Decision making under indeterminacy

J. Robert G. Williams*

First Draft. Last revised: 1st April 2011

***

Note for FEW-ers: I intend to concentrate on sections 1 and 3 at the workshop.

Contents

1 Introduction. **3**
- 1.1 The Cabinet . . . 4
- 1.2 Options for action . . . 5
- 1.3 Rationalizing action . . . 6

2 Scaling indeterminacy **7**
- 2.1 Reconciliation? . . . 8
- 2.2 Belief, desire and action under indeterminacy . . . 9
- 2.3 Affect and indeterminacy . . . 12
- 2.4 Intermediate summary . . . 13

3 Uncertainty and Inconstancy **14**
- 3.1 Uncertainty . . . 15
- 3.2 Caprice and Randomize . . . 15
- 3.3 Diachronic coherence . . . 17
- 3.4 Application . . . 19

*This paper grew out of work on indeterminate survival I worked on while on an AHRC research leave award 'Intrinsic survival, Multiple survival, Indeterminate survival'. Material relevant to this was presented in several places: in Oxford, Cambridge, Leeds, and MIT. I'd like to particularly thank Ofra Magidor and Cian Dorr for helping me see that the decision-theoretic side of that early version needed further study. Thanks to participants in a CMM seminar for feedback on part of this paper; and in particular to Elizabeth Barnes and Seamus Bradley. This work was supported by a British Academy Research Development Award (BARDA: 53286).

Abstract

This papers studies a decision puzzle with indeterminacy at its centre. It is indeterminate whether Alpha survives as Omega; but Alpha has the chance to invest at small cost to greatly benefit Omega. Furthermore, Alpha is entirely self-interested. Should he invest? What patterns emerge if we repeat the experiment?

Addressing such questions is a central challenge in explicating the cognitive role of indeterminacy. But there is little consensus in the literature about even such mundane questions as: what attitude to p is appropriate, when one knows that p is indeterminate? This paper explores two answers, both built on a 'supervaluational' treatment of indeterminacy. The first is drawn out from David Lewis's discussion of Parfit on what matters in survival, and is a view where the indeterminacy of the identity relation between Alpha and Omega *scales* the concern Alpha should feel. The second is developed on the model of imprecise credence treatments of indeterminacy, and generates some interesting and suprisingly successful predictions about the forced march sorites.

1 Introduction.

Sometimes, the facts run out. Faced with a reddish-orangish colour patch, it seems to many that it is *indeterminate*, or *indefinite*, or *borderline*, or *unsettled*, whether the patch is red. The notion of indeterminacy is deployed all over philosophy—in discussions of vague predicates, theory change, future contingents, incompleteness in mathematics, conditionals, presupposition failure, the paradoxes of self-reference, and more.

But what is this notion? Is there even a single unified concept that covers all the cases just mentioned? Much sophisticated work has been done outlining potential semantico-logical frameworks for theorizing about indeterminacy. These may provide the skeleton of an explanatory account. But such treatments do not directly address the most pressing question about describing something as indeterminate. Suppose we accept that p is indeterminate; how should this belief affect our other doxastic states and wider mental life—in particular, what attitude can we rationally adopt to p, while holding the belief that it is indeterminate? What is the cognitive role of indeterminacy-judgements?

The puzzling thing about this is that it can seem like we *must* already have answered this question (at least tacitly) if we understand the notion of indeterminacy at all. If someone didn't understand that (ceteris paribus) believing that the chance of p is 50/50 should lead to a 50/50 credence in p, you wouldn't credit them with a mastery of the notion of chance. Conversely, it's part of our tacit understanding of alethic contingency that (absent idiosyncratic background beliefs) accepting that p is contingent doesn't constrain our credence in p at all. Surely, if we understand indeterminacy, there must be something parallel to say about the role of indeterminacy judgements—whether highly constraining, like chance, or unconstrained, like contingency. Yet, in marked contrast to chance and contingency, there seems at the level of theory not even the beginnings of a consensus as to what this role might be. Some philosophers think indeterminacy induces attitudes just like everyday uncertainty; others think one should not assign any positive credence at all to the propositions judged indeterminate. Still others think there is a special 'quandary' state appropriate to indeterminate cases—though the characterization of this special state is then a matter of contention.

A complicating factor is that it's far from clear we should be looking for a unified answer to this cognitive role question across all cases that philosophers have described as indeterminate. The idea that the future is indeterminate has a long pedigree in philosophy. But unless something radically revisionary is being proposed, this kind of indeterminacy judgement better be compatible with the whole range of attitudes of uncertainty, partial belief, and confident expectation. Future contingents are paradigmatic exemplars of such attitudes. Contrast presupposition failure: it would be silly to be 50/50 over whether Homer wrote the Illiad, if I'm 0.9 confident that Homer never existed. Even when we refine our cognitive role question down—for example, to borderline cases of a vague predicate—the puzzles above persist. It seems like borderlineness-judgements must have some distinctive doxastic role. But we find nothing approaching consensus on what that is.

As presented so far, the focus for these inquiries is the rational relation between believing that p is indeterminate, and doxastic attitudes to p itself. But the general question is much wider. We believe, desire, hope, fear, and act under vague guises. Our best accounts of the interrelation between such attitudes (for example, the interrelation between belief and desire set out in decision theories) often presuppose a classical backdrop that is under fire in the literature on vagueness. The methodology below is to approach the direct cognitive role question indirectly, through looking at the interplay of beliefs, desires and action in the presence of indeterminacy.

1.1 The Cabinet

The survival of a person across time can be an indeterminate matter, as van Inwagen (1990) emphasizes:

> One's life may be disrupted in various ways. If a pin is stuck into one's finger, one's life goes on. If one is blown to bits by a bomb, then—even if God immediately puts the bits back together again—one's life has ended. ... If, at the extremes of a spectrum along the length of which are arranged more and more radical disruptions of lives, we can find definite cases of the end of a life and definite cases of the continuation of a life, then it seems reasonable to suppose that somewhere between the extremes will be found disruptions of which it is not definitely true or definitely false that they constitute the end of a life. And if this is so, then there are possible adventures of which it is not definitely true or definitely false that one would survive them. Let us call such episodes 'indeterminate adventures'. Not everyone—perhaps hardly anyone—will agree with my contention that one survives an adventure if and only if one's life persists through that adventure. But anyone who thinks that people are complex material organisms will be hard put to it to deny that possibility of indeterminate adventures. (van Inwagen, 1990, p.243)

Many cases of this kind involve episodes after (or before) which it is indeterminate whether there is any person at all around. But another class are those where it is definite that there are persons before and after the episodes, but it's indeterminate whether it is the same person throughout. To discuss such cases without getting into the nitty-gritty of how to set things up for this or that theory of persistence, we'll follow van Inwagen in appealing to 'the cabinet':

> Suppose that a person, Alpha, enters a certain infernal philosophical engine called the Cabinet. Suppose that a person later emerges from the Cabinet and we immediately name him 'Omega'. *Is* Alpha Omega? ... Let us suppose the dials on the Cabinet have been set to provide its inmates with indeterminate adventures. (We need not agree on what would constitute an indeterminate adventure to suppose this. Let each philosopher fill in for himself the part of the story that tells how the dials are set). Alpha has entered and Omega has left. It is, therefore, not definitely true or definitely false that Alpha is Omega. (van Inwagen, 1990, p.243-4)

If one wants to discuss the impact of indeterminacy on belief, desire and action, vague personal identity is a good place to start. The reason is that (at least at first pass) our own survival *matters* to us. One desires, inter alia, that good things happen to oneself in the future; and fears the bad things that may happen. And these fears and desires are intrinsic, rather than instrumental—it is not, in the most usual cases, that one fears pain because being in pain is a signal that something else fearful is happening. It is the pain itself one fears and desires to avoid. Of course, one might also desire that good things happen to loved ones, or that that the welfare of humanity in general is improved. But there's a particular kind of self-interested concern that's extremely psychologically salient.

We can imagine that Alpha takes this to extremes—all that he cares about is the good or bad things that are going to happen to him. Our question is: how should Alpha then think about known goods and evils that happen to Omega? To test this, let some broker offer Alpha an investment opportunity. He's certain, we'll assume, that he will soon be subject to the Cabinet. For a small investment now by Alpha (100 dollars perhaps), Omega will benefit to the tune of thousands. Alpha is entirely self-interested, but not exclusively present-self-interested: he's

prepared to give up goods now if doing so brings him great gains in the future. So if he was certain that Omega was him, this would be a no-brainer: he'd take the investment like a shot. But, caring only about his own benefit, if he was certain that he was not Omega, he'd keep his cash and spend it on a few final nights of partying. What should he do? What would *you* do, in his shoes?

Parallel cases without indeterminacy would make things easier. For example, perhaps the makers of the Cabinet also make a machine that gives people chancy, rather than indeterminate, adventures. With the dials suitably set, anyone who enters this second machine has a 50/50 chance of passing through unscathed. If the dice roll against him, then he undergoes destruction and reconstruction that (according to one's favoured view of personal identity) produces a person determinately distinct from the one who enters. Faced with the certainty he's about to enter this machine (and knowing the chances involved), the investment decision facing Alpha would be comparatively familiar. At first pass, he would calculate the expected utility of investing as opposed to not-investing, and choose whichever maximizes utility (if Alpha was not so coldly rational as to calculate in this way in that context, many of us would at least say that this procedure tells Alpha what he should do).

But this model seems to misfire in the case of the original Cabinet. Alpha knows exactly what the outcome of each available action (investing, not investing) is, under one natural mode of description—Omega either receives the money (with Alpha going short right now) or receives no money (with Alpha getting some extra partying done). The trouble is that this is not the mode of description we need to hook up to Alpha's intrinsic desires. Alpha desires outcomes to the extent he himself benefits—and by construction it's indeterminate whether this condition is or is not met.

1.2 Options for action

Assuming the decision problem facing Alpha is well-posed, the practical options available to him are either to invest, or to not invest. Arguably, a third (set of) option(s) should be on the table: Alpha may adopt a mixed strategy: to do something operationally equivalent to flipping a (potentially biased) coin and investing iff that coin lands heads.

Suppose that we repeat the experiment many times over (perhaps with duplicates of Alpha; or perhaps the Cabinet-makers present him with the investment decision, then take him aside, wipe his memory, and rerun the experiment). The three options above predict different *patterns* of responses will be elicited under repeat examination. Alpha might invariably invest, invariably not invest, or behave in an inconstant manner, sometimes investing and sometimes not investing (with the long-run frequency typically matching up to the chances involved in the mixed strategy).

The repetitions may involve variations in the parameters of the setup. We could, for example, raise or lower the prize given to Omega if Alpha invests (to avoid distractions, we'll suppose it's never lower than the real future value of the amount Alpha is asked to invest). One possible pattern of behaviour features a tipping point, prizes below which invariably lead to not-investing, prizes above which lead to investing. This is the sort of behaviour we'd expect to see if Alpha maximized expected utility when faced with the Cabinet's chancy cousin—he'd invest when the reward to the survivor is sufficiently high that the expected utility it promises is greater than the certain loss of utility from current investment; and decline to invest when the reward is lower than this. One option is that Alpha displays this kind of risk-like behaviour when faced with the Cabinet.

What are the other options? Alpha might reject the investment for any level of prize: call that the 'Universal rejector' option. Or he may accept the investment at any level of prize:

call that the 'Universal acceptance' option. Or there could be inconstancy involved—with the proportions potentially varying with the value of the prize on offer. These four—threshold; universal rejection; universal acceptance; inconstancy—are prima facie the leading options.

Any of these patterns could characterize Alpha's behaviour. Getting clear on which of them should characterize his behaviour would shed light on the cognitive role of indeterminacy: since we could then ask about what doxastic attitude to the known indeterminate proposition that Alpha is Omega would rationalize the pattern.

1.3 Rationalizing action

Parfit famously formulated—and rejected—the idea that personal identity is *what matters* in survival. Now, of course many things might matter to an individual, but we're to focus on the familiar phenomenon of self-interested concern. Personal identity matters, I take it, if my self-interested actions should take into account what's good for y iff I am the same person as y; and affective care and concern should be generated by what will happen to y. I'll focus for the time being on the first, practical characterization—but we'll return to the affective one at the end. We'll be assuming in what follows, pace Parfit, that identity is what matters in this sense.

Given this, there's a straightforward argument against two of the possible patterns identified above. My self-interested actions should be sensitive to events happening to y iff y is me. But suppose we are Universal Rejectors. Then (unless something more is said) it looks like my self-interested actions simply discount the interests of y, when it's indeterminate whether y is me. Modus tollens applies: y is not me. Thus, on the assumption that we're Universal Rejectors, and given the 'mattering' biconditional, indeterminate survival collapses into non-survival. There's a dual worry for Universal Acceptors—if what happens to indeterminate y always matters to self-interested action, then the other direction of the biconditional gives us that y is me. So indeterminate survival collapses to survival. Either of these conditionals is very odd—and under rather minimal assumptions can be turned into a reductio (if each premise is determinately true, the conclusions hold determinately; but that explicitly contradicts the starting assumption).

We are left with the threshold and inconstancy behaviours. Neither is directly susceptible to this concern. So we are particularly interested in what sorts of accounts of indeterminacy could rationalize these patterns of behaviour.

One option in particular should be mentioned straightaway—this is the predictions of a certain kind of *epistemicist*. Epistemicists of the Williamsonian stripe (Williamson, 1994) hold there is always a fact of the matter over whether or not Alpha is Omega in the case as we've set it up. To be sure, neither Alpha nor Omega will know one way or another. But the cognitive role of indeterminacy need be no more complicated than that acknowledging ignorance. In particular, there's nothing to stop Alpha and Omega having ordinary attitudes of uncertainty—perhaps degrees of belief in the ordinary sense—about the question. This would naturally lead to the kind of threshold behaviour mentioned above, granted Alpha and Omega maximize expected utility.

The epistemicist take on the cognitive role of indeterminacy allows us to subsume the case of Alpha and Omega under a familiar model. And I don't think its predictions for action in this case are bad. As a general description of the cognitive role of indeterminacy, however, I think it is starkly revisionary. For a central part of this conception, I think, is that when we know that it is indeterminate whether p, we take further inquiry into whether p to be pointless.[1] For example, Sider (2001) argues that the debate between physical and psychological accounts of indeterminacy has no answer—it's indeterminate which (if either) criterion is correct. But

[1] Compare Field (2003).

I take it the idea is that the project of putative evidence-gathering—intuition-mongering about recherché cases, or whatever—should stop (Sider characterizes the indeterminacy answer as one that allows the 'battle-weary' of metaphysics 'cast aside their swords'). This seems an absolutely central part of the role the concept plays, at least within philosophy. But if one were an epistemicist, this should seem like a non sequitur. Since (despite the indeterminacy) there's a fact of the matter, evidence should in principle be able to improve the accuracy of our credences on such matters (Alpha might get a better fix on what to pay). To be sure, knowledge in the matter is out of our reach. But knowledge is not the be all and end all of inquiry. Probabilistic judgement, informed by known evidence, may be equally valuable, and nothing in the epistemicist story, so far as I can see, rules this out.

The plan for the next few sections is to explore one central way of undergirding threshold behaviour without appeal to epistemicism—what I call the scaling model. I'll argue that David Lewis was committed to this kind of conception of indeterminacy—without it, his sophisticated attempt to reconcile Parfit's substantive account of what matters in survival with the commonsensical claim that *identity* is what matters, would fail. This delivers a threshold model of behaviour; and more generally an attractive account of belief, desire and action under indeterminacy. On balance, however, I think the account should be rejected—the trouble coming not from the significance of identity for action, but from the affective role.

In the final third of this paper I will provide the underpinnings for an *inconstant* pattern of behaviour in the Alpha-Omega case. This draws on the resources of the *imprecise probability* model of uncertainty—and we'll see that it has independent appeal.

2 Scaling indeterminacy

Derek Parfit (1984) gives a description of what matters in survival that many find appealing. For Parfit, the extent to which we should "self-interestedly" care about future individuals should vary with the strength of the (gradable) relations that underlie personal identity (e.g. psychological continuity and connectedness). Call this Parfit's *primary* claim. On this basis, Parfit goes on to claim one should *not* care intrinsically about *one's own* survival—as Alpha ex hypothesi does. Parfit proposes his account as a revision, not description, of our ordinary thinking about what matters in survival; so it wouldn't be surprising to the Parfittian to find an agent with Alpha's cognitive profile. But, says Parfit, Alpha should change his mind. Call this Parfit's *secondary* claim.

I'm going to explore the prospects for endorsing Parfit's primary claim while resisting the secondary one—in fact, I'm going to simply lift David Lewis's proposal for doing exactly this. What is of particular interest to us will be the *understanding of indeterminacy* that we must buy into, to make Lewis's response one that vindicates the primary claim.

Parfit has a number of arguments against identity being what matters. Two central cases are: (i) that psychological continuity and connectedness can hold between me-now and two distinct individuals existing at a future time (the fission cases); identity is Euclidean, and hence can hold between me now and individuals at a future time only if the latter are identical to one another; (ii) psychological continuity and connectedness comes in degrees, whereas identity is all-or-nothing. Alpha and Omega can be taken to illustrate the second point. We can suppose that their *degree* of R-relatedness is intermediate, whereas personal identity (so we are told) either obtains or does not.

Lewis (1976) responded to cases of type (i) by deploying a perdurantist metaphysics of persons. For me-now to survive as stage S is for there to be some perduring person (a transtemporal fusion) which contains both of us as maximal temporal parts. Fission is analyzed as a case

where two distinct persons both contain me-now as a current part. On Lewis's view pre-fission I am already a stage of two different people; but the qualitative histories of those people separate only after fission has occurred.[2]

Lewis's response to the second concern is to postulate indeterminacy in personal-identity across time—indeterminacy of dyadic survival that comes in degrees just as does psychological continuity and connectedness (the *R*-relation) does. On Lewis's metaphysics, this indeterminacy arises from indeterminacy in which transtemporal fusions are people. To model this, he supposes that there is some level of *R*-interrelatedness between person-stages that's necessary for the fusion of those stages to count overall as a person; the persons will be the maximal such *R*-interrelated fusions of person-stages. That by itself suggests a sharp-cut-off, not something that comes in degrees. But Lewis argues that it will be indeterminate how the 'boundary number' (the degree of *R*-relatedness required) is to be set. Extreme and implausible answers would be 0 (in which the stages would be *R*-related to the lowest degree) and 1 (in which the stages would be *R*-related to the highest degree). Lewis suggests that the degree to which it is determinate that some fusion is a person is given by the measure of the boundary numbers according to which it would count as a person (more generally, any statement is determinate to degree k iff it is true on measure k of the boundary-number 'sharpenings'). Two stages will stand in the personal-identity (*I*) relation iff there is some person of which they are each maximal temporal stages, and the degrees of determinacy of the two sides of this biconditional will match.

Let's see how this applies to the case of Alpha and Omega. The dials have been set, we can assume, so that Omega is psychologically connected to Alpha only to a reduced degree—0.5 say. According to resolutions of the vagueness of 'person' that set the boundary above 0.5, there will be no person that includes both the Alpha-stage and the Omega-stage—so Alpha will not survive as Omega. On more liberal resolutions, the connections are sufficient for there to be a person covering the two stages, and Alpha will be Omega. The overall degree of determinateness of Alpha being Omega is given by the measure of the sharpenings that give the second result— 0.5 again. So the degree of determinacy of Alpha being Omega exactly matches the degree of *R*-relatedness.

Lewis claims to be reconciling common-sense views of what matters in survival with Parfit's reductionist proposals for what *does* in fact. Part of the task he faces in both (i) and (ii) is to rebut Parfit's arguments that the 'logic of *R*-relatedness' and the 'logic of identity' conflict—that the *R* relation is Euclidean and personal identity non-Euclidean; that *R* comes in degrees and personal identity cannot. I don't want to reopen this familiar side of the debate—suppose we grant that Lewis's views on this front are at least internally consistent. But if Lewis is to truly reconcile reductionism and common-sense—if he's to convince someone like self-interested Alpha that the Parfittian *R*-relation is what he cares about under another name—then more is needed. Part of the task is to make sure that caring about personal identity and caring about *R*-relatedness really are consistent with one another.

2.1 Reconciliation?

The Parfittian position is that the extent to which what happens to Omega matters to Alpha should vary with the strength of the *R*-relations between them. But Lewis has said that it is indeterminate whether Alpha is Omega. For all we've said, we might conclude at this point that since it's indeterminate that Alpha is Omega, Alpha has no reason to care what happens to Omega; or we might think that inconstancy is the appropriate reaction to indeterminacy, and

[2]One difficulty in stating this position is whether to say that the referent of 'I' is a person, or a stage of a person. Only on the latter is what I've just said strictly true.

ask Alpha to make an arbitrary judgement-call to fix whether or not to factor Omega into his concerns on the model of full-survival, or on the model of full non-survival. So just by introducing indeterminacy into the picture, and hence a degree-like structure to match the degrees of *R*-relatedness, Lewis has not done enough to justify calling his position a reconciliation of the Parfittian theses with common-sense. He must in addition endorse a certain conception of one aspect of the cognitive role of indeterminacy, opposed to those just mentioned.

For his position to be reconciliatory, Lewis needs to maintain that intrinsic care for personal identity translates to *scaled* care for *indeterminate cases* of personal identity. With a *degree theory* of determinacy like Lewis's, there's a natural proposal to go for: the scaling of concern should be proportional to the degree of determinacy of the obtaining of the relation. More specifically: the utility appropriate to a situation where Alpha himself is F, should be scaled according by the relevant degree if Omega is F and it is indeterminate whether Alpha is Omega.

This model would underpin risk-like behaviour in the investment decision that faces Alpha. Alpha cares about what happens to Omega, but only to a certain degree—it takes proportionately more rewards to motivate him to take the investment opportunity. So Alpha will reject the investment opportunity—even given superficially favourable returns—until some threshold is reached. As we turn the dials on the cabinet to raise or lower the degree of determinate identity between the two, the appropriate threshold will get higher or lower respectively. The behaviour is very like what we would expect to see in the chance-variation of the Cabinet, as the chances are raised or lowered. But the explanation is very different. In chance-based reasoning, Alpha would care very intensely about the possible outcomes—but the difference in behaviour is explained by changing levels of belief in which outcome is to be actualized. In the case we're considering, the explanation for the changing threshold has really nothing to do with uncertainty or risk; it is simply that he cares about Omega to a reduced but non-negligable degree, and so only when the benefits for Omega are sufficiently great will he judge it worth a present sacrifice.

2.2 Belief, desire and action under indeterminacy

What we need to complete the reconciliatory ambitions of the Lewisian proposal is, in the first instance, a principle about how utilities are assigned to outcomes, on the basis of what an agent cares about. Suppose, for example, that we had a rule for assigning utilities to agents at times in terms of the bundles of commodities they possess. And suppose that x at t has commodity bundle c_t^x, which is assigned utility g; and the degree of determinacy of x being A is k. Then the contribution that x's situation makes to A's overall utility is gk. In this simple model, we can think of A's overall utility a as being the sum of the local utilities attaching to his "successors" situations, with the summands being weighted by degree of determinacy δ_t^x:

$$V(a) = \sum_{x,t} \delta_t^x c_t^x$$

This is an adaption of a standard model of intertemporal utility, in which context δ_t^x are known as the 'discount factors'.[3] The differences here are in the *interpretation* of the discounting, and also in that we sum (and let the discounting be dependent on) time-slices of individuals rather than time.

Sometimes, temporal discounting of utility can be justified as a fudge, justified as an approximation to account for the factors such as natural interest rates on commodities or risk that

[3] Various aspects of this model, not relevant for present purposes, may be questions—for example, the additivity of discounted utilities. See Mas-Colell *et al.* (1995, ch.x).

otherwise go unrepresented in a model (Broome, 1994). What we have here is a pure discounting that remains even once all these factors are accounted for.[4]

So we have a first-pass model about the way in which Omega's situation should be factored into the utility that Alpha attaches to various prospects. But originally we asked a different question: given that Alpha knows that it's determinate to degree k that he is Omega, what *doxastic* attitude should he adopt to the proposition that he is Omega, or that he himself is the one that receives any return from taking the investment?

There's a simple argument that an answer is implicit in what has gone before. Consider a situation where Alpha is certain that he will enter the Cabinet, and that Omega will thereby receive goodies C. Also, ex hypothesi, Alpha values a situation iff he himself gets goodies in it. Let A be the proposition that Alpha gets the goodies, O be the proposition that Omega gets goodies. Without loss of generality, we may assume that zero is chosen so that $U(\neg A) = 0$. Then there are, one would think, two equivalent ways for Alpha to evaluate the prospect of investing: the expected utility of investing, as calculated on the $A/\neg A$ partition; and the expected utility of investing, as calculated on the $O/\neg O$ partition. The natural thought is that these are:

1. $U(\textsc{invest}) = C(A)U(A) + C(\neg A)U(\neg A) = C(A)U(A)$
2. $U(\textsc{invest}) = C(O)U(O) + C(\neg O)U(\neg A) = U(O)$

The first identity follows from the assumption that $U(\neg A) = 0$; the second of the 1/0 distribution on the respective credences. From the equality of the two LHSs, we can derive that $C(A) = U(O)/U(A)$. But furthermore, we have our scaling claim about what fixes $U(O)$ itself: $U(O) = U(A)\delta_t^\Omega$. Hence $C(A) = \delta_t^\Omega$—that is, Alpha's credence that he himself will get goodies matches the degree of determinacy that he himself is Omega.

This motivates a generalization: just as the 'ideal aims' of credence in a purely classical setting are 1 (if the proposition in question is true) and 0 (if the proposition in question is false), the ideal aims of credence in a Lewisian degreed-determinacy setting should require that one's credence in a proposition match its degree of determinacy.[5] More specifically—these are the right degreed states to adopt to the propositions mentioned *insofar as their role in deliberating which acts to take*; they play a belief-like functional role with respect to the calculation of preference. But we should not assume that the whole of the ordinary functional role of 'credence' carries over to this setting. In particular, it's *not* clear that uncertainty itself carries over, if this is functionally connected to hope, fears, and other emotional attitudes, as well as desire and action. This will be important later.

The simple argument given earlier may be resisted. It presupposes, after all, that there is some kind of 'credence' that features in familiar expected utility calculations; that partitions over 'vague' propositions is an acceptable way to extend usual decision-theoretic calculations. One could, in principle, reject this model. For example, Lewis and Stalnaker have long insisted that we should think of the primary objects of attitudes as sets of possible worlds. And it's very unclear that there is a set of worlds associated with a claim like A above. So suspicions about

[4]The Parfittian rationalization for this, which we are here adopting, is well known in the literature on temporal discounting: a recent critical review calls it "the most compelling argument" for positive time preference (Frederick *et al.*, 2002).

One of the striking features of Parfit's account is that even the most intuitively imprudent (and one might claim: paradigmatically irrational) behaviour may not be classified as irrational. Thus, if discounting takes a *hyperbolic* form, then *preference reversals* are predicted—even if one's beliefs and basic preferences don't change, preferences for A and B can reverse, as A's discount rate decreases faster than B's. This is widely held to be an empirically verified phenomenon, and seems 'irrational'. But if the R-relation decays hyperbolically, Parfit would predict it. Since Lewis's account is built to replicate Parfit's substantive predictions, it will too.

[5]Compare Smith (2008).

the simple argument are reasonable; more needs to be said to convince ourselves that this sort of argument is legitimate.

My preferred framework is the following. Granted, we can model credences and utilities over a set of Lewisian worlds (and for the sake of argument we may grant that there's no indeterminacy in which such world is actual, and no indeterminacy in what credence and utilities are thereby assigned). But we can also model credences and utilities over a set of propositions, conceived in various other ways—as mentalese sentences, or structured Russellian propositions, or Fregean thoughts. In the classical case, there are natural ways of projecting credences and utilities between these representations—at first pass, the degree of belief and desirability of P matches the degree of belief/desirability of the set of worlds where P is true.

In a setting with Lewisian degree-determinacy, it is the credence and utility functions *distributed over the non-worldly propositions* that have non-classical truth-statuses—this is where we find propositions like 'Alpha gets C' living. The semantic relations between propositions in this sense and Lewisian worlds are degree-like; and we exploited this in the scaling hypothesis to assign utility to worlds on the basis of the degree of truth of a proposition ('I get C at t') at that world. If we spot ourselves credences over worlds, this allows us to carry out fully classical decision theoretic calculations over world-space. But it would be nice if we could also project back up—to allow decision-theoretic calculations to operate with mentalese sentences, Fregean thoughts or whatever (after all, on many theories—though not Lewis's—these will be the primary psychological reality).

The simple argument given earlier in effect presupposed that it was legitimate to apply decision-theoretic tools over a space of propositions taking degree-like non-classical truth values. But is the case? What should credences in this space look like? How should they interact with preferences?

The simple argument itself gives us our lead. It's conclusion identified an ideal psychological state to take to a proposition whose degree of truth was k—a credence of k. Just as accuracy in the classical case is measured by distance from the ideal psychological state of 1 (if the proposition is true) and 0 (if the proposition is false), our belief state is worse—less accurate—when the credence we adopt to the proposition is further away from its degree of truth.

What we can show (not only for this, but for many other possible non-classical settings) is that we can rerun Joyce's *accuracy-domination* arguments for probabilism (Joyce, 1998, 2009) in this generalized setting (indeed, if we stick with the central exemplar of an accuracy-measure—the Brier score—this was proved by De Finetti (1974), in giving a domination argument for expectations). I tackle this in detail elsewhere (Williams, forthcoming), but the executive summary is this: just as probabilism claims that rational credences in its setting should be convex combinations of classical truth value distributions, generalized probabilism claims that rational credences in its setting should be convex combinations of non-classical truth value distributions.[6] A generalized probability in this sense must be representable as follows:

$$P(q) = \sum_w \lambda_w |q|_w$$

where $1 = \sum_w \lambda_w$; and $|q|_w$ the truth value of q at w. For classical truth values, this of course gives a probability function. For detailed discussion of what this delivers in the degree-supervaluational settings, see (Williams, 2010) and compare (Paris, 2001).

Integrating generalized probabilities with generalized preferences is the next stage. But this again can be done smoothly. The crucial move in constructing a Jeffrey-style decision

[6]Intimately related to these generalization of Joyce-style arguments for probabilism are generalizations of dutch book arguments (Paris, 2001).

theory (Jeffrey, 1965) to this setting, is to get a grip on generalized conditional probabilities—which involves a not-very-obvious generalization of the ratio formula, using a strange "product" conjunction. This can be motivated by generalizations of familiar dutch-book arguments. But with this in place, desirabilities in arbitrary non-classical propositions are fixed by worldly credences and utilities, and also show that the kind of partition-relative calculations of utility that Jeffrey allows are legitimate in the non-classical setting. We can justify a Jeffrey-style characterization of the probability-desirability link of the following form:

$$D(q) = \sum_{\gamma \in \Gamma} P(q|\gamma) D(q \cdot \gamma)$$

Here $\cdot$ is the product conjunction (the rule being that the truth value of $A \cdot B$ is the product of the truth values of the conjuncts). The Γ is an arbitrary partition, where this is understood in the non-classical setting as a set of propositions whose truth values sum to 1 at every world.

This is of course the briefest of sketches, and I don't mean to do more than to indicate the direction in which the simple argument points. If it is made good, we will have substantiated the claim earlier: intermediate 'credences' in non-classical propositions interact with desirabilities and action in a way that directly generalizes the traditional classical functional role—Alpha should in this sense have an intermediate credence that he is Omega, and deliberate in this basis in familiar ways.

On this way of developing the Lewis degree-determinacy account, Alpha has intermediate credence that investing brings him the outcome he desires—himself getting the reward. The epistemicist would agree. And he should use this to calculate the expected utility of investing—holding fixed the credence, this means that he will invest only when the reward is sufficiently large to counteract any lack of confidence in the desired outcome coming about. Again the epistemicist agrees. Both pictures support threshold behaviour. But *outside* the sphere of belief-desire-action explanations, the two accounts differ markedly. For example, with an intermediate credence in himself getting the prizes, Alpha can know that he has full information—that his psychological state is a God-like one, ideally appropriate to the situation. The epistemicist should regard an intermediate credence as a compromise, and is on the look out for further evidence that would indicate how to shift credences towards the ideal ones: either 1 or 0. So the degree-like rationalization of threshold behaviour can account for the way in which information about indeterminacy *closes off inquiry* in a way that escapes the epistemicist.

2.3 Affect and indeterminacy

Thus far we have focused on the *practical* sense of personal identity mattering—mattering to action. We've sketched an account of desirability and belief in indeterminate propositions that aligns the practical significance of indeterminate personal identity with the underlying gradable R relation. But the practical import of indeterminacy is only one dimension of its overall cognitive role. I think it is here that the real worries for the account begin.

To take just one example (inspired by Bernard Williams' (1970) discussion of such cases), consider the phenomenon of *de se* fear—a distinctive emotional reaction when you suspect that you yourself are in harms' way. Of course, you needn't know or even be highly confident that you face danger in order for fear that you will be harmed to be intelligible—believing that you might be harmed seems sufficient.

Suppose that Alpha learns that Omega is (or has some chance) of undergoing severe pain. Is it rational for Alpha to be afraid that he himself will undergo pain in circumstances C? Williams' thought is that we face a dilemma: 'fear seems neither appropriate, nor inappropriate, nor appropriately equivocal'. Let's see if we can unpack this.

The propositional attitudes that seemingly need to be in place for de se fear to be appropriate are (i) Alpha believes that he might suffer pain, in circumstances C; (ii) Alpha strongly desires that this not come about. The second is satisfied. But how shall we fill out (i)? Two aspects of the above discussion are salient. First, that Alpha has an intermediate credence in the proposition that he himself will suffer pain. Second, that Alpha regards the question as closed to further inquiry—given the information he has, no further information could make it rational to have any higher credence that he himself will suffer pain.

How does this speak to the doxastic condition on fear? In standard settings, it's pretty clear that in circumstances where one has middling credences in p, the possibility of p is epistemically open—and so one can expect an agent to meet condition (i). But it's at all clear in present circumstances that this is sufficient. Perhaps we should characterize epistemic possibility in terms of whether the proposition in question can be rationally added to the information base. But if we go this way, then Alpha should not believe that he might suffer pain. Our familiar model of epistemic modality pulls in two directions in these non-classical settings.

But neither direction is satisfactory. Suppose we end up saying that the doxastic condition is not met, since the possibility that Alpha himself feels pain is not genuinely epistemically open. Then Alpha shouldn't fear the pain. But if identity is what matters, Alpha's de se affective states should focus on what happens to y iff y is Alpha. But it seems that (at least in this case) the left hand side is false—so unless this is somehow unrepresentative, we again get that Omega is not Alpha—which can be turned into a contradiction.

If we say that the doxastic condition is met, because Alpha at least has positive credences in the proposition that he will suffer pain. Then the fear will be permissible. But in this case, Alpha's affective state is unambiguously focused on what happens to Omega; arguably, by the mattering biconditional, this implies that Alpha is Omega, unless we can somehow make the case that this is unrepresentative—which can again can be turned into a contradiction.

Bernard Williams claimed that indeterminacy of my own survival is something that has no coherent representation in my emotions and expectations. On the latter front, I think we can argue on the basis of what's been said above that he was wrong on one count: expectations and desires can be coherently described, from what I can tell. But with respect to the wider cognitive setting—and in particular the affective correlates of propositional attitudes—he may yet be correct.

2.4 Intermediate summary

Where have we reached? We started with the puzzle of Alpha's indeterminate adventure. We asked: what would it be sensible for Alpha to do, faced with an investment decision that involved present sacrifice, for Omega's benefit. There are various patterns of behaviour a theory of the cognitive role of indeterminacy could countenance: universal acceptance, universal rejection, risk-like behaviour involving acceptance when above a threshold; and inconstant behaviour.

We then looked at one particular substantive account of the way that the interests of future individuals matter to current ones, from a self-interested point of view. This was Parfit's account. We argued that Lewis's famous strategy for reconciling Parfit's substantive theory of what matters in survival with the platitudes of common-sense required a distinctive take on the cognitive role of indeterminacy—in the the first instance, this is a scaling thesis about how the properties we intrinsically care about fix the utility of outcomes; but it translates to a principle about the credences in p ideally called for in cases where p is indeterminate. This principle about the 'psychological loading' of degrees of determinacy can be parlayed into a generalized probabilism about rational degrees of belief and an account of the interactions between

credences and desirabilities which supports the argument itself. As a theory that by design incorporates Parfit's substantive predictions about the pattern of care for individuals over time, Lewis's theory will deliver the same verdicts on paradigm cases of imprudence as does Parfit. But if anything Lewis seems better equipped to allay our initial suspicions over these predictions.

Is this the picture of personal identity that we should, all things considered, endorse? The case of Alpha and Omega was designed by van Inwagen to be maximally neutral over first-order accounts of what personal identity is. And though Lewis and Parfit themselves point to a psychological take on what plays the central role, any '*R*'-relation that comes in degrees (and which we care about in proportion to these degrees) can be fed into the Lewis account of degree-determinacy. More generally, this suggests a general rationale for the kind of scaling cognitive role of determinacy we've been using. If our underlying interest is in some quantitative phenomenon, but for whatever reason we find it convenient to represent it within our cognitive economy with an all-or-nothing predicate that makes no reference to degrees, degreed indeterminacy with the scaling interpretation allows us to *encode* our care for the underlying phenomenon in our attitudes to the all-or-nothing representational device. Personal identity is a plausible case for this; but not the only one.

But the whole point of the Lewis account is to 'save the phenomenon' of self-identity mattering, in the context of an underlying quantitative relation that seems to give a substantive account of what self-interestedly matters. While I think that for belief, desire and action a plausible case can be made for this, it's hard to see how to generalize the picture more widely; what to say about intensional emotional states like *de se* fear, for example.

I earlier distinguished four patterns of behaviour we might find in the Alpha/Omega situation. Eliminating universal acceptance and rejection leaves only threshold-centred, risk-like behaviour, or an inconstant pattern. The degree-theory just reconsidered is I think the best (non-epistemicist) route for a threshold view. In the remainder of the essay I want to look at a model of the mind supporting inconstancy.

3 Uncertainty and Inconstancy

Our goal is to find rational underpinnings for inconstant behaviour in the presence of indeterminacy.[7] The suggestion that borderline cases be associated with inconstancy is not a new one—Crispin Wright has long argued that the basic phenomenon of our interaction with borderline cases involves hesitant polar judgements, tolerance of competent speakers taking different views, and the like.

The idea of inconstancy I will work with is that our attitudes to indeterminacy are given by *partial* conflicting dispositions. Faced with the decision to invest or not in the situation we start with, Alpha must make a judgement-call. Will he treat Omega as the same person as himself or not? It is a genuine judgement, rather than a mere guess, since once Alpha has made it (on this view) he will act as if certain of its truth (not a characteristic of mere guesses). But nothing predetermines Alpha's action, and so under repeat trials, we expect inconstancy. But to endorse inconstancy is not to give up on prediction, as we'll see below.

[7]The idea that our attitude to borderline cases of vague predicates is characteristically inconstant is something that I draw from Crispin Wright's work (see especially (Wright, 1976, 2001, 2003); though the way I develop the idea is very different. My thinking below—especially the association of a degree theory with partial dispositions to act in conflicting ways, which grew into the weightings of Randomize—grew directly out of engagement with the characterization of vagueness-related degrees of partial belief in Schiffer (2003)—especially the brief characterizations of their characteristic *ambivalence*. Schiffer assures me, however, that his account should be intepreted as excluding the kind of inconstant behaviour I appeal to.

Inconstancy steers clear of trouble with the mattering biconditionals. To be sure, on occasion, Alpha treats Omega as the object of self-concern, both practically and affectively. But on other occasions, the opposite is true. There's no *generic* behaviour either way that would allow us to derive either the identity or non-identity via the biconditional.

3.1 Uncertainty

In the classical, bivalent setting, the 'ideal psychological states' associated with propositions are point-like credences: full confidence in p if its true; no confidence if it is false. Lewis's degree theory, I argued, can be fitted into this picture—with the ideal psychological state being one that matches an intermediate degree of truth. Cases of *uncertainty* in both classical and non-classical settings were similarly pointlike, represented by probabilities.

There is an alternative model of uncertainty: the 'imprecise credence' model favoured by Isaac Levi (1974) (among many others[8])—the idea here is to represent a person's doxastic state not by a single, precise credence, but by a whole *set* of them, represent the agents' 'open mindedness' or 'uncertainty' between the various more particular views. The view to be explored below is that the ideal psychological state associated with indeterminacy is that of *uncertainty* in the sense just mentioned. Following common usage, we will call the set of probabilities representing a belief state on the imprecise credence model the agent's representor.

But it's one thing to have a formal tool for representing belief states, quite another to say what the relevance of these states are for one's wider psychology or behaviour. There are many possible decision rules that may be associated with the imprecise-credences representation. One common basis for decision rules is the following 'dominance' principle: that if an action is impermissible[9] on every probability within the representor, then it is impermissible simpliciter. This means that each permissible action must be optimal on at least one probability within the representing set.

The decision-rule that Weatherson (Manuscript) labels 'Caprice' (and advocates) says that the dominance principle is both necessary and sufficient for permissibility. Decision-making under uncertainty, on this view, characteristically leaves the agent options. A different option is to supplement the dominance principle with more rules. Levi's preferred approach takes this form—once we have the arena of 'admissible' options, we select among them, by choosing the one (roughly) that guarantees us best 'worst case scenario'.[10]

I will not take a stand on whether the Caprice rule characterizes everyday decision-making under ordinary uncertainty. The conjecture here is rather that *a Caprice-supporting thick doxastic state is the appropriate attitude to take to borderline cases*.

3.2 Caprice and Randomize

The guiding idea of Caprice is the following. Your state of uncertainty consists in there being a set of options between which you have no discriminating evidence. The expected-utility maximizing action may differ according to which option is realized, but you have no obvious principled way of aggregating the option-relative verdicts. Caprice tells you simply to plump

[8]There are too many varieties to survey here, but see Joyce (2011) for the state of art within philosophy, and Weatherson (Manuscript) for references to some of the wider literature and the variety of approaches being pursued.

[9]does not maximize expected utility

[10]Notice that we can think of this second-stage rule in quite a few ways. It might be a rationally enforced choice rule, on a par with expected utility maximization in the point-like setting. It could on the other hand simply be a description of how we cope with the optionality Weatherson diagnoses. Or it might be something in between—like a convention. Actions not in accordance with the second-stage rule might be not irrational, but reckless.

for one or another option, and act in the way it recommends. More generally, any action that is recommended by some option is permissible. In Alpha's case, the relevant options might be: Alpha survives as Omega; Alpha doesn't survive. On the hypothesis that known indeterminacy produces uncertainty, Alpha should be totally uncertain which is realized, but sees that given the first, he should invest; and given the second, he shouldn't. Caprice tells him to simply go for one or the other. Hence the predicted inconstancy of behaviour.

It is not enough just to say that if p is borderline, then the correct attitude to p is that represented by a set of probabilities (whose functional role is given by the Caprice rule). We need to ask: what probabilities, in particular, make it into the set? In order to tackle this, I need to be more specific on the background semantic treatment of vagueness. I'll simply adopt the same Lewisian supervaluational model used in earlier sections. In particular, we'll be appealing to *sharpenings* (boundary numbers)—relative to which a classical interpretation is induced over our language (or, indeed, propositions, if propositions are the sort of things that themselves are vague, and have truth conditions—I'll speak this way below).

Let's go back to our worked example. Alpha, we may assume, has a definite credence assignment over a certain partition of possibilities—he knows exactly which particle-configurations will arise depending on how he acts. The difficulty is that 'Alpha survives as Omega' is true at one of these possibilities relative to one sharpening, and false at the same possibility relative to other sharpenings. Holding fixed a sharpening, one can read off a definite credence assignment over propositions (equal to the weight of the worlds at which the proposition is true, on that sharpening). In this way, a credence assignment over worlds generates a set of credence assignments over propositions—a state of uncertainty over those propositions—via the supervaluational machinery. This kind of projection works equally well for utility. If we desire P intrinsically, the rule might be that any world such that P gets $+1$ utility. If it's indeterminate whether P is true at w, we don't have a straightforward way to pick out a single utility function over worlds. Relative to a sharpening, however, we can construct a utility function in this way. Thus the supervaluational machinery, together with the vaguely specified recipe, generates a set of utility functions, the desirability-analogue of uncertainty.

The 'options' over which we are uncertain I suggest should be identified with the supervaluational sharpenings. Given a set of outcomes, the sharpenings allow us to pick out a definite credence function and utility function, and so to calculate the expected utility of each action on that sharpening. Caprice tells us that faced with a decision problem, it is permissible to pick any such pair (i.e. any sharpening) and act as it recommends.

(An interesting feature of this implementation of uncertainty is that while we should be uncertain about typical borderline propositions, there are outliers where we can adopt a point-like attitude. Suppose 'hails' is indeterminate in reference between 'heads' and 'tails'. I am 0.5 confident that the fair coin will lands heads, and 0.5 confident that it will land tails. On either sharpening, therefore, the probability attached to 'the coin will land hails' is 0.5. So here there is no uncertainty, but a pointy credence, in the proposition in question. Thus the cognitive role of indeterminacy needs to be formulated somewhat delicately: it is not a straightforward non-pointy attitude, but rather the sharpening-relative generation of attitudes (leading to the possibility of non-pointy attitudes) that is characteristic.)

It is clear why that the package of uncertainty, Caprice, and sharpening-generation of uncertainty fits nicely with Alpha's actions being inconstant when faced with the Cabinet. After all, investing looks good on one sharpening (that on which Alpha is Omega, and hence gets the returns) and seems terrible on the other. Inconstant acts are catered for, since we may interpret them as cases where Alpha capriciously opts to act on the basis of one, or the other, sharpening.

But while the package is consistent with inconstancy, it's not so clear that our package predicts it. Caprice is committed only to the permissibility of a certain range of actions. That

is compatible with an agent just dogmatically sticking to one particular sharpening no matter what the context. But such an agent doesn't act inconstantly. Furthermore, a capricious agent might capriciously opt for the survival sharpening if the reward is over a given amount, and opt for the non-survival sharpening iff the reward is lower. So, it appears, a capricious agent is consistent with threshold pattern of behaviour under repetitions. (Such examples could be multiplied: perhaps a pessimistic agent would capriciously opt for that sharpening that gives the lowest expected utility; an optimistic one for the highest: in our case these correspond to Universal rejection and acceptance respectively). If we allow such patterns, then it seems there is the potential for a whole body of different systematic dispositions to capriciously choose. In some contexts, this might be welcome flexibility. But we wanted an account that commits to inconstancy, and so shuts off these highly patterned options.

For simplicity, I assume we are working with a finite set of sharpenings. A local rule that would impose true arbitrariness would be to say that in selecting among the sharpenings to act under, we should randomize across sharpenings (imagine throwing a dart at random into the sharpenings, and then acting in accordance to the impaled sharpening). Rather than the bare Caprice rule, therefore, we go for a rule that instructs the agent to take a mixed act. More generally, we endorse a second-level selection principle to select among the permissible options: Randomize (compare Levi's maximizing security-level rule, mentioned earlier). Given a binary decision problem, if k of the sharpenings recommend the first option, and $1-k$ the second, then one should adopt a mixed strategy of acting on the first with chance k and the second with chance $1-k$. So a given piece of behaviour will be describable in two ways—as being in accordance with Caprice as originally characterized, with the operative sharpening selected truly arbitrary; or alternatively as being the result of Randomize, generating a mixed act uniquely determined by the proportions of sharpenings supporting one act over another.[11]

Randomize can't be quite correct however. Well known diachronic puzzles for Caprice cause problems for Randomize as presently stated. However, I will argue that those same puzzles motivate a more sophisticated, diachronic version of Randomize.

3.3 Diachronic coherence

Here is the diachronic puzzle for Caprice (well, not for Caprice as carefully formulated by Weatherson, but in the simple form we've been working with).[12] Suppose the broker offers Alpha the investment opportunity at t. Alpha opts to accept, so is 100 dollars down, pro tem. Now at $t+1$ Alpha is offered another deal: the broker will pay Alpha 99 dollars, so long as Alpha agrees to alter the earlier contract so nobody gets any money. At $t+1$, this looks a good deal to Alpha exactly on the sharpening where he does not survive—so construed, it's money at no cost! On the sharpening where he does survive, Alpha is giving up thousands for a trivial gain now. Thus, it's again a case of decision-making under uncertainty, and Caprice allows Alpha to opt for the non-survival sharpening; hence it is permissible for Alpha to accept the later deal. Suppose Alpha takes the later offer. The net effect is that Alpha is down a dollar whatever happens, and so is determinately operating at a loss. So it looks like we have an uncomfortable triad: permissibility of investing at t, permissibility of counter-investing at $t+1$,

[11] In a more general setting where there are infinitely many sharpenings, they could not straightforwardly be read off the set. However, just as Lewis's account utilized a measure over the boundary numbers, we will suppose that our 'degree supervaluational' semantics comes equipped with a measure over the space of sharpenings. This privileged measure (conditionalized on the live sharpenings where appropriate) can be used to characterize the appropriate mixed acts.

[12] See (Elga, 2010) for the puzzle in the case of ordinary uncertainty, and for arguments against a range of possible reactions—including Weatherson's 'caprice' rule (which Elga calls 'Sequence').

but the intuitive impermissibility of the conjunction of investing and counterinvesting together, which leads to a guaranteed (and determinate) loss.

One good question is what exactly the puzzle consists in. The permissibility of two actions taken individually doesn't generally entail the permissibility of their conjunction. Consider dating norms: it's permissible for A to be in a relationship with B, and permissible for her to be in a relationship with C, but typically is not permissible for her to be in relationships with both simultaneously. However, the striking thing about the current case is that the related events are *successive*, and we evaluate the later one as permissible *even bearing in mind that the earlier one has already taken place*. Given the temporal ordering, there's something weird about denying the conjunction is permissible though each conjunct is. For example: in signing the second contract, does Alpha do something permissible (qua counterinvesting) or something impermissible (qua investing-and-then-counterinvesting)?

The case is intuitively awkward, but no more, I think. Carefully described, there is nothing contradictory going on. But what we do learn is that if agents are to avoid violating norms (including norms governing conjunctions) then later actions will have to be in sync with earlier ones. Even though it's permissible in a one-off case to opt for whatever sharpening one likes, in extended chains of action, one needs to ensure that the product of the individual acts is jointly permissible. Caprice needs supplementation, so that it characterizes not just single actions, but whole sets, as permissible when optimal on one probability within the representor. This is indeed what Weatherson's version requires.

Randomize appears to be in more trouble. If Alpha Randomizes on the first offer, he must give a 50/50 chance (say) to investing. If Alpha Randomizes on the second offer, he again gives that a 50/50 chance. But since the first act has already taken place, the well-run agent will invest-and-then-counterinvest in a quarter of cases. But that overall course of action is uncontroversially impermissible. Randomize therefore leads to impermissible courses of action in a way that Caprice alone (due to its lack of predictive power!) does not.

The correct reaction is not to abandon Randomize completely, but (like Weatherson) to switch focus from single actions to sets of actions, and to distinguish two readings of Randomize. To say, ahead of time, that the chances in the two successive acts must be 50/50 in each case *does not* mean the chances need be independent of one another. Indeed, it is compatible with the prior chance of each being 50/50 that the chance of counterinvesting given one has already invested, is zero. To get a model for this, imagine that an agent must choose the sharpening on which to act at random at the beginning of an extended series of action, and subsequently must stick with that sharpening throughout. On one reading of Randomize, the letter of its recommendation will have been satisfied.

The trouble is that the more natural reading—on which it gives the chances of acting on each sharpening at the time the decisions is made—seems to be the practically relevant one. Fictions of 'choosing a sharpening' at the beginning of time are all very well, but how could ordinary agents actually follow such a rule?

Let's consider the same issue of diachronic coordination in the simpler case of Caprice. We shall assume that in the context of action (either purely practical, or linguistic) there is a contextual 'score', a set that initially contains all the sharpenings. When an action is carried out that is permissible on some but not all sharpenings, the score updates by eliminating those on which is it is not permissible. (As a special case, when a linguistic act is true on some but not all precisifications, those on which it is not true are eliminated). We call an action *dynamically permissible* at time t just in case it maximizes utility on some sharpening live at the score at t. The practical implementation of Caprice, is that agents should strive to make their actions dynamically permissible (as well as permissible in the absolute sense). If successful, they will have ensured that the product of their actions during the period in which the score is evolving is

permissible in the absolute sense.[13]

What of Randomize? The analogous idea is that the mixed act actually implemented at t should accord with chances that evolve dynamically through the action period. Suppose that our decisions are first, between A and $\neg A$, and second, between B and $\neg B$. Suppose that AB is a sure loss outcome, and so must be avoided. The proportion of sharpenings recommending AB, $\neg AB$, $A\neg B$, $\neg A\neg B$ respectively are: 0, 0.2,0.3,0.5.

The first choice is between A and $\neg A$, and in accordance with our original Randomize rule, the respective weightings are 0.3 and 0.7 (so on repetitions of the choice procedure, the relative frequency of A choices to $\neg A$ choices will be 3:7). But once the action has been carried out, just as on the Caprice model the score is updated to eliminate all sharpenings not in accord with the action taken. Suppose we end up choosing $\neg A$. Then we face another choice, and since all and only the A sharpenings have been eliminated, the relative proportion of Bs to $\neg B$s will be 2/7 and 5/7. We then locally Randomize at the new odds. On the other hand, suppose we chose A. Then all remaining sharpenings are $\neg B$, so applying the mixed strategy recipe here we go for $\neg B$ with chance 1.

This dynamic and local version of the Randomize rule accords perfectly with the original global Randomize rule, on the reading where it is applied once and for all before any act is taken. There's a 0.3 chance of getting A, and that then ensures we get $\neg B$, so the chance of ending up with $A\neg B$ is 0.3, as it should be. There's a 0.7 chance of getting $\neg A$, and given this a 2/7 chance of B, for a 0.2 chance of $\neg AB$ overall; and a 5/7 chance of B following $\neg A$, for a 0.5 chance overall. The overall chance of getting $\neg B$ through some route or other is 0.5, and likewise for B. So following the instructions for mixtures, but with respect to the live possibilities rather than all sharpenings whatsoever, we find a practical way of implementing the mixed strategies at their original ratios.[14] This is no coincidence. Effectively, given initial chances Ch, the chance of choosing A at the first choice point is $Ch(A)$, and the chance of choosing $\neg A$ is $Ch(\neg A)$. If the first is realized, the chance of choosing B at time 2 is $Ch(B|A)$. If the second is realized, the chance of choosing B at time 2 is $Ch(B|\neg A)$. The overall chance of B is therefore $Ch(A)Ch(B|A) + Ch(\neg A)Ch(B|\neg A)$, which by the law of total probability, is simply $Ch(B)$. So the diachronic version of Randomize is a way of implementing the original atemporal Randomize strategy.

3.4 Application

I've just been arguing that for practical purposes, at least, Randomize needs to be construed as invoking chances for judging this way or that, evolving over time. This is a prediction of the account: how well does it fit data? We can test this by thinking again about the special case of linguistic action.

[13]Compare Elga's discussion of the rule 'narrow'. The narrowing proposal says that one's doxastic change should update upon acting, to eliminate from the representor those probability functions that would not recommend the action one in fact takes. Elga complains that this is a case where one's doxastic state updates without relevant evidence being gained. In the current setting, I think of the doxastic state as being constant throughout the process, and a separate 'scoreboard' being updated—the sharpenings and their induced credal functions are still around in the doxastic state, but some are not 'live' in the sense that they have been eliminated from the scoreboard. It is the scoreboard, not the doxastic state, that updates and fixes what is dynamically permissible. But dynamic permissibility is only interesting because it's a way of practically ensuring that one's overall course of action is (atemporally) permissible by the the lights of the original Caprice or Randomize rule—and permissibility in this sense is fixed by the doxastic state alone. I don't see any discomfort in non-evidential updating of the scoreboard in this context. Thanks to Seamus Bradley for pressing me on these points.

[14]In the setting with infinitely many sharpenings, and consequently a measure over them, we simply need to conditionalize the measure to remove the sharpenings that have been eliminated, and then used this derived measure over the live score to set the 'local' chances for the mixed acts.

If we have just judged that a colour patch is red, then we are not likely the next second to judge it not-red. This local stability of judgements is predicted by the diachronic versions of Caprice and Randomize. Likewise, if we've judged borderline patch 1 to be red, and then are presented by borderline patch 2, which (though still borderline) is redder than it, we will feel constrained by our previous judgement to judge patch 2 to be red (even though there are, in the absolute sense, sharpenings on which patch 2 is not red—which we'd have been likely to opt for if we'd come to it cold). These all seem desirable features in an account of vagueness. Indeed, the account is strongly reminiscent of the scorekeeping account of context-sensitivity of vague terms argued for on independent grounds by Stewart Shapiro (2006).

Another interesting case to consider is that of a forced march sorites. Suppose that 100 borderline cases of redness are indexed from 1 to 99, from almost-determinately-red to almost-determinately-not-red. There are 100 ways of drawing the cut-off, and we treat these as the initial set of sharpenings. One is then asked, of patch 0, whether it is red. On 99/100 of these cutoffs, it counts as red, so almost certainly one will judge it red. The score is updated to remove the one cutoff inconsistent with this verdict. The forced marcher then moves to patch 2. On 98/99 of the remaining cutoffs, it counts as red. So chances are it'll be judged accordingly. The successive choices then give odds of finding it red 97/98, 96/97, and so on... in fact, even when one has gone almost completely through the sorites, and classified 90 of the patches as red, the chances of classify the (distinctly yellow-ish) 91st patch as red will still have a 9/10 chance of counting as red. The odds don't get very low till the last few patches: the 97th patch has a 3/4 chance of counting as red, the 98th patch 2/3; and the 99th 1/2. But of course, if the forced march had been run the other way, we could get into a situation where there's a 3/4's chance that patch 3 would be judged as not-red. So the account accounts for forced-march sorites in an extremely strong fashion (if anything, perhaps the predicted success of the forced march is too strong! But it'd be very natural, for example, if people got off the boat somewhat early—disconcerted by the vast discrepancy between the prior chance of judging a yellowish borderline patch red (low) and the dynamically updated chance of judging it red (high).)

Finally, let's consider the sorites as an argument. This can be set up in various ways. A 'long' form focuses on individual judgements of the form: $Fa \wedge \neg Fa'$, where a and a' are relevantly adjacent. Judging these false (or their negations true) throughout a sorites series allows us to derive a paradox, if neither F nor $\neg F$ is vacuous. On the current account, we predict that an arbitrary 'abominable conjunction' of this form will be judged false (and its negation true) almost all the time—it is only true on a vanishingly small percentage of the sharpenings (notice the essential appeal to Randomization, as opposed to Caprice, here: Caprice will only tell you it's permissible to judge this true, and permissible to judge it false; Randomization tells you that subjects overwhelmingly judge it false). Of course, on this theory one cannot coherently judge all premises of a sorites argument simultaneously false—if we consider them successively, and in order, we will in effect be back to the situation of a forced-march sorites. Further, the account predicts we will endorse as true the existential that some such cutoff exists (this is hard-wired into the semi-classical, supervaluational backdrop). Like contextualist theories of vagueness, we have the strong prediction that representative instances of the sorites premise in long form are judged certainly false (rather than simply being assigned very low credence)—and this is a promising explanatory starting point for accounting for the seductiveness of the sorites in argument-form.

Conclusion

The first third of the essay motivated the search for an account of action under indeterminacy by looking at an investment opportunity posed to van Inwagen's Alpha. In the second third I looked at Lewis's degree-like treatment of indeterminacy, and the model of decision making under indeterminacy (and the accompanying cognitive role) that this induced). In the final third, I drew on tools from the literature on uncertainty and imprecise credences ('representors') to develop a rationale for inconstant decision-making. I've highlighted some concerns with the what I argued was the Lewisian treatment, to do with extending the story from the practical to the affective. I've also highlighted some striking and attractive predictions of the 'uncertainty'-based account.

It's worth noting that the underlying semantic machinery required to run the story in the two cases is exactly the same—degree-supervaluationism. That we can develop similar machinery in two such radically different directions illustrates how much is left unsaid, if all we are given is the bare logico-semantic treatment of indeterminacy.

References

Broome, J. 1994. Discounting the future. *Philosophy and public affairs*, **23**(2), 128–156.

De Finetti, Bruno. 1974. *Theory of probability: vol.1*. New York: Wiley.

Elga, A. 2010. Subjective Probabilities Should be Sharp. *Philosophers imprint*, **10**(5).

Field, Hartry H. 2003. 'No fact of the matter'. *Australasian journal of philosophy*, **81**, 457–480.

Frederick, Shane, Loewenstein, George, & O'Donoghue, Ted. 2002. Time discounting and time preference: A critical review. *Journal of economic literature*, **40**(2), 351–401.

Jeffrey, Richard. 1965. *The logic of decision*. 2nd edn. Chicago and London: University of Chicago Press. Second edition published 1983.

Joyce, James M. 1998. 'A non-pragmatic vindication of probabilism'. *Philosophy of science*, **65**, 575–603.

Joyce, James M. 2009. 'Accuracy and coherence: prospects for an alethic epistemology of partial belief'. *Pages 263–297 of:* Huber, Franz, & Schmidt-Petri, Christoph (eds), *Degrees of belief*. Springer.

Joyce, James M. 2011. A Defense of Imprecise Credences in Inference and Decision Making. *In:* Gendler, Tamar Szabo (ed), *Oxford studies in epistemology*. Oxford University Press.

Levi, I. 1974. On indeterminate probabilities. *The journal of philosophy*, 391–418.

Lewis, David K. 1976. 'Survival and identity'. *In:* Rorty, Amélie O. (ed), *The identities of persons*. University of California Press. Reprinted in Lewis, *Philosophical Papers I* (Oxford University Press, 1983) 55–72.

Mas-Colell, A., Whinston, M.D., & Green, J.R. 1995. *Microeconomic theory*. Oxford university press.

Parfit, Derek. 1984. *Reasons and persons*. Oxford: OUP.

Paris, J.B. 2001. A note on the Dutch Book method. *Pages 301–306 of: Proceedings of the second international symposium on imprecise probabilities and their applications, isipta*. Ithaca, NY: Shaker.

Schiffer, Stephen R. 2003. *The things we mean*. Oxford: Oxford University Press.

Shapiro, S. 2006. *Vagueness in context*. Oxford University Press, USA.

Sider, Ted. 2001. 'Personal identity and the limits of conceptual analysis'. *Philosophical perspectives*, **15**.

Smith, Nick J. J. 2008. *Vagueness and degrees of truth*. Oxford: Oxford University Press.

van Inwagen, Peter. 1990. *Material beings*. Ithaca and London: Cornell University Press.

Weatherson, Brian. Manuscript. Decision Making with Imprecise Probabilities.

Williams, Bernard. 1970. 'The self and the future'. *Philosophical review*, **79**(2), 161–180. Reprinted in his *Problems of the Self*, CUP, 1973.

Williams, J. Robert G. 2010. 'Degree supervaluational logic'. *Review of symbolic logic*.

Williams, J. Robert G. forthcoming. 'Generalized probabilism'. *Journal of philosophical logic*.

Williamson, Timothy. 1994. *Vagueness*. London: Routledge.

Wright, Crispin. 1976. 'Language mastery and the sorites paradox'. *Pages 223–47 of:* Evans, Gareth, & McDowell, John (eds), *Truth and meaning*. Reprinted Keefe and Smith (eds) *Vagueness: A reader* (MIT Press, Cambridge MA: 1997) pp.151-173.

Wright, Crispin. 2001. 'On being in a quandry'. *Mind*, **110**.

Wright, Crispin. 2003. 'Vagueness: A fifth column approach'. *Pages 84–105 of:* Beall, J. C. (ed), *Liars and heaps: New essays on paradox*. Oxford: Oxford University Press.