensibility. It turns out that whenever $\Box\mathrm{NK}X$, $\widehat{X}$ obeys Inextensibility, i.e.

Proposition D25. $\vdash_{\mathsf{H}_{\mathsf{S4R}}} \Box\mathrm{NK}X \rightarrow \mathsf{R}C \rightarrow \forall\vec{z}(C\vec{z} \rightarrow \widehat{X}Y\vec{z}) \rightarrow \widehat{X}\forall\vec{z}(C\vec{z} \rightarrow Y\vec{z})$.

Proof: Suppose $\mathsf{R}C$, $\mathrm{W}_{@}\,w$, and $\forall\vec{z}(C\vec{z} \rightarrow \widehat{X}Y\vec{z})$. Then by Proposition D18, $\forall\vec{z}(C\vec{z} \rightarrow \Box(\neg X\neg w \rightarrow Y\vec{z}))$. So by Inextensibility, $\Box\forall\vec{z}(C\vec{z} \rightarrow \neg X\neg w \rightarrow Y\vec{z})$, or equivalently $\Box(\neg X\neg w \rightarrow \forall\vec{z}(C\vec{z} \rightarrow Y\vec{z}))$, which by D18 implies $\widehat{X}\forall\vec{z}(C\vec{z} \rightarrow Y\vec{z})$. □

Since $\Box\mathrm{NK}\Box$ and $\Box\mathrm{NK}\widehat{\Box}$, we can combine this with our $\widehat{\ }$-based characterizations of $\Box_{\neq}$ and $\Box_5$ (Propositions D17 and D21) to deduce that Inextensibility holds for them. And since the remaining rigidity axioms are unproblematic, we can conclude that:

Theorem D26. If $\vdash_{\mathsf{H}_{\mathsf{S4R}}} P$, then $\vdash_{\mathsf{H}_{\mathsf{S4R}}} P[\Box_{\neq}/\Box]$ and $\vdash_{\mathsf{H}_{\mathsf{S4R}}} P[\Box_5/\Box]$.

References

Adams, Robert Merrihew (1974). 'Theories of Actuality', *Noûs*, 8/3: 211–31.

Adams, Robert Merrihew (1979). 'Primitive Thisness and Primitive Identity', *Journal of Philosophy*, 76/1: 5–26.

Angell, R. Bradshaw (1977). 'Three Systems of First Degree Entailment', *Journal of Symbolic Logic*, 42/1:147.

Angell, R. Bradshaw (1989). 'Deducibility, Entailment and Analytic Containment', in J. Norman and R. Sylvan (eds.), *Directions in Relevant Logic* (Dordrecht: Kluwer): 119–43.

Arntzenius, Frank (2011). *Space, Time, and Stuff* (Oxford: Oxford University Press).

Arntzenius, Frank and Hawthorne, John (2005). 'Gunk and Continuous Variation', *The Monist*, 88: 441–65.

Bacon, Andrew (2013). 'Quantificational Logic and Empty Names', *Philosophers' Imprint*, 13.

Bacon, Andrew (2014). 'Representing Counterparts', *Australasian Journal of Logic*, 11/2: 90–113.

Bacon, Andrew (2018*a*). 'The Broadest Necessity', *Journal of Philosophical Logic*, 47/5: 733–83.

Bacon, Andrew (2018*b*). *Vagueness and Thought* (Oxford: Oxford University Press).

Bacon, Andrew (2019). 'Opacity and Paradox', in Carlo C. Nicolai and J. Johannes Stern (eds.), *Modes of Truth: The Unified Approach to Truth, Modalities, and Paradox* (London: Routledge) 231–65.

Bacon, Andrew (2020). 'Logical Combinatorialism', *Philosophical Review*, 129/4: 537–89.

Bacon, Andrew (unpublished). 'A Theory of Structured Propositions'.

Bacon, Andrew and Dorr, Cian (forthcoming). 'Classicism', in Fritz and Jones forthcoming.

Bacon, Andrew, Hawthorne, John, and Uzquiano, Gabriel (2016). 'Higher-Order Free Logic and the Prior-Kaplan Paradox', in Yli-Vakkuri and McCullagh 2017: 41–89. From *Canadian Journal of Philosophy*, 46/4–5 (2016): 493–541.

Bacon, Andrew and Russell, Jeffrey Sanford (2019). 'The Logic of Opacity', *Philosophy and Phenomenological Research*, 99/1: 81–114.

Bacon, Andrew and Uzquiano, Gabriel (2018). 'Some Results on the Limits of Thought', *Journal of Philosophical Logic*, 47/6: 991–9.

Ballarin, Roberta (2013). 'The Necessity of Origin: A Long and Winding Route', *Erkenntnis*, 78/2: 353–70.

Barcan, Ruth C. (= Marcus, Ruth Barcan) (1946). 'A Functional Calculus of First Order Based on Strict Implication', *Journal of Symbolic Logic*, 11/1:1–16.

Barnett, David (2005). 'The Problem of Material Origins', *Noûs*, 39: 529–40.

Bealer, George (1993). 'Universals', *Journal of Philosophy*, 90/1: 5–32.

Bealer, George (1996). 'On the Possibility of Philosophical Knowledge', *Philosophical Perspectives*, 10:1–34.

Belot, Gordon and Earman, John (2001). 'Pre-Socratic Quantum Gravity', in Craig Callender and Nick Huggett (eds.), *Physics Meets Philosophy at the Planck Scale* (Cambridge: Cambridge University Press): 213–55.

Benacerraf, Paul (1965). 'What Numbers Could Not Be', *Philosophical Review*, 74/1: 47–73.
Bennett, Karen (2004). 'Spatio-Temporal Coincidence and the Grounding Problem', *Philosophical Studies*, 118/3: 339–71.
Black, Max (1952). 'The Identity of Indiscernibles', *Mind*, 61/242:153–64.
Braun, David (1988). 'Understanding Belief Reports', *Philosophical Review*, 107/4: 555–95.
Bricker, Phillip (1996). 'Isolation and Unification: The Realist Analysis of Possible Worlds', *Philosophical Studies*, 84/2–3: 225–38.
Brighouse, Carolyn (1994). 'Spacetime and Holes', *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1994:117–25.
Brower-Toland, Susan (2007). 'Intuition, Externalism, and Direct Reference in Ockham', *History of Philosophy Quarterly*, 24/4: 317–36.
Brown, Jessica (2000). 'Against Temporal Externalism', *Analysis*, 60/2:178–88.
Burge, Tyler (1979). 'Individualism and the Mental', in Peter French, Theodore E. Uehling, Jr., and Howard K. Wettstein (eds.), *Midwest Studies in Philosophy*, iv: *Studies in Metaphysics* (Minneapolis, MN: University of Minnesota Press): 73–122.
Burgess, John (2008). 'Tarski's Tort', in *Mathematics, Models, and Modality: Selected Philosophical Essays* (Cambridge: Cambridge University Press): 149–68.
Burke, Michael B. (1992). 'Copper Statues and Pieces of Copper: A Challenge to the Standard Account', *Analysis*, 52/1:12–17.
Burke, Michael B. (1994). 'Dion and Theon: An Essentialist Solution to an Ancient Puzzle', *Journal of Philosophy*, 91/3:129–39.
Butler, Joseph (1736). 'The Analogy of Religion, Natural and Revealed, to the Constitution and Course of Nature', in *The Works of Bishop Butler*, ed. David E. White (new edition, Woodbridge: Boydell and Brewer, 2006): 147–314.
Butterfield, Jeremy (1989). 'The Hole Truth', *British Journal for the Philosophy of Science*, 40/1:1–28.
Byrne, Alex and Hilbert, David R. (1997). 'Colors and Reflectances', in A. Byrne and D. R. Hilbert (eds.), *Readings on Color*, i: *The Philosophy of Color* (Cambridge, MA: MIT Press): 263–88.
Caie, Michael, Goodman, Jeremy, and Lederman, Harvey (2020). 'Classical Opacity', *Philosophy and Phenomenological Research*, 101/3: 524–66.
Cappelen, Herman and Hawthorne, John (2009). *Relativism and Monadic Truth* (Oxford: Oxford University Press).
Carnap, Rudolf (1934). *The Logical Syntax of Language*, trans. Amethe Smeaton (London: Kegan Paul, Trench, Trubner, 1937) [Ger. orig., *Logische Syntax der Sprache* (Vienna: Springer, 1934)].
Carnap, Rudolf (1947). *Meaning and Necessity: A Study in Semantics and Modal Logic* (Chicago: University of Chicago Press).
Cartwright, Richard (1975). 'Scattered Objects', in Keith Lehrer (ed.), *Analysis and Metaphysics* (Dordrecht: D. Reidel): 153–71.
Casati, Roberto and Varzi, Achille C. (1999). *Parts and Places* (Cambridge, MA: MIT Press).
Chalmers, David J. (1996). *The Conscious Mind* (Oxford: Oxford University Press).
Chalmers, David J. (2005). 'The Matrix as Metaphysics', in C. Grau (ed.), *Philosophers Explore the Matrix* (Oxford: Oxford University Press).
Chalmers, David J. (2012*a*). *Constructing the World* (Oxford: Oxford University Press).
Chalmers, David J. (2012*b*). 'Twenty-First Excursus: Twin-Earthability and Internalism', http://consc.net/books/ctw/excursus21.pdf.
Chandler, Hugh S. (1975). 'Rigid Designation', *Journal of Philosophy*, 72/13: 363–9.
Chandler, Hugh S. (1976). 'Plantinga and the Contingently Possible', *Analysis*, 36/2:107–9.

Chisholm, Roderick (1967). 'Identity Through Possible Worlds: Some Questions', *Noûs*, 1: 1–8.
Chisholm, Roderick (1973). 'Parts as Essential to Their Wholes', *Review of Metaphysics*, 26: 581–603.
Chisholm, Roderick (1976). *Person and Object: A Metaphysical Study* (London: Routledge).
Chisholm, Roderick (1981). *The First Person* (Minneapolis, MN: University of Minnesota Press).
Church, Alonzo (1940). 'A Formulation of the Simple Theory of Types', *Journal of Symbolic Logic*, 5: 56–68.
Clark, Andy and Chalmers, David J. (1998). 'The Extended Mind', *Analysis*, 58: 7–19.
Clarke-Doane, Justin (2021). 'Metaphysical and Absolute Possibility', *Synthese*, 198: 1861–72.
Collins, John M. (2006). 'Temporal Externalism, Natural Kind Terms, and Scientifically Ignorant Communities', *Philosophical Papers*, 35/1: 55–68.
Correia, Fabrice (2010). 'Grounding and Truth-functions', *Logique et Analyse*, 53: 211–51.
Cover, J. A. and O'Leary-Hawthorne, John (= Hawthorne, John) (1999). *Substance and Individuation in Leibniz* (Cambridge: Cambridge University Press).
Craig, William (1953). 'On Axiomatizability Within a System', *Journal of Symbolic Logic*, 18/1: 30–2.
Cresswell, M. J. (1990). *Entities and Indices* (Dordrecht: Kluwer Academic Publishers).
Cresswell, M. J. and Hughes, G. E. (1996). *A New Introduction to Modal Logic* (London: Routledge).
Crossley, John N. and Humberstone, Lloyd (1977). 'The Logic of "Actually"', *Reports on Mathematical Logic*, 8: 11–29.
Dancy, Jonathan (2014). *Ethics Without Principles* (Oxford: Oxford University Press).
Deasy, Daniel (2015). 'The Moving Spotlight Theory', *Philosophical Studies*, 172: 2073–89.
Descartes, René (1644). *Principles of Philosophy*, trans. V. R. Miller and R. P. Miller (Dordrecht: Reidel, 1983).
Deutsch, Harry (2008). review of King 2007, in *Notre Dame Philosophical Reviews*, https://ndpr.nd.edu/news/the-nature-and-structure-of-content/-content/.
Ditter, Andreas (2020). 'The Metaphysics and Logic of Essence', PhD thesis (New York University).
Doepke, Frederick C. (1982). 'Spatially Coinciding Objects', *Ratio:* 10–24.
Dorr, Cian (2003). 'Vagueness Without Ignorance', *Philosophical Perspectives*, 17: 83–113.
Dorr, Cian (2005*a*). 'Counterparts and Propositions', *Analysis*, 65: 210–18.
Dorr, Cian (2005*b*). 'What We Disagree about When We Disagree about Ontology', in M. E. Kalderon (ed.), *Fictionalism in Metaphysics* (Oxford: Oxford University Press): 234–86.
Dorr, Cian (2010). 'Iterating Definiteness', in Richard Dietz and Sebastiano Moruzzi (eds.), *Cuts and Clouds: Vagueness, Its Nature, and Its Logic* (Oxford: Oxford University Press): 550–75.
Dorr, Cian (2011). 'Physical Geometry and Fundamental Metaphysics', *Aristotelian Society Supplementary Volume*, 111:135–59.
Dorr, Cian (2014*a*). 'Quantifier Variance and the Collapse Theorems', *The Monist*, 97: 503–70.
Dorr, Cian (2014*b*). 'Transparency and the Context-sensitivity of Attitude Reports', in Manuel García-Carpintero and Genoveva Martí (eds.), *Empty Representations: Reference and Non-existence* (Oxford: Oxford University Press), 25–66.
Dorr, Cian (2016*a*). 'Against Counterfactual Miracles', *Philosophical Review*, 125: 241–86.
Dorr, Cian (2016*b*). 'To Be F Is To Be G', *Philosophical Perspectives*, 30: 39–134.

Dorr, Cian (2019). 'Natural Properties', *Stanford Encyclopedia of Philosophy*, Fall 2019, ed. Edward N. Zalta, https://plato.stanford.edu/entries/natural-properties/.
Dorr, Cian (unpublished *a*). *Counterparts.*
Dorr, Cian (unpublished *b*). *The Multiplicity of Meaning.* The Hempel Lectures, Princeton, 2020.
Dorr, Cian and Goodman, Jeremy (2020). 'Diamonds Are Forever', *Noûs*, 54/3: 632–65.
Dorr, Cian, Goodman, Jeremy, and Hawthorne, John (2014). 'Knowing Against the Odds', *Philosophical Studies*, 170/2: 277–87.
Dorr, Cian and Hawthorne, John (2013). 'Naturalness', in Karen Bennett and Dean Zimmerman (eds.), *Oxford Studies in Metaphysics*, viii (Oxford: Oxford University Press): 4–75.
Dorr, Cian and Hawthorne, John (2014). 'Semantic Plasticity and Speech Reports', *Philosophical Review*, 123: 281–338.
Dorr, Cian and Rosen, Gideon (2002). 'Composition as a Fiction', in Richard Gale (ed.), *The Blackwell Companion to Metaphysics* (Oxford: Blackwell): 151–74.
Eagle, Antony (2011). 'Deterministic Chance', *Noûs*, 45/2: 269–99.
Earman, John (1989). *World Enough and Space-Time: Absolute Versus Relational Theories of Space and Time* (Cambridge, MA: MIT Press).
Earman, John and Norton, John (1987). 'What Price Spacetime Substantivalism? The Hole Story', *British Journal for the Philosophy of Science*, 38/4:515–25.
Eddon, Maya (2007). 'Parthood and Naturalness', *Philosophical Studies*, 174: 3163–80.
Egan, Andy, Hawthorne, John, and Weatherson, Brian (2005). 'Epistemic Modals in Context', in Gerhard Preyer and Georg Peter (eds.), *Contextualism in Philosophy* (Oxford: Oxford University Press), 131–70.
Elbourne, Paul (2005). *Situations and Individuals* (Cambridge, MA: MIT Press).
Elga, Adam (2014). 'Infinitesimal Chances and the Laws of Nature', *Australasian Journal of Philosophy*, 82: 67–76.
Fairchild, Maegan (2019). 'The Barest Flutter of the Smallest Leaf: Understanding Material Plenitude', *Philosophical Review*, 128:143–73.
Fairchild, Maegan and Hawthorne, John (2018). 'Against Conservatism in Metaphysics', *Royal Institute of Philosophy Supplement*, 82: 45–75.
Fara, Delia Graff (2008). 'Relative-Sameness Counterpart Theory', *Review of Symbolic Logic*, 1/2:167–89.
Fara, Michael and Williamson, Timothy (2005). 'Counterparts and Actuality', *Mind*, 114: 1–30.
Field, Hartry (1973). 'Theory Change and the Indeterminacy of Reference', *Journal of Philosophy*, 70/14: 462–81.
Field, Hartry (2003). 'No Fact of the Matter', *Australasian Journal of Philosophy*, 81: 457–80.
Fine, Kit (1975). 'Vagueness, Truth, and Logic', *Synthese*, 30: 265–300.
Fine, Kit (1977*a*). 'Prior on the Construction of Possible Worlds and Instants', in Fine 2005*a*: 133–75. From Prior and Fine 1977: 116–61.
Fine, Kit (1977*b*). 'Properties, Propositions and Sets', *Journal of Philosophical Logic*, 6/1: 135–91.
Fine, Kit (1982). 'Acts, Events and Things', in Werner Leinfellner, Eric Kraemer, and Jeffrey Schank (eds.), *Language and Ontology*, Proceedings of the Sixth International Wittgenstein Symposium (Vienna: Holder-Pichler-Tempsky), 97–105.
Fine, Kit (1994). 'Essence and Modality', in James Tomberlin (ed.), *Philosophical Perspectives 8: Logic and Language* (Atascadero: Ridgeview): 1–16.
Fine, Kit (1995). 'The Logic of Essence', *Journal of Philosophical Logic*, 24/3: 241–73.

Fine, Kit (1999). 'Things and Their Parts', in French and Wettstein 1999: 61–74.
Fine, Kit (2000). 'A Counter-Example to Locke's Thesis', *The Monist*, 83: 357–61.
Fine, Kit (2002). 'The Varieties of Necessity', in Fine 2005*a*: 235–60. From Tamar Gendler and John Hawthorne (eds.), *Conceivability and Possibility* (Oxford: Oxford University Press, 2002), 253–81.
Fine, Kit (2003). 'The Non-identity of a Material Thing and Its Matter', *Mind*, 112: 195–234.
Fine, Kit (2005*a*). *Modality and Tense: Philosophical Papers* (Oxford: Oxford University Press).
Fine, Kit (2005*b*). 'Necessity and Non-existence', in Fine 2005*a*: 321–57.
Fine, Kit (2005*c*). 'The Problem of Possibilia', in Fine 2005*a*: 214–34. From *The Oxford Handbook of Metaphysics*, ed. Michael J. Loux and Dean Zimmerman (Oxford: Oxford University Press): 161–79.
Fine, Kit (2012). 'Guide to Ground', in Fabrice Correia and Benjamin Schneider (eds.), *Metaphysical Grounding* (Cambridge: Cambridge University Press): 37–80.
Fine, Kit (2016*a*). 'Angellic Content', *Journal of Philosophical Logic*, 45/2:199–226.
Fine, Kit (2016*b*). 'Williamson on Fine on Prior on the Reduction of Possibilist Discourse', in Yli-Vakkuri and McCullagh 2017: 96–118. From *Canadian Journal of Philosophy* 46/4–5 (2016): 548–70.
Fine, Kit (2017*a*). 'A Theory of Truthmaker Content I: Conjunction, Disjunction and Negation', *Journal of Philosophical Logic*, 46/6: 625–74.
Fine, Kit (2017*b*). 'The Possibility of Vagueness', *Synthese*, 194: 3699–725.
von Fintel, Kai (2001). 'Counterfactuals in a Dynamic Context', in Michael Kenstowicz (ed.), *Ken Hale: A Life in Language* (Cambridge, MA: MIT Press) 123–52.
Forbes, Graeme (1982). 'Canonical Counterpart Theory', *Analysis*, 42/1: 33–7.
Forbes, Graeme (1984). 'Two Solutions to Chisholm's Paradox', *Philosophical Studies*, 46: 171–87.
Forbes, Graeme (1985). *The Metaphysics of Modality* (Oxford: Oxford University Press), https://spot.colorado.edu/~forbesg/metmodpage.html.
Frege, Gottlob (1879). 'Begriffsschrift: A Formula Language, Modeled Upon That of Arithmetic, for Pure Thought', in *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, ed. and trans. Jean van Heijenoort (1967), 1–82 [Ger. orig., *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens* (Halle: Nebert, 1879)].
Frege, Gottlob (1884). *The Foundations of Arithmetic: A Logic-mathematical Enquiry Into the Concept of Number*, ed. and trans. J. L. Austin (2nd edn., Oxford: Blackwell, 1974) [Ger. orig., *Die Grundlagen der Arithmetik: eine logisch mathematische Untersuchung über den Begriff der Zahl* (Breslau: W. Koebner, 1884)].
Frege, Gottlob (1892). 'On Concept and Object', trans. P. T. Geach and Max Black, *Mind*, 60 (1951): 168–80 [Ger. orig., 'Über Begriff und Gegenstand', *Vierteljahrsschrift für wissenschaftliche Philosophie*, 16 (1892): 192–205].
Frege, Gottlob (1918–9). 'Thoughts: A Logical Inquiry', in Peter Geach (ed.), *Logical Investigations* (Oxford: Basil Blackwell, 1977), 1–30 [Ger. orig., 'Der Gedanke: Eine logische Untersuchung', *Beiträge zur Philosophie des deutschen Idealismus, 1* (1918–9): 58–77].
French, Peter, Uehling, Jr., Theodore E., and Wettstein, Howard K. (1986) (eds.). *Studies in Essentialism*, [vol. xi of *Midwest Studies in Philosophy*] (Minneapolis, MN: University of Minnesota Press).
French, Peter and Wettstein, Howard K. (1999) (eds.). *New Directions in Philosophy* [vol. xxiii of *Midwest Studies in Philosophy*] (Minneapolis, MN: University of Minnesota Press).

Fritz, Peter (2017*a*). 'A Purely Recombinatorial Puzzle', *Noûs*, 51/3: 547–64.
Fritz, Peter (2017*b*). 'How Fine-Grained Is Reality?', *Filosofisk Supplement*, 13: 52–7.
Fritz, Peter (2020). 'On Higher-Order Logical Grounds', *Analysis*, 80: 656–66.
Fritz, Peter and Goodman, Jeremy (2017). 'Counting Incompossibles', *Mind*, 126: 1063–108.
Fritz, Peter and Jones, Nicholas K. (eds.) (forthcoming). *Higher-Order Metaphysics* (Oxford: Oxford University Press).
Gallin, Daniel (1975). *Intensional and Higher-Order Modal Logic: With Applications to Montague Semantics* (Amsterdam: North-Holland).
Gallois, Andre (2016). 'Identity Over Time', *The Stanford Encyclopedia of Philosophy*, Winter 2016, ed. Edward N. Zalta (Metaphysics Research Lab, Stanford University), https://plato.stanford.edu/archives/win2016/entries/identity-time/.
Gallow, J. Dmitri (2021). 'A Subjectivist's Guide to Determinstic Chance', *Synthese*, 198: 4339–72.
Garson, James W. (2001). 'Quantification in Modal Logic', in D. M. Gabbay and F. Guenther (eds.), *Handbook of Philosophical Logic*, iii (2nd edn.): 267–324.
Garson, James W. (2006). *Modal Logic for Philosophers* (2nd edn., Cambridge: Cambridge University Press, 2013).
Geach, P. T. (1962). *Reference and Generality* (Ithaca, NY: Cornell University Press).
Gendler, Tamar Szabó and Hawthorne, John (2000). 'Origin Essentialism: The Arguments Reconsidered', *Mind*, 109: 285–98.
Goguen, J. A. (1969). 'The Logic of Inexact Concepts', *Synthese*, 19: 325–73.
Goodman, Jeremy (2017). 'Reality Is Not Structured', *Analysis*, 77: 43–53.
Goodman, Jeremy (2018). 'Agglomerative Algebras', *Journal of Philosophical Logic*, 48/4: 631–48.
Goodman, Jeremy (unpublished *a*). 'Consequences of Conditional Excluded Middle', https://jeremy-goodman.com/ConsequencesCEM.pdf.
Goodman, Jeremy (unpublished *b*). 'Deep Structure'.
Goodman, Jeremy (unpublished *c*). 'Matter and Mereology', https://jeremy-goodman.com/MatterMereology.pdf.
Goodman, Jeremy (unpublished *d*). 'Theories of Aboutness'.
Goodman, Jeremy (unpublished *e*). 'Williamson's Earrings', https://jeremy-goodman.com/WilliamsonEarrings.pdf.
Goodman, Jeremy and Lederman, Harvey (2021). 'Perspectivism', *Noûs*, 55/3: 623–48.
Goodman, Jeremy and Salow, Bernhard (2018). 'Taking a Chance on KK', *Philosophical Studies*, 175/1:183–96.
Govan, Fiona (2012). '19th Century Fresco Destroyed by Rogue DIY Pensioner', *Daily Telegraph* (22 Aug. 2012), https://www.telegraph.co.uk/news/worldnews/europe/spain/9492805/19th-century-church-fresco-destroyed-by-rogue-DIY-pensioner.html.
Graff, Delia (= Fara, Delia Graff) (2000). 'Shifting Sands: An Interest-Relative Theory of Vagueness', *Philosophical Topics*, 28: 45–81.
Grandy, Richard E. and Freund, Max A. (2020). 'Sortals', *The Stanford Encyclopedia of Philosophy*, Winter 2020, ed. Edward N. Zalta (Metaphysics Research Lab, Stanford University), https://plato.stanford.edu/archives/win2020/entries/sortals/.
Grice, H. P. (1961). 'The Causal Theory of Perception', *Aristotelian Society Supplementary Volume*, 35/1:121–52.
Hájek, Alan (1996). '"Mises Redux"—Redux: Fifteen Arguments Against Finite Frequentism', *Erkenntnis*, 45/2–3: 209–27.
Hart, W. D. (1992). 'Hat-Tricks and Heaps', *Philosophical Studies (Maynooth)*, 33:1–24.
Hawley, Katherine (2001). *How Things Persist* (Oxford: Oxford University Press).

Hawthorne, John (2002). 'Deeply Contingent A Priori Knowledge', *Philosophy and Phenomenological Research*, 65/2: 247–69.
Hawthorne, John (2006*a*). 'Determinism De Re', in Hawthorne 2006*c*: 234–44.
Hawthorne, John (2006*b*). 'Epistemicism and Semantic Plasticity', in Hawthorne 2006*c*: 185–210.
Hawthorne, John (2006*c*). *Metaphysical Essays* (Oxford: Oxford University Press).
Hawthorne, John (2006*d*). 'Motion and Plenitude', in Hawthorne 2006*c*: 111–44.
Hawthorne, John (2006*e*). 'Plenitude, Convention, and Ontology', in Hawthorne 2006*c*: 53–70.
Hawthorne, John (2007). 'A Priority and Externalism', in Sanford C. Goldberg (ed.), *Internalism and Externalism in Semantics and Epistemology* (Oxford: Oxford University Press): 201–18.
Hawthorne, John (2009). 'Superficialism in Ontology', in David J. Chalmers, David Manley, and Ryan Wasserman (eds.), *Metametaphysics: New Essays in the Foundations of Ontology* (Oxford: Clarendon Press): 213–30.
Hawthorne, John and Lasonen-Aarnio, Maria (2009). 'Knowledge and Objective Chance', in Patrick Greenough and Duncan Pritchard (eds.), *Williamson on Knowledge* (Oxford: Oxford University Press): 92–108.
Hazen, Allen (1979). 'Counterpart-Theoretic Semantics for Modal Logic', *Journal of Philosophy*, 76: 319–38.
Hellie, Benj, Murray, Adam, and Wilson, Jessica (2021). 'Relativized Metaphysical Modality: Index and Context', in Scott A. Shalkowski and Otávio Bueno (eds.), *The Routledge Handbook of Modality* (London: Routledge): 82–99.
Henkin, Leon (1950). 'Completeness in the Theory of Types', *Journal of Symbolic Logic*, 15/2: 81–91.
Hirsch, Eli (1976). 'Physical Identity', *Philosophical Review*, 85: 357–89.
Hirsch, Eli (2002). 'Quantifier Variance and Realism', *Philosophical Issues*, 12: 51–73.
Hirsch, Eli (2010). *Quantifier Variance and Realism* (Oxford: Oxford University Press).
Hobbes, Thomas (1655). 'Concerning Body', in William Molesworth (ed.), *The English Works of Thomas Hobbes* (London: John Bohn, 1839).
Hodes, Harold T. (2015). 'Why Ramify?', *Notre Dame Journal of Formal Logic*, 56/2: 379–415.
Hoefer, Carl (1996). 'The Metaphysics of Space-Time Substantivalism', *Journal of Philosophy*, 93/1: 5–27.
Hovda, Paul (2013). 'Tensed Mereology', *Journal of Philosophical Logic*, 42/2: 241–83.
Hudson, Hud (2006). *The Metaphysics of Hyperspace* (Oxford: Oxford University Press).
Hughes, Christopher (1997). 'Same-Kind Coincidence and the Ship of Theseus', *Mind*, 106/421: 53–67.
Inman, Ross (2014). 'Neo-Aristotelian Plenitude', *Philosophical Studies*, 168: 583–97.
Jackman, Henry (1999). 'We Live Forwards but Understand Backwards: Linguistic Practices and Future Behaviour', *Pacific Philosophical Quarterly*, 80: 157–77.
Jackman, Henry (2005). 'Temporal Externalism and Our Ordinary Linguistic Practices', *Pacific Philosophical Quarterly*, 86/3: 365–80.
Jackson, Frank (1996). 'The Primary Quality View of Color', *Philosophical Perspectives*, 10:199–219.
Jago, Mark (2016). 'Essence and the Grounding Problem', in *Reality Making* (Oxford: Oxford University Press): 99–120.
Johnston, Mark (1989). 'Relativism and the Self', in Michael Krausz (ed.), *Relativism: Interpretation and Confrontation* (South Bend, IN: Notre Dame University Press).

Johnston, Mark (2002). 'Parts and Principles: False Axioms in Mereology', *Philosophical Topics*, 30:129–66.
Johnston, Mark (2006). 'Hylomorphism', *Journal of Philosophy*, 103: 652–98.
Johnston, Mark (2010). *Surviving Death* (Princeton, NJ: Princeton University Press).
Johnston, Mark (2016). 'Personites, Maximality and Ontological Trash', *Philosophical Perspectives*, 30: 198–228.
Johnston, Mark (2017). 'The Personite Problem: Should Practical Reason Be Tabled?', *Noûs*, 51: 617–44.
Kaplan, David (1968). 'Quantifying In', *Synthese*, 19:178–214.
Kaplan, David (1975). 'How to Russell a Frege-Church', *Journal of Philosophy*, 72/19: 716–29.
Kaplan, David (1989). 'Demonstratives', in Joseph Almog, John Perry, and Howard Wettstein (eds.), *Themes from Kaplan* (Oxford: Oxford University Press): 481–566.
Keefe, Rosanna (2000). *Theories of Vagueness* (Cambridge: Cambridge University Press).
Khamara, Edward J. (1988). 'Indiscernibles and the Absolute Theory of Space and Time', *Studia Leibnitiana*, 20:140–59.
Kim, Jaegwon (1984). 'Concepts of Supervenience', *Philosophy and Phenomenological Research*, 45/December: 153–76.
King, Jeffrey C. (2007). *The Nature and Structure of Content* (Oxford: Oxford University Press).
Kment, Boris (2012). 'Haecceitism, Chance, and Counterfactuals', *Philosophical Review*, 121/4: 573–609.
Kment, Boris (2014). *Modality and Explanatory Reasoning* (Oxford: Oxford University Press).
Kment, Boris (2018). 'Chance and the Structure of Modal Space', *Mind*, 127/507: 633–65.
Kölbel, Max (2002). *Truth Without Objectivity* (London: Routledge).
Korman, Daniel Z. (2015). *Objects: Nothing Out of the Ordinary* (Oxford: Oxford University Press).
Krämer, Stephan (2013). 'A Simpler Puzzle of Ground', *Thought*, 2/2: 85–9.
Kratzer, Angelika (1977). 'What "Must" and "Can" Must and Can Mean', *Linguistics and Philosophy*, 1: 337–55.
Kripke, Saul A. (1959). 'A Completeness Theorem in Modal Logic', *Journal of Symbolic Logic:* 1–14.
Kripke, Saul A. (1963). 'Semantical Considerations on Modal Logic', *Acta Philosophica Fennica*, 16: 83–94.
Kripke, Saul A. (1969). Letter to David Lewis: 11 Aug. 1969, in Yale Weiss and Romina Padro (eds.), *Saul Kripke on Modal Logic* (New York: Springer) (forthcoming).
Kripke, Saul A. (1971). 'Identity and Necessity', in M. K. Munitz (ed.), *Identity and Individuation* (New York: New York University Press).
Kripke, Saul A. (1972). *Naming and Necessity* (Cambridge, MA: Harvard University Press, 1980). From Donald Davidson and Gilbert Harman (eds.), *Semantics of Natural Language* (Dordrecht: D. Reidel, 1972): 253–355.
Kripke, Saul A. (2005). 'Russell's Notion of Scope', in *Collected Papers*, i: *Philosophical Troubles* (New York: Oxford University Press, 2011): 225–53. From *Mind*, 114 (2005): 1005–37.
Lambert, Karel (1963). 'Existential Import Revisited', *Notre Dame Journal of Formal Logic*, 4/4: 288–92.
Lange, Marc (2005). 'A Counterfactual Analysis of the Concepts of Logical Truth and Necessity', *Philosophical Studies*, 125/3: 277–303.

Leibniz, Gottfried Wilhelm (1710). *Theodicy: Essays on the Goodness of God, the Freedom on Man and the Origin of Evil*, trans. E. M. Huggard (La Salle, IL: Open Court, 1985).
Leibniz, Gottfried Wilhelm (1956). *Philosophical Papers and Letters*, ed. and trans. Leroy E. Loemker (Chicago: University of Chicago Press).
Leibniz, Gottfried Wilhelm (1989). *Philosophical Essays*, ed. and trans. Roger Ariew and Daniel Garber (Indianapolis, IN: Hackett).
Leslie, Sarah-Jane (2011). 'Essence, Plenitude, and Paradox', *Philosophical Perspectives*, 25: 277–96.
Levin, Janet (2000). 'Dispositional Theories of Color and the Claims of Common Sense', *Philosophical Studies*, 100/2:151–74.
Lewis, David K. (1968). 'Counterpart Theory and Quantified Modal Logic', in Lewis 1983*b*: 26–46. From *Journal of Philosophy*, 65 (1968): 113–26.
Lewis, David K. (1971). 'Counterparts of Persons and Their Bodies', in Lewis 1983*b*: 47–54. From *Journal of Philosophy*, 68 (1971): 203–11.
Lewis, David K. (1973). *Counterfactuals* (Cambridge, MA: Harvard University Press).
Lewis, David K. (1974). 'Counterpart Theory, Mk. II', manuscript shared by Steffi Lewis (unpublished).
Lewis, David K. (1976). 'Survival and Identity', in Lewis 1986*b*: 55–77. From Amelie Oksenberg Rorty (ed.), *The Identities of Persons* (Berkeley, CA: University of California Press, 1976): 17–40.
Lewis, David K. (1979). 'Counterfactual Dependence and Time's Arrow', in Lewis 1986*b*: 32–66. From *Noûs*, 13 (1979): 455–76.
Lewis, David K. (1980*a*). 'A Subjectivist's Guide to Objective Chance', in Lewis 1983*b*: 83–133. From R. C. Jeffrey (ed.), *Studies in Inductive Logic and Probability*, ii (Berkeley, CA: University of California Press, 1980): 263–93.
Lewis, David K. (1980*b*). 'Veridical Hallucination and Prosthetic Vision', in Lewis 1986*b*: 273–90. From *Australasian Journal of Philosophy*, 58/3 (1980): 239–49.
Lewis, David K. (1982). 'Logic for Equivocators', in *Papers in Philosophical Logic* (Cambridge: Cambridge University Press, 1998), 97–110. From *Noûs*, 16 (1982): 431–41.
Lewis, David K. (1983*a*). 'New Work for a Theory of Universals', in Lewis 1999: 8–55. From *Australasian Journal of Philosophy*, 61 (1983): 343–77.
Lewis, David K. (1983*b*). *Philosophical Papers*, i (Oxford: Oxford University Press).
Lewis, David K. (1986*a*). *On the Plurality of Worlds* (Oxford: Blackwell).
Lewis, David K. (1986b). *Philosophical Papers*, ii (Oxford: Oxford University Press).
Lewis, David K. (1988). 'Vague Identity: Evans Misunderstood', *Analysis*, 48:128–30.
Lewis, David K. (1993). 'Many, But Almost One', in Lewis 1999: 164–82. From Keith Campbell, John Bacon, and Lloyd Reinhardt (eds.), *Ontology, Causality, and Mind: Essays on the Philosophy of D. M. Armstrong* (Cambridge: Cambridge University Press, 1993) 23–37.
Lewis, David K. (1994). 'Humean Supervenience Debugged', in Lewis 1999: 224–47. From *Mind*, 103 (1994): 473–90.
Lewis, David K. (1999). *Papers in Metaphysics and Epistemology* (Cambridge: Cambridge University Press).
Linnebo, Øystein (2013). 'The Potential Hierarchy of Sets', *Review of Symbolic Logic*, 6: 205–28.
Locke, John (1689). *An Essay concerning Human Understanding*, ed. P. H. Nidditch (Oxford: Oxford University Press, 1975).
Loewer, Barry (2001). 'Determinism and Chance', *Studies in the History and Philosophy of Modern Physics*, 32: 609–20.

Look, Brandon C. (2013). 'Leibniz's Modal Metaphysics', *The Stanford Encyclopedia of Philosophy*, Spring 2013, ed. Edward N. Zalta (Metaphysics Research Lab, Stanford University), https://plato.stanford.edu/archives/spr2013/entries/leibniz-modal/.

Lovett, Adam (2020*a*). 'The Logic of Ground', *Journal of Philosophical Logic*, 49/1: 13–49.

Lovett, Adam (2020*b*). 'The Puzzles of Ground', *Philosophical Studies*, 177/9: 2541–64.

Lowe, E. J. (1983). 'On the Identity of Artifacts', *Journal of Philosophy*, 80:220–32.

Łukasiewicz, Jan and Tarski, Alfred (1930). 'Investigations into the Sentential Calculus', in Tarski 1956: 38–59 [Ger. orig., 'Untersuchungen über den Aussagenkalkül', *Comptes rendus des séances de la Société des Sciences et des Lettres de Varsovie*, 3, 23 (1930): 1–21, 30–50].

MacFarlane, John (2015). *Assessment Sensitivity: Relative Truth and its Applications* (Oxford: Oxford University Press).

McGee, Vann (1985). 'A Counterexample to Modus Ponens', *Journal of Philosophy*, 82: 462–71.

McGee, Vann and McLaughlin, Brian P. (1995). 'Distinctions Without a Difference', *Southern Journal of Philosophy*, 33: 203–51.

McGee, Vann and McLaughlin, Brian P. (2000). 'Lessons of the Many', *Philosophical Topics*, 28:129–52.

McKay, Thomas J. (1999). 'Against Constitutional Sufficiency Principles', in French and Wettstein 1999: 295–304.

Mackie, Penelope (2006). *How Things Might Have Been: Individuals, Kinds, and Essential Properties* (Oxford: Oxford University Press).

Mackie, Penelope and Jago, Mark (2017). 'Transworld Identity', *The Stanford Encyclopedia of Philosophy*, Winter 2017, ed. Edward N. Zalta (Metaphysics Research Lab, Stanford University), https://plato.stanford.edu/archives/win2017/entries/identity-transworld/.

McNamara, Paul (2019). 'Deontic Logic', *The Stanford Encyclopedia of Philosophy*, Summer 2019, ed. Edward N. Zalta (Metaphysics Research Lab, Stanford University), https://plato.stanford.edu/archives/sum2019/entries/logic-deontic/.

Maidens, Anna (1992), review of Earman 1989, in *British Journal for the Philosophy of Science*, 43/1:129.

Manley, David and Wasserman, Ryan (2008). 'On Linking Dispositions and Conditionals', *Mind*, 117/465: 59–84.

Markosian, Ned (1998). 'Brutal Composition', *Philosophical Studies*, 92: 211–49.

Maudlin, Tim (2010). 'Time, Topology and Physical Geometry', *Aristotelian Society Supplementary Volume*, 84: 63–78.

Maudlin, Tim (2014). *New Foundations for Physical Geometry: The Theory of Linear Structures* (Oxford: Oxford University Press).

Merricks, Trenton (2003*a*). 'Maximality and Consciousness', *Philosophy and Phenomenological Research*, 66:150–8.

Merricks, Trenton (2003*b*). 'The End of Counterpart Theory', *Journal of Philosophy*, 100/10: 521–49.

Mondadori, Fabrizio (1975). 'Leibniz and the Doctrine of Inter-World Identity', *Studia Leibnitiana*, 7/1: 21–57.

Montague, Richard (1974). *Formal Philosophy* (New Haven, CT: Yale University Press).

Moore, G. E. (1959). 'Four Forms of Scepticism', in *Philosophical Papers* (London: Allen and Unwin), 196–226.

Murray, Adam and Wilson, Jessica (2012). 'Relativized Metaphysical Modality', in Karen Bennett and Dean Zimmerman (eds.), *Oxford Studies in Metaphysics*, vii (Oxford: Oxford University Press): 189–224.

Muskens, Reinhard (2007). 'Intensional Models for the Theory of Types', *Journal of Symbolic Logic*, 72: 98–118.
Myhill, John (1958). 'Problems Arising in the Formalization of Intensional Logic', *Logique et Analyse*, 1: 78–83.
Nolan, Daniel (1997). 'Impossible Worlds: A Modest Approach', *Notre Dame Journal of Formal Logic*, 38: 535–72.
Nolan, Daniel (2011). 'The Extent of Metaphysical Necessity', *Philosophical Perspectives*, 25/1: 313–39.
Noonan, Harold W. (1988). 'Reply to Lowe on Ships and Structures', *Analysis*, 48/4: 221–3.
Noonan, Harold W. (2010). 'The Thinking Animal Problem and Personal Pronoun Revisionism', *Analysis*, 70: 93–8.
William of Ockham (1317–18). 'In libros Sententiarum', in *Opera Philosophica et Theologica*, v, ed. Gedeon Gál et al. (St. Bonaventure, NY: The Franciscan Institute, 1967–88).
Olson, Eric T. (2001). 'Material Coincidence and the Indiscernibility Problem', *Philosophical Quarterly*, 51/204: 337–55.
Parfit, Derek (1971). 'Personal Identity', *Philosophical Review*, 80: 3–27.
Parfit, Derek (1984). *Reasons and Persons* (Oxford: Oxford University Press).
Parsons, Josh (2007). 'Theories of Location', in Dean Zimmerman (ed.), *Oxford Studies in Metaphysics*, iii (Oxford: Oxford University Press): 201–32.
Plantinga, Alvin (1974). *The Nature of Necessity* (Oxford: Oxford University Press).
Poincaré, Henri (1905–6). 'Les mathematiques et la logique', *Révue de Metaphysique et de Morale*, 13: 815–35; 14:17–34, 294–317.
Pooley, Oliver (2005). 'Points, Particles, and Structural Realism', in Dean Rickles, Steven French, and Juha T. Saatsi (eds.), *The Structural Foundations of Quantum Gravity* (Oxford: Oxford University Press): 83–120.
Priest, Graham (1987). *In Contradiction* (Dordrecht: Martinus Nijhoff; 2nd edn., Oxford: Oxford University Press, 2006).
Priest, Graham (2021). 'Metaphysical Necessity: A Skeptical Perspective', *Synthese*, 198: 1873–85.
Prior, A. N. (1957). 'Modality and Quantification in S5', *Journal of Symbolic Logic*, 21/1: 60–2.
Prior, A. N. (1963). *Formal Logic* (Oxford: Oxford University Press).
Prior, A. N. (1971). *Objects of Thought* (Oxford: Oxford University Press).
Prior, A. N. and Fine, Kit (1977). *Worlds, Times, and Selves* (London: Duckworth).
Putnam, Hilary (1975). 'The Meaning of "Meaning"', *Minnesota Studies in the Philosophy of Science*, 7:131–93.
Quine, Willard van Orman (1948). 'On What There Is', in Quine 1953*a*: 1–19. From *Review of Metaphysics*, 2 (1948): 21–38.
Quine, Willard van Orman (1953*a*). *From a Logical Point of View* (2nd rev. edn., Cambridge, MA: Harvard University Press, 1980).
Quine, Willard van Orman (1953*b*). 'Reference and Modality', in Quine 1953*a*: 139–59.
Quine, Willard van Orman (1953*c*). 'Three Grades of Modal Involvment', *Proceedings of the XIth International Congress of Philosophy*, 14: 65–81.
Quine, Willard van Orman (1956). 'Quantifiers and Propositional Attitudes', *Journal of Philosophy*, 53:177–87.
Quine, Willard van Orman (1970). *Philosophy of Logic* (2nd edn., Cambridge, MA: Harvard University Press, 1986).
Raffman, Diana (1994). 'Vagueness Without Paradox', *Philosophical Review*, 103: 41–74.

Ramachandran, Murali (1989). 'An Alternative Translation Scheme for Counterpart Theory', *Analysis*, 49:131–41.
Ramachandran, Murali (1990). 'Contingent Identity in Counterpart Theory', *Analysis*, 50/3: 163–6.
Ramachandran, Murali (2020). 'Chisholm's Modal Paradox(es) and Counterpart Theory 50 Years On', *Logic and Logical Philosophy*, 29/4: 571–92.
Ramsey, F. P. (1925). 'The Foundations of Mathematics', *Proceedings of the Mathematical Society*, 25: 338–84.
Rayo, Agustín (2013). *The Construction of Logical Space* (Oxford: Oxford University Press).
Rayo, Agustín (2020). 'On the Open-Endedness of Logical Space', *Philosophers' Imprint*, 20.
Rea, Michael C. (1995). 'The Problem of Material Constitution', *Philosophical Review*, 104/4: 525–52.
Roberts, Alexander (2019). 'Modal Expansionism', *Journal of Philosophical Logic*, 48/6:1145–70.
Robertson, Teresa (= Robertson Ishii, Teresa) (1998). 'Possibilities and the Arguments for Origin Essentialism', *Mind*, 107/428: 729–49.
Robertson, Teresa (= Robertson Ishii, Teresa) and Forbes, Graeme (2006). 'Does the New Route Reach its Destination?', *Mind*, 115/458: 367–74.
Robertson Ishii, Teresa (2018). 'A Puzzle About Kinds', *Philosophical Perspectives*, 32/1: 352–64.
Robertson Ishii, Teresa and Salmón, Nathan (2020). 'Some Highs and Lows of Hylomorphism: On a Paradox About Property Abstraction', *Philosophical Studies*, 177/6:1549–63.
Roca-Royes, Sonia and Cameron, Ross (2006). 'Rohrbaugh and deRosset on the Necessity of Origin', *Mind*, 115/458: 361–6.
Rohrbaugh, Guy and deRosset, Louis (2004). 'A New Route to the Necessity of Origin', *Mind*, 113/452: 705–25.
Rohrbaugh, Guy and deRosset, Louis (2006). 'Prevention, Independence, and Origin', *Mind*, 115/458: 375–86.
Rosen, Gideon (2006). 'The Limits of Contingency', in Fraser MacBride (ed.), *Identity and Modality* (Oxford: Oxford University Press): 13–39.
Russell, Bertrand (1903). *The Principles of Mathematics* (Cambridge: Cambridge University Press).
Russell, Bertrand (1905). 'On Denoting', in *Logic and Knowledge*, ed. Robert Charles Marsh (New York: G.P. Putnam's Sons, 1956): 41–56. From *Mind* 14 (1905): 479–93.
Russell, Bertrand (1906). 'On "Insolubilia" and their Solution by Symbolic Logic', in *Essays in Analysis*, ed. Douglas Lackey (London: Allen and Unwin, 1973), 190–214 [Fr. orig., 'Les Paradoxes de la Logique', *Révue de Métaphysique et de Morale*, 14 (1906): 627–50].
Russell, Bertrand (1908). 'Mathematical Logic as Based on the Theory of Types', *American Journal of Mathematics*, 30: 222–62.
Russell, Bertrand (1918–19). *The Philosophy of Logical Atomism*, ed. David F. Pears (La Salle, IL: Open Court, 1985). From *The Monist*, 28 (1918): 495–527; 29 (1919): 32–63, 190–222, 345–380.
Russell, Jeffrey Sanford (2013). 'Actuality for Counterpart Theorists', *Mind*, 122: 85–134.
Russell, Jeffrey Sanford (2018). 'Quality and Quantifiers', *Australasian Journal of Philosophy*, 96/3: 562–77.
Salmon, Nathan (1986*a*). 'Modal Paradox: Parts and Counterparts, Points and Counterpoints', in Salmon 1981: 255–344. From French, Uehling, and Wettstein 1986: 75–120.
Salmon, Nathan (1987). 'Existence', *Philosophical Perspectives*, 1: 49–108.
Salmon, Nathan (1989*a*). 'Illogical Belief', *Philosophical Perspectives*, 3: 243–85.

Salmon, Nathan (1989*b*). 'The Logic of What Might Have Been', *Philosophical Review*, 98: 3–34.
Salmon, Nathan (1993). 'This Side of Paradox', *Philosophical Topics*, 21/2:187–97.
Salmon, Nathan Ucuzoglu (1979). 'How Not to Derive Essentialism From the Theory of Reference', *Journal of Philosophy*, 76/12: 703–25.
Salmon, Nathan U. (1981). *Reference and Essence* (Oxford: Blackwell; 2nd edn., Amherst, New York: Prometheus Books, 2006).
Salmon, Nathan U. (1986*b*). *Frege's Puzzle* (Cambridge, MA: MIT Press).
Salmón, Nathan (2019). 'Impossible Odds', *Philosophy and Phenomenological Research*, 99: 644–62.
Schlesinger, George (1985). 'Spatial, Temporal and Cosmic Parts', *Southern Journal of Philosophy*, 23: 255–71.
Shoemaker, Sydney (1998). 'Causal and Metaphysical Necessity', *Pacific Philosophical Quarterly*, 79: 59–77.
Sider, Theodore (1996). 'All the World's a Stage', *Australasian Journal of Philosophy*, 74: 433–53.
Sider, Theodore (2002). *Four-Dimensionalism: An Ontology of Persistence and Time* (Oxford: Oxford University Press).
Sider, Theodore (2003). 'Maximality and Microphysical Supervenience', *Philosophy and Phenomenological Research*, 66:139–49.
Sider, Theodore (2009). 'Beyond the Humphrey Objection', http://www.tedsider.org/papers/counterpart_theory.pdf (unpublished).
Sider, Theodore (2011). *Writing the Book of the World* (Oxford: Oxford University Press).
Sider, Theodore (2020). *The Tools of Metaphysics and the Metaphysics of Science* (Oxford: Oxford University Press).
Simons, Peter (1987). *Parts: A Study in Ontology* (Oxford: Clarendon Press).
Skow, Bradford (2008). 'Haecceitism, Anti-haecceitism, and Possible Worlds: A Case Study', *Philosophical Quarterly*, 58: 98–107.
Smart, Brian (1972). 'How to Reidentify the Ship of Theseus', *Analysis*, 32/5: 145–48.
Smart, Brian (1973). 'The Ship of Theseus, the Parthenon and Disassembled Objects', *Analysis*, 34/1: 24–7.
Soames, Scott (1985). 'Lost Innocence', *Linguistics and Philosophy*, 8: 59–71.
Soames, Scott (1987). 'Direct Reference, Propositional Attitudes and Semantic Content', *Philosophical Topics*, 15: 47–87.
Soames, Scott (1989). 'Semantics and Semantic Competence', *Philosophical Perspectives*, 3: 575–96.
Soames, Scott (1999). *Understanding Truth* (Oxford: Oxford University Press).
Soames, Scott (2002). *Beyond Rigidity: The Unfinished Semantic Agenda of* Naming and Necessity (Oxford: Oxford University Press).
Soames, Scott (2010). *What Is Meaning?* (Princeton, NJ: Princeton University Press).
Sorensen, Roy A. (1988). *Blindspots* (Oxford: Oxford University Press).
Sosa, Ernest (1999). 'Existential Relativity', in French and Wettstein 1999: 132–43.
Stalnaker, Robert C. (1968). 'A Theory of Conditionals', in *Knowledge and Conditionals: Essays on the Structure of Inquiry* (Oxford: Oxford University Press, 2019): 151–62. From Nicholas Rescher (ed.), *Studies in Logical Theory* (American Philosophical Quarterly Monograph Series, 2; Oxford: Blackwell, 1968): 98–112.
Stalnaker, Robert C. (1978). 'Assertion', in *Context and Content: Essays on Intentionality in Speech and Thought* (Oxford: Oxford University Press, 1999): 78–95. From P. Cole (ed.), *Syntax and Semantics*, ix: *Pragmatics* (New York: New York Academic Press, 1978): 312–32.

Stalnaker, Robert C. (1980). 'A Defense of Conditional Excluded Middle', in W. L. Harper, R. Stalnaker, and G. Pearce (eds.), *Ifs: Conditionals, Belief, Decision, Chance and Time* (Dordrecht: D. Reidel).

Stalnaker, Robert C. (1986). 'Counterparts and Identity', in Stalnaker 2003*b*: 111–32. From French, Uehling, and Wettstein 1986: 121–40.

Stalnaker, Robert C. (1994). 'The Interaction of Modality with Quantification and Identity', in Stalnaker 2003*b*: 144–62. From W. Sinnott-Armstrong, D. Raffman, and N. Asher (eds.), *Modality, Morality, and Belief* (Oxford: Oxford University Press, 1994): 12–28.

Stalnaker, Robert C. (2003*a*). 'Conceptual Truth and Metaphysical Necessity', in Stalnaker 2003*b*: 201–16.

Stalnaker, Robert C. (2003*b*). *Ways a World Might Be* (Oxford: Oxford University Press).

Stalnaker, Robert C. (2012). *Mere Possibilities* (Oxford: Oxford University Press).

Stanley, Jason and Szabó, Zoltán Gendler (2000). 'On Quantifier Domain Restriction', *Mind and Language*, 15: 219–61.

Strawson, P. F. (1979). 'May Bes and Might Have Beeens', in Avishai Margalit (ed.), *Meaning and Use* (Dordrecht: Reidel): 229–38.

Tarski, Alfred (1956). *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*, trans. J. H. Woodger (Oxford: Clarendon Press).

Tarski, Alfred (1936). 'On the Concept of Logical Consequence', in Tarski 1956*a*: 409–20 [Pol. orig., 'O pojciu wynikania logicznego', *Przeglqd Filozoficzny*, 39 (1936): 58–68].

Thomson, Judith Jarvis (1998). 'The Statue and the Clay', *Noûs*, 32:149–73.

Unger, Peter (1980). 'The Problem of the Many', *Midwest Studies in Philosophy*, 5: 411–67.

Uzquiano, Gabriel (2006). 'Receptacles', *Philosophical Perspectives*, 20:427–51.

Uzquiano, Gabriel (2015). 'A Neglected Resolution of Russell's Paradox of Propositions', *Review of Symbolic Logic*, 8: 328–44.

van Inwagen, Peter (1981*b*). 'The Doctrine of Arbitrary Undetached Parts', in van Inwagen 2001: 75–94. From *Pacific Philosophical Quarterly*, 62 (1981): 123–37.

van Inwagen, Peter (1986). *An Essay on Free Will* (Oxford: Oxford University Press).

van Inwagen, Peter (1990). *Material Beings* (Ithaca, NY: Cornell University Press).

van Inwagen, Peter (1998). 'Modal Epistemology', in van Inwagen 2001: 243–58. From *Philosophical Studies*, 92 (1998): 67–84.

van Inwagen, Peter (2001). *Ontology, Identity and Modality* (Cambridge: Cambridge University Press).

Vlach, Frank (1973). '"Now" and "Then": A Formal Study in the Logic of Tense Anaphora', PhD thesis (Department of Philosophy, University of California, Los Angeles).

Wasserman, Ryan (2018). 'Material Constitution', *The Stanford Encyclopedia of Philosophy*, Fall 2018, ed. Edward N. Zalta (Metaphysics Research Lab, Stanford University), https://plato.stanford.edu/archives/fall2018/entries/material-constitution/.

Weatherson, Brian (2003). 'Many Many Problems', *Philosophical Quarterly*, 53: 481–501.

Weatherson, Brian (2016). 'The Problem of the Many', *Stanford Encyclopedia of Philosophy*, Winter 2016, ed. Edward N. Zalta, https://plato.stanford.edu/archives/win2016/entries/problem-of-many/.

Weber, Zach (2010). 'A Paraconsistent Model of Vagueness', *Mind*, 119/476: 1025–45.

Whitehead, Alfred North (1920). *The Concept of Nature* (Cambridge: Cambridge University Press).

Whitehead, Alfred North and Russell, Bertrand (1910). *Principia Mathematica* (2nd edn., Cambridge: Cambridge University Press).

Wiggins, David (1967). *Identity and Spatio-Temporal Continuity* (Oxford: Blackwell).

Wiggins, David (1968). 'On Being in the Same Place at the Same Time', *Philosophical Review*, 77/1: 90–5.

Wiggins, David (1980). *Sameness and Substance* (Cambridge, MA: Harvard University Press).

Wikipedia contributors (2020). 'Ecce Homo (Martínez and Giménez)', https://en.wikipedia.org/wiki/Ecce_Homo_(Mart%C3%ADnez_and_Gim%C3%A9nez), accessed 26 Dec. 2020.

Williamson, Timothy (1990). *Identity and Discrimination* (Oxford: Blackwell).

Williamson, Timothy (1994). *Vagueness* (London: Routledge).

Williamson, Timothy (1996). 'The Necessity and Determinacy of Distinctness', in Savina Lovibond and Stephen G. Williams (eds.), *Essays for David Wiggins: Identity, Truth and Value* (Oxford: Blackwell): 1–17.

Williamson, Timothy (2000). *Knowledge and Its Limits* (Oxford: Oxford University Press).

Williamson, Timothy (2003). 'Everything', *Philosophical Perspectives*, 17: 415–65.

Williamson, Timothy (2007). *The Philosophy of Philosophy* (Oxford: Oxford University Press).

Williamson, Timothy (2013*a*). 'How Deep Is the Distinction Between A Priori and A Posteriori Knowledge?', in Albert Casullo and Joshua C. Thurow (eds.), *The A Priori in Philosophy* (Oxford: Oxford University Press): 291–312.

Williamson, Timothy (2013*b*). *Modal Logic as Metaphysics* (Oxford: Oxford University Press).

Williamson, Timothy (2016). 'Modal Science', in Yli-Vakkuri and McCullagh 2017: 1–39. From *Canadian Journal of Philosophy*, 46/4–5: 453–92.

Williamson, Timothy (2018). 'Alternative Logics and Applied Mathematics', *Philosophical Issues*, 28/1: 399–424.

Williamson, Timothy (2020). *Suppose and Tell: The Semantics and Heuristics of Conditionals* (Oxford: Oxford University Press).

Wittgenstein, Ludwig (1921). *Tractatus Logico-Philosophicus*, trans. C. K. Ogden (London: Kegan Paul, Trench, Trübner & Co., 1922) [Ger. orig., 'Logisch-Philosophische Abhandlung', *Annalen der Naturphilosophische*, 14 (1921)].

von Wright, G. H. (1951). 'Deontic Logic', *Mind*, 60:1–15.

Yablo, Stephen (1987). 'Identity, Essence, and Indiscernibility', *Journal of Philosophy*, 84: 293–314.

Yablo, Stephen (forthcoming). 'Everything Exists: Plenitude, Kinds, and Paradox', in Javier Cumpa (ed.), *The Question of Ontology: The Contemporary Debate* (Oxford: Oxford University Press).

Yalcin, Seth S. (2015). *'Actually*, Actually', *Analysis*, 75/2:185–91.

Yli-Vakkuri, Juhani (unpublished *a*). 'Counterpart Theory and Modal Logic'.

Yli-Vakkuri, Juhani (unpublished *b*). 'Rigid Designation'.

Yli-Vakkuri, Juhani and Hawthorne, John (2018). *Narrow Content* (Oxford: Oxford University Press).

Yli-Vakkuri, Juhani and Hawthorne, John (forthcoming a). 'Counterpart Theory and Counterfactuals' in Karen Bennett and Dean Zimmerman (eds.), *Oxford Studies in Metaphysics*, xiii (Oxford: Oxford University Press).

Yli-Vakkuri, Juhani and Hawthorne, John (forthcoming b). 'Intensionalism and Propositional Attitudes', in Uriah Kriegel (ed.), *Oxford Studies in Philosophy of Mind*, ii (Oxford: Oxford University Press).

Yli-Vakkuri, Juhani and McCullagh, Mark (2017) (eds.). *Williamson on Modality* (London: Routledge).

Zhang, John et al. (2017). 'Live Birth Derived from Oocyte Spindle Transfer to Prevent Mitochondrial Disease', *Reproductive Biomedicine Online*, 34/4: 361–8.

Index of Names

Adams, Robert Merrihew 48, 125, 328, 329, 355, 358
Ali, Muhammed 184, 188
Angell, R. Bradshaw 222
Arntzenius, Frank 269

Bacon, Andrew 16, 22, 30ff, 40, 45, 50, 52, 186, 191, 193, 195, 200, 211ff, 215–19, 257, 331, 337, 379f, 399
Ballarin, Roberta 131
Barcan, Ruth C. 37, 39
Barnett, David 130, 142ff
Bealer, George 289, 376
Belot, Gordon 351
Benacerraf, Paul 196
Bennett, Karen 100
Black, Max 358
Blair, Tony 214
Braun, David 183, 207
Bricker, Phillip 306
Brighouse, Carolyn 351
Brower-Toland, Susan 289
Brown, Jessica 290
Burge, Tyler 289
Burgess, John 46
Burke, Michael B. 100, 107
Butler, Joseph 10f
Butterfield, Jeremy 351
Byrne, Alex 304

Caie, Michael 22, 30, 115, 205
Cameron, Ross 131
Cappelen, Herman 93
Carnap, Rudolf 5, 22, 273
Cartwright, Richard 269
Casati, Roberto 270
Chalmers, David J. 165f, 289, 322
Chandler, Hugh S. 67f, 84, 86, 107, 109, 136, 171, 233
Charles I 128
Chippendale, Thomas 135
Chisholm, Roderick 10, 49, 59, 122–9, 153, 156, 259
Church, Alonzo 22, 26, 36, 353
Clark, Andy 326
Clarke-Doane, Justin 202
Cleopatra 190f, 193
Coe, Sebastian 367
Collins, John M. 290
Correia, Fabrice 222
Cover, J.A. 10
Craig, William 36
Cresswell, M. J. 51, 273
Crossley, John N. 113

Dancy, Jonathan 277
Deasy, Daniel 267
deRosset, Louis 131
Descartes, René 100
Deutsch, Harry 189
Ditter, Andreas 26
Doepke, Frederick C. 100

Eagle, Antony 229, 349
Earman, John 269, 351
Eddon, Maya 306
Egan, Andy 93
Elbourne, Paul 62
Elga, Adam 235
Erlich, Franz 134f

Fairchild, Maegan 266ff, 272, 274f, 277
Fara, Delia Graff 16, 85, 247, 252, 309
Fara, Michael 257
Field, Hartry 17, 218
Fine, Kit 16, 17, 26, 38, 43, 49, 100f, 131, 149, 199, 202, 205, 214, 222, 266, 268, 276, 283, 285, 297, 306, 313
Forbes, Graeme 17, 84, 88, 247, 252, 254ff
Frege, Gottlob 3, 21f, 56, 66, 178, 182, 196, 198
Freund, Max A. 297
Fritz, Peter 22, 26, 38, 189, 211, 217

Gallin, Daniel 40, 46, 51f, 398
Gallois, Andre 8
Gallow, J. Dmitri 229
Garson, James W. 273f
Geach, P. T. 265, 283
Gendler, Tamar Szabó 133, 137, 142

Goguen, J. A. 17
Goodman, Jeremy 22, 30, 38, 96, 115, 176, 180, 183, 189–94, 205, 207, 209, 219, 222, 266, 268, 306, 329, 331, 343, 352, 367, 382, 407
Govan, Fiona 68
Grandy, Richard E. 297
Grice, H. P. 305

Hart, W. D. 80
Hawley, Katherine 249
Hazen, Allen 247, 249, 252, 257
Hellie, Benj 93, 260, 273f, 309
Henkin, Leon 26
Henry VIII 138
Hilbert, David R. 304
Hirsch, Eli 276, 311f
Hobbes, Thomas 67f, 106, 131f
Hodes, Harold T. 189
Hoefer, Carl 351
Hovda, Paul 305
Hudson, Hud 269f
Hughes, Christopher 106, 121
Hughes, G. E. 273
Humberstone, Lloyd 113
Hájek, Alan 236

Inman, Ross 266

Jackman, Henry 290
Jackson, Frank 304
Jago, Mark 59, 247, 266
Johnston, Mark 101, 300, 321, 323
Jones, Nicholas K. 22

Kaplan, David 93, 124, 287, 290, 300, 309
Keefe, Rosanna 16
Khamara, Edward J. 329, 331, 369
Kim, Jaegwon 148
King, Jeffrey C. 189
Kment, Boris 88, 101, 103, 108, 214, 227, 232f, 261
Korman, Daniel Z. 268, 277f
Kratzer, Angelika 11, 276
Kripke, Saul 3, 5f, 10, 28, 33, 37f, 47, 62, 71, 77, 87, 110, 130, 200ff, 209f, 214, 216, 247f, 250, 252, 255, 300, 349
Krämer, Stephan 26
Kölbel, Max 93

Lambert, Karel 31
Lange, Marc 214
Lederman, Harvey 22, 30, 115, 183, 205, 207, 209
Leibniz, Gottfried Wilhelm 5, 9, 169
Leslie, Sarah-Jane 101, 115, 117, 199, 266, 272, 283f, 310
Levin, Janet 304
Lewis, David K. 9, 16, 42, 59, 85, 100, 103, 106, 108, 110, 114–9, 121, 124, 150, 167, 171, 218f, 228f, 235f, 246–54, 258f, 265, 276, 283, 305f, 313, 323, 328, 343, 347, 351
Linnebo, Øystein 42
Locke, John 100, 304
Loewer, Barry 229, 349
Look, Brandon C. 8
Lovett, Adam 222
Lowe, E. J. 67
Łukasiewicz, Jan 17

MacFarlane, John 93
McGee, Vann 283
McKay, Thomas 88, 168
Mackie, Penelope 59, 126, 247
McLaughlin, Brian 283
McNamara, Paul 202
Magidor, Ofra 94
Maidens, Anna 351
Manley, David 304
Markosian, Ned 166
Maudlin, Tim 269
McGee, Vann 17
Merricks, Trenton 346, 320
Mondadori, Fabrizio 10
Montague, Richard 22
Moore, G. E. 6
Murray, Adam 93, 260, 273f, 309
Muskens, Reinhard 21, 46
Myhill, John 189

Nolan, Daniel 103, 108, 201
Noonan, Harold W. 100, 320
Norton, John 351

Olson, Eric T. 100
Ostertag, Gary 250

Parfit, Derek 106, 323
Parsons, Josh 270
Philip IV 128
Plantinga, Alvin 48, 111, 202, 376
Plato 184, 188f
Poincaré, Henri 32
Pooley, Oliver 351
Priest, Graham 15, 202
Prior, A. N. 19, 22, 40, 49, 110, 196, 374, 403
Putnam, Hilary 289

Quine, Willard van Orman 4, 9, 121, 206, 225

Raffman, Diana 85
Ramachandran, Murali 252, 254ff, 260
Ramsey, F. P. 22
Rayo, Agustín 195, 211
Rea, Michael C. 101
Roberts, Alexander 211
Robertson Ishii, Teresa 29, 32, 131, 136f, 189, 208
Roca-Royes, Sonia 131
Rohrbaugh, Guy 131
Rosen, Gideon 205, 214
Roth, Richard 66
Russell, Bertrand 18, 22, 28, 32, 189, 194, 196, 271
Russell, Jeffrey Sanford 22, 30, 32, 257, 330

Salmón, Nathan 29, 32, 39, 48, 68–72, 75, 84, 86, 107, 111, 114, 130f, 133–7, 146, 166, 171, 178, 180, 183, 189, 201ff, 207f, 233, 247, 258, 300, 380
Salow, Bernhard 96
Schlesinger, George 266
Shoemaker, Sydney 75
Sider, Theodore 3, 172, 225, 246f, 249, 251, 271, 280, 312, 320
Simons, Peter 68, 121
Skow, Bradford 124
Smart, Brian 67
Soames, Scott 31, 85, 183, 189, 207f
Sorensen, Roy A. 16
Sosa, Ernest 276
Spencer, Jack 282
Stalnaker, Robert C. 38, 49, 51, 103, 108, 131, 200, 209, 246f
Stanley, Jason 284
Strawson, P. F. 139–41
Szabó, Zoltán Gendler 284

Tarski, Alfred 17, 27
Thomson, Judith Jarvis 100, 305

Unger, Peter 265, 283
Uzquiano, Gabriel 22, 32, 189, 269

van Inwagen, Peter 100, 107, 200, 276, 348
Varzi, Achille C. 270
Virén, Lasse 367
Vlach, Frank 51
von Fintel, Kai 103
von Wright, G. H. 200

Wasserman, Ryan 100, 304
Weatherson, Brian 93, 283
Weber, Zach 17
Whitehead, Alfred North 196, 269
Wiggins, David 67, 100f, 121, 300
William of Ockham 289
Williamson, Timothy 16, 19, 22, 27, 37f, 40, 51, 64ff, 68, 80, 87, 103, 108, 112ff, 117, 153, 171–8, 199, 205, 217–20, 223, 257, 264, 287, 301, 313, 374, 379, 401
Wilson, Jessica 93, 260, 273f, 309
Wittgenstein, Ludwig 48, 194, 269, 379

Yablo, Stephen 76, 266, 268, 272f, 285f, 296, 319, 362, 381
Yalcin, S. 113

Zhang, John 128

Index of Topics and Named Principles

4 (axiom) *see* Iteration
5 (axiom) **39**, 40, 51, 74, 111, 125, 134, 148, 177, 219, 339, 357, 398

aboutness 185, 328
  intensional 369
  singular 369
  see also Basic Theory of Aboutness.
Aboutness Closure 367ff, 371
Aboutness Monotonicity 367ff
actual world 48f
  existence of 45
  specialness of 88, 117, 164, 279, 294, 382
Actuality **45**, 50f, 395, 397ff
actuality operator 74, 257, 374
Alien Compatibility **376**, 379–82
Alien Matter **376**, 379f
Ancestral Microphysical Supervenience **181**, 182, 185, 190, 195, 198, 233, 259
Anti-haecceitism **124**, 125, 147f, 351, 359
  see also Weak Anti-haecceitism
arbitrariness 85, 178, 186, 276, 280, 291, 299, 324, 343
Atom Permutability **336**, 336ff, 345f
atomic number 72, 333f, 336f, 339, 361
Atomic Structure **190**, 191, 194
  see also Generalized Atomic Structure; Weak Qualitative Atomic Structure; Generalized Qualitative Atomic Structure
Atomicity **50**, 51f, 395, 398f
atoms 72, 94, 125, 184, 186, 334f, 361
  Democratean 337, 347, 352, 358, 362
  in distant galaxies 74, 336
  permutations of 125, 336ff, 344ff, 356
  xenon 74, 371
attitudes, propositional 19, 182, 205
axes 115, 284

B (axiom) **39**, 40, 51, 110f, 224, 257, 374, 379, 389, 395ff, 400, 403, 407
Barcan Formula (BF) **39**, 40, 42f, 49, 51, 158, 168, 175, 219–22, 269f, 302, 357, 373f, 376–9, 396–403
Basic Theory of Aboutness **368**
beards 184f, 188, 279
Being Constraint 38
belief *see* attitudes, propositional
Beta-conversion **23**, 26, 190ff, 214, 329
  see also Extensional Beta-conversion (Eβ); Beta-identity
Beta-identity **191**, 192f
  see also Non-vacuous Beta-identity
bicycles 86f, 90, 171, 274f
Bodleian Library 139f, 276f, 292
Boolean Completeness **44**, 45, 395, 397
Booleanism 211, 223f, 329, 377
Bucket Argument 60, 82
buckets 59, 82, 108f, 117f, 121, 243, 286, 294, 296

cars 67, 275f, 321
Chance Easiness **95**, 97
Chance Fixity **228**, 229–37, 239, 241ff, 245, 261
Chance Hypertolerance 227f, 232, 244
Chance Independence **95**, 262
Chance Iteration **228**, 231f, 234f, 237ff, 242, 245
Chance Non-contingency **228**, 238f, 244
Chance Persistent Closeness 227f, 238, 243ff
Chance Reflection **231**, 232, 235
Chance Security **95**, 96ff
Chance Tolerance 227f, 233, 237ff, 244
Chance Tolerance Argument **228**
chance, epistemic 232f
chance, objective 39, 42, 65, 105, 121, 163, 166, 170, 260, 310, 324
  Humean theories of 245
Chance/Possibility **232**, 233f, 244
Chance-MP **238**
chess 1, 91, 173, 228, 286
Choice **26**
Classicism 26, 330, 400
closeness
  ancestral 56
  of origins 71
  swap- **353**
Closure Under Composition **223**, 224
Closure Under Conjunction **223**, 223f
C-moderation **370**, 371f
C-non-aboutness **370**, 372

coincidence 265, 275, 377
  necessary 268, 377f
  possibility of 100
  same-kind 100, 285, 297, 381
  see also plenitude
Coincidence Arguments *see* Coincidence Puzzles
Coincidence Plenitude **271**, 272f, 377f, 380
Coincidence Puzzles 249, 255
  involving objective chance 105, 121
  temporal 105ff, 118
Collection Aboutness **368**
collections 43
  see also properties, rigid
colour 215
combinatorialism 215, 280, 331, 337
Compossibility **133**, 135
  see also Weak Compossibility; Weak Plan Compossibility
concreteness 72, 148, 168, 269, 272, 279, 313f, 379
Conditional BF **175**
Conditional Excluded Middle (CEM) **175**, 352
consciousness 306, 319, 364
Constant Aboutness 367f
context-sensitivity 67, 257, 260, 275ff, 297, 303, 320
  in attitude reports 207, 318
  in common nouns 121, 287
  in counterfactuals 120
  in modal operators 10, 276
  in quantifiers 116
  in tense operators 121
  in variables 309
  see also plasticity, semantic
contingentism
  see also necessitism
Converse Barcan Formula (CBF) **37**, 40, 51, 57, 64, 71, 175, 272, 387, 392, 399, 406
Countable Additivity 230
Counterfactual ND **103**, 104, 108, 112
Counterfactual NI **109**, 110
counterfactuals 4, 51, 94, 103, 120, 199, 349, 352
counterpart theory 10, 59, 119, 171, 276
  temporal 249
counting 31, 111, 115, 118, 284, 312
credence 218, 230ff

Death Is Bad **322**, 323
demonstratives 5, 281, 305, 314
descriptions, definite 5, 31, 284
determinism 175, 229
Distinctness **102**, 103, 108, 111f, 114, 121
  see also Forest Distinctness; Ship Distinctness; Modal Ship Distinctness
distinctness, necessity of *see* necessity of distinctness (ND)
donkeys, talking 117, 246, 254
dualism 165, 324ff, 364
Duality **34**, 36, 40, 65, 385, 388, 393, 396

Earring Gradualness **160**, 161
  see also Weak Earring Gradualness
earrings 166, 170, 243, 291, 302
Easiness **90**, 93ff, 97
  see also Chance Easiness
eggs
  mitochondrial transfer in 128
  poached 127, 150, 180
  scrambled 214
Equivalence (rule) **212**, 214
essence 185, 199
essentialism *see* origins, essentiality 126
  extreme 10, 125, 332
Eta-conversion **26**, 214
eternal recurrence 173
existence 279, 379f
  see also concreteness
Existential Distribution (ED) **25**, 408
Existential Generalization (EG) 24ff, 30–3, 38, 205, 406
Extensional Beta-conversion (Eβ) **23**, 24f, 27–31, 37f, 58, 182, 189, 191, 204f, 212, 251, 385f, 388, 391f
externalism 289
  temporal 289
Extreme Anti-essentialism **126**, 128, 150f, 216, 222, 227

fineness of grain 221f, 316
fission 118, 294
fists 185, 188
Forest Distinctness **106**
Forest Non-coincidence **106**
Forest Resiliency **106**
Forest Vacuity **106**
forests 105f, 117f, 285
Four Worlds Paradox 70, 72, 134
four-dimensionalism 280, 323
Fregean Axiom **26**, 33
Functionality **26**, 27
Fundamental Part Plenitude **271**, 273
fundamentality 167, 185, 195, 215
  see also properties, fundamental; objects, fundamental

Generalization (GEN) **25**, 33f, 36f, 220, 385–389, 393f, 403, 405, 408
Generalized Atomic Structure **190**
Generalized Qualitative Atomic Structure 190, **192**, 193–6, 198, 216, 218, 222
Generalized Structure **188**, 189f, 192f
God 124, 330
gold 200, 209, 214f, 344
grounding 26, 149, 185

H_0 (logic)
Heap Hypertolerance **79**, 80ff
  see also Unspecific Heap Hypertolerance
Heap Possibility Transfer **82**
Heap Tolerance **79**, 80ff
  see also Unspecific Heap Tolerance
Hesperus (planet) 29f, 200, 203–7
History Fixity 96, **229**, 231, 236, 347
H_K (logic) **34**
H_{K4} (logic) **40**
H_{K4} (logic) 389
H_{K4R} (logic) **42**
H_{KA} (logic) 385–391
H_{KA} (logic) **35**
H_{KR} (logic) **42**
hole argument 351
holes 305
houses 275, 316
H_{S4} (logic) **40**
H_{S4} (logic) 389
H_{S4R} (logic) **42**
H_{S4R} (logic) 403
H_{S5} (logic) **40**, 223, 270
H_{S5} (logic) 124, 134, 148, 220f, 223, 270, 339, 379f, 389, 395, 397, 403
H_{S5R} (logic) **42**, 271
H_{S5R} (logic) 403
Hyperrobustness **242**, 243ff, 292
Hypertolerance 2, 6f, 9, 11, 54, **56**, 58f, 61, 64f, 67f, 73f, 79f, 84ff, 90f, 93, 99, 101, 108, 171, 173f, 225, 227, 232, 238, 240, 243f, 255f, 262, 280f, 287, 291, 309f, 318f, 324, 335, 338, 342ff, 352, 364
  Ancestral 178–98, 225, 259, 309
  Broad 225
  coarse-grained 324
  temporal 173
  see also Heap Hypertolerance; Chance Hypertolerance; Strong Chance Hypertolerance; Indiscernible Hypertolerance; Swap Hypertolerance

Identity **108**
identity
  confusion 225
  higher order 33, 208f, 259, 277, 316, 368
  necessity of 33, 38, 109, 123
  personal 106, 118, 184, 320–23
  qualitative *see* indiscernibility, qualitative
  transitivity of 110
  see also Leibniz's Law (LL); opacity; necessity of distinctness (ND); counting
Identity of Indiscernibles 8, 352
impredicativity 31
incompleteness theorem 35
Independence **90**, 91, 93ff, 99, 262ff, 266, 281
  see also Chance Independence
indiscernibility, qualitative 123
Indiscernible Hypertolerance 337ff, 346, 351f, 355f, 359, 361f, 379, 382
Indiscernible Iteration **338**, 339
Indiscernible Non-contingency **338**, 339f, 344, 360f, 364
Indiscernible Persistent Closeness **338**, 339
Indiscernible Tolerance 337f, 339–344, 346, 348–51, 353, 355, 359, 361ff, 371, 382
Inextensibility **42**, 43, 45, 49, 125, 129, 391ff, 395–8, 403, 409
Infinity **26**
Intensionalism **40**, 51, 111, 183, 186f, 195, 214, 268, 277f, 316f, 331, 369
  ancestral 40
Irrelevance of U **376**
Iteration 2, 39, **55**, 57ff, 61, 63, 65ff, 69ff, 73f, 80ff, 84ff, 92, 97, 99, 101, 114, 125, 134, 165, 170f, 232, 262, 291, 297, 302f, 361, 366, 370ff, 378, 382, 388f, 398ff, 407
  Ancestral 187, 309
  ancestral 44
  for tense operators 172
  see also KK principle; Chance Iteration; Loose Chance Iteration; Indiscernible Iteration

K (axiom) **34**ff, 35–8, 40, 57, 65, 104, 206, 212f, 219f, 252f, 357, 385–390, 393, 396, 399–402, 407
KK principle 65
knives 269, 276, 279, 373, 379
knowledge
  a priori 204f
  and objective chance 96, 227
  of modality 210, 224
  safety requirement on 92, 164
  self-knowledge 319
  see also KK principle
Kripke frames 47

languages 68, 93, 172, 286
Leibnizian Necessity **50**, 52, 59, 270
Leibnizian Possibility **50**, 52, 59
Leibniz's Law (LL) 8f, 24f, 29ff, 33, 40, 111, 114, 182, 189, 205, 207, 212, 214, 250f, 278
  see also opacity

location 37, 267–270, 325
Location Plenitude **267**, 268–273, 277f, 320, 330
logic
  consistency 202ff
  free 31
  higher order 385–99
  higher-order 3
  logical consequence 14
  logical truth 14, 27, 34
  modal 32–52, 385–399
  model theory 51, 273f
  of rigidity 391–9
  propositional 14
  tense 67, 172, 217, 249, 405
Logical Equivalence **211**, 212, 214
Loose Chance Iteration 238f,
Loose Chance Non-contingency 238f
Loose Chance Persistent Closeness 238f
loose talk 247
lumps of clay *see* statues

Marmite 93
mass 304
matter 70f
  alien 374ff
  see also Tolerance Arguments, originating matter
MaxCon **50**, 51, 395, 398f
metasemantics 296, 299
  self-reference 319
methodology, philosophical 6, 16, 180, 193, 277, 280f
Micro-Indiscernible Tolerance **342**, 343
Micro-Orderlininess **342**, 343f
Microphysical Supervenience **148**, 148–52, 156ff, 162, 165–70, 173f, 177, 180f, 185, 187, 216, 237, 259, 291, 301, 324, 342f, 352, 357, 359, 361, 375, 382
  see also Weak Microphysical Supervenience; Property Microphysical Supervenience; Weak Property Microphysical Supervenience; Qualitative Microphysical Supervenience
microstories **151**, 155, 158, 164, 173, 175, 181, 352
Modal Ship Distinctness **107**
Modal Ship Non-coincidence **107**
Modal Ship Resiliency **107**
Modal Ship Vacuity **107**
modality
  ancestral **35**, 44, 258ff, 309, 385–94
  atomic **73**, 74, 77, 94, 97, 129f, 333f
  broadest 258f, 310, 331
  context sensitivity of 93
  C-indiscernible context sensitivity
  de re 1
  de re versus *de dicto* 5
  deontic 65, 202
  easy 5, 171, 176, 346–51
  epistemic 33, 65, 206
  identity-based 221, 400f
  indiscernible 361
  inner sphere 343
  logic of see logic, modal
  logical 202f
  mathematical 205
  metaphysical 5, 32, 39, 57, 70, 72ff, 79, 87, 93f, 97, 108, 113, 126, 129, 146, 148, 150, 152f, 155, 157, 163, 166, 170f, 178, 227, 232ff, 246, 257, 267–74, 309f, 316, 334, 336, 361, 382, 401
  microphsyical 167
  nomic 65, 94, 166, 210, 349
  U-indiscernible 373
Modus Ponens (MP) 11, **25**
modus ponens (MP) 17
molecules 184f, 187
Mona Lisa, 1, 146, 172

N (axiom) **36**, 65, 104, 213, 253, 385–8, 390, 393, 396, 400, 406
naïve formalization **27**
Naïve Property Conversion **18**, 19, 23, 189
names, proper 5
natural deduction 23
Necessitated Robustness **242**
Necessitated Tolerance **57**, 59, 63f, 66, 79f, 88, 123, 128, 171, 176, 178, 180, 254, 297
necessitation **34**, 212, 385–9, 406
  ancestral 389f
necessitism 57, 195, 279, 373f, 379f
necessity *see* modality
necessity of distinctness (ND) **39**, 40, 43, 49ff, 70f, 73, 108–14, 119, 126, 129–35, 218–22, 249f, 257, 270, 330, 377, 379, 396ff, 400–3
  see also Counterfactual ND
Negative Persistence 43, **396**
nihilism, mereological 197, 276, 278
No Tight Connection **377**, 378, 380
nominalism 197
Non-contingent Robustness 242, 243ff, 292
Non-vacuous Beta-identity **191**, 192f
Non-coincidence **102**, 103, 107f, 114, 117ff, 121, 294ff, 381
  see also Forest Non-coincidence; Ship Non-coincidence; Modal Ship Non-coincidence

Non-contingency 2, **55**, 56, 58, 60f, 63, 66, 73f, 77ff, 82, 85f, 88f, 94, 99, 119, 121, 127f, 136, 152–5, 158, 164f, 170, 176, 178f, 227, 241, 245, 253, 256, 260, 262–7, 280f, 287, 290–3, 307, 309, 324, 327, 332f, 361
  Ancestral 179f, 187
  C-indiscernible 372
  Indiscernible 334
  see also Security Argument; Chance Non-contingency; Loose Chance Non-contingency; Indiscernible Non-contingency
numbers 194, 198
  see also objects, abstract

Object Aboutness 367f
objects 194
  abstract 194ff
  fundamental 167, 185, 197, 280, 313, 350
  higher order reconstruction 196ff
  new 168
  qualitative 330, 368f
  worldbound 378
One Earring Per Point **176**, 301
One Point Per Earring **176**
opacity 32, 183, 203–10, 218, 231
ordered pairs 194
Orderlinineess **340**, 341ff, 349
Origin Uniqueness 131f, 133ff
origins
  essentiality of 75, 129, 186, 349
  sufficiency of 71, 105, 147
Overlap Essentialism **129**, 130–3, 135, 137–44, 293f, 299, 338f, 344ff, 352, 372, 381f
  see also Strong Overlap Essentialism

paintings 1, 68, 286
pairs, ordered 312
panpsychism 165
parthood 197, 268, 270, 313f, 325, 334, 361
particles, fundamental 173
  in value theory 322f
permutation-indifference **357**
Persistence **42**, 43, 45, 48f, 391, 393, 395–8, 403, 409
Persistent Aboutness 367f, 372f, 376
Persistent Closeness **55**, 57f, 63f, 70–4, 77, 80, 85, 154, 179, 253, 255, 262, 291, 335, 371f
  see also Persistent Closeness; Chance Persistent Closeness; Loose Chance Persistent Closeness; Indiscernible Persistent Closeness
Persistent Closeness **57**, 58, 64
Persistent Qualitiveness **329**, 330f, 367
Persistent Rigidity **42**, 45, 49, 391, 393, 395f, 403, 409
personites 323
Phosphorus (planet) 29f, 200, 203–7
physicalism 185
Plan Compossibility
  see Weak Plan Compossibility
Plan Necessity **135**
Plan Sufficiency 107, **135**, 136f
plans 70, 72, 135–8
plasticity, semantic 91, 121, 381
  and speech reporting 315
  in 'categorical' vocabulary 323
  in 'red' 304
  in "categorical" predicates 313
  in common nouns 290
  in model operators 276, 309
  in names and demonstratives 263
  in quantifiers 310f
  in semantic vocabulary 314
plates 299
plenitude 377ff
plural signification 318
possibility *see* modality
Possibility Preservation **103**, 107f
Possibility Transfer **58**, 63f, 84, 238, 243
  see also Unspecific Heap Possibility Transfer; Heap Possibility Transfer
Possible *C*-moderation **372**
Possible *U*-moderation **376**, 380
possible worlds *see* worlds
predicativism 31
Principal Principle 229ff
Principle of Sufficient Reason 124
Prior's theorem 32
problem of the many 115, 265
proper names *see* names, proper
properties 17
  alien 166, 343
  closeness among 66, 70, 152, 333
  forgetful **192**
  fundamental 166f, 185, 215, 328, 331
  higher-order formalization 18f, 195
  intrinsic 160, 320
  microphysical 149, 167
  microspecific **151**, 152–158, 352
  natural 216, 235, 279, 306, 328
  naïve theory of 17, 189
  qualitative 246
  rigid 41–6, 148, 179
  undiscriminating **272**
  unrepeatable **272**
Property Microphysical Supervenience **149**

propositions 19
  higher-order formalization 19
  microphysical 148f, 167, 181, 185
  natural 235
  structured theories of 187
  see also attitudes, propositional; properties
pyramids 1f, 4, 11, 68, 122, 146, 185, 297

qualia 165f, 323
Qualitative Atomic Structure **192**, 194
  see also Weak Qualitative Atomic Structure; Generalized Qualitative Atomic Structure
Qualitative Closure **329**, 330ff, 334
Qualitative Collections **329**, 330, 335
Qualitative Constants **329**, 330, 332, 334f
Qualitative Microphysical Supervenience 168, **341**, 342ff, 349
qualitativeness *see* properties, qualitative 123, 335
quantification 28
  contextual domain restriction 116, 284
  logic of 23ff, 28, 31f, 36–9, 250ff, 254ff
  unrestricted 311
quantifiers
  contextual domain restriction 310
  quantifier variance 310
Queen Elizabeth II (ship) 139, 277
quotation 203, 207

Reflexivity of identity (Ref) 24f, 30, 38, 40, 189, 278
relations
  alien 167
  qualitative 192ff, 328f, 336
    see also properties relativism
Relativity (in physics) 350
Resiliency **102**, 103, 105, 107–10, 114, 117–21
  see also Forest Resiliency; Ship Resiliency; Modal Ship Resiliency
restaurants 162
Rigid Comprehension **42**, 45, 48, 51, 223, 330, 391ff, 395–9, 403
Rigid Extensionality **44**, 270, 330
  rigidity *see also* properties, rigid
roads 104, 117, 140, 275, 277, 281, 292
Robustness **242**, 319, 324, 381
robustness 170, 346, 353, 362
Robustness Arguments **242**, 287, 291f, 324
Robustness Iteration **242**, 243ff
Robustness Transfer **242**, 243f
Russell-Myhill theorem 32, 189–92
Russell's paradox 189

salad 317f
Salmón Tolerance **69**, 70ff, 75f
sandcastles 276, 287, 292f, 312, 316
sausages 277, 281
Security **90**, 91ff
  see also Chance Security
Security Argument **90**, 117, 121, 126, 164, 170, 179, 262–7, 280
  Chance **95**
sets 41
Ship Distinctness **106**
Ship Non-coincidence **106**
Ship of Theseus 106, 131, 142, 172, 280, 286, 295, 300
Ship Resiliency **106**, 107
Ship Vacuity **106**
ships 173f, 274, 285f
Simple Theory of Types 26
Sorites Paradox 91, 159, 176, 178, 280
Sorites paradox 59
spacetime 148, 169, 173, 197, 267–70, 350f
  see also hole argument
Spherical Zone **76**
spoons, nanoscale 364
stamps, nanoscale 94
statues 100, 149, 168ff
Stepwise Necessitated Robustness **242**, 242ff
Stepwise Necessitated Tolerance **57**, 58, 63f, 84, 238
stools 313
Strong Chance Hypertolerance 238f
Strong Chance Tolerance 238f
Strong Overlap Essentialism **133**
Structure **187**, 188f, 191ff, 209, 221, 277
  see also Generalized Structure; Atomic Structure; Generalized Atomic Structure; Generalized Qualitative Atomic Structure; Qualitative Atomic Structure; Weak Qualitative Atomic Structure
Sufficiency **130**, 131–5, 144, 147
  see also Weak Sufficiency; Plan Sufficiency
Super-strong Compossibility **131**
supernovas 74, 336
supervenience 124
Swap Hypertolerance **354**, 355f
Swap Tolerance **353**, 354, 356, 362, 371, 382
Sweet Spot **76**, 285
symmetry 173, 328, 336, 382

T (axiom) **39**, 40, 65, 74, 111, 133f, 177, 212f, 221, 224, 389, 399ff, 407

tables 1, 4, 89, 94, 98, 120, 129, 146, 166, 181, 185f, 228, 272, 284, 293, 316, 333f, 339, 361–4, 370–6, 380–3
  3D printed 141
  bucket-shaped 134
  golden 344ff
  IKEA 84, 140f, 286, 333
  indiscernible 354, 358
  made of smaller tables 285, 294
  metal, stamped 144
  nanoscale 84
  nearly indiscernible 355
  one-legged 84
  recycled 134, 136, 355
  replacement of parts of 136f
time 65, 105ff
  A- and B-theories of 66, 169, 173, 267, 271
  see also Tolerance Arguments, temporal
Tolerance 2, 54ff, 58, 60–4
  iterated 85
  see also Salmón Tolerance; Heap Tolerance; Chance Tolerance; Strong Chance Tolerance; Indiscernible Tolerance; Swap Tolerance
Tolerance Arguments
  chance 95f, 227
  de re **57**
  Indiscernible 337
  originating matter 91, 181, 307
  temporal 93, 286
Tolerance Puzzles *see* Tolerance Arguments
triangles 236
Tri-State Area 215, 259, 368f
tsunamis 140
Twin-Earthability 289, 322
two-dimensionalism 289

Uniform Tolerance **76**, 285
Universal Distribution (UD) **25**, 36, 386f, 408
Universal Instantiation (UI) 23f, **25**, 29, 31ff, 36f, 40, 49, 189, 203, 248, 250f, 253–6, 309, 386ff, 391, 403
Unspecific Heap Hypertolerance **81**
Unspecific Heap Possibility Transfer **81**, 82
Unspecific Heap Tolerance **81**, 82

Vacuity **102**, 103, 108
  see also Forest Vacuity; Ship Vacuity; Modal Ship Vacuity
vagueness 159, 178, 186, 217f, 291, 299, 303, 324, 343, 380, 382
Vienna Circle 1, 4
vision 268

Weak Anti-haecceitism **124**, 124ff, 128, 147, 332f, 335–8, 350f, 356, 358f
Weak Compossibility **132**, 133f
Weak Earring Gradualness **161**
Weak Microphysical Supervenience **148**, 150, 152f, 177, 354ff
Weak Plan Compossibility **135**, 136
Weak Property Microphysical Supervenience **149**
Weak Qualitative Atomic Stucture **192**
Weak Sufficiency **133**, 134, 147
Witnessing **175**
worlds 8f, 58, 201, 247
  as collections of propositions 201
  as propositions 49

zombies 166