1
Introduction

1. SOME QUESTIONS

'If the American ambassador had understood her instructions, Iraq would not have invaded Kuwait.' 'If Shackleton had known how to ski, then he would have reached the South Pole in 1909.' 'If rabbits had not been deliberately introduced into New Zealand, there would be none there today.' These are probably all true, and we can hear, think, and say such things without intellectual discomfort.

Philosophy teaches us how to inspect familiar things from an angle that makes them look disturbing and problematic. What Locke wrote of stones and iron also holds for thoughts and statements: 'Though the familiar use of things about us take off our wonder; yet it cures not our ignorance' (*Essay Concerning Human Understanding* III.vi.9). What first drew philosophers' attention to conditionals like the three above was the question: how can we know whether they are true? If any of the three is true, it is contingently so: it is a truth about what happens to be, not what must be, the case. We discover contingencies by examining the actual world; but how are we to do that for conditionals that concern what *would* have ensued if something *had* been the case that in fact *was not* the case? How can we learn about such matters by investigating how things stand in actuality?

It easy to gesture vaguely towards the right answer to this question. But getting the details right, and finding the best theory in which to express and explain them, proves to be difficult and rewarding.

Other conditionals are not openly and explicitly about what would have ensued if things had been different. 'If they have chains on their tyres, there must be snow in the pass.' 'If he learned that from Alice, then she is smarter than I thought.' 'If they scheduled the meeting for Wednesday, I must be losing my memory.' These also raise a problem about verification: when we beam our senses on the actual world, we learn unconditional things—P, Q, and R—and it is not obvious how these discoveries guide our judgements on conditional matters, how they tell us that if S, then T.

To answer this we must become clear about what these conditionals mean—a formidable task. Any conditional in this second group is false if its antecedent (the part between 'If' and the comma) is true and the consequent (the part after the comma or after 'then' if that occurs) is false. Does that yield the whole story about what such a conditional means? If so, then

> If he learned about that from Alice, then she broke a promise to me

simply means

> It is not the case that: he learned about that from Alice and she did not break a promise to me.

Some philosophers have accepted this account of what the conditional means, but nearly everyone now rejects it (§9). Well, then, what *else* does the meaning of such a conditional contain? We shall explore that in Chapters 3 and 6–7.

The first three examples all contain '. . . had . . . would . . .', whereas the second trio do not. This reflects the fact that conditionals are of two principal types. To see how different they are, consider this pair: (1) 'If Booth did not shoot Lincoln, someone else did.' (2) 'If Booth had not shot Lincoln, someone else would have.' These are both of the form 'If A, C'; each A concerns Booth's not shooting Lincoln, and each C concerns someone else's doing so. Yet every informed person accepts 1, whereas 2 is disputable.

It has not been easy to understand exactly where the line falls between the two types, and what puts it there; and philosophers have disagreed about this. In Chapter 22 below, I shall support one of the views by disproving the only rival that has seemed plausible to anyone; but that proof lies on the far side of a large philosophical terrain.

Some difficult questions that an analytic philosopher might tackle are not worth the trouble. Paul Grice told me that a discussion group of J. L. Austin's wrestled with this question:

> For some adjectives, the phrase 'very [adjective]' can be changed, with no significant change of meaning, to 'highly [adjective]'; for other adjectives the 'highly' version sounds wrong. What generates or explains the line between the two classes of adjectives?

In the first class we have 'excited' and 'intelligent', in the second 'bored' and 'stupid'. It may occur to you that some adjectives have an *up* or *down* element in their meanings, suggesting that 'highly' can stand in for 'very' with the up adjectives but not the down ones. Not so, however; 'highly conceited' is all right, but 'highly vain' is not. Nobody has found the answer to this question, it seems, and I know of no reason to care. One could tackle such a problem only for the mountaineer's reason, *because it is there*.

The pursuit of a good analytic understanding of conditionals is doubly unlike that. In many philosophical areas we *use* conditionals as load-bearing parts of theoretical structures, and sometimes our use of them reflects assumptions about how they work—assumptions that could be challenged by a good analysis. Also, working through the central problems about conditionals forces one to learn things in many philosophical areas: the metaphysics of possible worlds, probability and belief-change, probability and logic, the pragmatics of conversation, determinism, ambiguity, vagueness, the law of excluded middle, facts versus events, and more. After each of the five graduate seminars on conditionals that have I taught, several students told me, in almost exactly these words, 'This was the hardest seminar I have ever taken, and the most profitable.' I report this in praise of my topic, not of my teaching.

Having admitted difficulty, I add that this book ought to be manageable by any student who has some solid analytic philosophy behind him or her. The literature on conditionals is harder than it needs to be: many contributors to it take such work in their stride, and neglect the needs of the less well muscled, who must struggle to keep up. I have needed much time, effort, and help from friends and colleagues, to get on top of various things that others manage more easily. My travails warn me against cutting corners fast when explaining things to others.

I aim to work through the main parts of the fairly recent philosophical literature on conditionals, showing what goes on in them and helping the reader to see that this topic is—like philosophy in general—tough but manageable. For some account of earlier work, see Sanford 1989: 13–45.

2. DEFINING 'CONDITIONAL'

We encounter a conditional through a sentence expressing it, that is, a sentence whose meaning it is. How can we tell whether a given sentence expresses a conditional? One might look to its superficial features for an answer:

> S: An item is a conditional if it is expressed by an English sentence consisting of 'If' followed by an English sentence followed by 'then' followed by an English sentence.

Conditionals are often expressed without 'if'—for example, 'Had the civil war not been fought, American slavery would have continued into the twentieth century'. Still, the meaning of that sentence could be expressed in the 'If . . .' form; so we might replace S by:

S*: An item is a conditional if it is *expressible* by an English sentence . . . etc.

This does not read off conditionalness purely from the surface of the sentence, but it comes close to that. The changes needed to get something into the 'If . . .' form are trivial and obvious.

S* fails to cover all the ground, because some conditionals have 'If' but not 'then', and some of those differ significantly from all that do contain 'then' (§91). Also, conversely, S and S* both count as conditionals some items that are not so. In response to someone's assertion of S, James Cargile devised a counterexample (unpublished). Let 'Only Joe would come to the party bringing his wife who always says at first that she won't come' be the first sentence, and let 'She does come' be the second. Assemble those according to S and you get:

If only Joe would come to the party bringing his wife who always says at first that she won't come then she does come.

This undistinguished English sentence is plainly not a conditional, but rather an optative, like 'If only she would let me kiss her!'

To explain why it is not a conditional, we have to go further down from the surface. Every conditional applies a binary (= two-place) conditional operator to a pair of propositions; that is, it means something of the form $O_2(A,C)$. Cargile's optative, on the other hand, applies the singulary (= one-place) operator 'If only [it were the case that] . . . ' to a single proposition; that is, it means something of the form $O_1(P)$. It may look like 'If A, then C', but that is an illusion.

Perhaps, then, we can define 'conditional' as 'item expressible in a sentence of the form "If [sentence A], then [sentence C]", the effect of the whole being to apply a binary operator to propositions expressed by those two contained sentences'. If this were the best we could do, our recognition of conditionals in other languages would depend upon our translating items in those languages as meaning what 'if' does in English. That seems to go backwards: the best way to discover that some Turkish word means the same as our 'if', one would think, is to learn how the Turks use it in expressing conditionals. Still, this English-dependent account might be about the best we can do at this preliminary stage. It lets me capture in my net swarms of obviously genuine conditionals, perhaps indeed all of them (if I adjust it to cover conditionals that do not contain 'then'); and with luck it does not capture too much else. A stronger, deeper account can emerge from the analysis (or analyses) that we eventually come up with. In philosophy we often have to start with rough criteria—or even a mere ostensive

list—to pick out a lot of members of some class, develop an analytic account of how they work, and treat that as implicitly defining the class.

Admittedly, that preliminary account does capture some items that we would rather exclude. For example, 'She was always home before midnight: if she missed the bus, then she would walk' (Dudman 1984*a*), where 'if' means about the same as 'whenever'. Again, when my colleague hears me say that I have applied for leave, he may say 'If you have applied, I'm going to apply too' (Akatsuka 1985). This is not a conditional, either. My colleague means 'Because you have applied, I'm going to apply', and he uses 'if' because he is still digesting the news about my application. Had he known about it for a month, that use of 'if' would be unnatural. Having noted these two kinds of sentence, we can avoid tripping over them; but they do imply a further imperfection in my preliminary account of what a conditional is.

My rough, preliminary account of what I take 'conditional' to mean fails to fit one large class of conditionals unless it is pulled into shape around them. That is because of a fact that was (so far as I know) first described in Gibbard 1981*b*: 222–6. In the conditional:

If the British did not attack Suez, the Israelis rigged the situation to make it look as though they had,

we find the two sentences

The British did not attack Suez
The Israelis rigged the situation to make it look as though they had,

each of which occurs in the conditional with the same meaning that it has when it stands alone. Contrast that with the following:

Con: If the British had not attacked Suez, Soviet troops would not have entered Hungary a few weeks later.

This has nested within it these two sentences:

Con_1: The British had not attacked Suez
Con_2: The Soviet troops would not have entered Hungary a few weeks later.

Though Con_2 is a complete sentence, it is hardly usable except in compounds with other sentences. Con_1 clearly has a stand-alone meaning, but not the one it bears as an ingredient in that conditional. Standing alone, it involves the notion of a *past past*—a time that was already past at some understood past time. For example: 'Before you criticize his naivety, remember how things stood when he entered the Foreign Office. The British had not attacked Suez

. . .' and so on. The conditional does not use this *past past* tense. (See also Dudman 1984*a*, *b*.)

I share the common opinion that Con should be represented as having the form $O_2(A,C)$, but Gibbard's point warns us that its A and C are not Con_1 and Con_2, the sentences appearing on its surface. I take Con to mean something of the form

> O_2(The British did not attack Suez, Soviet troops did not enter Hungary a few weeks later).

We can even use those sentences to express what Con expresses, by writing 'If it had been the case that . . . it would have been the case that . . .'. I shall revert to this matter in §6.

When I speak of the 'antecedent' (or 'consequent') of a conditional, I mean the first (or second) sentence that occurs *when we represent it in the form $O_2(A,C)$*. So I shall say that the above two conditionals have the same antecedent, though it does not appear on the second conditional's surface.

3. CHALLENGING THE TERNARY STRUCTURE

My loose account of what conditionals are would be rejected by Gilbert Harman (1979), who has argued that 'if' never stands for an operator on a pair of propositions. His alternative account of it, significant and challenging though it is, seems not to have won acceptance. What motivates it is Harman's desire for a unified treatment of 'if' that covers its use in conditionals and also in sentences like 'He asked her if she would marry him'. Unity is always worth wanting, but in this case, I believe, it cannot be had.

V. H. Dudman (forthcoming) also denies that 'if'-sentences (as he calls them) mean something of the form $O_2(A,C)$. Rather than seeing 'If the bough breaks, the cradle will fall' as having a 'ternary' structure in which one item ('if') links two other items (constituent sentences), Dudman argues for its being a two-part item, the first part of which is 'If the bough breaks', which he calls 'a string beginning with "if"', or an '"if"-string', for short:

> Examples like the following demonstrate that an 'if'-sentence has a *string beginning with 'if'* for a constituent:
>
> > . . . the moment for such frivolities, if it had ever existed, was now past. . . .
> >
> > The knowledge that there was a leak, if it became public, could be more damaging than the leak itself. . . .
>
> In them we descry that string interrupting a *prior sentence* ('The moment for such frivolities was now past', 'The knowledge that there was a leak could be more damaging than the leak itself') which it could synonymously have preceded or followed: there is no doubting its integrity. And then the point is simply that the commonplace ternary structure would sunder this undoubted constituent. (Dudman (forthcoming))

This is part of Dudman's basis for his theory about how 'if'-sentences work grammatically, and what their role is in our lives. I share the fairly general disbelief in this theory, and can devise no way of engaging with it that would profit you or me. However, you may take a sceptical view of my attitude, for either of two reasons. If Dudman's theory were correct, most of this book would be wrong; and Dudman sometimes decorates his theoretical themes with a descant of amused contempt directed at other writers on conditionals. So you may suspect that my rejection of his theory arises from fear or from indignation.

Perhaps in part it does, but it also has rational support. In §137 I shall briefly cast doubt on Dudman's account of what the two kinds of 'if'-sentences are *for*. Chapter 22 as a whole will constitute a decisive argument against a main pillar of his position, namely a thesis about where to draw the line dividing 'if'-sentences into two main groups (§6). And right now I shall comment briefly on his grammatical 'demonstration', reported above.

Dudman writes that 'the commonplace ternary structure would sunder this undoubted constituent', namely the 'if'-string. Would it? That depends on what you mean by 'sunder'. When we regard the string as consisting of two constituents—'if' and a sentence we call the 'antecedent'—we need not deny that in certain transformations the string holds together. However strong its integrity, there are things to be said about its 'if' component, and others to be said about the remainder of the string; it would be astonishing if there were not. The things that are separately sayable constitute much of the literature on 'conditionals', as we call them; all of that could not be destroyed by Dudman's point about the 'integrity' of 'if'-strings.

Lycan 2001: 6–15 is like Dudman's work in adducing syntax in partial defence of a semantics of conditionals. Lycan does not contend that any rival semantic analysis conflicts with the syntactic facts, but I think he holds that some of the latter can be explained by his analysis and not by any other. I am not convinced of that, but those pages are worth study.

The syntactic facts that Lycan emphasizes seem to have no overlap with the ones highlighted by Dudman.

4. TWO TYPES OF CONDITIONAL

Like most students of them, I hold that conditionals fall into two importantly different species. We have met one sample pair, and here is another:

Did-did: If Shakespeare did not write *Hamlet*, then some aristocrat did.

Had-would: If Shakespeare had not written *Hamlet*, then some aristocrat would have.

Without having a full analysis of either, I have committed myself to understanding Had-would as meaning something of the form:

O_2(Shakespeare did not write *Hamlet*, Some aristocrat wrote *Hamlet*);

and Did-did also means something of that form, with the same antecedent and consequent. Since the two are plainly not equivalent—one being probably true and the other daft—they must use different conditional operators (thus Adams 1970).

There are countless other such pairs, for example: 'If this tyre did not get a puncture, it had a blowout'; 'If this tyre hadn't got a puncture, it would have had a blowout'. The members of each pair are strikingly alike, yet they differ in what it takes to make them true or acceptable. In my examples so far, the Did-did conditional is acceptable while the Had-would one is not, but it can be the other way around. The Department wants the eulogy for our honorary graduand to be written by either Alston or Bennett, and we two decide that he will do it. Asked later who wrote the eulogy, I reply that I think Alston did, adding 'If not, I don't know who wrote it.' I would reject 'If he didn't, I did', because I know that I didn't; but I may well accept 'If he hadn't, I would have'.

We want a good analytic understanding of each of the two kinds of conditional. One might hope to find a Y-shaped analysis of them—first stating what is common to the two kinds and then bifurcating in order to describe the differences. That they have much in common seems clear. Each involves a binary operator on propositions; and we shall see that the two operators share most of their logical properties (§69). This nourishes the hope for a Y-shaped analysis with a long, thick stem.

In my final chapter I shall expound a Y-shaped account by Wayne Davis, and another by Robert Stalnaker. Each gives for the two types of conditional a unified account that is nearly complete, lacking only a choice between two settings of a single parameter, one for each type. These are rival analyses, which cannot both be right; and I shall argue that each is wrong.

Can some other Y-shaped analysis succeed? Allan Gibbard, in an important paper, answers No. He goes further: the two types of conditional, he startlingly suggests, differ so deeply that their logical similarities are 'little more than a coincidence' (1981*b*: 211). Robert Stalnaker has argued powerfully against that (1984: 111–13), and I agree with him. In §69 I shall *explain* the logical likenesses between the two, not generating a Y-shaped analysis, but at least exorcizing the 'coincidence' spectre.

One kind of conditional is the topic of my Chapters 2–9, the other of 10–21. So the main dividing line through conditionals, with Did-did on one side of it and Had-would on the other, is this book's central organizing feature. I need, therefore, to place the line correctly. If I put some conditionals on the wrong side of it, I may try to make my analyses do impossible things.

Here is an attractive project: (1) develop criteria for sorting conditionals into two groups, then (2) sort them on the basis of those criteria, then finally (3) analyse the two sorts of conditionals. But we ought not to assume that this can be done. One important writer on conditionals has written that a sorting criterion 'might more properly be expected to *accompany* and not to precede the formulation of an adequate theory' of one of the sorts of conditionals (Adams 1975: 104). If so, then our (2) sorting of individual conditionals—our decisions about which belong to which type—cannot look back to (1) superficial recognitional criteria, and so must look ahead to (3) the final analyses. The procedure will be to start with what we hopefully take to be paradigm members of each sort of conditional, get analytic ideas about how they work, and then find out how far each analysis extends—that is, which other conditionals it fits.

Even if we must proceed in that way, it would be nice if our launching paradigms were at least marked off by some reliable rules of thumb. Here is an example of what I mean. H. P. Grice (1957) offered an analysis of the concept of 'non-natural meaning', as he called it. That is the concept that is most relevant to language, and is at work in

Those three rings on the bell (of the bus) mean that the bus is full

and not in

Those spots mean that he has measles.

We have an intuitive sense of the difference, but Grice did more than offer examples and appeal to our intuitions. He pointed out that in the first example, (1) one could go on to say 'But in fact it is not; the conductor made a mistake'; (2) it is all right to talk about 'what was meant' by the three rings; (3) the statement implies that by making the three rings *someone* meant that the bus is full; (4) the statement could reasonably be expressed in the form 'Those three rings mean "The bus is full"'; (5) the statement is not equivalent to 'The fact that the bell has rung three times means that the bus is full' (the bell might have rung three times because of a short-circuit). Those five truths about the first example—which are also, *mutatis mutandis*, true of most examples that Grice wanted to group with it—are all false of the second. You cannot properly say 'Those spots mean that he has measles, but he does not have measles', or 'What is meant by those spots is that he has measles', or 'By those spots someone means that he has measles', or 'Those spots mean "He has measles"'; and you *can* say 'The fact that he has those spots means that he has measles'. Grice offered those five differences as drawing the line separating the uses of 'mean' that his analysis aimed to cover from ones it did not. Then he proceeded to the analysis. His way of demarcating his topic is not verbal and mechanical—it does go into meaning, but not far enough to be controversial.

I have nothing as substantial as that. For initially segregating conditionals into two main groups I can offer only something almost purely verbal and superficial:

> In every conditional of one group, and in no conditional of the other, the sentence expressing the consequent has 'would' as the auxiliary of its main verb.

In my examples up to here, the verb 'had' has occurred in the antecedent, but that is not essential. Sometimes it is 'were': 'If he were to reduce taxes by that much, the country would be bankrupt within a decade.' Sometimes neither word occurs: 'If the Palestinians declared statehood now, the Israelis would retaliate'. The common thread through all these is just 'would' in the consequent (strictly, in *expressing* the consequent, but I shall sometimes cut that corner). This is not purely verbal, because it requires recognizing an operator on two propositions, the antecedent and consequent.

The criterion as stated is not quite correct, because some uses of 'would' do not put a conditional in the 'would' group, and recognizing those requires going below the surface a little. Some involve the idiom 'would like'. What is the word 'would' doing in 'I would like to go for a swim'? The thought might be a conditional one: 'If I were to go for a swim, I would like it', but more often it will categorically express a desire, meaning 'I want to go for a swim'. Now, it could be used in this way in the consequent of a conditional, such as 'If you will come with me, I would like to go for a swim'; and I do not classify this along with the main bulk of 'would' conditionals. I expect agreement about this. I wish I had a neat way of saying why it is right.

There can also be trouble from uses of 'would' to express the notion of a *past future*, as in 'Some day he too would stand upon that scaffold' and 'She announced that they would be married in the following month'. That use of 'would' could also be built into the consequent of a conditional that does not belong in the 'would' group. Not knowing how Ambrose Bierce died after he disappeared into Mexico in 1913, someone might write: 'If he was jauntily setting off for Mexico intending to spy on Pancho Villa, then he would die a violent death within the next two months.' That is the past-future use of 'would', not the use that puts a conditional into the 'would' group.

The imperfections in my gesture towards a line of demarcation are not fatal. The final line will emerge out of the analyses of the two types of conditional.

5. LABELS FOR THE TWO TYPES

We need labels for the two kinds of conditional. It would be good to label each in a way that helpfully describes it, rather than being stipulated and meaningless ('Type One', 'Type Two'), or meaningful and false. Unfortunately, we have no labels with this virtue.

One popular nomenclature calls the 'would' group *subjunctive conditionals* and the other *indicative conditionals*, with those adjectives often used as nouns. John Pollock wrote a book about the former, entitled *Subjunctive Reasoning*.

The conditionals that are called 'indicative' under this proposal are indeed all in the indicative mood, but that does not make the label a good one because most and perhaps all of the others are in the indicative mood also. The subjunctive mood has some slight use in English, in expressions of wishes such as 'God help you', and 'Would that he were here!', but English has never worked it hard, as do Latin and French. In 1860 a textbook on English said: 'The subjunctive is evidently passing out of use, and there is good reason to suppose that it will soon become obsolete altogether' (see *OED*, 'subjunctive'). Seventy years later the prediction was almost fulfilled: 'The word "subjunctive" is almost meaningless to Englishmen, the thing having so nearly perished'; 'As a matter of grammar, the instinct for using subjunctives rightly is dying with the subjunctive' (Fowler 1931: 163, 166; see also Zandvoort 1963). I have found no grammatical authority supporting the claim that the conditionals of the 'would' type (as I am provisionally calling them) employ the subjunctive mood. When we move to nearby languages that do have a robust subjunctive, we find that in German the equivalents of the 'would' conditionals are indeed expressed with the *subjunctive mood* but that French and Spanish do the work with something different, the *conditional tense*.

The other popular label for conditionals of the 'would' type is 'counterfactual', used both as an adjective and as a noun. David Lewis's book on them bears the title *Counterfactuals*. This comes from the idea of the antecedent's being false (counter to fact), but exactly *how* that idea is involved is not obvious. You can easily satisfy yourself that it is not that someone who asserts the conditional thereby commits himself to the antecedent's being false, or that the truth of the conditional requires the falsity of its antecedent. It may be nearer the mark to say that someone asserting the conditional ordinarily expresses his confidence that the antecedent is false without outright saying so. 'If you had called 911 when I asked you to, the ambulance would be here by now' would normally be said only by someone who thought the hearer had not made the call. Without his actually *saying* this, his form of words does somehow imply or suggest his confidence in the falsity of the antecedent.

Even if that were right for all conditionals in the 'would' camp, it would not make 'counterfactual' a good label for them. It would be based on a feature that has nothing to do with the antecedent's being contrary to fact, but only with the speaker's thinking that it is so. Also, secondly, 'counterfactual' is not matched by a corresponding label for the other type of conditional, which nobody has called 'factual' or 'profactual' or the like. Thirdly, it is unsatisfactory to base a label for a class of conditionals on a fact about what is merely implied or suggested rather than asserted outright in its instances. (In this I agree with Woods 1997: 8 n.)

A better basis for 'counterfactual' comes from the fact that conditionals in the 'would' camp can properly be accepted and asserted by someone who is absolutely certain that the antecedent is false. Even, indeed, by someone who would not accept the conditional if he thought its antecedent to be true; that will not always be so, but the 'would' form of words leaves the door open to it. Nothing like that holds for conditionals of the other type, in which 'would' does not occur.

Although this is solider than the simple 'speaker's disbelief' basis, it still does not make the label 'counterfactual' a good one; and we still have no good matching label for the other type of conditional. Furthermore, much analytic work is needed properly to explain and justify this latest contrast between the two kinds of conditional (§88).

Holding my nose, I adopt the labels 'subjunctive' and 'indicative'. Fortunately, their defenders never claim that 'subjunctive' helps us to understand conditionals of the type to which they apply it—what its analysis is, the roles of such conditionals in inferences and in our lives. The word serves only to remind them of the primacy of 'would' in (most of) the conditionals to which the label is applied, and that reminder works well enough, despite reflecting an error about what the subjunctive is.

In Chapter 12 and thereafter, I shall sometimes need to focus on subjunctive conditionals whose antecedent is false—not assumed to be false by someone, but actually false. I shall use 'counterfactual' as a label for those. When others use the word in passages I shall quote, they always mean what I mean by 'subjunctive'.

6. THE RELOCATION THESIS

Some students of conditionals deny that the word 'would' helps to draw the principal line. They group the conditionals that I call 'subjunctive' together with many that I count as 'indicative', namely ones like:

Does-will: If you swim in the sea today, your cold will get worse.

If conditionals of this Does-will type belonged in the same hopper as the subjunctives, my use of the label 'subjunctive' would be indefensible, and my preliminary drawing of the line would be wrong.

I used to be a *relocator*: I wanted to relocate the main line through conditionals so as to put the 'would' conditionals and also the Does-will ones together on one side of it (Bennett: 1988*b*). I have since seen the error of my ways. My first defence of my switch (1995) went off at half-cock, but I got it right in 2001. In §§134–5 below I shall, developing material in the latter paper, prove that the relocation thesis is wrong. (Because the labels 'indicative' and 'subjunctive' are intolerable if the relocation thesis is correct, I once coined the terms 'straight' and 'corner' for the two kinds of conditional, on a hint from symbols that are often used to express them, namely → and >. A few writers have followed me in this. I apologize.)

It will help me, tactically, to expose the relocation thesis if I first expound two of the most persuasive arguments in its favour.

The first has to do with conditions of acceptability. Suppose the facts are such that I am right when I say to you:

Does-will: If you swim in the sea today, your cold will get worse.

Suppose further that you do not swim in the sea today, and that tomorrow I tell you:

Had-would: If you had swum in the sea yesterday, your cold would have got worse.

It seems reasonable to think that my Does-will conditional was acceptable at the earlier time if, and only if, my Had-would conditional is true at the later time: the two stand or fall together. So perhaps they are merely two wordings for a single conditional, differing only in tense, in which case Does-will belongs with the subjunctives, despite its not containing 'would'.

That plausible line of thought has seduced many of us. In §§134–5 I shall show what is wrong with it.

The other argument cannot be so briefly stated, and is harder to dislodge (Dudman 1983, 1984*a*, *b*). Consider four ways for a statement to pick out a moment or period of time. (1) *Past past.* In the sentence 'We were rolling up the tent, and Charles had doused the camp fire', the verb phrase 'had doused' puts the action at a time earlier than the past time that is primarily being spoken of. (2) *Simple past.* In 'We were rolling up the tent' or 'We rolled up the tent', reference is made to a past time, but not as past relative to some other past time that is also spoken of. (3) *Present.* (4) *Future.*

The verb forms that normally refer to past past time in simple sentences can also do so in the antecedents of conditionals. For example: 'We were wondering whether Charles had doused the camp fire, and I still do not know whether he had. *If he had doused the fire, he had been very quiet about it.*' In some antecedents, on the other hand, the 'had'-plus-participle form pertains not to (1) the past past but to either (2) the simple past: 'If Hitler had invaded England instead of Russia, he would have won the war'; or (3) the present, as in 'If Antoinette had been here, we would not be drinking ale'; or (4) the future: 'If the auditors had come tomorrow, they would have found the books in perfect order.'

The verb forms that normally refer to simple past time can also do so in the antecedents of conditionals. For example: 'I do not know whether it was Charles who made the anonymous gift. *If it was, I have misread his character.*' But in some antecedents those same verbal forms take us not to (2) the simple past but rather to (3) the present, as in 'If the Provost was [or: were] plotting against the President, he would be on campus now rather than sunning himself in Florida'; or (4) the future, as in 'If the Provost plotted against the President, he would lose'. In this last, other readings are possible, but 'plotted' *can* be understood as taking us to the future.

Finally, the verb forms that in simple contexts take us to the time at which the utterance is being made can also do so in antecedents of conditionals, as in: '*If John loves Mary, he has a funny way of showing it*', where 'loves' clearly has to do with John's feelings about Mary at the time of utterance. Present-tense forms, however, serve in certain other antecedents to refer not to (3) the present but only to (4) the future, as in: 'If it rains tonight, the river will be in flood tomorrow': 'If you swim in the sea today, your cold will get worse'. Each of those antecedents contains a present-tense verb—'rains', 'swim'—referring to the future.

This is the phenomenon of *forward time shift*, in which a 1 verb form does the work of 2, 3 or 4, a 2 form does the work of 3 or 4, and a 3 form does the work of 4. Some conditionals force a forward time shift on the principal verb in the antecedent, others do not.

Dale and Tanesini (1989) argue for taking this lightly. The time-shift criterion, they contend, appears to be a localized English phenomenon, not appearing in Italian or—they assume—in other languages. They conjecture that it is explained by the idiosyncratic willingness of English to 'allow the present tense in simple non-compound sentences which clearly express propositions about the future: John is taking his exam tomorrow, John takes his exam tomorrow'. Yet Spanish, French, German, and Turkish all have this feature, as no doubt do other languages; though not all of these have a time shift in some conditionals. Gibbard may be right in suggesting that time-shift conditionals 'form a significant grammatical class' (1981*b*: 226).

Dudman has proposed the line between time-shift conditionals and others as a basis for separating the two great classes of conditionals. It has the effect, as did the previous point about what conditionals stand or fall with what others, of grouping the Does-wills with the 'would' conditionals that I call 'subjunctive'. (Or rather: grouping the time-shift Does-wills in that way. Some Does-wills lack the time-shift feature—for example 'If he loves her, he will marry her' (Thomason and Gupta 1980: 299).) Whether the grammatical distinction has that semantic effect, however, is controversial. Among those who think it does not, Edgington (1995*b*: 314) is alone in offering a different and milder suggestion for what the time shift signifies.

The forward time shift was first described in detail by Gibbard (1981*b*: 222–6). He uses it as a basis for calling Does-will conditionals 'grammatically subjunctive', while apologizing for the label, and reserves 'grammatically indicative' for conditionals that lack the time-shift feature. But he does not use that to reorder the domain of semantic enquiry. On the contrary, he announces that he will 'leave aside' the Does-wills (229), while defending one semantic analysis for the remaining indicatives and another for subjunctives. Despite this caution, at p. 228 he gives a good reason for putting the Does-wills semantically with the ones he calls 'indicative'.

Without having got to the bottom of the time-shift phenomenon, I decline to follow Dudman in allowing it to force a change in how we classify conditionals.

For the next eight chapters my topic will be indicative conditionals, considered on their own. I shall regularly use → to express the operator of such a conditional, so that 'If you swim in the sea today, your cold will get worse' can be expressed as 'You will swim in the sea today → Your cold will get worse', or as 'Swim → Cold', for short. Mostly, though, I shall combine the arrow with dummy letters, writing 'A→C' to stand for any indicative conditional.

7. INDEPENDENT CONDITIONALS

A different line through conditionals also helps to shape this book. It has both subjunctives and indicatives on each side of it. Here are two subjunctives, one on each side of this new line:

(1) If the river were to rise another two feet, the subway system would be flooded.

(2) If the river were to rise another two feet, it would be two feet higher than it is now.

Of these, 1 is contingent, 2 is necessary; but that is not the essence of the distinction I am drawing. What I care about is that in 1 the consequent is reachable from the antecedent only with help from unstated particular matters of fact, while in 2 one can get the consequent from the antecedent without input from any matters of particular fact. I shall say that 1 is a *dependent* conditional, 2 an *independent* one.

An independent conditional may not be logically necessary; it may depend on causal laws alone. In contrast to

If that cyclist had been on the other side of the mountain two hours ago, he would not have been here now,

which depends on unstated particular matters of fact, we have the independent conditional

If that particle had been two light years from here a month ago, it would not have been here now,

which holds as a matter of sheer physics. The other basis for independent conditionals is morality. As well as fact-dependent conditionals such as

If she had told her husband about his prognosis, she'd have acted very badly,

we have independent ones like

If she had tortured a child just for the fun of it, she'd have acted very badly.

Those are the three categories of independent conditionals: logical, causal, and moral. Most examples in the conditionals literature are logical.

Those contrasting pairs are all subjunctives, but the line also cuts across the indicatives. Thus, logical:

If the closing date is Tuesday the 14th, then we shall have to hurry with the paper work.

If the closing date is Tuesday the 14th, then the closing date is a Tuesday.

Causal:

If two more people get into the dinghy, it will sink.

If the dinghy and its contents come to weigh more than the same volume of the water they are floating in, the dinghy will sink.

Moral:

If you lied to him about that, you did him the worst disservice you possibly could.

If you lied to him about that, you did something morally questionable.

Solidly independent causal and moral indicative conditionals are hard to come by, but that does not matter. What matters is to grasp the difference between dependent and independent, and not to identify it simply with the line between contingent and necessary.

Whereas the indicative/subjunctive line cuts through the middle of this book, the dependent/independent line runs around its perimeter. I plan to keep independent conditionals out of sight, and out of mind except as unloved exiles, occasionally mentioning them at places where I think one might be tempted to readmit them to the arena.

One reason for this attitude is that in various parts of our philosophical terrain, useful truths about dependent conditionals do not fit the independent ones. For example, many philosophers believe, in my opinion correctly, that an indicative conditional is useful, acceptable, worth asserting or at least considering, only to someone who regards its antecedent as having some chance of being true. 'If I am at this moment sitting in the visitors' gallery of the House of Commons . . .'—I can do nothing with this, because I am dead certain that I am sitting at my desk looking out at Bowen Bay. But now consider: 'If I am at this moment sitting in the visitors' gallery of the House of Commons, then I am sitting somewhere.' Whatever my doubts and certainties, doesn't that one take me by the throat and command assent, on the grounds that logic warrants it? Yes. I suppose it does. But it is a nuisance, spoiling an otherwise good bit of philosophy, and offering nothing by way of recompense.

That brings in a second reason for snubbing independent conditionals: outside the realm of logic (broadly construed), they are not useful or interesting; any work they do can as well be done in other ways. The examples I have given can be replaced by: rising by two feet makes a river two feet higher; no particle can travel faster than the speed of light; it would always be wrong to torture a child for fun; Tuesday 14th is a Tuesday; things sink in liquids that weigh less than they do; lying is always morally questionable. Nothing in these gives the conditional form serious work to do, whereas in dependent conditionals that form does work that we have no other convenient way of doing.

With one class of exceptions, which I shall note in §113, virtually every conditional that occurs outside logic texts, philosophy books, and classrooms depends for its truth or acceptability on beliefs about particular matters of fact. This fact-dependence is what makes them useful to us; it is also the source of the provocative problems that they raise, and is central to the most powerful solutions of them.

I do not hold that as between dependent and independent conditionals 'if' is ambiguous; that would be absurd. So I owe you some other account of how those two sorts relate to one another, and I shall pay that debt when the ground has been prepared—for indicative conditionals in §62, for subjunctives in §95.

8. IDIOMS

The conditional 'if' combines in regular ways with certain other words, and one may be tempted to treat such a phrase as an idiom, that is, an expression whose meaning cannot be read off from the meanings of its constituents. In every case the pull should be resisted.

One philosopher offered to explain the meaning of 'Even if . . .' in an account that makes that an idiom; but all recent workers on that topic have been sure—rightly, in my opinion—that the meaning of 'Even if' comes directly and regularly from the meanings of 'even' and 'if'. This will appear when we explore that topic in Chapter 17.

Similarly, 'Only if . . .' gets its meaning from 'only' and from 'if'. Very roughly: put an 'Only' in front of 'The boys were allowed to play soccer' and you get a statement meaning that no one who was not a boy was allowed to play soccer. Similarly, put an 'Only' in front of 'If the board asks me to resign, I will resign' and you get a statement meaning that no condition that does not involve the board's asking me to resign is one in which I shall resign. No one could doubt that the very same 'only' operates in both of these.

It is one thing to be sure that in the meaning of an 'Only if . . .' conditional the meaning of 'only' intersects with that of 'if'; it is quite another to understand the exact geography of the intersection. The latter is tricky, and I shall not discuss it. For three resolute shots at it, see Appiah 1993, Barker 1993, and Lycan 2001. I have pondered these materials, but not to the point where I feel entitled to report on them.

The case of 'If . . . then . . .' is peculiar. David Sanford wrote a book entitled *If P, Then Q*, presumably intending it to be about all conditionals, not just ones that put 'then' between antecedent and consequent. Many of us used to assume—if we thought about it at all—that the decision to include 'then' was stylistic, with no effect on meaning; but Wayne Davis (1983, listed in Sanford's bibliography but not mentioned in his text) has taught us otherwise. Any good conditional containing 'then' between its antecedent and its consequent remains good if 'then' is dropped; but the converse does not hold, because some good conditionals lacking 'then' turn bad when 'then' is inserted into them (§91). 'If war breaks out tomorrow, the tides will continue to rise and fall' is acceptable. You might prefer the wording '*Even* if war breaks out tomorrow, the tides will *still* continue to rise and fall', but the conditional passes muster just as it stands. On the other hand, 'If war breaks out tomorrow, then the tides will continue to rise and fall' implies that the tidal rise and fall depends somehow on the outbreak of war, which is absurd.

That gives significance to 'then' between antecedent and consequent, but leaves open the question of whether 'If . . . then . . .' is an idiom. Lycan (2001: 8) convincingly argues that it is not. He contends that 'If A, then, C' means something that could be expanded into 'If A, *in that eventuality* C'. This is comparable with 'When she leaves, then I will leave', which can be expanded to 'When she leaves, *at that time* I will leave'. In each sentence, 'then' does similar work.

2

The Material Conditional: Grice

9. THE HORSESHOE ANALYSIS: → IS ⊃

We understand perfectly the truth-functional or 'material' conditional operator that is standardly expressed by the horseshoe symbol. We define this so as to make P⊃Q equivalent to ¬(P&¬Q), that is, to

It is not the case that: P and it is not the case that Q.

This operator is *truth-functional*, meaning that the truth value of P⊃Q is determined solely by the values of P and Q; the horseshoe stands for a function to single truth values from pairs of them; feed in values for P and Q and out rolls the value of P⊃Q.

We know exactly what the material conditional is, or what operator ⊃ is; its truth-functional properties constitute its whole intrinsic nature. Its verbal role in plain English and everyday thought is another question, however. Some philosophers have held that it shows up in informal thought and speech in the indicative conditional, because → is ⊃. I shall call this 'the horseshoe analysis' of indicative conditionals.

According to the horseshoe analysis, 'If Booth didn't shoot Lincoln, someone else did' means the same as 'Either Booth shot Lincoln or someone else did'. This offers us a comfortingly secure hold on conditionals of that sort. We understand ⊃ as well as we do anything in our repertoire; if we found it at work in ordinary speech and thought, firmly linked to one major way of using 'If . . .', that would be a large step towards understanding our conceptual structures. So we have reason to want the horseshoe analysis to be right. There are also reasons to think it is.

The superficially most persuasive of them has occurred to many people and has been presented with helpful clarity by Jackson (1987: 4–6). It concerns something that I call 'the or-to-if inference', and it runs as follows.

You believed Vladimir when he told you 'Either they drew or it was a win for white'; which made it all right for you to tell Natalya 'If they didn't draw, it was a win for white'. That was all right because what Vladimir told you entailed what you told Natalya. Quite generally:

(1) PvQ entails ¬P→Q.

If 1 is correct, then so is the horseshoe analysis, as the following shows. In 1 substitute ¬A for P and C for Q, and you get:

(2) ¬AvC entails ¬¬A→C,

which is equivalent by definition to:

(3) A⊃C entails A→C.

Furthermore, → is at least as strong as ⊃, that is,

(4) A→C entails A⊃C.

The conjunction of 3 with 4 is equivalent to:

(5) A→C is logically equivalent to A⊃C,

which is the horseshoe analysis.

Most of this argument is elementary formal logic, and unquestionable. One might challenge the second premiss (line 4) by suggesting that A→C could be true while A⊃C is false; but this looks like a forlorn hope. With one exception (§61), nobody believes that A→C can be true when A is true and C false. All that remains is to challenge the first premiss (line 1). Your or-to-if inference, from what Vladimir told you to what you told Natalya, was clearly acceptable, and line 1 embodies a *theory* about that—a conjecture about why the inference was all right. In §18, when the time is ripe, I shall argue that the facts are better explained by something that does not imply the horseshoe analysis.

Famously, the latter is open to hosts of seeming counterexamples. They come from the fact that P⊃Q is true for *every* false P and for *every* true Q, so that the horseshoe analysis implies that the likes of these:

If I ate an egg for breakfast this morning, you ate a million eggs, and
If there are no planets anywhere, the solar system has at least eight planets,

are true, though each would be a silly thing to say. Most writers in this area have declared them false, contending that the meaning of A⊃C lacks, while the meaning of A→C includes, the notion of A's being suitably *connected* with C. But specifying the kind of connection has not been easy, and some friends of the horseshoe analysis have stood firm, arguing that the apparent counterexamples can be explained away, leaving their preferred analysis standing.

Let us look first at Grice's way of explaining them away, and then, in the next chapter, at Jackson's different one. Neither succeeds (I shall argue); but each leads through territory that is worth exploring on its own merits, apart from its relevance to our present concerns.

10. CONVERSATIONAL IMPLICATURE

H. P. Grice held that → is ⊃, so that A→C can be true even when A is not connected with C in any way. The two conditionals displayed above, he would say, are unsatisfactory but nonetheless true; and he sought to explain why they strike us as defective, through a powerful theory which everyone now sees to have much truth in it. (He began this work in his 1961, and developed it further in some 1967 lectures that were later published in his 1989. See his 1967*a* for the theory in general, and his 1967*b* for its application to indicative conditionals. Some of it seems to have been arrived at independently by James Thomson (1990: 67–8; written in about 1963).)

The theory concerns *conversational implicature*—a phrase in which the noun points to a certain linguistic phenomenon, the adjective to a way of explaining it. The phenomenon, implicature, occurs when a statement conveys, suggests, signals, or implies something without outright asserting it. I now confess: *In February 1952, in the faculty common room at Auckland University College, I disconcerted my colleagues by spilling hot tea into the lap of the newest assistant lecturer.* You picture me spilling tea on someone else, but you are wrong. Nervous and shaky in my first day in the job, I fumblingly dumped tea into my own lap. You thought otherwise because my report suggested (signalled, implied) that my victim was someone else, and Grice's theory explains how it did so—this being where the adjective 'conversational' comes into play. It is unusual—even a touch peculiar—to refer to oneself through a definite description rather than a pronoun; you assumed that I was not speaking in an off-beat fashion; so you took my phrase 'the newest assistant lecturer' to refer to someone other than myself. Still, I told you the truth. I planned to mislead you, and succeeded; but what I said was true.

So we do distinguish what a statement says from what it 'implicates' (as Grice put it), that is, what it more weakly implies or signals or conveys other than by outright assertion. Of the many sources of such implicatures, Grice focused on one cluster, namely some broad rules of conduct governing civilized discourse:

Be appropriately informative (give enough news but not too much).
Be truthful (say only what you believe, and try to have only true beliefs).

Be relevant.
Be orderly, brief, clear, etc.

These, Grice said, fall under the super-principle 'Be helpful'. They create implicatures because when someone asserts something, we can draw conclusions not only from what he outright asserts but also from other things that must be true if he is playing by the normal rules of civilized discourse. I now give four examples, three of them uncontroversial.

If I say to you 'IBM shares will go up' you will infer that I believe they will go up. Why? Not because 'IBM shares will go up' entails that the speaker thinks they will, for obviously it does not. Grice's ideas provide an explanation (though not the whole story). When I say 'IBM shares will go up', you are entitled to assume that I am playing by the rules, including the one enjoining us not to say what we do not believe; so you can reasonably infer that I believe IBM shares will go up, inferring this not from *the proposition I assert* but rather from *the fact that I assert it*. The proposition itself remains chaste, unsullied by any content about my beliefs.

If someone says 'He saw Lobatchewsky's proof of the theorem, and published his own', this conveys that he saw Lobatchewsky's proof first. How? Some philosophers used to attribute it to the sentence's meaning, contending that 'and' in this sentence means 'and then'. The sentence-joining 'and' sometimes conveys no thought of temporal sequence—'Nine is three squared and there are nine planets'—so these philosophers had to call 'and' ambiguous. It is usually bad in philosophy to postulate a multiplicity of senses of a word, and Grice offered an escape from this. He held that the sentence-joining 'and' is truth-functional: a sentence using it is true if each of the joined sentences is true, and otherwise false; so its meaning contains nothing temporal, and thus the Lobatchewsky sentence does not mean anything about temporal order. It *suggests* to us that the person saw the other proof before publishing his own because we assume that the speaker is presenting his narrative in an orderly manner, which usually involves making *its* order correspond to that of the reported events. Because this is a general rule of good conversational conduct, we are entitled to expect a speaker not to depart from it without signalling the departure ('Meanwhile, back at the ranch . . .', 'Before all this . . .'). In the absence of such a signal, we are inclined and entitled to infer that the narrative order matches the order of the narrated events, which explains the temporal suggestion of the Lobatchewsky sentence, and of 'They got married and they had a baby' and the like. (This kind of orderliness may be flouted for artistic purposes, as happens in unsignalled flashbacks in some of Vargas Llosa's novels.)

Some philosophers have thought that the sentence-joining 'or' is not v, because it is not strictly truth-functional in its meaning. If it is, then 'P or Q' is true just so long as P is true or Q is true, so that I am speaking truly when I say 'Either my father was F. O. Bennett or my father was Stafford Cripps'. Some would hold that this statement is not true, given that I know my father to be F. O. Bennett. It is part of the meaning of the sentence-joining 'or', they have said, that the speaker does not confidently believe, or confidently disbelieve, either disjunct.

This might explain why we wrinkle our noses at a disjunction that someone asserts just because she believes one disjunct. It has drawbacks, though, including a renewed threat of ambiguity. In playing games, giving tests, teasing, etc., it can be proper to assert a disjunction when you know which disjunct is true. For example, an acquaintance cannot remember when we first met, and I tease him with a hint: 'It was either at the Eastern Division meetings in 1956 or in Cambridge in 1963.' This is wholly proper; yet it involves asserting a (slightly compacted) disjunction when one knows which disjunct is true. So even if 'or' sometimes means that the speaker is not sure which disjunct is true, it plainly sometimes does not; so it must be ambiguous.

Grice explained the facts differently. The injunctions 'Be informative' and 'Be brief' tend to pull in opposite directions, and sometimes we have to compromise. But if someone asserts 'P or Q' when she is sure that P, she offends against both rules at once: she could be more informative *and* briefer; or, if she believes both disjuncts, she could say 'P and Q', thereby saying much more at no greater length. This entitles us to expect, normally, that someone who asserts a disjunction is not confident of either disjunct taken on its own; so we do in general expect this; so in asserting a disjunction one implies or signals that one is not confident of either disjunct. If the context provides a special reason to be less informative than one could be—e.g. because one is testing, teasing, playing, or the like—the implication of uncertainty drops out. So the sentence-joining 'or' has only one meaning, namely the truth-functional one, and we can explain all the intuitive evidence seeming to go against this.

Now, fourthly, we come to the thesis that → is ⊃. You can guess how a Gricean defence of that will go. We generally think it absurd to assert A→C purely on the grounds that one is sure that ¬A or sure that C; but this is consistent with (A→C)'s being true in such a case (and thus with →'s being ⊃), its unsatisfactoriness coming from a different source. Grice based this on the same points about brevity and informativeness that we saw at work in arguing that 'or' is v. That is to be expected, because according to the horseshoe analysis A→C really is a disjunction. On this account, it would be absurd but not untruthful to say 'If the speed of light is finite, then bananas are usually yellow'. This conditional is true because its consequent is true; but it would ordinarily be a silly thing to say because one could say something stronger yet shorter: 'Bananas are usually yellow.'

I shall need §12 to set the scene for explaining in §13 why the theory of conversational implicature fails in the task of reconciling the horseshoe analysis with our intuitions. Before all that, I shall discuss an aspect of Grice's theory that is too important to neglect.

11. SEMANTIC OCCAMISM

In all but the first of my four examples, Grice's *general* theory of conversational implicature offers to explain facts that would otherwise have to be explained through the meanings of *individual* words. So if we have grounds for attributing thin meanings to those words—taking 'or' to mean v, and 'and' to mean &—Gricean theory enables us to defend those semantic views against seeming counter-evidence, thus keeping the meanings thin. In some of his work, Grice contended that if we *can* keep the meaning of a word thin then we *should*. His theory of conversational implicature, he held, can do more than merely defend something arrived at on other grounds; it provides a positive reason for holding that the meanings of 'or' and 'and' are purely truth-functional, and thus thin.

He based this on a variant on Occam's Razor: *Sensus non sunt multiplicandi nec magnificandi praeter necessitatem*; don't postulate more senses, or thicker ones, than you have to. The two constraints are connected: if you put too much into the meaning of a word in some of its uses, you will have to plead ambiguity—'multiplying senses'—to cope with other uses from which some of this meaning is absent (Grice 1987: 47–9). We have already seen this illustrated. If 'and' sometimes means 'and then', then it must be ambiguous because sometimes it plainly does not mean that. Similarly with 'or' and uncertainty. The unspoken premiss here is that ambiguity claims in philosophy and semantics are a source of danger, and should be avoided as far as possible.

When Grice urged us to assign thin meanings to words, this was not only so as to keep down attributions of ambiguity. *Sensus non sunt magnificandi . . .*—he meant this injunction to stand on its own feet. Suppose we have two rival accounts—call them Lean and Fat—of the meaning of some word W. Fat attributes to W all the meaning that Lean does, plus some more. Lean explains certain facts about W's role in discourse partly through the meaning it attributes to W and partly through general principles of language use. Fat, on the other hand, explains all those facts through the meaning it attributes to W. In Grice's view, this difference counts in favour of Lean, because it makes less appeal to the highly specific—the idiosyncrasies of the individual word—and handles more of the data in terms of what is highly general; and that leads towards greater understanding and intellectual control. A full account of any language must, of course, include facts about individual words. How people might use or respond to 'There's a snake in that bush' depends in part on what 'snake' means, as distinct from 'steak' or 'rhinoceros'; you cannot get the whole story out of general principles. This need to come down to the level of specificity marked by individual words makes language study unlike physics, say. But the scientific spirit commands us to keep such specificity to a minimum; which encourages and perhaps justifies Grice's preference for thin single-word meanings, assigning much of the work to general rules governing civilized discourse.

('Why do you contrast the relatively *specific* facts about individual words with *more general* facts about language use, rather than contrasting *particular* facts about individual words with *general* facts about language use?' Because the words in question are universals, not particulars. The facts about what 'snake' means are facts about what many tokens of the word—instances of that universal—mean. So the needed contrast is not of particular with general, but of less with more general.)

Strawson (1986) has criticized Grice's treatment of indicative conditionals by attacking the Occamism it involves. The criticism, though wrong, is instructive. In its background is Strawson's conjecture that A→C means something like:

> There is a connection between A and C which ensures that: A⊃C.

Because of what comes before the colon, this gives to → a stronger meaning than Grice accorded it. Strawson calls his 'if' a first cousin to 'therefore'. At the end of the paper he offers two counter-arguments of which one—in the final paragraph—seems to relate wrongly to Grice's theory and to Strawson's other argument. The latter goes as follows. Whatever the actual truth about 'if', there *could* be a connective that means what Strawson thinks 'if' means, whereas Grice's line of thought implies that there *could not*. Grice should predict that if a language contained a Strawsonian conditional operator, all its extra strength would be drained off into the principles 'Be informative' and 'Be brief', leaving the operator itself with no need to carry (and therefore, according to semantic Occamism, *not* carrying) any more meaning than ⊃. Grice's position is guilty of overkill, Strawson concludes, and so must be wrong.

Strawson does not remark that if his point is sound, it counts not just against Grice's account of → but against his Occamism generally. It could equally well

have been brought against the Gricean cases for equating 'or' with v and 'and' with &. It generates things like this:

> There could be a connective 'uns' such that 'P uns Q' conventionally meant 'P and Q are not both false, but I am not sure about the truth value of either taken separately'. Grice's Occamism, however, implies that no connective could retain such a thick meaning. All the meaning of 'uns' beyond its truth-functional v component could be—and therefore according to Grice should be—explained in general conversational terms, and not assigned to 'uns' in particular.

This is disturbing, because Grice was plainly right about 'or', and yet 'uns' as described seems to be *possible*.

Fortunately, Strawson is wrong about Grice's commitments. Grice might have to say that we, with our actual practices and forms of life, could not have a Strawsonian 'if'; but he could comfortably allow that there could be societies that had it. Suppose a society where people often give disjunctive information—something meaning 'Either A is false or C is true'—although they knew which disjunct was true. They might do this, for example, in games, intelligence tests, initiation rites, or teasing. Given enough of this kind of disjoining, work could be done by a connective whose conventional meaning was that of 'It is not the case that A is true and C false, *and this is not one of those deliberate withholdings of information*'. That could be 'uns', or the Strawsonian 'if'.

Strawson may have meant his argument to have the premiss that we—in our actual language, with our actual ways of life—could have a conditional connective that is a first cousin to 'therefore'. But *that* premiss is far from self-evident, and Grice gave reasons for thinking it false.

Semantic Occamism, important and true as I think it is, turns out to have little direct bearing on the horseshoe analysis of indicative conditionals. The Occamism debate turns on this question:

> If the facts about the use of expression E—including ones about what uses of it would be found peculiar or unsatisfactory—can be explained either by (1) attributing to E a fat meaning or by (2) attributing a thin meaning and bringing in Gricean conversational principles, are we intellectually obliged to adopt 2?

This question gets a bite on the horseshoe analysis of → only if the facts about the latter *can* be explained by equating it with ⊃ and bringing in Gricean principles. They cannot. Grice thought otherwise because he had not considered enough of the data. A certain thesis about the propriety of indicative conditionals, now widely accepted, gives us a sounder idea than Grice had of what he needed to explain; and in the light of this we shall see that his theory of conversational implicature falls short.

12. THE RAMSEY TEST

The thesis in question was first presented by Frank Ramsey in 1929. Ernest Adams has greatly developed it in the past twenty years, and Frank Jackson—one of many who accept it—has actually called it 'Adams'. What happens when one considers whether to accept an indicative conditional? In a famous footnote Ramsey said this:

> If two people are arguing 'If A will C?' and are both in doubt as to A, they are adding A hypothetically to their stock of knowledge and arguing on that basis about C. . . . We can say they are fixing their degrees of belief in C given A. (Ramsey 1929: 143)

The core of what is going on here has been compactly stated by Gibbard, who attributes to Ramsey

> the thesis that, in whatever ways the acceptability, assertability, and the like of a proposition depend on its subjective probability, the acceptability, assertability, and the like of an indicative conditional A→C depend upon the corresponding subjective conditional probability . . . (Gibbard 1981*a*: 253)

. . . and then he uses a formula that I shall introduce later, meaning 'the amount of credence one gives to C on the supposition of A'. The phrase 'acceptability, assertability, and the like' is deliberately open and vague; we shall later pin it down.

So the core idea is that of conditional probability: the probability one assigns to C on the supposition of A. This is what Ramsey's phrase 'degrees of belief' points to, and we shall see that the concept of conditional probability has dominated most thinking about indicative conditionals since Grice.

Gibbard, like many others, calls the Ramseyan procedure for evaluating a conditional 'the Ramsey test'. Some of us have been encouraged in this usage by thinking that in Ramsey's procedure, as one writer has put it, 'we take our stock of beliefs, add the belief that A to the stock, and see whether our new stock of beliefs contains C'. This is a *test*, all right. Drop some of that liquid into this, stir, and see whether it turns blue; drop A into your belief system, stir, and see whether you turn C. However, it gives the wrong answer for some indicative conditionals, and it does not quite capture the spirit of Ramsey's remark.

This is shown by a certain class of examples that van Fraassen (1980: 503) says were first adduced by Richmond Thomason. I might accept 'If my business partner is cheating me, I will never realize that he is'; but when I pretend to accept 'My partner is cheating me' and whatever flows from that in my belief system, I find myself also pretending to accept 'I am aware that my partner is cheating me'. So the conditional fails the quoted version of the Ramsey test, yet the conditional may be perfectly all right. What has gone wrong, obviously, is that in this case my pretended belief that I am aware that he cheats has been produced as a by-product of my thought-experimental method, and not as a result of inferring 'I believe he cheats' from 'He cheats'. So it has no bearing on the evaluation of the conditional, and we should state the Ramsey test in a way that allows for this.

The statement of it that I have criticized has been common in the literature, and I have been guilty of it myself. There have also been better versions, but I have not found any that with perfect clarity steer around the difficulty created by the Thomason examples, though Harper (1981: 5) comes close. Here is my attempt:

> To evaluate A→C, I should (1) take the set of probabilities that constitutes my present belief system, and add to it a probability = 1 for A; (2) allow this addition to influence the rest of the system in the most natural, conservative manner; and then (3) see whether what results from this includes a high probability for C.

This does not involve pretending to believe A. Rather, it is a matter of seeing what results when A is added to a certain system of assignments of probabilities to propositions. The word 'test' is not entirely inappropriate even now, and I shall retain it for old time's sake.

The Ramsey test, as well as not being a matter of pretending to believe anything, is also not a matter of considering what you *would* believe if . . . It is easy to get this wrong, and to think that the Ramsey test amounts to this: to evaluate A→C, consider what probability you would accord to C if you became certain of A. We were first warned against this by Ramsey himself (1926: 82). If it were right, according a high probability to C on the supposition of A would be—roughly speaking—being such that if one came to accept A one would also come to accept C. It is not hard to see the flaw in this. As an atheist I accord a low probability to the proposition that God exists; and my probability for *God exists* on the supposition that I have terminal cancer is equally low, of course. Yet if I came to be sure I had terminal cancer, perhaps my weakness and fear would seduce me into religious belief. If so, then what I would believe if I came to believe A is irrelevant to my probability for A→C. I shall return to this in §49.

Most theorists of conditionals accept the Ramsey test thesis for indicatives. Two dissenting voices should be mentioned.

Peter Gärdenfors has argued against a version of the thesis: he conjoined it with some assumptions about belief revision generally, and derived a contradiction. Dorothy Edgington (1995*a*: 73–4) has challenged one of the assumptions, namely:

> (P) If a proposition B is accepted in a given state of belief K, and A is consistent with the beliefs in K, then B is still accepted in the minimal [viz. rational] change of K needed to accept A. (Gärdenfors 1986: 82)

Gärdenfors accepts this, Edgington says, only because he is thinking only of all-or-nothing acceptance and rejection, without giving play to degrees of acceptance or subjective probability. When the latter comes in, thesis P loses its plausibility. It can easily happen that you learn something A that is consistent with your present belief state, and that this discovery brings your confidence in B from ≈ 1 to ≈ 0. I shall not pursue this matter here, but for a certain light on it see my explanation of 'monotonic' in §32. For helpful discussion, and references to more of the Gärdenfors literature, see Hansson 1992, 1995.

Levi (1996: 8–15) contends that most of us have misunderstood what Ramsey meant. This is part of a larger concern of Levi's with different things that may be going on when one reasons from premisses that are supposed 'for the sake of argument'. His formidable work on this topic has defeated me. I hope my main conclusions in this book are not undercut by it.

The Ramsey test thesis does not hold for subjunctive conditionals. I think that if Yeltsin had been in control of Russia and of himself, Chechnya would have achieved independence peacefully; but for me this conditional does not pass the Ramsey test. When I take my present system of beliefs, add to it the proposition that Yeltsin was firmly in control of Russia and of himself at the time in question, and allow this to ramify through the rest of the system in the most natural and conservative manner, the result does *not* accord a high probability to 'Chechnya achieved independence peacefully'. The supposition about Yeltsin will make differences, but not that one. Rather, it will lead to changing my views about the unreliability of the media, the subtlety of the concept of control, and so on.

13. RAMSEY AND GRICE

The literature on indicative conditionals is a parade of attempts to explain why the Ramsey test is a valid criterion for their acceptability. This—the Ramsey test thesis—is not explained by the conjunction of the horseshoe analysis and Grice's theory of conversational implicature. Indeed, it conflicts with that conjunction, serving to refute it. Because the Ramsey test thesis is true, we cannot equate → with ⊃ and explain the apparent counterexamples through conversational implicature.

The failure of match between Grice and Ramsey is not total. Both condemn the asserting of A→C *just* on the grounds that one believes C, or disbelieves A, or both, without giving much credence to C on the supposition of A. My confidence that Polynesians didn't come from India convinces me that *Polynesians originally came from India ⊃ Most Maori speak Sanskrit*, and thus, according to the horseshoe analysis, that *If Polynesians originally came from India, then most Maori speak Sanskrit*. Grice would frown on my asserting this, because I could in fewer words say something stronger, namely: *Polynesians did not originally come from India*. The Ramsey test frowns also, because the supposition that Polynesians originally came from India, when added to the rest of my belief system with suitable adjustments made, does not generate a high probability for *Most Maori speak Sanskrit*. In this case, then, Grice and Ramsey pass the same judgement.

But they do not coincide across the board, and the discrepancies all suggest that Grice's theory of conversational implicature, though shiningly true, cannot defend the horseshoe analysis against apparent counterexamples. Perhaps the earliest solid attack on this project was that of Brian Ellis (1978: 114–19); it was followed up by Frank Jackson (1979: 112–19), Brian Skyrms (1980: 83–7), and Dorothy Edgington (1986: 181–3). I shall be guided by Jackson's attack, presenting five of its highlights in my own words.

(1) If I am sure that ¬A, Gricean principles automatically proceed to frown on my asserting A→C, because I could say more in fewer words. In fact, however, it can be acceptable to assert an indicative conditional whose antecedent one is pretty sure is false. I am virtually certain that the Polynesians didn't originally come from India; but it is all right for me to think and say 'If the Polynesians did come from India, there have been inhabitants of India whose language was not Indo-European'. What makes this all right is its passing the Ramsey test: in my belief system, the consequent is highly probable on the supposition of the antecedent. Score one for the Ramsey test over the Gricean approach.

(2) Grice is also committed to saying that A→C is not very assertible by someone who is sure that C is true, because it would be more informative and less wordy for him to assert C outright; but no such thing results from the Ramsey test. Suppose I am sure that C is true, and am also sure of C given A; that means that for me A→C passes the Ramsey test. In at least some examples of this divergence Ramsey is clearly right and Grice wrong. 'I'm sure he means well by me.' 'Even if it was he who persuaded them not to promote you?' 'Even then.' This reply here means 'Even if he persuaded them not to promote me, he means well

by me'; the respondent (implicitly) says this so as to indicate that the consequent is, for him, highly probable even on the supposition of the antecedent. Ramsey takes account of this, while Grice does not, so once again Grice's approach does not do justice to the data.

If Grice's regulative principles failed to condemn something bad, a Gricean might be able to amplify them, making the theory more condemnatory. But when, as in these two objections, Gricean theory condemns something innocent, there is no rescue.

(3) Jackson's third argument against the Gricean defence of the Ramsey test amounts to the point that every logical truth is entirely uninformative, telling us nothing about the actual world, and yet some logical truths are assertible while others are not. This point has no special relevance to conditionals, and draws little blood from Gricean theory, which does not offer 'Be informative' as the whole truth about how discourse should be conducted. In my opinion, *any* theory of speech and communication must handle the assertibility of logical truths with cautious delicacy—this challenges all of us.

(4) Jackson usefully compresses some of Grice's principles into this: 'Assert the stronger instead of the weaker (when probabilities are close).' He remarks that this does not distinguish among logically equivalent sentences so far as assertibility is concerned, since they are all equally weak; yet they can differ in assertibility. He instances a pair of statements related as ¬A & (A→C) is to ¬A & (A→B), which are logically equivalent according to the horseshoe analysis—because each is equivalent to the shared antecedent ¬A—but which differ in how assertible they are. Jackson cites the examples 'The sun will come up tomorrow, but if it doesn't it won't matter' and 'The sun will come up tomorrow, but if it doesn't that will be the end of the world'. This point of Jackson's seems to be sound.

Some of the best evidence for it has to do with contraposition, that is, the relation that holds between A→C and ¬C→¬A. According to the horseshoe analysis these are strictly equivalent, because A⊃C and ¬C⊃¬A are so. But it often happens that A→C is acceptable or assertible for someone for whom its contrapositive is not. I accept that *even if the Bible is divinely inspired, it is not literally true*; but I do not accept that if it is literally true, it is not divinely inspired (§59).

Grice's 'Assert the stronger' cannot explain the difference between the members of a contrapositive pair; and, although his account of conversational implicature contains other elements, none of them do—none *could*—explain how a conditional can be preferable to its contrapositive. This explanatory job, it seems clear, requires us to credit 'if' with more meaning of its own than ⊃ has; general principles could not do it.

(5) Jackson also has an argument from classroom experience. Gricean theory offers parallel explanations for two supposed facts: that 'if' as used in indicative conditionals seems not to be truth-functional though really it is, and that the sentence-joining 'or' seems . . . etc. Students easily accept the story about 'or', Jackson remarks, whereas most of them strenuously resist the story about 'if'. This does indeed suggest that Gricean theory is less than the whole story, and that a convincing answer to 'Why does → seem not to be truth-functional?' cannot be expected from general principles, and must owe something to the meaning of → in particular.

Jackson's attack on Grice's attempt to defend the horseshoe analysis wholly succeeds, I think. Among those it converted was David Lewis. He offered a partial explanation of the validity of the Ramsey test through Gricean principles, but later said that for the test's basis he no longer looked to Grice but rather to a theory of Jackson's to which we now turn. (Lewis 1976: 142–3; 1986*e*: 152–4. Like Appiah (1985: 178–9), I have had to struggle to grasp how Lewis meant his explanation to work.)

3
The Material Conditional: Jackson

14. SETTING THE SCENE

Jackson also accepts the horseshoe analysis according to which → is ⊃. Or, more carefully now, he holds that the truth conditions of A→C are those of A⊃C, and what is outright asserted by someone who says one is the same as what is asserted by someone who says the other. Grice said this too, but Jackson handles the apparent counterexamples differently, giving primacy to the Ramsey test thesis (1987: 22–32).

In presenting these materials, I shall follow Jackson's use of the technical term *robust*. This word stands for a concept that is present in the Ramsey test: to say that for me Q is fairly (very) robust with respect to P is to say that I accord Q a fairly (very) high probability on the supposition that P is true. In these terms, the Ramsey test thesis says that A→C is acceptable or assertible by me to the extent that for me C is robust with respect to A.

When Jackson first launched this concept he had in mind cases where someone assigns a high probability to C as well as to C-given-A; this person's fairly confident belief in C is 'robust' because the confidence can survive his also coming to believe A (Jackson 1979: 115). In later work Jackson expressed respect for 'a more general account' in which the robustness of C with respect to A requires only a high probability for C-given-A (Jackson 1987: 22). That is what we need; the restricted version does not belong in any general account of indicative conditionals. Although 'robust' does not carry well the meaning the more general account gives it, I shall continue to use it; but I shall always be working with the more general account.

Jackson helps us to get the Ramsey test in focus by exhibiting the role of robustness in our linguistic lives generally, and not only in relation to conditionals. For example, he remarks that usually a disjunction is assertible only if it is robust with respect to the falsity of each disjunct, because disjunctions are

commonly used in inferences of the form 'P or Q; not P; so Q', which are useless if one accepts the first premiss only because one rejects the second. Through an elegant example on p. 23 he shows how his approach—emphasizing the need for robustness so that certain inferences will go through—can explain facts which Gricean principles about brevity and informativeness cannot.

We mainly want indicative conditionals, Jackson says, for use in Modus Ponens—that is, arguments of the form:

A→C, A ∴ C

—but a given instance of A→C fails for this purpose if one accepts it mainly because one rejects A or accepts C. Think of A→C as serving in Modus Ponens like a ticket for a particular rail journey: while you reject A, you are not at the place where the journey starts; while you accept C, you are already where it ends. Either way, the ticket gets you nowhere. So it suits our purposes to frown on indicative conditionals, even true ones, if their consequents are not robust with respect to their antecedents.

Robustness is needed for an indicative conditional to be acceptable, but, Jackson points out, it is not in itself sufficient (pp. 15–16). Some assertings of conditionals that pass the Ramsey test are nevertheless unsatisfactory for Gricean reasons; so the Ramsey test thesis does not make the Gricean approach irrelevant to indicative conditionals—it merely blocks it from reconciling the horseshoe analysis with all the data. Up to here, I entirely agree with Jackson.

15. CONVENTIONAL IMPLICATURE

Why does the Ramsey test hold good for indicative conditionals? Having shown that this cannot be answered through purely general principles of discourse, Jackson concludes that the test's validity must come from the meaning of 'if' as used in those conditionals. The semantic truth about the 'if' of indicatives, he holds, is not exhausted by the thesis that → is ⊃; there is more to its meaning than this. That will be warmly endorsed by those who reject the horseshoe analysis, but Jackson *accepts* that analysis: according to him, someone who asserts A→C *asserts* only A⊃C, so that if the latter is true he has spoken truly. But, he adds, the speaker also conveys to his hearers something further that he does not assert but merely implicates—suggests or signals or implies. Grice said that much; but the two disagree about the source of this further implicature or suggestion. Grice traces it to the hearers' expecting the speaker to abide by certain general rules; Jackson traces it to the conventional meaning of the indicative 'if' in particular. It is, he says, borrowing a term from Grice, a matter of '*conventional* implicature'.

This phrase names a real phenomenon. For an uncontroversial instance of it, compare these two sentences:

(1) Noam Chomsky would be a good Commencement speaker, and he is the country's most famous radical left-winger.

(2) Noam Chomsky would be a good Commencement speaker, but he is the country's most famous radical left-winger.

Each conjunction is true so long as both its conjuncts are true. They differ, however, because 2 suggests, as 1 does not, that the two conjuncts stand in some kind of contrast. (One might have to work out what contrast it is. Many of us followed Frege in thinking that '. . . but . . .' always suggests that the first conjunct makes the second surprising, but Dummett (1973: 86) has shown this to be wrong. Someone who utters 2 may mean to contrast a good feature of a Chomsky visit with a bad one; or—thinking that Chomsky's fame would prevent his acceptance—to contrast the visit's being desirable with its being unachievable.)

Other words in our language also serve to suggest things without their being outright asserted. One such is the word 'even', according to the majority view about it. If someone says 'Bertrand Russell was an even more boldly athletic thinker than G. E. Moore', this is defective because it falsely suggests that Moore was a boldly athletic thinker (Russell was 'even more' so), but what it actually says is true, for Russell was a more boldly athletic thinker than Moore. (Not all students of 'even' take that view of it: for reasoned dissent, see Lycan 1991 and 2001. I shall return to 'even' in Chapter 17.)

So we can distinguish what is said from what is more weakly implied; and if a speaker implies something false, we characterize his statement not as false but as infelicitous, potentially misleading, or the like. Grice has called our attention to things that an assertion may weakly imply because of general principles of discourse—these are *conversational* implicatures. Now we encounter weak implications arising from special facts about the conventional meanings of individual words such as 'but', 'even', 'although' and so on—these are *conventional* implicatures.

How do we decide that the two Chomsky sentences have the same assertive force? If in 2 you do not see any significant contrast between the conjuncts, you will find 2 inappropriate or misleading, but you should not call it false unless you reject one of the conjuncts. Or so I say, but on what evidence? What *shows* that the contrastive element in the meaning of 'but' is a matter of implication rather than outright assertion? When a Gricean theorist declares a given utterance to be false in what it *conversationally* implicates but not in what it outright says, he sometimes has a firm, structural answer to the question: 'On what basis do you decide that the communicated falsehood is only implied and not said?' He can reply:

> If you assert something false, the falsehood comes from a relation between reality and the sentence you have uttered; it depends strictly on what *that one sentence* means. But when I say that the utterance of an indicative conditional involves falsehood by way of conversational implicature rather than assertion, that is because there falsehood enters the picture only through certain *general principles* to which the speaker is taken to be subject.

This is clear-cut and objective. Now return to my question about 'but'. The obvious basis for maintaining that 'She was poor but honest' is defective only in what it implies, not in what it says, is the fact that the most natural dissent would not be 'That is not true' but rather 'You may be right that she was poor and honest, but I wouldn't put it the way you did'. Similarly with sentences that do not have truth values—questions, for example. I might ask 'Did John make a donation?' or 'Even John made a donation, did he?'; each asks the same question though the latter insinuates something that the former does not.

These intuitive responses, which Jackson acknowledges to be the only evidence for judgements about conventional implicature (1981: 133; 1987: 40–1), are a fragile basis for theoretically distinguishing assertion from conventional implicature (§106). That the distinction *is* a matter of theory is shown by the fact that one theorist has questioned it (Appiah 1985: 190–1). But Jackson puts it beyond serious question, in my opinion, in the convincing fifth chapter of his 1987 book, where he explains why we should have conventional implicature at all—what role it plays in our linguistic lives. This chapter establishes that conventional implicature does exist and that the standard account of it describes it correctly.

In discourse, says Jackson, we primarily aim to affect one another's beliefs; the asserted content of what someone utters fixes the belief she wants to communicate, and conventional implications can help her to achieve this by removing obstacles:

> If I say 'Hugo is bad at mathematics; nevertheless, he is a fine chess player', what I want you to believe is that Hugo is a fine chess player, not something else. The role of 'nevertheless' is to guard against your refusing to accept my word because you think I am ignorant of the general, but not universal, connection between ability at mathematics and at chess, or even perhaps your thinking that I am in the process of revamping my system of beliefs. (Jackson 1987: 94)

What belief I want to communicate determines what I outright say or assert, according to Jackson; and what my statement conventionally implies or signals helps me to get this belief across smoothly and without needless fuss. This is in the spirit of Locke, who wrote that two of the three 'ends of language in our discourse with others' are: 'First, to make known one man's thoughts or ideas to another. Secondly, to do it with as much ease and quickness as is possible' (*Essay Concerning Human Understanding* III.x.23).

These ideas of Jackson's help to round out and solidify our notion of conventional implicature. Let us now see whether it helps Jackson in his use of that concept to explain why the Ramsey test holds for indicative conditionals.

(Jackson says that 'but' is governed by a special rule of *assertibility*, against which Woods (1997: 61 n) made the point that there is more to it than that. The special flavour of 'but' is at work in 'If she was born in Turin, but left when she was three, she doesn't know Italy well', though the clause containing it is not asserted. Apparently unlike Woods and Edgington (1997*a*: 103–4), I do not see this as greatly harming Jackson's basic position, but a related point by Read (1992: 11–12) has power. Kent Bach (1999) emphasizes the behaviour of 'but' and its kin in indirect quotation ('He said that she was poor but honest') in an attack on the entire category of conventional implicature. The attack, though considerable, does not convince me that my dissent from Jackson should start earlier than it does.)

16. THE CASE AGAINST JACKSON'S THEORY

When someone asserts A→C, Jackson maintains, he *says* only that A⊃C but he *implies* that for him C is robust with respect to A. This is a conventional implicature, he contends, belonging to the class of phenomena he describes so well in his fifth chapter. Speaking of the parallel between how A→C relates to A⊃C and how 'but' relates to 'and', Jackson says that 'the parallel . . . is intended to be exact' (1987: 9, 37).

He needs it to be exact. We have asked what connects indicative conditionals with the Ramsey test. What fact *about* them makes it the case that A→C is satisfactory only for someone for whom the probability of C given A is high? Jackson answers that the two are linked by the meanings of the conditionals. The Ramsey test need not be deduced from general principles: it sits there, in a lump, in the meanings of the conditionals for whose assertibility it is a valid test. Jackson's explanation of *how* it sits there brings in a concept of conventional implicature for which he offers a general theory. Without the latter, his account of how indicative conditionals come to have the Ramseyan property would be

empty, or at least ad hoc and unexplanatory. So he really needs there to be an 'exact' or at least a close parallel between his treatments of 'if' and of the likes of 'but' and 'even'.

Jackson does acknowledge one awkwardness of fit, though he does not describe it as such (pp. 38–9). With each of his other examples of conventional implicature, he acknowledges, what a speaker conventionally implies may be true even if what she asserts is improbable: 'Even my cousin could easily beat Jimmy Connors at tennis'; 'Wittgenstein was not a deep thinker; however, he had a strong influence on thinkers who knew him well'. These are unassertible by me, because I regard each as false; but they satisfy the special 'conventional implicature' requirements that they involve: my cousin is a duffer at tennis, so that 'even my cousin' is apt; those who influence other thinkers tend to be deep thinkers themselves, so 'however . . .' is apt. In contrast with this, Jackson acknowledges, the Ramsey test gives 'the whole story' about the assertibility of A→C: such a conditional cannot pass the test but be unassertible on grounds of improbability or falsehood. Jackson calls this 'an unusual property' of the Ramsey test, an 'exception' to what 'usually' happens with conventional implicature.

I shall now describe four other ways in which Jackson's account of conventional implicature as the source of the Ramsey test's validity fails to fit his account of conventional implicature generally. The resulting quintet—his one and my four—undermines his theory of indicative conditionals.

(1) Jackson explains conventional implicature as helping a speaker to get truths to glide smoothly into people's souls, but the supposed Ramseyan conventional implicature of A→C cannot be doing that. If it were, the following would be right:

> When I tell you 'If (A) Nixon's top aides were not blackmailing him into defending them, then (C) he gave them more loyalty than they deserved', I signal to you that my probability for C on the supposition of A is high. I signal this to you, choosing words that conventionally imply it, because this may help me to get across the belief I am primarily trying to communicate, namely that *either Nixon's top aides were blackmailing him into defending them or he gave them more loyalty than they deserved.* The conventionally implied robustness will improve your chances of acquiring precisely this belief rather than being distracted by irrelevant side-issues.

What irrelevant side-issues? What are the threatening distractions, and how does the implication of robustness remove them? I can find no answer to this. The above story is wholly incredible; yet it needs to be true if A→C is to relate to

A⊃C as 'but' does to 'and'. Jackson illustrates his thesis about the transfer of belief with examples using 'even' and 'nevertheless' but not with any using 'if'.

David Lewis, aiming to present Jackson's ideas on this topic, offered a more complex and believable story about the point of implying that for the speaker C is robust with respect to A. It does not, however, concern the removal of obstacles to getting the hearer to believe A⊃C. In his account, the speaker disbelieves A, believes C, and asserts A→C because he wants the hearer to accept C even if he, unlike the speaker, believes A:

> Maybe you (or I in future) know something that now seems to me improbable. I would like to say . . . something that will not need to be given up, that will remain useful, even if a certain hypothesis that I now take to be improbable should turn out to be the case. If I say something that I would continue to believe even if I should learn that the improbable hypothesis is true, then that will be something that I think you can take my word for even if you already believe the hypothesis. (Lewis 1986e: 153)

In confining himself to cases where the person who asserts A→C believes C and disbelieves A, Lewis doubly narrowed the range. Also, his account does not concern getting the hearer to acquire the belief one supposedly wants him to acquire, namely that A⊃C. Lewis's account bears little resemblance to Jackson's comparison of ⊃/→ with 'and'/'but', being concerned rather with the durability of the belief in various vicissitudes. It is a good story; but it leads away from Jackson and towards the theory of Adams that I shall come to in Chapters 6–7.

Something like this also occurs in Jackson 1981: 135. Having recounted how 'but', 'even', and the rest help the hearer to absorb the speaker's message, Jackson moves on to a Lewis-like story about the point of the implication of robustness of an indicative conditional. He does not note how greatly it differs from what has gone before, and merely introduces it as 'a second example of the problems attendant on an apparently simple speaker-hearer exchange'.

(2) The misfit shows up also in a formal way. In his fifth chapter Jackson mentions these vehicles of conventional implicatures: 'but', 'nevertheless', 'yet', 'anyhow', 'however'. In the relevant senses of these, the following holds for each:

> When W links two sentences, it can be replaced by 'and' without affecting the truth conditions of what is asserted; when used as an operator on a single sentence, it can be deleted without affecting the truth conditions of what is asserted.

Because a sentence-joining 'and' can always be deleted in favour of a full-stop, without affecting the truth conditions, it follows that each of those five words can be dropped without altering the asserted content. Jackson's only other example in his fifth chapter is 'even'. When used in the relevant sense, that too can be deleted, as can his two earlier examples (p. 49), namely 'anyway' and 'at least' as used in 'He is a communist, or at least left-wing'. All of Jackson's examples are deletable: each can be simply omitted (perhaps with a little repunctuation) without affecting what is asserted. Nothing like that holds for 'if'. In indicative conditionals, as everywhere else, 'if' is structural: delete it at your peril! 'If' is not alone in this, Jackson has written to me, because it holds also for 'unless'. I am not persuaded by this defence, which relies on Jackson's associating 'unless' with conventional implicature in the same kind of way that he associates 'if' with it. 'P unless Q' conventionally means the same as 'not-Q ⊃ P', he holds, and conventionally implicates that for the speaker P is robust with respect to not-Q. (Thus, I am not to say 'I'll be miserable unless she kisses me' purely because I am sure she will kiss me.) I reply that the awkwardness of fit between Jackson's account of 'if' and his chapter 5 account of conventional implicature applies equally to 'unless' on this view of the latter.

(3) In his general account of conventional implicature, but before offering to apply it to indicative conditionals, Jackson asks '. . . why tone, why conventional implicature?' (p. 91), and speaks of '. . . the words that are responsible for conventional implicatures, that carry tone . . .' (p. 93). Dummett brought the word 'tone' into this, replacing words of Frege's that mean 'colouring' and 'illumination' (1973: 2, 83–8). It fits some of his examples—'dead' and 'deceased', 'sweat' and 'perspiration'—and countless others, such as 'defecate' and 'shit', 'intellectually challenged' and 'mentally retarded', and so on. These do perhaps involve a difference in what is implied or suggested, but that is not the heart of them; and Jackson was right to ignore them in his account of conventional implicature. As for 'but', 'although', 'even', and the others that he does mention, 'tone' is a less apt label for what they add to the asserted content, but it is still a possible one. If someone said 'Even teaching assistants don't get paid a million dollars a year', it would not be absurd to remark that he had said something true with a wrong tone (because of its implication that TAs are notably well paid). In contrast with this, when a materially true conditional fails the Ramsey test, as when someone says 'If snow did not fall on Mount Rainier last year, the US national debt was halved', the diagnosis 'True assertion, wrong tone' misses the mark.

(4) When a true assertion conventionally implies something false, how should we characterize this? In the context of his theory of indicative conditionals, Jackson repeatedly implies that in such a case the assertion is not 'justified or warranted', explaining that he means epistemic rather than pragmatic justification (pp. 8–10). An assertion may be pragmatically unjustified because pointlessly hurtful or in breach of a promise, or the like, to all of which the Ramsey test is irrelevant. The test does involve something like epistemic justification, as

Jackson says, but what does the latter have to do with 'but', 'even', 'although', and the rest, and with Jackson's general theory of conventional implicature? Nothing, so far as I can see, though Jackson evidently thinks otherwise. In claiming to be employing a single concept of implicature or signalling, he implies that 'He wasn't educated at Eton but he is a civilized human being' is defective because not epistemically justified. This strikes me as untenable. When Jackson uses the phrase 'epistemic and semantic considerations, *widely construed*' (p. 19, my emphasis), he may be countering this difficulty by backing off from linking 'but' and 'even' etc. with epistemic justification ordinarily construed. But he does not tell us *how* he means to widen the construal. This point connects with an objection that Edgington (1986: 186) brings against both Jackson's theory and Grice's, namely that they purport to explain the *assertibility* of something when their topic ought to be its *acceptability* or *believability*. She writes elsewhere: 'There simply is no evidence that one *believes* a conditional whenever one believes the corresponding material implication, and then is prepared to *assert* it only if some further condition is satisfied' (see also 1995*b*: 287 n. 50). Actually, Jackson knew that 'assertibility' is wrong quite early in the piece: 'It is, indeed, better labelled "assentability"—but it is too late to change now' (Jackson 1984: 72).

In five ways, then, Jackson's general account of conventional implicature misfits his application of that concept to indicative conditionals. The last two failures may not matter much; but the first three—Jackson's own and my 1 and 2—are structural, serious, and in my view fatal. A sixth will be presented in §39.

17. THE UNITY POINT

At one place Jackson hedges his claims for his account of conventional implicature. Having asked 'Why tone, why conventional implicature?', he writes: 'Perhaps it is wrong to expect the answer to be the same for each example, but, in many cases at least, it seems that the reason . . .'—which launches him into his general theory of conventional implicature (p. 91). This is cautious; it creates wiggle room. Perhaps Jackson means to allow that indicative conditionals may not be among the 'many cases', and may thus not fall under his general theory. If so, one wonders why a book entitled *Conditionals* should devote a chapter to an account of conventional implicature that applies to some parts of language but not to conditionals.

A little later, Jackson seems poised to confront the problem. Right after completing his (general?) account of conventional implicature, he writes: 'We now have answers to why conventional implicature exists in natural languages . . . and to why it affects meaning without affecting content . . . But why did we need to turn to conventional implicature . . . in our . . . theory?' (1987: 95–6). Put like that, it is a good question, but Jackson does not put it just like that. More fully, he asks: 'Why did we need to turn to conventional implicature, *rather than conversational*, in our . . . theory?' (my emphasis). Instead of considering for the first time how conventional implicature succeeds in explaining how indicatives work, he considers for the second time why conversational implicature fails in this. The earlier point that the Gricean approach must accept Contraposition for indicative conditionals, that is, must regard $\neg C \rightarrow \neg A$ as being no less assertible than $A \rightarrow C$ (§10, argument 4), now becomes the point that the Gricean approach must endorse not only Modus Ponens:

$$A \rightarrow C, A \therefore C$$

but also Modus Tollens:

$$A \rightarrow C, \neg C \therefore \neg A.$$

It is a sharp point against Grice; perhaps it deserves to be presented twice in these two guises. But it serves here merely to displace the question Jackson should be asking: '*Did* we really "turn to conventional implicature" as this has just been described?' I answer, No, we did not.

Why did Jackson come at things in this way? Given the story as I have told it, one might think:

> Jackson rightly says that the Ramsey test is valid because of the meaning of 'if' as used in indicative conditionals. He ran into trouble because of *where* in the meaning of 'if' he located the Ramseyan element. If he had put it into the core of asserted content, rather than the conventionally implied penumbra, he would have escaped the troubles exhibited here.

I agree with this, but the matter is tricky. As we shall see in chapter 6, the two most obvious ways of building the Ramseyan property of indicative conditionals into their conventional meaning are demonstrably wrong. The right way to do it is somewhat elusive, and still a matter of controversy. So Jackson had reason to want conventional implicature to provide him with a solution.

18. THE OR-TO-IF INFERENCE

The horseshoe analysis of → should be rejected, because of the failure of the only two attempts (it seems) that can be made to reconcile it with the intuitive data. Before finally turning away from it, we should revisit the or-to-if inference (§9) in order to see how feebly it supports the analysis. The attempt to get support from it went like this:

> You believed Vladimir when he told you 'Either they drew or it was a win for white'; which made it all right for you to tell Natalya 'If they didn't draw, it was a win for white'. Why was this all right? The explanation is that what Vladimir told you entailed what you told Natalya, because quite generally P∨Q entails ¬P→Q.

If this is right, then A⊃C entails A→C, which secures the horseshoe analysis.

Considered as support for the horseshoe analysis, this fails twice: the analysis is not needed, and does not suffice, to explain why it was all right for you to say what you did to Natalya.

Here is why it is not needed. Vladimir was behaving badly unless he was more confident of the disjunction than of either disjunct; and Grice's theory about conversational implicature explains why (§9). If he was not misbehaving, therefore, he accepted the disjunction independently of whether one disjunct (either one) turned out to be false; so for him Q is robust with respect to ¬P. That would make it all right by the Ramsey test for him to assert ¬P→Q; and your trust in him makes it all right for you to assert this also. This explanation has nothing to do with the horseshoe analysis. It is given by Stalnaker (1975), who calls the transaction a 'reasonable inference' of one assertion from another, not the entailment of one proposition by another. In §58 we shall see that the or-to-if inference can also be explained in another way.

Anyway, whether or not some rival to the horseshoe analysis *does* explain the acceptability of the or-to-if inference, the analysis itself *does not*. It contributes only the thesis that P∨Q entails ¬P→Q, and thus that the *truth* of what Vladimir told you guarantees the *truth* of what you told Natalya. But it is a famous fact that a true material conditional may be an absurd thing to say; so this entailment thesis does not imply or explain the fact that if you *accepted* what Vladimir told you, then it was *all right for you to say* what you did to Natalya.

The or-to-if inference haunts the literature on indicative conditionals. In §41 we shall see a valid special case of it being used in a powerful argument for a significant conclusion.

4
The Equation

19. OTHER APPROACHES

The horseshoe analysis having failed, we must look further. Four other avenues of approach to indicative conditionals have been proposed.

One is indicated by the view, attributed to Strawson in §11, that A→C means something like:

> Because of a connection between A and C: A⊃C.

This arises from the natural thought that the horseshoe analysis fails because it does not provide for a link between A and C. It certainly *is* defective; A→C may entail A⊃C but is not entailed by it. However, although it is plausible to suppose that the missing ingredient is the idea of A's being connected with C, there are obstacles in the way of developing this into something solid. For one thing, many respectable indicative conditionals involve no such link. Not just jokes like 'If he repays that debt, I'm a monkey's uncle', but sober conditionals like 'If she apologized to him, then he lied to me', which would not ordinarily be based on a view about a direct link between her apology and his lie (§133); and ones like '(Even) if he apologizes, I shall (still) be angry', which rests on the *lack* of connection between his apology and my anger. Unsurprisingly, nobody has worked hard on trying to turn this 'connection' idea into a semantic analysis of indicative conditionals.

A second avenue of approach is lined with possible worlds. These have enjoyed much success in analyses of subjunctive conditionals (Chapters 10–13); so it is natural to hope that they can also cope with indicatives, through the idea that A→C means that C obtains at a certain possible world at which A obtains. This hope is encouraged not only by a desire for theoretical economy, but also by the hope for a Y-shaped analysis of conditionals—one that first sets out what the two kinds of conditional have in common and then goes on to say what differentiates them (§4). Wayne Davis and Robert Stalnaker have both approached indicative conditionals in this second way; this might be a natural place to discuss those endeavours, but for reasons that will appear I choose to postpone them until §§138–9.

A third avenue has been explored by William Lycan (2001). I shall say what I have to say about it in §37 and §84.

The fourth avenue leads, I believe, to the truth about indicative conditionals. It starts from the Ramsey-test thesis that the assertibility or acceptability of A→C for a person at a time is governed by the probability the person then assigns to C on the supposition of A. Jackson used this to marshal facts about usage that are not explained by the horseshoe analysis aided by Gricean conversational implicature. It served that negative purpose well, but with the right soil and climate it can grow into a positive theory.

First, we should focus on acceptability rather than assertibility. They are linked, because you ought not to assert what you do not accept; but still they are different, and acceptability is the concept we need. If I do not think that C is probable given A, I ought not even to *think* A→C let alone to assert it. So we should attend to the Ramsey test thesis considered as saying that A→C is *acceptable* by you in proportion as your probability for C on the supposition of A is high.

Perhaps acceptability is just subjective probability: for something to be acceptable by you is for you to find it probable. If not, then at least we can agree that acceptability depends upon probability and nothing else: you ought not to accept what you do not find probable, and there is no obstacle to your accepting what you do find probable. So we can move to the thesis that the *probability* for you of A→C is proportional to your probability for C on the supposition of A.

20. KINDS OF PROBABILITY

Before moving further into conditionals, I should provide a setting for the Ramsey test by distinguishing different kinds of probability. This will be routine stuff, probably bringing you no news; but I choose to play safe by laying it out explicitly.

Objective probabilities hold independently of what anybody thinks about them. They divide into two kinds.

Absolute: This kind attaches to a single proposition, considered in itself; the proposition's absolute probability is often called its 'chance' of being true. If determinism holds, then the objective absolute probability of any proposition at any time is either 0 or 1. (Or so it is often and plausibly said, but Loewer (2001) disagrees. He holds that a deterministic theory of statistical mechanics makes room for intermediate probabilities that are not merely subjective. His arguments take me out of my depth.) If on the other hand some basic laws of physics

are probabilistic, then there are objective chances between those extremes; it may be an objective, absolute, fundamental fact about the proposition that *this atom will decay during the next seven years* (referring to a particular atom) that it has a 50 per cent chance of being true. If such chances also show up on the macroscopic scale, there may be such a thing as the objective, absolute chance that exactly two branches will fall from my trees during the next million seconds, or that more water will flow through the Grand Canyon in 2009 than did in 1987. A statement about *objective* absolute probability has the form $R_3(P,t,n)$, with triadic R_3 relating proposition P to a time t and to a number n (or other measure of probability); it says what probability of truth P has at that time. If P concerns a time earlier than T, then the probability of P at T is either 0 or 1; intermediate objective probabilities all concern the future.

Relative: Even if determinism is true and there are no non-extreme objective chances, there are objective facts about what probability a proposition has relative to certain others—for example, about what probability a body of evidence confers on an hypothesis. These are reportable in statements of the form: 'Given Q, the probability of P's being true is . . .'. It is objective, but whereas chance is absolute, this is relative; it accords P a probability relative to Q. Given that he will fairly deal seven cards from a normal pack, the probability that he will deal at least one ace is 0.538. A statement about *relative* objective probability has the form $R_3(P,Q,n)$, relating proposition P to a second proposition Q and to a measure of probability. The mention of time ('t') drops out because true propositions of this kind are logically or causally necessary, and so do not change truth values over time.

Often enough we will speak of Q's being made probable by . . . and then we express only a part of the total P relative to which Q is probable. 'Given that he will deal seven cards from a normal pack, the probability that he will deal at least one ace is 0.538', with nothing said about the fairness of the deal; or 'Given that he will fairly deal seven cards from this pack, the probability that he will deal at least one ace is 0.538', with nothing said about the pack's being normal. Taking such statements just as they stand, we judge them to be neither causally nor logically necessary; but nor, taken just as they stand, are they true. They are stand-ins for something that is true (and, indeed, necessary), the remainder of their antecedents being supplied by the charitable hearer or reader.

Subjective probabilities are people's degrees of belief. A subjective *absolute* probability of a proposition P is some person's degree of belief in P's truth at a particular time. A statement about P's subjective absolute probability has the form $R_4(P,x,t,n)$, which relates P to a person x, a time t, and a number n. Even if there are no objective chances between 0 and 1, their subjective counterparts will exist so long as people are uncertain. A subjective *relative* probability is the probability or belief level that someone accords to P on the supposition that some other proposition Q is true. The basic form of a statement about P's subjective relative probability has the form R5(P,Q,x,t,n), which relates P to a second proposition Q and to a person x, a time t, and a number n. Here we retain the mention of time ('t') because the subjective relation between one proposition and another depends upon the person's intellectual frame of mind, which can change.

Isn't this madly unrealistic? How often are we in a position to express our levels of credence in numbers? 'Are you sure they will make the film?' 'No, one cannot be *sure* of such a thing; but I give it a probability of 0.827.' Absurd! Even such approximate values as 'about 0.3' and 'somewhere near 0.6' have little place in our everyday subjective probabilities. We can realistically assign 0 (or 1) to something we know for sure to be false (or to be true); but apart from those two termini, and apart from the contrived set-ups of betting on cards or the like, we have little use for the numerical apparatus. Yet the formal, numerical idea of probability informs many of our thought processes. The formal logic of probability (§21) needs to be able to handle probabilities arithmetically, working on the assumption that they all have numerical values; and this logic is normative for us all—it implies things that *do* apply directly to our actual ways of thinking. Without being able to assign precise numbers to our levels of credence, we have thoughts of the forms:

I have more faith in P's truth than in Q's.
P now strikes me as more likely than it did a week ago.
I am almost entirely sure that P is true (≈ 1).
The more likely P is, the less likely Q is.

These and their like are familiar elements in the life of the mind; and in a rational mind they will be governed by the principles of probability logic.

Of these everyday thoughts involving probability, the kind that will chiefly concern us is (subjective) relative probability. I imagine you don't think either of these is likely:

Public: Not later than the year 2084, public executions will be common in the United Kingdom.

Fascism: Some time in the twenty-first century, an outright fascist government will come into power in the UK.

But I am sure that your (subjective) probability for *Public given Fascism*, or *Public on the assumption of Fascism*, is higher than your (subjective, absolute) probability for Public.

Your probability for Public is absolute—not a probabilifying relation between some other proposition and Public. You do base it on other propositions that you accept, and if you are thinking impeccably it is equivalent to your relative probability for *Public given everything else you believe*. Still, it is not itself a relative probability. Your thought in assigning it has the form 'It is not likely that . . .' and not the form 'The evidence that I have does not support . . .'. Similarly, if you say 'It will probably rain tonight' your topic is rain, not evidence.

21. ELEMENTS OF PROBABILITY LOGIC

I shall now expound some elements of probability logic. A knowledge of them will be needed at certain points in the next five chapters, and we should prepare for it now. Even if you dislike logic and shrink from symbols, it will be worth your while to go patiently through this material. There is little of it; it is not formidable; and it could vivify what is to follow. Throughout these chapters, I use the form 'P(x)' to name the probability—objective or subjective—of x.

We start with three axioms:

(1) If Q is logically equivalent to R, $P(Q) = P(R)$.
(2) If Q is inconsistent with R, then $P(Q \vee R) = P(Q) + P(R)$.
(3) If Q is necessarily true, then $P(Q) = 1$.

(Alan Hájek reminds me that P must be understood as a function, so that it cannot assign different values to the same proposition; and that we should somehow secure that no probability lies outside the interval from 0 to 1.) A little reflection shows these axioms to be reasonable when understood in terms of objective probability or the subjective probabilities of an idealized thinker.

These axioms require us to pair $\neg(A \& \neg A)$ with 1, to pair A with 0 if $\neg A$ is absolutely necessary, to pair *It is not the case that sea-water is usually salty* with 0.93 if we pair *Sea-water is usually salty* with 0.07. And so on.

The 'idealized thinker' to whom I have referred does not represent a humanly achievable ideal. Mark Lance, in warning me about this, gives a good example:

> Consider the claim that *if the position on the chess board is P and the rules of chess R, then S: white has a winning strategy*. Say P is the position, R are the rules, and white does have a winning strategy. Then since chess is a finite determinate game, P&R entails S. But there are many such P and R where the winning strategy is unknown and indeed, in which, given our best understanding of chess theory, we are led to think the position drawn.

The fact remains that someone who assigns a probability < 1 to what is in fact a necessary truth thereby makes a mistake. My idealized thinker is simply someone who knows all the logical truths and makes no logical mistakes.

Of the theorems derivable from axioms 1–3, I shall present seven, numbered 4–10. Of these, 7 and 10 will be prominent in Chapter 9; the others are needed to prove those two.

Because Qv¬Q is a necessary proposition, we get

(4) P(Qv¬Q) = 1 (from 3).

This, combined with axiom 2, tells us that if P(Q) = 0.7 then P(¬Q) = 0.3, and quite generally that

(5) P(Q) = 1 − P(¬Q).

Because Q&R is inconsistent with Q&¬R, we get

(6) P((Q&R) v (Q&¬R)) = P(Q&R) + P(Q&¬R) (from 2).

Because (Q&R v Q&¬R) is logically equivalent to Q, and so (by 1) has the same probability as Q, we get the Addition Theorem:

(7) P(Q) = P(Q&R) + P(Q&¬R) (from 6).

This says that the probability of Q's being true is the probability of its being true while R is also true plus the probability of its being true while R is false. This obvious truth is important, as we shall see.

Hartry Field has warned me that (7) the Addition Theorem fails if the Law of Excluded Middle fails. If it can happen that neither R nor ¬R is true, then the probability of Q will be equal to P(Q&R) + P(Q&¬R) + P(Q & Neither-R-nor-¬R). I believe that all my uses of 7 could be reconstructed so as to survive the denial of the law of excluded middle; but this is only a guess, and I am not competent to put it to the test.

Continuing: if Q entails R, then Q is inconsistent with ¬R; and so we get

(8) If Q entails R, then P(Qv¬R) = P(Q) + P(¬R) (by 2),

from which it follows that

(9) If Q entails R, then P(Qv¬R) = P(Q) + 1 − P(R) (by 5).

Now suppose that Q entails R and that P(Q) > P(R). Then it follows (by 9) that P(Qv¬R) is equal to *P(Q) plus 1 minus something smaller than P(Q)*, which means that P(Qv¬R) > 1. But no probability can be > 1; this upshot is absurd. What led to it, namely the supposition that P(Q) > P(R), must be false. So we get:

(10) If Q entails R, then P(Q) ≤ P(R).

Considered as a thesis about subjective probability, 10 might be false of an individual person. Someone might be thick-headed enough to be more sure that Smith is a cannibal than that he is a carnivore, not realizing that being a cannibal entails being a carnivore. But a probability *logic* is normative: it sets constraints on how people's degrees of closeness to certainty should behave and combine, as do the laws of ordinary logic and arithmetic. When Dummkopf grasps that being a cannibal entails being a carnivore, he *ought* to stop being more sure of Smith's being a cannibal than he is of his being carnivorous. This is the spirit in which we must understand all the axioms and theorems. (Thus, to someone who has no ideas about what probability to assign to Q, which entails R, theorem 10 says: 'If you do arrive at an opinion about Q, you had better not accord it a higher probability than you give at that time to R.')

Other parts of the foregoing fragment of probability logic can also be normative for a thinker whose probabilities are not expressible in precise numbers. The Addition Theorem, for instance, passes judgement on anyone who is almost certain of P while giving low probabilities to P&Q and to P&¬Q.

22. THE RATIO FORMULA

That fragment of probability logic concerns only absolute probabilities, whether objective or subjective; I have not yet presented any logic of relative probabilities—or, as I shall henceforth call them, *conditional* probabilities—though these will be our chief concern. In discussing them we need a shorthand for 'the probability of C given A'. I shall express this by π(C/A), which can be pronounced pi-C-on-A. In this notation, π(/) is a binary operator on two propositions, C and A. Do not slip into reading π as a singulary operator on a proposition C/A, for there is no such proposition. Some writers use P( ) for unconditional and P(/) for conditional probability, but I join with those who think it safer to have a greater notational difference. In quoting others I shall silently bring their usage in line with mine.

How should subjective conditional probability be explained? You might think: 'Subjective relative probability is simply what someone believes about objective relative probability. Explain the latter, and the former will fall into place.' Not so! Consider a geographer who, as evidence accumulates, becomes ever more confident that *the Ross ice shelf will lose at least half of its surface area by 2025*; his subjective probability for—or degree of credence in—that proposition has grown steadily for the past decade. Throughout this time, he has firmly believed that the objective probability of the shrinkage proposition is either 1 or 0, this being a belief that is not responsive to evidence, because it comes straight from his meteorological determinism. Similarly with subjective conditional probability, which is not a firm opinion about an objective conditional probability, but rather a degree of credence accorded to one proposition on the supposition of another.

You can explore your value for $\pi(C/A)$ through the Ramsey test (§12). According to this, your value of $\pi(C/A)$ is high to the extent that the result of adding $P(A) = 1$ to your system of beliefs, and adjusting conservatively to make room for it, generates a high value for P(C).

Thus understood, conditional *probabilities* are properly so called, because they can be proved to conform to the logic of unconditional probabilities, and for other reasons. For a helpful development of this point, see Edgington 1996: 620–3. Responding to an ill-aimed attack, Edgington writes memorably: 'Like any technical notion in logic, mathematics, science or philosophy, you come best to understand [conditional probability] by working with it, not by reading about it' (p. 618). When you work with it, you find that it really is probability properly so called.

The most useful bridge between conditional and absolute probabilities is a formula that has been known and employed for more than two centuries, but which seems to have no standard name. I adopt the name Hájek gives it (forthcoming):

the Ratio Formula: $\pi(C/A) = P(A\&C) \div P(A)$,

subject to a proviso which I shall explain at the start of §23. Throughout this book, I use the form 'n ÷ m', which strictly names an arithmetical procedure, as though it named the fraction resulting from the procedure. The reason is aesthetic.

According to the Ratio Formula, the probability of C given A is the probability of A&C divided by the probability of A—it is not an unconditional probability, but rather a ratio of two such probabilities. This view was presented by Ramsey (1926: 82). I learn from Edgington (1995*b*: 262) that it was presented by Bayes and then by Laplace in the eighteenth century.

We are, as I have noted, hardly ever in a position to express our subjective probabilities in numerical terms, and few of my uses of the Ratio Formula will require me to do so. Its real thrust is its implication that $\pi(C/A)$ is large in proportion as P(A&C) is large and in proportion as P(A) is small.

Let us pause to see intuitively why this looks true. For a reasonable person like you, P(A&C) cannot be higher than P(A)—you will not regard A&C as more likely to be true than A is. Now let us consider three cases.

Suppose P(A&C) is only slightly lower than P(A), and look at the two sides of the Formula on that basis. On the left: your regarding A&C as almost as probable as A is means that you think A's truth is likely to bring C's truth with it, or to leave C's truth standing, which means that for you $\pi(C/A)$, as understood through the Ramsey test, is close to 1. On the right: a certain number, namely P(A&C), is divided by something nearly as great as it is, namely P(A), the result being close to 1.

Now suppose that P(A&C) is much lower than P(A), and look again at the Formula. On the left, $\pi(C/A)$ as understood through the Ramsey test comes out as having a low value, and so does the division on the right.

If we try it out again for the case where P(A&C) = P(A), it is easy to see that we get the value 1 on the left and on the right. The Ratio Formula looks right, does it not? For a more rigorous and grounded defence of it, see Edgington 1997*a*: 108–9.

I do not offer the Ratio Formula as *defining* conditional probability, for it does not. The best definition we have is the one provided by the Ramsey test: your conditional probability for C given A is the probability for C that results from adding $P(A) = 1$ to your belief system and conservatively adjusting to make room for it.

Given that in the Ratio Formula the right-hand side does not define the left, one would not predict that it is always one's route to the left; and clearly it is not. A person's value for $\pi(C/A)$ can often not be estimated by finding her values for P(A&C) and P(A) and dividing one by the other. Ellis, Edgington, and Mellor have all presented convincing cases of someone's having a value for $\pi(C/A)$ but none for P(A&C).

Indeed, to arrive at a value for P(A&C) one often has to proceed *through* a value for $\pi(C/A)$ or for $\pi(A/C)$, which means that we cannot use the right-hand side of the Ratio Formula to calculate the left (thus Blackburn 1986: 227–8). A better guide to evaluation is given by something which, though arithmetically equivalent to the Ratio Formula, starts in a different place. I shall call it the Multiplying Variant of the Ratio Formula:

MVRF: $P(A\&C) = \pi(C/A) \times P(A)$.

This will come to the fore in §25. In the meantime, we shall employ the Ratio Formula—not as an analysis or definition or procedural guide, just as a truth.

Hájek (forthcoming) argues fiercely and lavishly against the Ratio Formula. His arguments, however, do not cast doubt on any of my dealings with it. (These include my approval of its use in various technical bits of work, including the beautiful arguments of Hájek 1994 and 1989, which I shall expound in §28 and §31 respectively.) I now explain why that is so.

In his new paper Hájek relies upon examples of four kinds. Two involve probabilities of zero and infinitesimal probabilities respectively, and concern things like a mathematically continuous surface at which an infinitely fine dart is thrown. The probability that the dart will land on a given point on the surface is smaller than any assignable fraction, so that although such a landing is possible its occurrence has a probability that is either zero or infinitesimal (your choice). I contend that no such theoretical possibility could have any bearing on our lives; we could not possibly have a practical use for the notion of a *point* on a physical surface. If you think otherwise, try to take an infinitely less bold step: try to think of having a practical use for the difference, on a physical surface, between two regions which differed only at the 99th decimal place when measured in square centimetres.

The third of Hájek's four kinds of example concerns vague probabilities. When for a given person P(A) and P(C) are both vague, the right-hand side of the Ratio Formula must also be vague; and yet, for that person at that time, π(C/A) may be precise. In some of Hájek's examples A entails C, so that π(C/A) is 1, exactly and precisely, however vague P(A) and P(C) are. This counts against the Ratio Formula considered as an item of probability theory; but it is negligible when our main focus is on conditionals that are not 'independent' in the sense laid down in §7. That the Formula fails for those is a reason not for dropping it but rather for keeping independent conditionals off our screen. I also set aside conditionals that are *falsified* without help from any particular facts, as in Hájek's other examples, where π(C/A) = 0 because A logically contradicts C. (These remarks are not confined to conditionals whose truth or falsity is settled by logic alone. This challenge of Hájek's could arise in cases where A leads to C purely through causal laws.)

Fourthly, Hájek examines cases where the person has no value for either P(A) or P(C), using these to show that the Ratio Formula cannot be correct as an *analysis* of conditional probability. With this, as I have said, I agree.

23. INDICATIVE CONDITIONALS ARE ZERO-INTOLERANT

The Ratio Formula must be understood as confined to cases where P(A) > 0. If P(A) = 0, then the right-hand side of the Formula has the form n ÷ 0, which has no solution. The right hand side of MVRF—as Alan Hájek has pointed out to me—has the form n × 0, which = 0 for every n. Thus, π(C/A) is undefined by one formula and unconstrained by the other, which means that neither formula has any use when P(A) = 0.

Rather than being a flaw in the Formula, this result sheds light on its left-hand side, bringing into view an important consequence of the fact that indicative conditionals are devices for intellectually managing states of partial information, and for preparing for the advent of beliefs that one does not currently have. For an A that you regard as utterly ruled out, so that for you P(A) = 0, you have no disciplined way of making such preparations, no way of conducting the Ramsey test; you cannot say what the upshot is of adding to your belief system something you actually regard as having *no* chance of being true; so you have no value for π(C/A). Given the Ramsey test thesis, tying your value of π(C/A) to your value for P(A→C), it follows that someone for whom P(A) = 0 cannot find A→C in any degree acceptable, whatever C may be. There is abundant intuitive evidence that *nobody has any use for A→C when for him P(A) = 0*. I call this property of indicative conditionals their *zero-intolerance*. The point is widely known; for a good exposition of it see Warmbröd 1983: 250–1.

To see its rightness, combine the Ramsey test thesis with the Ratio Formula, implying that the acceptability to a person of A→C is given by his probability for A&C divided by his probability for A: 'Believe A→C to the extent that you think A&C is nearly as likely as A' (Edgington 1991: 189). You can do nothing with this in the case where your P(A) = 0.

Having boldly said that indicative conditionals are (always) zero-intolerant, I now take it back, acknowledging the existence of three sorts of counterexample to that. Each can be explained without detriment to the Ramsey test thesis.

(1) There are conversational stretches, as I call them. I know I visited Spain last year. If someone expressed doubts about this, I might say 'If I didn't visit Spain, then I am obviously a shameless liar, so if you doubt me, why should you believe anything I say?' This is a conversational stretch: although for me P(I didn't go to Spain) = 0, I assert a conditional with that as its antecedent in order to accommodate my speech to the doubter's beliefs. In this case, where I make the stretch for the sake of argument, I could as well have said '*You ought to be sure that* if I didn't visit Spain then I am a shameless liar, so . . . etc.'. In other cases I might make a stretch out of courtesy, tact, or embarrassment, and in many of those it would not suit my turn to state my whole thought explicitly. But there too I silently base the value of π(C/A) on the upshot of adding A not to my belief system but rather to that of the person(s) to whom I am adapting my speech. This counterexample succeeds because in it I base A→C on the Ramsey procedure as applied not to my system of beliefs but to someone else's.

(Stalnaker, to whom I owe that example and much of my understanding of stretches, holds that A→C is put out of business for me not only when for me P(A) = 0 but also when I *presuppose* A's falsity, being in a frame of mind where its truth is not in question for me. That might be because for me P(A) = 0, but it might have other sources and rationales. I agree about this, but it will suit my purposes to forgo this bit of generality, and attend to the special case where for the speaker P(A) = 0. So I shall retain the term 'zero-intolerance'.)

(2) Even if for me P(A) = 0, I may be forced to accept A→C—or anyway forbidden to write it off as a non-starter—because it follows from something that is more directly acceptable. This is an 'inferential exception' to the zero-intolerance of indicative conditionals. I believe that if *anyone* admires the Rolling Stones, he (or she!) will never be Pope; and this belief does not run into trouble from zero-intolerance. But it pretty clearly entails that if any member of the Roman Curia admires the Rolling Stones, he will never become Pope, so I am committed to this too. Yet it offends against the unqualified zero-intolerance thesis, on the most natural construal of the latter, because I give a zero probability to the proposition that some member of the Roman Curia admires the Rolling Stones.

The thesis that indicative conditionals are zero-intolerant should be confined to ones that stand on their own feet, and not applied to inferred ones. An 'inferred conditional' in my sense is one that someone accepts only because he accepts a general conditional that entails it. It is easy to explain the inferential exception to zero-intolerance. The latter's source is the Ramsey test: if for me P(A) = 0, I have no disciplined way of adding A to my belief system and conservatively adjusting the rest to make room for this. But the Ramsey test is not needed for inferred conditionals; their basis does not involve adding P(A) = 1 to one's belief system, but rather inferring A→C from a different conditional whose antecedent does not have a zero probability.

The moon is always more than ten thousand miles from Detroit, so if anything is ever on the moon then at that time it is that far from Detroit, so *if I have been on the moon I have been over ten thousand miles from Detroit*. This conditional is one of the 'problems for the Ramsey Test' adduced by Sanford (1989: 143–4); but it is plausible only considered as an inferred conditional, and so it does not score off the Ramsey test, properly understood.

(3) Then there are counterexamples that succeed because in them the Ramsey test is at work in a simplified fashion which skirts around the usual trouble of supposing P(A) = 1 when in fact for you P(A) = 0. All these cases have special features which enable one to go from A directly to C without having to consider how the rest of one's beliefs should be adjusted to make room for P(A) = 1. This genus has three species that I know of.

(*a*) One involves conditionals that are *independent*, meaning that in them the route from A to C needs no help from matters of particular fact. I said in §7 that these conflict with various useful general principles concerning conditionals, and now we see that one such conflict concerns the zero-intolerance of indicatives. If you can get from A to C purely through logic or general causal laws, or through general morality, your Ramsey test for A→C is a simplified affair: rather than having to accommodate P(A) = 0 by adjusting your other probabilities, you can just dump it in there and go straight to P(C) = 1.

(*b*) Secondly, there are *non-interference* conditionals (§50), which are accepted on the ground that the person holds C to be true and thinks that A's being true would not interfere with that. In *some* of these cases, the irrelevance of A to C is so obvious that one can arrive at A→C in one fell swoop, without having to look into changes that might need to be made in one's belief system if A were to be added to it. Even if I didn't go to Spain last year, Gibraltar still faces the Mediterranean.

(*c*) A third species was brought to my attention by Alan Hájek. An utterly convinced theist, for whom P(atheism) = 0, can accept '(Even) if there has never been a God, the first sentence of the Book of Genesis (still) says that in the beginning God created the heaven and the earth', this being a non-interference conditional. Without having the faintest idea of how he would change his doxastic scheme if it had to include atheism, he is sure that his belief about how Genesis begins would not alter. Then on that basis he can establish a conditional that does not fall into the non-interference category because its consequent is actually false (he thinks)—namely, the conditional 'If there has never been a God, then the opening sentence of Genesis is false'. Another example (for me): 'If there are Hobbits, then Tolkien's most famous novel refers to real beings as its heroes.' In every conditional A→C belonging to *c*, the person accepts A→B as a non-interference conditional, where B asserts the existence of some item having a certain meaning or content, and C assigns to B a relational property ('is true', 'does not refer to anything', etc.) which it must have if A is true.

Several writers have offered purported examples of acceptable indicative conditionals whose antecedents are known to be false, but their examples have all been suspect. A typical one is this: 'If I put my hand on this stove it will be burned.' One naturally thinks of this as said with reference to a hot stove, and as based on the belief that hot stoves can be depended on to burn hands that touch them. This is starting to sound like a subjunctive conditional—'If I were to . . . it would . . .'—which makes it risky to use as a counterexample to a thesis about indicatives. There is a complex story involved here, which I shall tell in §134.

24. THE EQUATION

Back in §12 I said that we ought to agree that your probability for A→C is measured by your value for $\pi(C/A)$. Thus:

For any person x at any time t, x's probability for A→C at t = x's value for $\pi(C/A)$ at t,

or, for short:

$P(A \rightarrow C) = \pi(C/A)$, where $P(A) > 0$.

This powerful, simple, and attractive thesis was first made widely known by Stalnaker (1970), and it has been called 'Stalnaker's Hypothesis'. But he tells me that Ernest Adams and Richard Jeffrey propounded it before he did; and Adams says that Brian Ellis deserves equal credit; so I shall leave personal names out of it, and follow Edgington in calling it the Equation.

On the face of it, the Equation looks like a possible starting point for an analysis of indicative conditionals. If we know that someone's level of credence in the truth of A→C is to be measured by his credence level for C on the supposition of A, this might be a first step towards finding out what proposition A→C is. It is not a bad start: it has → on the left and not on the right, and it strikes one as being central and deep, not marginal and shallow. An apparent drawback is its having on the left not A→C but P(A→C); and we have no way of getting rid of that 'P( )' and arriving at a solidly old-fashioned analytic formula: A→C is true if and only if . . .

Stalnaker showed one way to go: 'Conditional propositions are introduced as propositions whose absolute probability is equal to the conditional probability of the consequent on the antecedent. An axiom system for this conditional connective is recovered from the probabilistic definition' (1970: 107). So we are to start with the Equation, and work out what the logic of A→C would have to be for the Equation to be right. So much for the logic of →, but what about the account of what → is? the truth-conditions for A→C? the desired analytic biconditional with A→C on the left and something helpful on the right? In 1970 Stalnaker thought he was on the way to having such an analysis—the one involving 'worlds', mentioned early in §19—and hoped to be able to show that this analysis and the equation are in harmony, so that each could illuminate the other. Those of us who reject the 'worlds' approach to indicative conditionals can still build on the Equation, though in a different spirit. We can contend that A→C means *whatever it has to mean* such that P(A→C) is $\pi(C/A)$—that it means *whatever* satisfies the logic that has been built on this foundation. Do not dismiss this as evasive; it might be the entire fundamental truth about indicative conditionals. We like helpful analytic biconditionals, but nothing guarantees that they are always to be found.

In the case of indicative conditionals, there is a positive reason why they cannot be found. The ultimate obstacle to laying out the truth conditions for A→C in an analytic biconditional is that it does not have truth conditions. A central thesis of this book—one in which I am in agreement with many contemporary workers in the field—is that indicative conditionals are not ordinary propositions that are, except when vagueness or ambiguity infects them, always true or false.

A recognizably philosophical case for this will be made in Chapters 6 and 7. Before coming to that, however, I shall devote a chapter to laying out reasons of a different and more technical kind for the same conclusion. They all take the form of arguments against the Equation, purporting to show that the probability of a conditional cannot always be the relevant conditional probability; but they all assume that indicative conditionals are normal propositions whose probability is a probability of being true. Some make other assumptions that might also be challenged, but one thing whose denial undercuts them all is that A→C is a proposition with a truth value.

Some of Chapter 5 is fairly tough going, but I advise you to engage with this material. If you pass it by you will be left—as I was until I knuckled down to it—with a sense of skulking around the periphery of exciting but forbidden territory. My guide to it, though hard enough, is easier than anything the previous literature offered to me.