Belief and Indeterminacy

Michael Caie
University of Rochester

1. Introduction

What attitude should a rational agent have toward a proposition expressed by a paradoxical sentence such as the liar? In this paper, I argue that consideration of a paradox concerning doxastic rationality motivates a surprising answer to this question.

The question that I'm interested in is not the same as the question about what semantic status(es) we should assign to such paradoxical sentences. But the two questions are intimately related. Sometimes, given an account answering the latter question, the answer to the former is obvious.[1] But this is not the case for every account of semantic paradox. For the type of account that I'll be interested in here, the answer to our question is far from obvious.

This account traces back to Kripke 1975.[2] According to the account, ϕ and $T\ulcorner\phi\urcorner$ are intersubstitutable (at least within extensional

For helpful discussion of this material, thanks to Justin Bledin, Colin Caret, Fabrizio Cariani, Cian Dorr, Jessica Gelber, Niko Kolodny, Geoff Lee, John MacFarlane, Mike Martin, Dave Ripley, Sherri Roush, Robbie Williams, Seth Yalcin, the referees for the *Philosophical Review*, and audiences at the 2011 Formal Epistemology Workshop (FEW), University of California, Berkeley, the University of Rochester, and the University of Toronto. Thanks finally to Hartry Field, whose comments on an earlier draft of this material at FEW improved this article in numerous ways.

1. For example, some theories hold that the liar sentence is not true. But this, of course, is just what the liar sentence says. So the proponent of such a theory will hold that one should believe the proposition expressed by the liar sentence.

2. Kripke's paper has given rise to a number of theories. The relevant theory for our purposes is called *KFS*. See also Soames 1999; Field 2008; Richard 2008; Yablo 2003.

Philosophical Review, Vol. 121, No. 1, 2012
DOI 10.1215/00318108-1426355

contexts).[3] In order to achieve this desirable goal, however, certain elements of classical logic must be abandoned. According to this theory, the logic governing the "Boolean" connectives and quantifiers is K_3, the logic induced by the Strong Kleene valuation scheme.[4] Notably, in this logic excluded-middle fails to be unrestrictedly valid. Such theories are labeled *paracomplete*.

This type of account has been extended by Hartry Field. Field has provided a model-theory that, like Kripke's, allows for the intersubstitutability of ϕ and $T\ulcorner\phi\urcorner$ by allowing for the failure of excluded-middle.[5] The models developed by Field, however, are defined for a language containing a reasonable conditional, $\rightarrow$, for which the T-schema holds, and an indeterminacy operator, I. Using the latter we can characterize the status of those paradoxical sentences that lead to failures of excluded-middle. While we cannot, without committing ourselves to a contradiction, say either that the liar sentence λ or its negation are not true, we can say that neither λ nor its negation are *determinately* true; that is, we can say that it is *indeterminate* whether λ is true.

The arguments that follow are directed at anyone who advocates a paracomplete treatment of the semantic paradoxes. In what follows, I will, however, help myself to the expressive resources that Field's theory offers. Because Field's theory contains an object language operator I, which can be used to express the paradoxical status of the liar sentence, working with this theory will allow us to reason more easily about this status and our attitudes toward sentences having this status. In principle, however, the main arguments that follow could be reframed in terms acceptable to advocates of alternative paracomplete accounts.[6]

3. I'm going to be rather loose about use and mention with the quasi-formal language that I'll be employing in this article. Sometimes I'll *use* expressions involving symbols such as 'T'. Other times I'll be talking about such expressions. In this case, I'll (typically) omit quotes to ease readability.

4. See the technical appendix for the details of this valuation.

5. See Field 2008, 2007 for a development of this position.

6. I should also note that the arguments given here could be applied, mutatis mutandis, to other nonclassical frameworks. The claim to be defended is that the rational response to the failures of classical logic that a paracomplete theorist claims arise in cases of semantic paradox is for one to have a mirroring nonclassicality in one's doxastic state. The arguments that follow can be easily recast to provide arguments for analogous claims within a paraconsistent supervaluationist or revision-theoretic framework. I leave detailed treatment of these alternative nonclassical accounts for another occasion, but the interested reader shouldn't find it too difficult to see how these variant cases work.

Letting ϕ be some proposition that one ought to believe is indeterminate, the question then arises for one who is attracted to this sort of account of semantic paradox: what attitude should one have toward ϕ? Call this the *Normative Question*.

The orthodox answer to this question is:

> **REJECTION** For any proposition ϕ, it is a consequence of the claim that one ought to believe that ϕ is indeterminate, that one ought to reject both ϕ and its negation.[7]

To say that this answer is orthodox is in some ways to undersell how wide is the agreement on this point. To my knowledge, *all* prominent defenders of a paracomplete theory have either explicitly or implicitly endorsed the view that rejection is the correct attitude to take toward the proposition expressed by the liar sentence.[8]

In what follows, I'll argue, instead, that the correct answer to the Normative Question is:

> **INDETERMINACY** For any proposition ϕ, it is a consequence of the claim that one ought to believe that ϕ is indeterminate that one ought to be such that it is indeterminate whether one believes ϕ.[9]

For rational agents, indeterminacy in the objects of doxastic states will filter up to the doxastic states themselves. That this is so is certainly not obvious. Nonetheless, I'll try to show that there are good reasons to think it is true.

The argument for this will proceed as follows.

7. Read: $OBI\phi \vDash OR\phi \wedge OR\neg\phi$. Rejection of a proposition ϕ can be thought of as having an appropriately low credence in ϕ. Note that for the proponent of REJECTION, rejection of ϕ does not require that one have a high credence in $\neg\phi$.

8. See Field 2008, 2003a for this sort of view. Parsons 1984 is the first explicit endorsement that I know of of the rejectionist line. The view can, however, be seen as being implicit in parts of Kripke 1975. See also Soames 1999 and Richard 2008. I should note that despite the popularity of the rejectionist line, some murmurs of dissent have started to emerge. In particular, Robbie Williams has recently provided arguments against the claim that rejection is the correct cognitive response to certain failures of classical logic that one might think emerge in domains besides that of semantic paradox. See, for example, Williams 2011, 2008. Williams, however, doesn't discuss cases of semantic pathology, and the positive view that he sketches is rather different from that which I'll endorse here.

9. Read: $OBI\phi \vDash OIB\phi$.

In section 2.1, I develop a normative paradox. I show that three plausible principles concerning doxastic rationality are classically inconsistent.

In section 2.2, I show how the prima facie inconsistent triad can be whittled down to a prima facie inconsistent pair.

In section 3, I show that the structure of the paradox developed in section 2.1 is formally identical to a modal liar paradox. A paracomplete solution to the liar paradox can be straightforwardly extended to deal with the modal liar sentence. Given the structural identity of these two paradoxes, the normative paradox can also be resolved by appeal to paracomplete resources. If, as paracomplete theorists, we allow that an agent's doxastic state may be a locus of certain failures of excluded-middle, then we can hold on to our plausible normative principles. The same solution applies to the paradox developed in section 2.2.

The possibility of doxastic states giving rise to the appropriate failures of classical logic is defended in section 4.

In section 5, I argue that resolving the normative paradoxes in this way requires rejecting REJECTION. Acceptance of this norm leads us back into normative paradox. As paracomplete theorists, we can hold on to some very plausible principles concerning doxastic rationality. But only if we give up REJECTION. Whatever confidence we have in these principles should make us correspondingly less confident in REJECTION. Since these principles are much more intuitively compelling than REJECTION, this provides a strong argument against REJECTION.

In section 6, I show how the argument in section 5 generalizes to provide an argument against other potentially attractive answers to the Normative Question.

In section 7, I argue that INDETERMINACY is independently motivated and that, unlike REJECTION, it is compatible with the proposed resolution of the normative paradox.

In section 8, I provide further considerations in favor of INDETERMINACY, and in section 9 I assess the state of the dialectic between the proponent of REJECTION and the proponent of INDETERMINACY, given the preceding arguments.

Finally, in section 10, I consider three prima facie compelling arguments in favor of REJECTION and show how they can be resisted.[10]

10. Included is also an appendix that outlines the underlying technical machinery and proves various results that I appeal to in the main body of the article. Readers who wish

2. Normative Paradoxes

2.1.

Consider the following sentence:

> I do not believe that this sentence is true.[11]

What should your attitude be toward this sentence? You're likely to be puzzled. You know the following facts:

- If you believe that it's true, then it's false.
- If you don't believe that it's true, then it's true.

Assuming that you'll know whether or not you believe that it's true, you'll then either be in the position of knowing that you believe that the sentence is true, and knowing that your so believing makes it false, or knowing that you fail to believe that the sentence is true, and knowing that your so failing to believe makes it true. Neither of these seems like a rational state for an agent to be in.

Our puzzlement at this case can be sharpened into a paradox. In a moment, I'll show how this works in precise detail, but let me first give you a sense of the form that the paradox takes.

Using the type of sentence above, we can argue that there is a possible agent who, without being guilty of any antecedent rational failing, is unable to satisfy the following two plausible normative principles:

> **CONSISTENCY** For any proposition ϕ, it is a rational requirement that if one believes ϕ, then one does not believe its negation $\neg\phi$.[12]

> **EVIDENCE** For any proposition ϕ, if an agent's evidence makes ϕ certain, then the agent is rationally required to believe ϕ.[13]

---

to skip this will not, however, miss anything essential to understanding the main arguments of the paper.

11. This type of sentence and some of its odd features are discussed in Burge 1978, 1984; Conee 1987; and Sorensen 1988.

12. Read: $O(B\phi \rightarrow \neg B\neg\phi)$.

13. Two points. (i) Note that EVIDENCE is a *synchronic* norm. If, at a particular time t, an agent has evidence that makes ϕ certain and fails to believe it, then the agent is thereby subject to rational criticism. (ii) I take it that there are weaker levels of evidential support that also rationally mandate belief. It is, however, neater to work with this (logically) weaker rational constraint. But note that if one is uncomfortable with the idea that

If our agent believes that the above sentence is true, then it will either fail to satisfy CONSISTENCY or EVIDENCE. And, if our agent doesn't believe that the above sentence is true, then it will fail to satisfy EVIDENCE. In either case, our agent will be guilty of a rational failure. What this would seem to show is that CONSISTENCY and EVIDENCE are incompatible with the following general constraint on principles of rationality:

> **POSSIBILITY** It must always be possible for an antecedently rational agent to continue to meet the requirements imposed on it by rationality.

Later in this article, I will show how this paradox can be resolved and what lessons we can extract from its resolution. But first let us see how this paradox works in detail.

Let B_α be an operator meaning *Alpha believes that*. Let β name the following sentence: $\neg B_\alpha T(\beta)$.[14] Then, as an instance of the T-schema, we have:

(1) $T(\beta) \leftrightarrow \neg B_\alpha T(\beta)$[15]

Actual agents are good at detecting their own doxastic states. This works in two directions. First, when one believes something, one often believes that one believes it; similarly when one does not believe something, one often believes that one does not believe it. Second, when one believes that one believes something, for the most part one is right; similarly when one believes that one does not believe something. Still, actual agents are fallible in both directions. There are plenty of beliefs of mine of which I am unaware and that would remain hidden to me even after a thorough introspective search, and the same is, I take it, true of you. There are also beliefs that I have about my own belief state that are false. While it may have once been common to suppose that each agent's mind is transparent to herself, this thought now seems indefensible.

---

one's evidence ever makes anything certain, the following puzzle could be recreated by appeal to a plausible normative constraint to the effect that there is some less than conclusive evidential threshold beyond which belief is rationally mandated.

14. Here, following Kripke 1975, sentential self-reference is achieved by stipulation. This could also, of course, be achieved by a technique such as Gödel numbering.

15. Note that $\leftrightarrow$ is the Field biconditional. All of the inferences involving the conditional that I'll be appealing to are valid in the class of Field models. See the technical appendix for an outline of this model theory and Field 2008 for a more in-depth treatment.

Many of our limitations in this respect would, however, seem to be medical in nature, not metaphysical. Consider an ideal agent. Call it *Agent Alpha*. The following seems, at the very least, metaphysically possible. Whenever Alpha believes that β is true, then Alpha also believes that it believes β. And whenever Alpha does not believe that β is true, then Alpha believes that it does not believe β. Moreover, Alpha is, overall, *perfectly* reliable in the higher-order beliefs that it has with respect to whether or not it believes that β is true. Alpha believes that it believes that β is true only if it does believe that β is true, and it believes that it does not believe that β is true only if it does not believe that β is true.

More perspicuously, then, we have the following:

(2) $B_\alpha T(\beta) \leftrightarrow B_\alpha B_\alpha T(\beta)$
(3) $\neg B_\alpha T(\beta) \leftrightarrow B_\alpha \neg B_\alpha T(\beta)$[16]

We may further suppose that our agent believes the truth expressed in (1). We have then:

(4) $B_\alpha(B_\alpha T(\beta) \rightarrow \neg T(\beta))$
(5) $B_\alpha(\neg B_\alpha T(\beta) \rightarrow T(\beta))$

The possibility of an agent, such as Alpha, who satisfies (2)–(5) raises problems for the conjunction of CONSISTENCY, EVIDENCE, and POSSIBILITY. To see this, first consider the following two cases.

Case 1: On the assumption that Alpha does not believe that β is true, it follows that Alpha ought to believe that β is true.

Assume that Alpha does not believe that β is true. By (3), it follows that it believes that it does not believe this. Alpha also believes that if it does not believe this, then β is true. This is (5). The setup of the case is such that the status of both of these beliefs is superlative. The first belief is *perfectly* reliable. The second proposition that it believes is a theorem, and we can assume that its grounds for believing this are the same as ours. Given their bona fides, these beliefs, I claim, form part of the agent's total body of evidence.[17] We can assume, further, that Alpha has no

16. We should think of B_α as involving a first-personal mode of presentation for Alpha.

17. Exactly what sort of relation one must bear to a proposition in order for the latter to be part of one's evidence is a topic of some controversy. The case, however, is set up so that Alpha should meet any reasonable standards. Alpha, for example, knows that it does not believe that β is true, and that if it does not believe that β is true, then β is true. Our assumption, then, is justified if one holds that a proposition ϕ counts as part of an agent's evidence just in case the agent knows that ϕ. See, for example, Williamson 2000. In

other evidence that bears one way or the other on the question of whether β is true. Alpha, then, would seem to be in the position in which its evidence makes it certain that β is true. By EVIDENCE, it follows that Alpha ought to believe that β is true.

Case 2: On the assumption that Alpha does believe that β is true, it follows that Alpha ought not to believe that β is true.

Assume that Alpha does believe that β is true. By (2), it follows that it believes that it believes that β is true. Alpha also believes that if it believes that β is true, then β is not true. This is (4). Again the evidential status of these beliefs is, by the setup of the case, superlative. We can again assume that Alpha has no other evidence that bears on whether or not β is true. Alpha, then, is such that its evidence makes it certain that β is not true. By EVIDENCE, it follows that Alpha ought to believe that β is not true. That is, we have $OB_\alpha \neg T(\beta)$. As an instance of CONSISTENCY, we have: $O(B_\alpha \neg T(\beta) \rightarrow \neg B_\alpha T(\beta))$. I assume that rational obligations are such that if a proposition γ is a consequence of a set of propositions Γ and the members of Γ are all rationally obligatory, then so is γ.[18] Given this, $OB_\alpha \neg T(\beta)$ follows from the previous two claims.

Cases 1 and 2 show that CONSISTENCY and EVIDENCE are *classically* inconsistent with POSSIBILITY. If Alpha does not believe that β is true, then, according to case 1, it ought to believe that β is true. While if Alpha does believe that β is true, then, according to case 2, it ought not to believe that β is true. As a theorem of classical logic, we have $B_\alpha T(\beta) \vee \neg B_\alpha T(\beta)$. What cases 1 and 2 show is that if CONSISTENCY and EVIDENCE hold, then whichever disjunct is realized, Alpha will be guilty of a rational failure. In neither case, however, will Alpha be guilty of

addition, the agent's knowledge in both cases need not be inferential. Our assumption is justified if one thinks that a proposition ϕ counts as part of an agent's evidence just in case the agent knows that ϕ, and ϕ is not inferred from other known premises. And, of course, our assumption is justified a fortiori if one thinks that a proposition ϕ that one believes counts as evidence just in case one satisfies some less-demanding criteria, for example, having a justified belief in ϕ. See, for example, Feldman 2004.

18. This sort of multipremise closure principle is not completely uncontroversial. In particular, those who think that there are rational dilemmas, that is, cases in which $O\phi$ and $O \neg \phi$, will want to reject such a closure principle since rational dilemmas together with this closure principle lead to deontic trivialization. For this type of worry see, for example, Van Fraassen 1973. Let me note, then, that the cases in which I will be appealing to multipremise closure for rational obligations are all cases in which such trivialization is avoided. So if one is inclined to be suspect of such a closure principle due to rational dilemmas, there are restricted closure principles that would avoid such worries and suffice for my purposes.

any initial rational failing. All that we require is that Alpha have knowledge of a theorem and that Alpha be sensitive to its own doxastic states. By proof-by-cases reasoning, then, we can establish that if CONSISTENCY and EVIDENCE hold, it isn't possible for Alpha, an antecedently rational agent, to meet all of the requirements imposed by rationality.[19] But this is exactly what POSSIBILITY denies.

CONSISTENCY and EVIDENCE are, at least prima facie, fairly central principles of epistemic rationality. Perhaps, then, one might think that the lesson to draw is that we should reject POSSIBILITY. We should not, however, underestimate the intuitive costs of this response. It is, I think, initially quite implausible that an agent could do everything that rationality requires and yet nonetheless wind up in a situation in which it cannot continue to meet the requirements of rationality.[20]

This intuition can be bolstered by considering the sorts of conditions under which rational criticism seems to be appropriate. An agent may be subject to rational criticism given the set of doxastic options that it has realized. Let Γ be this set. The following seems to me to be a plausible constraint on the conditions under which such criticism is appropriate:

19. Note that I am assuming that, on the intended reading of POSSIBILITY, the modality is restricted to situations in which we hold fixed the facts about the agent's actual situation that are relevant to the rationality of particular options were they to be realized by the agent. In the case we are concerned with, in applying POSSIBILITY, we must hold fixed the facts about Alpha that are relevant to its evidential situation. But these include all the facts that were appealed to in establishing cases 1 and 2, namely, that (2), (3), (4), and (5) all hold and that the relevant beliefs were arrived at in a particular manner. We can thus take cases 1 and 2 for granted in applying POSSIBILITY.

20. I should note that not everyone is inclined to accept this. There are those who think that moral and rational dilemmas—cases in which you are damned if you do and damned if you don't—may arise even if an agent is not already subject to such criticism. See, for example, Lemmon 1962 and Marcus 1980 for the existence of moral dilemmas. For the existence of rational dilemmas, see the case of Death in Damascus in Gibbard and Harper 1978 and Priest 2002. For resistance to the idea of rational and moral dilemmas, see Conee 1982 and Arntzenius 2008. I do not have the space here to deal with this substantial literature, but it suffices to say that my sympathies are with those who want to reject the possibility of such cases. As a minimal point, let me note that all parties to the dispute should, I think, agree that it would be ideal if we could find a way for the putative normative principles to coexist without conflict. In game-theory, a standard move to deal with putative cases of rational dilemmas is to expand the space of possible options and allow for so-called "mixed-decisions." See, for example, Osborne and Rubenstein 1994. To foreshadow somewhat, my plan is to offer a similar strategy for the doxastic case. It is, I hope, enough to motivate this response that one sees the pull of POSSIBILITY.

APPROPRIATENESS If an agent is to be subject to rational criticism for realizing Γ, then there is some set of sets of options Δ meeting the following conditions:

(i) Each member of Δ is incompatible with Γ.
(ii) The agent should have realized some member of Δ (although there need not be any particular member that it should have realized).
(iii) Each member of Δ is such that had the agent realized this set of doxastic options, rational criticism would have been inappropriate.

Justification: If an agent is subject to rational criticism given the total set of doxastic options Γ that it has realized, then it should not have realized Γ. In such a case the agent ought to have realized some other set of doxastic options (although there need not be a specific set of options that the agent ought to have realized). That is, there will be a set Δ of sets of options incompatible with Γ such that the agent ought to have realized one of the members of Δ. (Indeed, there will typically be many such sets.) That the agent ought to have realized one of the members of Δ can serve as the grounds for rational criticism for the agent's having instead realized Γ. If, however, failure to realize some member of Δ is to serve as an adequate ground for rational criticism, it must, I think, be the case that the agent's realizing some member of Δ would have made rational criticism inappropriate. Rational criticism for an agent's doxastic situation is grounded in the idea that the agent should be some other way that would have made such criticism inappropriate. It is this plausible intuition that APPROPRIATENESS captures.

It can be shown that if POSSIBILITY fails, then so must APPROPRIATENESS. If one wants to maintain what is, I think, a natural principle about the conditions for appropriate rational criticism, then one should endorse POSSIBILITY.[21]

21. Here's the argument for this claim: Let Γ pick out the set of doxastic options that a rationally blameless agent has realized. Assume that POSSIBILITY fails. This means that there must be some jointly exhaustive set of options Σ such that for every $\sigma' \in \Sigma$, the agent is rationally culpable if it realizes $\Gamma \cup \sigma'$. We pick some arbitrary member σ of Σ. Let Δ be an arbitrary set of sets of options incompatible with $\Gamma \cup \sigma$, such that the agent should realize one of these sets. I'll argue that there are members of Δ such that were an agent to realize that option, then it would be rationally culpable. This shows that a violation of POSSIBILITY leads to a violation of APPROPRIATENESS.

Among the members of Δ must be some set containing Γ since if one ought to realize

Rejecting CONSISTENCY, EVIDENCE, or POSSIBILITY, then, each brings with it significant intuitive costs. And yet the case of Agent Alpha would seem to show that we cannot accept each of these plausible normative principles. We are faced with a normative paradox.

2.2.

In this section, I provide a different case that whittles the apparent inconsistency down to the pair CONSISTENCY and POSSIBILITY. Now I'm not inclined to reject EVIDENCE, but I can imagine the following response seeming attractive. Although, prima facie, it would seem that evidence of a certain strength rationally mandates belief, what the above paradox shows is that this is not always the case. For what the above paradox shows is that in certain cases one can have evidence that makes it certain that a particular proposition is true, but in such a case one's having that evidence essentially depends on one's *not* responding to the evidence by believing the proposition in question. At least in such cases, according to this line of thought, evidence does not rationally mandate belief. What the following case shows is that rejecting EVIDENCE will not get us out of trouble.

Belief, it is common to assume, is a relation that holds between an agent and an abstract object, a proposition. Assuming this picture of belief, we can show that CONSISTENCY and POSSIBILITY are classically inconsistent. Consider the following propositional analogue of β:

(∗) Alpha doesn't believe the proposition expressed by (∗).[22]

Let's abbreviate *the proposition expressed by* as ρ. The above can, then, be represented as:

(∗) $\neg B_{\alpha}\rho(*)$

---

some set among a collection of sets all of which are incompatible with Γ, then one would, contrary to hypothesis, be rationally blameworthy in realizing Γ. However, since Γ is not itself incompatible with $\Gamma \cup \sigma$, any set in Δ containing Γ must also contain some other doxastic option(s) Γ', in addition to those options in Γ. By hypothesis, $\Gamma \cup \Gamma'$ is incompatible with $\Gamma \cup \sigma$, that is, $\Gamma \cup \Gamma' \cup \sigma \vDash \bot$. It follows that $\Gamma \cup \Gamma' \vDash \neg\sigma$. (Note that $\{\sigma, \neg\sigma\}$ are, by the setup of the case, exhaustive, and so this particular use of reductio is fine by both classical and paracomplete lights.) But given this, our agent will be rationally culpable in realizing $\Gamma \cup \Gamma'$ since this will involve realizing $\Gamma \cup \neg\sigma$, and so realizing Γ together with some member of $\Sigma - \sigma$.

22. Again we should think of *Alpha* as having a first-personal mode of presentation for Alpha.

Note that since both (*) and '$\neg B_\alpha \rho(*)$' name the same sentence, the following holds:

(r1) $\rho(*) = \rho$'$\neg B_\alpha \rho(*)$'

Our transparency assumptions can be captured by the following analogues of (2) and (3):

(6) $B_\alpha \rho(*) \leftrightarrow B_\alpha \rho$'$B_\alpha \rho(*)$'
(7) $\neg B_\alpha \rho(*) \leftrightarrow B_\alpha \rho$'$\neg B_\alpha \rho(*)$'

It can now easily be shown that the assumption that Alpha does not believe the proposition expressed by (*) leads to a contradiction.

(1f)	$\neg B_\alpha \rho(*)$	Assumption
(2f)	$B_\alpha \rho$'$\neg B_\alpha \rho(*)$'	(1f) and (7)
(3f)	$B_\alpha \rho(*)$	(2f), (r1), substitution of equivalents
	$\bot$	

It follows that Alpha cannot fail to believe the proposition expressed by (*). However, when Alpha believes this proposition, given (6), Alpha is doomed to inconsistency. Thus:

(1g)	$B_\alpha \rho(*)$	Assumption
(2g)	$B_\alpha \rho$'$B_\alpha \rho(*)$'	(1g) and (6)
(3g)	$B_\alpha \rho$' $\neg B_\alpha \rho(*)$'	(1g), (r1), substitution of equivalents

Given CONSISTENCY, we will once again have a violation of POSSIBILITY. Holding fixed (6) and (7), it follows that Alpha's only option is to believe the proposition expressed by (*). But in doing so, Alpha will be in violation of CONSISTENCY. It follows that it is not possible for Alpha to meet the rational requirements imposed by CONSISTENCY. Since Alpha need not be guilty of any antecedent rational failing, this is a violation of POSSIBILITY.

Let me say a little about how this case is related to our earlier case. The key difference is the replacement of $T(\beta) \leftrightarrow \neg B_\alpha T(\beta)$ by $\rho(*) = \rho$' $\neg B_\alpha \rho(*)$'. Changing a conditional linking the truth-values of propositions to an identity between propositions has the same effect as assuming conformity to EVIDENCE. If we assume that Alpha meets EVIDENCE, we can provide parallel derivations to (1f) – (3f) and (1g) – (3g) involving the sentence β.

Corresponding to (1f) – (3f) we have:

(1h)	$\neg B_\alpha T(\beta)$	Assumption

(2h)	$B_\alpha \neg B_\alpha T(\beta)$	(1h) and (3)
(3h)	$B_\alpha T(\beta)$	(2h), (5), EVIDENCE
$\bot$		

Corresponding to (1g)–(3g) we have:

(1i)	$B_\alpha T(\beta)$	Assumption
(2i)	$B_\alpha B_\alpha T(\beta)$	(1i) and (2)
(3i)	$B_\alpha \neg T(\beta)$	(2i), (4), EVIDENCE

Where in (1f)–(3f) appeal is made to the propositional identity $\rho(*) = \rho$‘ $\neg B_\alpha \rho(*)$’, in (1h)–(3h) we must appeal to Alpha's justified belief in $T(\beta) \leftrightarrow \neg B_\alpha T(\beta)$, together with the assumption that Alpha meets the evidential norm EVIDENCE. The same is true of (1g)–(3g) and (1i)–(3i). Appeal to propositions such as that expressed by (*) obviates the need for an appeal to EVIDENCE. The conflict between CONSISTENCY, EVIDENCE, and POSSIBILITY can thereby be reduced to a conflict between CONSISTENCY and POSSIBILITY.

3. The Solution

We've seen that, given classical logical assumptions, we can't hold on to CONSISTENCY, EVIDENCE, and POSSIBILITY. I'll now show how if we are paracomplete theorists, we can resolve the above two normative paradoxes. In particular, I'll show that Alpha can satisfy the requirements imposed by both CONSISTENCY and EVIDENCE if we allow that excluded-middle may fail for the claim that Alpha believes that β is true. I'll focus primarily on the first paradox, noting later how the same treatment can be applied to the second.

It will help to first take a brief look at a related semantic paradox. Let η name the following sentence: $\neg \Box T(\eta)$. On the assumption that the logic governing $\Box$ is S5, we can derive a contradiction from this sentence as follows:

(1η)	$T(\eta) \leftrightarrow \neg \Box T(\eta)$	T-schema
(2η)	$\Box(\Box T(\eta) \rightarrow \neg T(\eta))$	(1η), Nec.
(3η)	$\Box(\neg \Box T(\eta) \rightarrow T(\eta))$	(1η), Nec.
(4η)	$\Box\Box T(\eta) \rightarrow \Box \neg T(\eta)$	(2η), K
(5η)	$\Box \neg \Box T(\eta) \rightarrow \Box T(\eta)$	(3η), K
(6η)	$\Box T(\eta) \rightarrow \Box\Box T(\eta)$	4
(7η)	$\neg \Box T(\eta) \rightarrow \Box \neg \Box T(\eta)$	5

(8η) $\Box\neg T(\eta)\rightarrow\neg\Box T(\eta)$	S5 theorem
(9η) $\Box T(\eta)\vee\neg\Box T(\eta)$	Classical Theorem

Subproof 1

(10η) $\Box T(\eta)$	Assumption
(11η) $\Box\Box T(\eta)$	(6η), (10η)
(12η) $\Box\neg T(\eta)$	(4η), (11η)
(13η) $\neg\Box T(\eta)$	(8η), (12η)
(14η) $\bot$	(10η), (13η)

Subproof 2

(15η) $\neg\Box T(\eta)$	Assumption
(16η) $\Box\neg\Box T(\eta)$	(7η), (15η)
(17η) $\Box T(\eta)$	(5η), (16η)
(18η) $\bot$	(15η), (17η)
(19η) $\bot$	(9η), (10η)–(14η), (15η)–(18η)

(1η) is an instance of the T-schema. (2η) and (3η) follow from (1η) on the assumption that the logic governing $\Box$ is a normal modal logic, and so obeys the rule of necessitation. (4η) follows from (2η), and (5η) from (3η), on the assumption that the logic for $\Box$ is a normal modal logic, and so obeys axiom K: $\Box(\phi\rightarrow\psi)\rightarrow(\Box\phi\rightarrow\Box\psi)$. (6$\eta$) holds if the logic governing $\Box$ obeys axiom 4: $\Box\phi\rightarrow\Box\Box\phi$. (7$\eta$) holds if the logic governing $\Box$ obeys axiom 5: $\neg\Box\phi\rightarrow\Box\neg\Box\phi$. (8$\eta$) holds given that axiom T: $\Box\phi\rightarrow\phi$ holds. (9η) is a theorem of classical logic.

If we take $\Box$ to express metaphysical necessity, then it is very plausible that S5 is the correct logic for the operator, and so all of the above modal axioms hold. Given these principles, we can derive a contradiction on the assumption $\Box T(\eta)$ and on the assumption $\neg\Box T(\eta)$ using simply *modus ponens*. A contradiction can then be derived outright from (9η) by proof-by-cases reasoning.

The approach to the liar paradox that I'm interested in holds that excluded-middle is not valid for paradox-inducing sentences. The above derivation is blocked at (9η). It is a fairly straightforward exercise to extend the class of models used to treat standard paradox-inducing sentences such as the liar sentence to models for languages containing $\Box$.[23] In any such paracomplete possible-worlds model in which accessibility between worlds is an equivalence relation, (1η)–(8η) all hold, but

23. See the technical appendix for the formal details.

excluded-middle fails for $\Box T(\eta)$. Although we cannot say that η is *not* necessarily true, we can say that it is neither determinate that it is necessarily true nor determinate that it is not necessarily true; that is, we can say that it is indeterminate whether η is necessarily true. $I \Box T(\eta)$ is valid in the class of models.

The paradox developed in section 2.1 can be represented in the form of a derivation that parallels the modal liar paradox:

(1β)	$T(\beta) \leftrightarrow \neg B_\alpha T(\beta)$	T-schema
(2β)	$B_\alpha(B_\alpha T(\beta) \rightarrow \neg T(\beta))$	Assumption
(3β)	$B_\alpha(\neg B_\alpha T(\beta) \rightarrow T(\beta))$	Assumption
(4β)	$B_\alpha B_\alpha T(\beta) \rightarrow B_\alpha \neg T(\beta)$	(2β), EVIDENCE
(5β)	$B_\alpha \neg B_\alpha T(\beta) \rightarrow B_\alpha T(\beta)$	(3β), EVIDENCE
(6β)	$B_\alpha T(\beta) \rightarrow B_\alpha B_\alpha T(\beta)$	Assumption
(7β)	$\neg B_\alpha T(\beta) \rightarrow B_\alpha \neg B_\alpha T(\beta)$	Assumption
(8β)	$B_\alpha \neg T(\beta) \rightarrow \neg B_\alpha T(\beta)$	CONSISTENCY
(9β)	$B_\alpha T(\beta) \vee \neg B_\alpha T(\beta)$	Classical Theorem

Subproof 1

(10β)	$B_\alpha T(\beta)$	Assumption
(11β)	$B_\alpha B_\alpha T(\beta)$	(6β), (10β)
(12β)	$B_\alpha \neg T(\beta)$	(4β), (11β)
(13β)	$\neg B_\alpha T(\beta)$	(8β), (12β)
(14β)	$\perp$	(10β), (13β)

Subproof 2

(15β)	$\neg B_\alpha T(\beta)$	Assumption
(16β)	$B_\alpha \neg B_\alpha T(\beta)$	(7β), (15β)
(17β)	$B_\alpha T(\beta)$	(5β), (16β)
(18β)	$\perp$	(15β), (17β)
(19β)	$\perp$	(9β), (10β)–(14β), (15β)–(18β)

Formally this derivation is almost identical to the first; where they differ are in the justifications of certain steps.

As above, (1β) is an instance of the T-schema. (2β) and (3β) correspond to our assumption that the agent believes the theorem expressed at (1β). Given the setup of the case, (4β) is a consequence of the assumption that Alpha meets the requirements imposed by EVIDENCE. Why? Because as we set up the case, it follows from the evidential status of the belief codified in (2β) that if Alpha believes that it believes that β is true, then Alpha will have evidence that makes it certain that β is not true. So, assuming that Alpha meets the requirements imposed by EVIDENCE, it follows that if Alpha believes that it believes that β is true, then it will believe that β is not true.[24] This is (4β). A similar story explains how (5β) follows from the assumption that Alpha meets the requirements imposed by EVIDENCE. (6β) and (7β) are assumptions that we made about Alpha. (8β) corresponds to our assumption that the agent meets the normative condition specified in CONSISTENCY. From these we can derive a contradiction, using *modus ponens*, on the assumption $B_\alpha T(\beta)$ and on the assumption $\neg B_\alpha T(\beta)$. Given the assumption that the classical validity $B_\alpha T(\beta) \vee \neg B_\alpha T(\beta)$ obtains, a contradiction can be derived outright by proof-by-cases reasoning.

Since the two derivations proceed in parallel, the strategy for blocking the contradiction in the former case will work equally well in the latter. Just as we can block the derivation of a contradiction in the first case by giving up excluded-middle for η, so too can we block the derivation of a contradiction in the second case by giving up excluded-middle for β.

Allowing that Alpha's doxastic state may be such that excluded-middle fails for the claim that Alpha believes that β is true allows us to defuse the argument given in section 2.1 that CONSISTENCY, EVIDENCE, and POSSIBILITY are incompatible. It is true that, given the assumption that Alpha meets the requirements imposed by both CONSISTENCY and EVIDENCE, we can derive a contradiction on the assumption that Alpha believes that β is true and on the assumption that Alpha doesn't believe that β is true. This was taken to show that CONSISTENCY and EVIDENCE were in conflict with POSSIBILITY. Crucially, however, this requires that we be able to help ourselves to the claim that excluded-middle holds for the proposition that Alpha believes that β is true. If, then, excluded-middle fails for this proposition, it will not follow from the fact that we can derive a contradiction from the assumption that Alpha meets the requirements imposed by CONSISTENCY and EVIDENCE, *given* that Alpha believes that β is true and *given* that Alpha does not believe that β is true, that we can derive a contradiction from this assump-

24. This reasoning assumes that the deduction theorem holds. In fact, this is not so in the logic that I'll be adopting. Such reasoning is, however, "pretheoretically valid," and so can be used in setting up a prima facie paradox. And, as we'll see, the conditionals do in fact hold in the class of models in which I'm interested.

tion outright. We need not infer from the possibility of an agent such as Alpha that CONSISTENCY and EVIDENCE entail a violation of POSSIBILITY.

Of course, given the explicit appeal to excluded-middle in the development of the normative paradox, it is obvious that a paracomplete theorist can block the paradox as stated. But importantly the paracomplete theorist can do more than this. What the paradox developed in section 2.1 purported to show is that there is a possible case in which an antecedently rational agent cannot meet the requirements imposed by CONSISTENCY and EVIDENCE. Using paracomplete resources, however, we can assure ourselves that an agent can in fact satisfy the stipulations made about Alpha and also meet all the requirements imposed by CONSISTENCY and EVIDENCE.

One way to think about the paradox developed in section 2.1 is as follows. Given that Alpha knows that $T(\beta) \leftrightarrow \neg B_\alpha T(\beta)$, it follows from the transparency assumptions (2) and (3) that whether or not Alpha believes that β is true, EVIDENCE will impose a requirement that some of Alpha's beliefs be closed under logical consequence. But meeting this local closure requirement is either impossible (in the case in which Alpha does not believe that β is true) or leads to a violation of the requirement that the agent satisfy CONSISTENCY (in the case in which Alpha does believe that β is true).

Viewing the problem in terms of a local closure requirement lets us see more clearly how an appeal to paracomplete resources can help resolve the problem. For, using the paracomplete model-theory developed to deal with the modal liar sentence to interpret Alpha's belief operator, we can provide a model in which the following all hold:[25]

(8) $B_\alpha T(\beta) \leftrightarrow B_\alpha B_\alpha T(\beta)$

(9) $\neg B_\alpha T(\beta) \leftrightarrow B_\alpha \neg B_\alpha T(\beta)$

(10) $B_\alpha(T(\beta) \leftrightarrow \neg B_\alpha T(\beta))$

(11) Every instance of the following schema is satisfied:
$B_\alpha \phi \rightarrow \neg B_\alpha \neg \phi$

(12) Alpha's beliefs are closed under logical consequence

In this model, excluded-middle fails for the claim that Alpha believes that β is true. Indeed, in this model it is indeterminate whether Alpha believes that β is true.

25. See the technical appendix for proofs.

We can show, then, that if excluded-middle fails for the claim Alpha believes that β is true, Alpha can satisfy the transparency assumptions, know that $T(\beta)\leftrightarrow \neg B_\alpha T(\beta)$, and yet not be forced into violating CONSISTENCY in order to meet EVIDENCE. If, then, we allow for this failure of excluded-middle, we can hold on to what are some seemingly quite basic normative principles.

Importantly for what will follow, it is also clear that (at least if one is a paracomplete theorist) this is *the only way* to hold on to CONSISTENCY, EVIDENCE, and POSSIBILITY. For, on the assumption that Alpha meets CONSISTENCY and EVIDENCE, a paracomplete theorist will accept (1β)–(8β). In addition, this theorist will accept reasoning by *modus ponens* and disjunction elimination. But, as the derivation makes clear, given these commitments, the only way to avoid a contradiction is to allow that excluded-middle fails for β.

Appeal to paracomplete resources, then, allows us to hold on to CONSISTENCY, EVIDENCE, and POSSIBILITY in the face of the paradox developed in section 2.1. Such resources are also similarly effective in dealing with the paradox developed in section 2.2. There we derived a contradiction on the assumption that Alpha failed to believe the proposition expressed by (*). We then inferred that Alpha must believe the proposition expressed by (*). If, however, Alpha believes the proposition expressed by (*), then Alpha will also believe its negation, and so be in violation of CONSISTENCY. Alpha seemed doomed to irrationality, which by POSSIBILITY should not be possible.

The problem with this argument is to be located in the appeal to reductio. In the logics we are dealing with, reductio fails as a valid metarule. Where excluded-middle fails for ϕ, one cannot validly infer $\neg\phi$ given the derivation of a contradiction from ϕ. If, then, we allow that certain doxastic states may give rise to the same failures of classical logic as cases of semantic paradox, it follows that while we can derive a contradiction from the assumption that Alpha fails to believe the proposition expressed by (*), we cannot infer from this that Alpha does believe the proposition expressed by (*). This suffices to block the conclusion that Alpha must be in violation of CONSISTENCY.

4. Belief and Excluded-Middle

I've argued that the normative paradoxes developed in section 2 can be resolved if we allow that an agent's doxastic state may be such that excluded-middle fails for certain claims about whether or not it believes certain propositions. But one may worry: is it really possible for an agent to have such a doxastic state? The argument in this paper is addressed to one who accepts failures of excluded-middle at least with respect to cases of semantic pathology. So the worry is not germane if it stems from a general skepticism about such putative failures of classical logic. The question, then, is whether one who accepts a paracomplete treatment of the semantic paradoxes should also allow that such failures of classical logic may arise in the doxastic realm. In response to this question, let me say two things.

(i) If one thinks that a paracomplete treatment is appropriate not only for cases of semantic pathology but also for cases of vagueness, then there should be no worry at all about failures of excluded-middle in the doxastic realm. For vagueness is ubiquitous, and there are certainly cases in which it is vague whether or not an agent has certain beliefs.

The question of whether one should treat the semantic paradoxes and the paradoxes of vagueness in the same manner is subtle, and it would take us too far afield at this point to make serious inroads on that question. But this unified approach does seem to me to be an attractive option for one who supports a paracomplete treatment of the semantic paradoxes.[26]

(ii) Although the nature of doxastic states is an area of controversy, there are a number of attractive accounts of doxastic states that would make it intelligible, given what I've said so far, that such states could give rise to certain failures of classical logic. A full development of any one of these accounts and how it is able to make sense of paracomplete treatments of doxastic states would require much more space than I have here. I can, however, briefly say something about the feature common to these views that lets us understand how doxastic states could be a source of appropriate failures of classical logic.

The key feature of the accounts I have in mind is that they hold that principles of rationality are constitutive of intentional mental states such as belief and desire.[27] Given such an account, one should take doxastic states to be capable of having whatever properties are required by

26. This is the position taken by Field. See, for example, Field 2003b. Unified approaches to vagueness and semantic paradoxes are also developed in McGee 1991 and Soames 1999.

27. One attractive view of this type has been developed by David Lewis. See Lewis 1974, 1999. According to Lewis, belief and desire states are defined by a tacit theory of folk-psychology, and this tacit theory takes such states to conform to various rational principles. See also Stalnaker 1984 for a slightly different account that takes conditions of

rationality, at least insofar as these properties are logically possible. CONSISTENCY and EVIDENCE provide rational constraints on beliefs, and, as we've seen, in certain cases in order for an agent to conform to CONSISTENCY and EVIDENCE, the agent's doxastic state must be such that excluded-middle fails for claims about whether or not it believes certain propositions. Rationality, thus, requires that an agent's doxastic states be correctly described in paracomplete terms. If one holds that principles of rationality are constitutive of doxastic states, then, if one is also a paracomplete theorist, one should take such states to be, in principle, capable of giving rise to the appropriate failures of excluded-middle. Given such an account, it is a *discovery* afforded by the preceding arguments that doxastic states may give rise to the same failures of classical logic as arise in cases of semantic pathology.[28]

Note that the thesis that doxastic states may give rise to such failures of classical logic poses no threat to physicalism. The account is perfectly compatible with the thesis that an agent's total doxastic state is identical to, say, a particular neural state of the agent. The point is just that for certain propositions, such as that expressed by β, that neural state will be such that excluded-middle fails for the claim that it is a state of believing the proposition in question.

There is much more to be said about paracomplete treatments of doxastic states. Our focus now, however, will be on the ways in which the paracomplete solution to the normative paradox can provide insight into the correct answer to the Normative Question. I take it that this discussion is sufficiently motivated by the existence of at least two attractive views that countenance paracomplete treatments of doxastic states.

5. Against Rejection

We've seen that a paracomplete theorist can block the arguments given in section 2 that purported to show the incompatibility of CONSISTENCY, EVIDENCE, and POSSIBILITY. If one is a paracomplete theorist, and one accepts CONSISTENCY and EVIDENCE (as, I think, one ought to), then one will hold that if Alpha is rational, Alpha will be such that excluded-

rationality to be definitional of belief states. For another such account, see Davidson 1980a, 1980b.

28. Note that I don't claim that theories of this type are the *only* theories that can admit nonclassical doxastic states. The claim is just that, *given the preceding arguments*, such states are intelligible given this type of theory.

middle fails for the claim that it believes that β is true. Having arrived at this conclusion, we are now in a position to provide an argument against the standard answer to the Normative Question. What I'll now argue is that although a paracomplete theorist can resolve the normative paradoxes developed in section 2, doing so requires that one reject:

> **REJECTION** For any proposition ϕ, it is a consequence of the claim that one ought to believe that ϕ is indeterminate that one ought to reject both ϕ and its negation.

The argument for this claim requires a further premise. What we have seen is that meeting the requirements imposed by CONSISTENCY and EVIDENCE demands that excluded-middle fail for the claim that Alpha believes that β is true. Now *one way* for this to be the case is that it be *indeterminate* whether Alpha believes that β is true (indeed, this is the status of Alpha's doxastic state concerning β in the simple model mentioned in the previous section and outlined in the technical appendix). But this is not the only status that is compatible with the failure of excluded-middle. An adequate treatment of the semantic paradoxes that appeals to indeterminacy requires that the indeterminacy operator iterate in a nontrivial manner; we must not, for example, have $II\phi \vDash I\phi$. In particular, this is required in order to adequately treat higher-order paradoxical sentences that employ the determinacy operator.[29] In addition to first-order indeterminacy, then, there is indeterminate indeterminacy, and indeterminate indeterminate indeterminacy, and so forth. Each of these is such that, when a proposition has the status in question, excluded-middle fails for that proposition. It doesn't *follow*, then, from the fact that Alpha ought to be such that excluded-middle fails for the claim that it

29. Consider, for example, the sentence λ^*, which is provably equivalent to $\neg DT\ulcorner\lambda^*\urcorner$. Given excluded-middle, we can derive a contradiction using λ^*. First, assume λ^*. In general $\phi \vDash D\phi$, and so in particular $\lambda^* \vDash D\lambda^*$. We have, then, $D\lambda^*$. Given the intersubstitutivity of ϕ and $T\ulcorner\phi\urcorner$, we also have $DT\ulcorner\lambda^*\urcorner$. But it also follows from λ^* that $\neg DT\ulcorner\lambda^*\urcorner$. So a contradiction can be derived from λ^*. Now assume $\neg\lambda^*$. This entails $DT\ulcorner\lambda^*\urcorner$, which entails $T\ulcorner\lambda^*\urcorner$, which in turn entails λ^*. Again we have a contradiction. Assuming that we have $\lambda^* \vee \neg\lambda^*$, we can derive a contradiction outright. If one wants to treat the liar sentence by rejecting excluded-middle, one should extend the same treatment to this case as well. But note that in this case we cannot characterize λ^* as being indeterminate. For this entails $\neg D\lambda^*$, which is provably equivalent to, and so entails, λ^*, which of course lands us right back in paradox. If we want a way of characterizing this sentence's paradoxical status, the indeterminacy operator must iterate in a nontrivial manner. We can characterize λ^* as being indeterminately indeterminate; but only if it's not the case that $II\phi \vDash I\phi$.

believes that β is true that Alpha ought to be such that it is indeterminate whether it believes that β is true. Luckily we don't need such a strong claim to mount an argument against REJECTION. All that we need is the claim that a *rational* way for Alpha to meet the requirement that it be such that excluded-middle fails for the proposition that it believes that β is true is for it to be *indeterminate* whether it believes that β is true.

Later (in section 8) I'll consider how to reframe the argument against REJECTION if we reject the claim that it is rational for Alpha to be such that it is indeterminate whether it believes that β is true. But I think that it's very hard to see what sort of principled reason there could be for denying this. Given that Alpha is required to be such that excluded-middle fails for the claim that it believes that β is true, if we are to avoid falling back into normative paradox, there should be some way for Alpha to meet this requirement without incurring rational criticism. The question is, then: what are the rational ways for Alpha to meet this requirement if its being indeterminate whether it believes that β is true isn't one of them? Is its being indeterminately indeterminate whether Alpha believes that β is true a rational way to meet the relevant requirement, or its being indeterminately indeterminately indeterminate whether Alpha believes that β is true, or some other status? It is very hard to see why some such higher-order status(es) should be rational, while it is irrational for it to be simply indeterminate. Perhaps there is some deep surprising explanation for this. But at present I can't see what that would be, and so I'll proceed for now on the assumption that Alpha may rationally meet the requirements imposed by CONSISTENCY and EVIDENCE by being such that it is indeterminate whether Alpha believes that β is true.

Given this assumption, we can now argue as follows. We can show that although the paracomplete theorist can resolve our earlier normative paradox, given acceptance of REJECTION, we can resurrect this paradox in a way that is not amenable to a similar solution.

We have assumed that it is indeterminate whether Alpha believes that β is true, and that this is a rational way for Alpha to meet the obligations imposed by CONSISTENCY and EVIDENCE. This is logically equivalent to the claim that it is indeterminate whether Alpha does not believe that β is true, that is $I\neg B_\alpha T(\beta)$. Let us add to our story about Agent Alpha. We are now allowing Alpha's doxastic states to be indeterminate. We should, therefore, extend our transparency assumptions to take account of this possibility:

(13) $I \neg B_\alpha T(\beta) \leftrightarrow B_\alpha I \neg B_\alpha T(\beta)$

(13) holds in the class of models in which we have represented Alpha's doxastic state. The assumption that an agent with indeterminate doxastic states may at least in principle satisfy this condition is therefore reasonable.[30]

Given that it is indeterminate whether Alpha does not believe that β is true, by (13) it follows that Alpha believes this, that is, $B_\alpha I \neg B_\alpha T(\beta)$. As in the earlier cases, we assume that Alpha is perfectly reliable in this belief. The following is a theorem: $I \neg B_\alpha T(\beta) \rightarrow IT(\beta)$.[31] As in the earlier cases, we can assume that Alpha believes this on the basis of the same superlative grounds as we do. Given these assumptions, it follows that Alpha's evidence makes it certain that $IT(\beta)$. By EVIDENCE, it follows that $OB_\alpha IT(\beta)$. By REJECTION, it follows that $OR_\alpha T(\beta)$. If one rejects ϕ, it follows that one does not believe ϕ. Assuming, then, that Alpha meets the rational requirement imposed on it by REJECTION, we have $\neg B_\alpha T(\beta)$.

This, however, lands us back into normative paradox. We need simply rehearse case 1. By (3), we have $B_\alpha \neg B_\alpha T(\beta)$. We also have that $\neg B_\alpha T(\beta) \rightarrow T(\beta)$ is a theorem, and that it is believed by Alpha on excellent grounds. It follows that Alpha's evidence makes it certain that $T(\beta)$. Assuming compliance with the normative demands imposed by EVIDENCE, we have $B_\alpha T(\beta)$. But this, of course, is impossible since we already have $\neg B_\alpha T(\beta)$.

We have derived a contradiction on the assumption that Alpha, an antecedently rational agent, meets all of the requirements imposed by EVIDENCE, CONSISTENCY, and REJECTION. Note that no appeal was made to excluded-middle.[32] The same moves that were available to us to reconcile the seeming incompatibility of CONSISTENCY and EVIDENCE are not available in this case. If we are to hold on to CONSISTENCY, EVIDENCE, and POSSIBILITY by allowing for doxastic states to be indeterminate we must reject REJECTION.

CONSISTENCY, EVIDENCE, and POSSIBILITY seem to me individually and jointly much more plausible normative conditions than REJECTION. Faced with the choice between holding on to CONSISTENCY,

30. See the technical appendix for the proof of this claim.

31. In general, where $\phi \leftrightarrow \psi$ is a theorem, so is $I\phi \leftrightarrow I\psi$.

32. Nor was any use made of reductio or other forms of proof that fail given the approach to the liar under consideration.

EVIDENCE, and POSSIBILITY and holding on to REJECTION, it seems clear to me that the former course is preferable.

It is far from obvious what answer one should give to the Normative Question. It is difficult to get an independent grip on this issue. Given this, we should, I think, take seriously an argument that shows how the answer to this question is constrained by our acceptance of other clearer normative conditions. The incompatibility between CONSISTENCY, EVIDENCE, POSSIBILITY, and REJECTION provides a good reason to give up REJECTION.

6. Generalizing the Argument

I've argued that if we are to avoid resurrecting our normative paradox, we can't accept REJECTION. Now, it isn't hard to see that the argument I've presented easily generalizes in the following two ways.

Claim 1: If we want to hold on to CONSISTENCY, EVIDENCE, and POSSIBILITY, then we should reject any view that endorses $OBI\phi \vDash O\neg B\phi$.

In arguing that acceptance of REJECTION leads to the resurrection of our normative paradox, I argued that, given CONSISTENCY and EVIDENCE, Alpha was required to believe that it is indeterminate whether it believes that β is true. But, given this, in order to meet the requirement imposed by REJECTION, Alpha was forced to *not* believe that β is true. And, as demonstrated by case 1, this resulted in a failure to meet a requirement imposed by EVIDENCE. The important feature of REJECTION for the argument was that meeting the requirement it imposes demands that one not believe a proposition that one ought to believe is indeterminate. Thus any view that endorses $OBI\phi \vDash O\neg B\phi$ will be subject to the same argument.

Claim 2: If we want to hold on to CONSISTENCY, EVIDENCE, and POSSIBILITY, then we should reject any view that endorses $OBI\phi \vDash OB\phi$.

The argument against REJECTION showed that Alpha must believe that it is indeterminate whether it believes that β is true in order to meet certain requirements imposed by CONSISTENCY and EVIDENCE. Case 2 earlier showed that if Alpha does believe that β is true, then Alpha will fail to meet either a requirement imposed by CONSISTENCY or a requirement imposed by EVIDENCE. If, then, we were to accept that one ought to believe a proposition that one ought to believe is indeterminate, we would be able to resurrect the normative paradox developed in section 2.1. Any view that endorses $OBI\phi \vDash OB\phi$ is therefore subject to a variant of the argument against REJECTION.

The generalization of the argument against REJECTION serves to rule out a number of interesting potential answers to the Normative Question. Perhaps it isn't terribly surprising to find that belief is not the correct response to indeterminacy. However, one might be tempted to think that if rejection isn't the correct attitude toward cases of indeterminacy, then perhaps agnosticism is. Or perhaps one might be tempted by the thought that, in response to perceived cases of indeterminacy, one should simply opt out of having any normal doxastic attitude toward such a proposition. That is, one might be tempted by the thought that, in response to indeterminacy, one shouldn't believe, be agnostic about, or reject such a proposition but have some other attitude incompatible with these. What the argument I've provided shows is that these options are ultimately just as unacceptable as REJECTION.

7. Indeterminacy

I've argued that we shouldn't accept REJECTION as providing the correct answer to the Normative Question. Indeed, if the argument I've given against REJECTION is right, then we shouldn't accept any answer to the Normative Question that holds that one shouldn't believe a proposition that one ought to believe is indeterminate, nor should we accept any answer that holds that one should believe such a proposition.

What then is the correct answer to the Normative Question? One option would be to argue that there is no general normative condition connecting the indeterminacy of propositions and our doxastic states concerning those propositions. According to this line, in standard cases of indeterminacy, such as the liar sentence, REJECTION does give us the right story, but in other cases, such as β, another story is appropriate. Indeterminacy would in this way be like contingency.[33] There is no single attitude one should have toward propositions that one takes to be contingent. Some we should believe, some we should reject, and others we should simply be agnositic about; it depends on what our evidence tells us.

This is a consistent position, but I don't see that it has much to recommend it. How exactly should we restrict REJECTION? To say that we simply restrict it for those cases in which it leads to normative paradox seems hopelessly ad hoc. But what other principled distinction can we draw between, say, the case in which an agent believes that the prop-

33. Contingency is understood as follows: $C\phi \leftrightarrow_{df} \Diamond\phi \wedge \Diamond\neg\phi$.

osition expressed by the liar sentence is indeterminate and the case in which it believes that the proposition expressed by β is indeterminate? In the case of contingency we can say something about why an agent may believe both that ϕ is contingent and that ψ is contingent and yet rationally take different attitudes toward the two propositions; the agent may, for example, have conclusive evidence that one is true and the other false. In the case of indeterminacy, however, I have no idea what sort of analogous story one could tell that would make REJECTION deliver the correct verdict in all but the problematic cases.

What I'll argue in this section is that we should instead accept:

> **INDETERMINACY** For any proposition ϕ, it is a consequence of the claim that one ought to believe that ϕ is indeterminate that one ought to be such that it is indeterminate whether one believes ϕ.

The argument has two parts. First I'll argue that INDETERMINACY is independently motivated. Then I'll show that, unlike the demands imposed by REJECTION, Alpha is able to satisfy the demands imposed by INDETERMINACY, in addition to those imposed by CONSISTENCY and EVIDENCE.

7.1.

First, the argument for independent motivation.

It is very easy to be puzzled about what answer to give to the Normative Question. For, prima facie, the following three claims are all plausible:

(14) If one ought to believe that ϕ is indeterminate, it follows that one ought not believe ϕ.
(15) If one ought to believe that ϕ is indeterminate, it follows that one ought not be agnostic about ϕ.
(16) If one ought to believe that ϕ is indeterminate, it follows that one ought not reject ϕ.

(14) will, I suspect, strike you as immediately plausible. Consider, for example, a paradigmatic indeterminate proposition such as that expressed by the liar sentence. In this case, belief would certainly seem to be an inappropriate attitude.

It would also seem, as (15) maintains, to be inappropriate to be agnostic toward this proposition. After all, agnosticism is the correct attitude to take toward a proposition about which one takes oneself to be ignorant. One who thinks that the proposition expressed by the liar sentence is indeterminate would not, however, seem to think that there is some fact of the matter concerning the truth-value of this proposition about which he or she is ignorant.

(16) may be less immediately compelling, but one can argue for it by appeal to the following principle:

> **NEGATION** One ought to be such that one rejects a proposition ϕ just in case one believes its negation $\neg\phi$.[34]

ϕ is indeterminate just in case $\neg\phi$ is indeterminate. If one ought to believe that ϕ is indeterminate, then one ought to believe that $\neg\phi$ is indeterminate. By (14), then, one ought not believe $\neg\phi$. By NEGATION, one ought to be such that if one does not believe $\neg\phi$, then one does not reject ϕ. Given that doxastic obligations are closed under consequence, it follows that one ought not reject ϕ. This gives us (16).

The problem is that the following claim is also quite plausible:

> (17) If one ought to believe that ϕ is indeterminate, it follows that one should not fail to have *some* positive doxastic attitude toward ϕ.

To have a positive doxastic attitude toward a proposition ϕ is (as I'm using the term) to either believe ϕ, be agnostic about ϕ, or reject ϕ. Of course, in certain situations it may indeed be permissible to fail to have some such positive doxastic attitude toward a proposition. In particular, where one lacks certain relevant concepts, it would, indeed, be wrong to say that rationality demands that the agent form opinions that require the deployment of such concepts. But, given this, if one is required to believe that ϕ is indeterminate, then one will have all the relevant concepts required for forming judgments as to whether or not ϕ obtains. But, given that the agent is equipped with the relevant concepts, it is at least prima facie plausible that an agent should have some doxastic attitude toward ϕ.

A paracomplete theorist will, presumably, think that there are cases in which an antecedently rational agent ought to believe that a certain proposition is indeterminate. But, given this and (14)–(17), this agent will be saddled with a set of obligations that it is impossible

34. Read: $O(R\phi \leftrightarrow B\neg\phi)$. This is a principle that a proponent of REJECTION will reject. But it should be conceded that this is prima facie quite plausible.

to meet. (14) – (17), then, are incompatible with POSSIBILITY. What (14) – (17) amount to is the claim that there is simply no rational response to cases of perceived indeterminacy. This is something that, I think, a paracomplete theorist should clearly not accept.

Call the prima facie plausibility of (14) – (17) the *Normative Problem*. An adequate response to the Normative Problem should identify which of (14) – (17) we should give up, and *in addition* it should provide a plausible error-theory that can account for the prima facie plausibility of (14) – (17).

A proponent of REJECTION can provide the following response to the Normative Problem. Such a proponent will say that we should give up (16) but hold on to (14), (15), and (17). To this end, the proponent of REJECTION will reject NEGATION. In support of this response, she could offer the following plausible error-theory. First, it is worth noting that while NEGATION does not hold unrestrictedly, it does hold in those cases in which excluded-middle holds. But then it would seem quite reasonable that we could mistakenly find (16) plausible. For, of course, we are naturally prone to overgeneralize from those cases in which classical logic holds. Thus while (16) is not correct, we can account for its prima facie plausibility.

At first glance, this is an attractive response to the Normative Problem. However, what we have now seen is that it is ultimately unacceptable. For accepting (14), (15), and (17) while giving up (16) commits one to the acceptance of REJECTION. But this, we've seen, saddles us with normative paradox, and so is ultimately unacceptable.

Indeed, the generalizations of our argument against REJECTION show that one can't solve the Normative Problem by simply rejecting one of (14) – (17). If one rejects (17) but holds on to (14) – (16), then one is committed to the view that if one ought to believe that a certain proposition ϕ is indeterminate, then one ought to not believe ϕ. Similarly if one rejects (15) but holds on to (14), (16), and (17). And if one rejects (14) but holds on to (15) – (17), then one is committed to the view that if one ought to believe that a certain proposition ϕ is indeterminate, then one ought to believe ϕ. Any of these minimal responses to the Normative Problem, then, leads to an ultimately unacceptable view.

What I will now do is outline an alternative response to the Normative Problem that is able to do justice to the intuitions that motivate (14) – (17) while avoiding the untoward consequences of accepting these claims. I'll then show how INDETERMINACY is a consequence of this response. Given the paucity of reasonable responses to the Normative Problem, that INDETERMINACY follows from an elegant error-theoretic response gives us a reason to take INDETERMINACY seriously as the answer to the Normative Question.

The response to the Normative Problem that I advocate involves rejecting each of (14)–(17). In their stead, we should accept the following closely related principles:

(14^d) If one ought to believe that ϕ is indeterminate, it follows that one ought not *determinately* believe ϕ.

(15^d) If one ought to believe that ϕ is indeterminate, it follows that one ought not be *determinately* agnostic about ϕ.

(16^d) If one ought to believe that ϕ is indeterminate, it follows that one ought not *determinately* reject ϕ.

(17^d) If one ought to believe that ϕ is indeterminate, it follows that one ought not *determinately* fail to have some positive doxastic attitude toward ϕ.

Unlike with (14)–(17), an agent can meet each of the requirements imposed by (14^d) –(17^d). And a proponent of (14^d)–(17^d) can provide the following simple error-theory to account for the prima facie plausibility of (14)–(17). We are not terribly good at distinguishing between something being the case and its *determinately* being the case. Indeed, insofar as we are able to make this distinction, it is only as the result of significant theoretical work; that there is a distinction becomes clear only when we see how it is necessary in order to resolve certain paradoxes, such as that raised by the liar sentence. We should not expect, then, that in advance of this work our intuitions should be finely attuned to this distinction. If there are true principles, such as (14^d)–(17^d), that concern certain conditions obtaining determinately, it should not be unexpected that we would confuse such principles for other principles, such as (14)–(17), that concern those conditions simply obtaining whether determinately or not. By accepting (14^d)–(17^d), then, we can account for the plausibility of (14)–(17) while avoiding their undesirable consequences.

Next I'll show that INDETERMINACY is a consequence of (14^d)–(17^d); that is, I'll show that from (14^d)–(17^d) it follows that $OBI\phi \vDash OIB\phi$. To show that this is so, it will suffice to show that both

(i) $OBI\phi \vDash O\neg DB\phi$ and (ii) $OBI\phi \vDash O\neg D\neg B\phi$ follow from (14^d)–(17^d).[35]

Now, (i) just is (14^d), and so it trivially follows from (14^d)–(16^d).

To show that (ii) is a consequence of (14^d)–(16^d), we can argue as follows. We want to establish $OBI\phi \vDash O\neg D\neg B\phi$. One fails to believe ϕ just in case one is either agnostic about ϕ or one rejects ϕ or one simply has no positive doxastic attitude toward ϕ. That is, we have $\vDash\neg B\phi \leftrightarrow (A\phi \vee R\phi \vee \neg P\phi)$. So what we want to establish is $OBI\phi \vDash O\neg D(A\phi \vee R\phi \vee \neg P\phi)$. (15^d) tells us that $OBI\phi \vDash O\neg DA\phi$. (16^d) tells us that $OBI\phi \vDash O\neg DR\phi$. And (17^d) tells us that $OBI\phi \vDash O\neg D\neg P\phi$. Since, I take it, obligations are closed under consequence, we have $OBI\phi \vDash O(\neg DA\phi \wedge \neg DR\phi \wedge \neg D\neg P\phi)$. Now, in general we have:

(18) $\neg D\phi \wedge \neg D\psi \wedge \neg D\xi \vDash \neg D(\phi \vee \psi \vee \xi)$[36]

So in particular we have $\neg DA\phi \wedge \neg DR\phi \wedge \neg D\neg P\phi \vDash \neg D(A\phi \vee R\phi \vee \neg P\phi)$. And so given the closure condition on rational obligations, we have $OBI\phi \vDash O\neg D(A\phi \vee R\phi \vee \neg P\phi)$. This suffices to establish (ii).

Since (i) and (ii) follow from (14^d)–(17^d), it follows that INDETERMINACY is a consequence of (14^d)–(17^d). The latter, I've argued, provide an attractive response to the Normative Problem. This is a reason to think that INDETERMINACY gives the correct answer to the Normative Question.

7.2.

Having argued that INDETERMINACY is independently motivated, the next point to make is that, unlike REJECTION, Alpha can satisfy the demands imposed by INDETERMINACY while satisfying the demands imposed by CONSISTENCY and EVIDENCE.

Alpha is an agent who believes the theorem $T(\beta) \leftrightarrow \neg B_\alpha T(\beta)$ and is, in addition, doxastically self-transparent with respect to the proposition that it believes that β is true. I noted that we could represent the doxastic state of an agent satisfying these stipulations by a paracom-

35. To see that this will suffice, note that, since $IB\phi$ is equivalent to $\neg DB\phi \wedge \neg D\neg B\phi$, it follows that $OBI\phi \vDash OIB\phi$ is equivalent to $OBI\phi \vDash O(\neg DB\phi \wedge \neg D\neg B\phi)$. And as an instance of a general closure principle, we have: $O\neg DB\phi,\ O\neg D\neg B\phi \vDash O(\neg DB\phi \wedge \neg D\neg B\phi)$. Thus if we can show that (i) and (ii) hold, then we can show that $OBI\phi \vDash O(\neg DB\phi \wedge \neg D\neg B\phi)$, and so $OBI\phi \vDash OIB\phi$.

36. See the technical appendix.

plete possible-worlds model in which the accessibility relation is an equivalence relation. In such a model, the agent's beliefs will be consistent and closed under logical consequence. This assured us that an agent such as Alpha could in principle meet the demands imposed by CONSISTENCY and EVIDENCE. In this model, however, it is indeterminate whether Alpha believes that β is true.

This model is also sufficient to assure us that Alpha is able to meet whatever additional demands might be imposed by INDETERMINACY. First note that, in general, for any class of paracomplete possible-worlds models $\mathcal{M}$ the following holds:[37]

(19) $B_\alpha I\phi \vDash_{\mathcal{M}} IB_\alpha\phi$

Given (19), it follows that in any of the models in which Alpha meets the demands imposed by CONSISTENCY and EVIDENCE, Alpha will also meet the demands imposed by the combination of CONSISTENCY, EVIDENCE, and INDETERMINACY.

To see this, first note that INDETERMINACY, which states $OB_\alpha I\phi \vDash OIB_\alpha\phi$, only issues in an obligation given an input of the form $OB_\alpha I\phi$. We can think about the obligations that result from CONSISTENCY, EVIDENCE, and INDETERMINACY as the result of the following iterated process. We start with the obligations that result from CONSISTENCY and EVIDENCE. These are the obligations that result from CONSISTENCY, together with the obligations that result from EVIDENCE, together with whatever obligations follow from these given general principles of deontic logic. (In what has preceded, and in what follows, the only general principle of deontic logic I'm assuming is that the logical consequences of a set of rational obligations are themselves rationally obligatory.) Label this set $\mathbf{\Omega}^0$. This delivers possible inputs to INDETERMINACY. This delivers further obligations, which, together with general principles of deontic logic, gives us a set of obligations $\mathbf{\Omega}^1$. And so on. The end result of this process is the set of obligations entailed by CONSISTENCY, EVIDENCE, and INDETERMINACY.

Given a model M in which our agent meets CONSISTENCY and EVIDENCE, (19) assures us that, at each stage of this process, the agent will, in M, satisfy the obligations that result at that stage. By hypothesis the agent in M meets $\mathbf{\Omega}^0$. By (19), in M the agent will meet whatever obligations result from $\mathbf{\Omega}^0$ together with INDETERMINACY. In such a

37. Assuming that B_α is treated in the models as a universal quantifier over possible worlds. See the technical appendix for the proof of this claim.

model, moreover, if a set of obligations is met, then so is any logical consequence of this set. So the agent will meet $\mathbf{\Omega}^1$. And it is clear that the reasoning here generalizes. For each stage γ, Alpha will meet the obligations imposed by $\mathbf{\Omega}^\gamma$.

We can be assured, then, given that Alpha is able to meet the obligations that follow from CONSISTENCY and EVIDENCE, that it can meet any additional obligations that result from the endorsement of INDETERMINACY. INDETERMINACY, unlike REJECTION, does not land us back in normative paradox.

8. Further Considerations

A key premise in the preceding argument was that a *rational* way for Alpha to meet the requirements imposed by CONSISTENCY and EVIDENCE was for Alpha to be such that it was indeterminate whether it believed that β is true. I noted that I couldn't see any principled reason for denying this claim. Nonetheless, it is, I think, worth considering how matters stand if one does not rely on this assumption.

As I noted earlier, not every paradoxical proposition for which excluded-middle fails can be characterized as being indeterminate. Some we can only characterize as being indeterminately indeterminate, others as being indeterminately indeterminately indeterminate, and so on. And, indeed, there will be some very complicated paradoxical propositions for which there is no appropriate indeterminacy-type operator. The Normative Question that I've been focused on, then, can be seen as a member of a tightly knit family of questions: Given that one ought to believe that a proposition ϕ is indeterminate^n, what attitude should one have toward ϕ? Or if ϕ is a proposition for which excluded-middle fails but whose status cannot be characterized by any such indeterminacy-type operator, then what attitude should one have toward ϕ in this case?

Prima facie, it is natural for the proponent of REJECTION to give the same answer to each of these questions.[38] However, I think that the *minimal* lesson that we should draw from the earlier discussion of the normative paradoxes is that if one is a paracomplete theorist, then one should not accept that rejection is *always* the correct attitude to take toward propositions for which excluded-middle fails. For, in order to resolve the normative paradoxes developed in section 2, we required that Alpha be such that excluded-middle fails for the claim that it be-

38. Field, for example, has endorsed this type of general rejectionism.

lieves that β is true. However, if rejection were always the correct attitude to have in response to such propositions, then we would be able to resurrect our normative paradox. Commitment, then, to CONSISTENCY, EVIDENCE, and POSSIBILITY precludes us from saying that rejection is the correct cognitive response to all cases in which excluded-middle fails.

Given this fact, the proponent of REJECTION needs to provide a story about when it's okay to reject such a proposition and when some other attitude is appropriate. Here are some options. Although the proponent of REJECTION wants to say that rejection is always the correct cognitive response to cases of indeterminacy, perhaps she will say that rejection is not always the correct response to cases of indeterminacy^n, for some higher-order indeterminacy-type status. This option bifurcates. Perhaps the proponent of REJECTION will want to say that rejection is *never* the correct response to cases of indeterminacy^n, or perhaps it is only in certain cases like β that rejection is not the correct response. Or the proponent of REJECTION may want to say, instead, that it is only in cases where excluded-middle fails but the status of the proposition cannot be characterized using one of the indeterminacy-type operators that rejection is not always the correct cognitive response. This option again bifurcates into a version that holds that in such cases rejection is never appropriate and a version that holds that it is only for certain cases such as β that rejection is inappropriate.

Obviously this catalogue of options is rather schematic at this point. However, I think it's hard to see how the details could be filled in in a way that wouldn't involve rather ad hoc stipulations. Prima facie it seems quite implausible to hold that the correct response to failures of excluded-middle should be rejection in all cases *except* where this would lead us into normative paradox. Of course, this would get us out of trouble, but I think it's reasonable to ask for some independent reason that we should treat cases like β differently from other cases in which excluded-middle fails, and as I've already noted, it's far from obvious what reason could be given here. This sort of worry doesn't apply to views that hold that rejection is never the correct response to certain types of higher-order indeterminacy (or to cases that cannot be characterized by appeal to indeterminacy-type operators). However, such views still raise the rather difficult question about why it is that rejection is the correct response to indeterminacy but not to, say, indeterminacy^n? What sort of plausible general principle could explain the difference in the cognitive significance of these statuses? This seems to me to be a rather tricky question for proponents of REJECTION.

These considerations are clearly far from decisive. What they amount to is a challenge to the proponent of REJECTION to provide some plausible principled story about how we should respond to various cases in which excluded-middle fails that is compatible with resolving the normative paradoxes. Perhaps this can be done, but I think it's reasonable to be skeptical that any such account is available.

In contrast, the proponent of INDETERMINACY can provide what seems to me to be a simple and principled general story about how one should respond to different cases in which excluded-middle fails. The general principle underlying INDETERMINACY is, I suggest, the following: the correct cognitive response to any indeterminacy-type status is for there to be a mirroring indeterminacy-type status in one's doxastic state. Given this, we should accept the following schematic principle:

> **INDETERMINACY^n** For any proposition ϕ, it is a consequence of the claim that one ought to believe that ϕ is indeterminate^n, that one ought to be such that it is indeterminate^n whether one believes ϕ.[39]

While it's true that we shouldn't respond to all cases in which excluded-middle fails in the same way, the proponent of INDETERMINACY, unlike proponents of REJECTION, has what seems to me to be a principled story to tell about what attitudes one should have in different cases that is compatible with resolving the normative paradoxes. This seems to me to provide good prima facie reason to prefer INDETERMINACY to REJECTION.

Even, then, if we don't help ourselves to the claim that it's rational for it to be indeterminate whether Alpha believes that β is true, our earlier reflections on the normative paradoxes can be seen as providing support for giving up REJECTION in favor of INDETERMINACY.

9. Taking Stock

I have argued, so far, that Agent Alpha can satisfy the demands imposed by CONSISTENCY and EVIDENCE, given that its doxastic state is a source

39. And if ϕ is a proposition for which excluded-middle fails but whose status cannot be characterized by any such higher-order indeterminacy operator, then the correct cognitive response is for one to be such that excluded-middle fails for the proposition that one believes ϕ in such a way that it too cannot be characterized by any such indeterminacy-type operator.

of certain failures of excluded-middle. Alpha cannot, however, satisfy the demands imposed by CONSISTENCY, EVIDENCE, and REJECTION. This, I've claimed, gives us good reason to reject REJECTION. Further, I've argued that Alpha can satisfy the demands imposed by CONSISTENCY, EVIDENCE, and INDETERMINACY. The same problem that besets REJECTION does not beset INDETERMINACY. Given that INDETERMINACY is independently motivated, we have good reason to hold that it provides the correct answer to the Normative Question.

It is worth noting, however, certain limitations of this argument. What has not been shown is that, in *any* possible case in which an agent is not already guilty of a rational failure, the agent may in principle satisfy CONSISTENCY, EVIDENCE, and INDETERMINACY. Indeed, this is, I think, something that is not (at least in our present state of knowledge) amenable to proof. For both the conditions under which a proposition counts as part of an agent's evidence and the conditions under which an agent might count as being antecedently rational are in various ways unclear—the latter in particular depending on what other correct principles there are governing doxastic rationality. Trying, then, to prove in general that there are no cases in which an antecedently rational agent cannot meet CONSISTENCY, EVIDENCE, and INDETERMINACY seems hopeless to me.

Given this limitation, it must be allowed that the arguments here are not unassailable. Perhaps there are cases in which an agent, not guilty of any antecedent rational failing, cannot satisfy CONSISTENCY and EVIDENCE despite the resources afforded by our paracomplete theory. Or perhaps the addition of INDETERMINACY may lead to cases in which an antecedently rational agent is condemned to irrationality. If the former were true, this would undermine our claim that we should reject REJECTION given the incompatibility of this principle with CONSISTENCY, EVIDENCE, and POSSIBILITY. If the latter were true, this would undermine the claim that there is an important asymmetry between REJECTION and INDETERMINACY.

So far as I can see, however, we have no good reason to think that cases of either sort exist. And here it is worth emphasizing that despite the fact that we cannot prove *in general* that there are no such problematic cases, the arguments given above do generalize in important ways. In particular, the paracomplete model theory to which I have appealed is sufficient to assure us that there can be no case in which an agent is forced to violate CONSISTENCY, EVIDENCE, and INDETERMINACY simply given knowledge of theorems and doxastic transparency. For in the models in which we have represented Alpha's doxastic state, Alpha will count as believing *all* theorems and as being completely doxastically self-transparent. Such models assure us, then, that such cognitive achievements are compatible with the agent satisfying CONSISTENCY, EVIDENCE, and INDETERMINACY. We can be assured, then, not only that the case of Alpha provides no problem for the combination of CONSISTENCY, EVIDENCE, and INDETERMINACY but that, in general, there is no similar, but less obvious case, in which an agent is doomed to irrationality, by the lights of these principles, simply due to its knowledge of theorems and its own sensitivity to its doxastic states. This takes care of a large class of potentially problematic cases.

The burden of proof, at this point, seems to me to be squarely on the opponent of INDETERMINACY. Perhaps there are problematic cases of the sort described. However, until such a case is produced, we should, I think, invest a good amount of credence in INDETERMINACY.

10. Rejection Again

I've argued that we have good reason to prefer INDETERMINACY to REJECTION. REJECTION, however, is not without its own positive motivations. In these final sections, I'll take up what are, I think, the three clearest arguments in favor of REJECTION and show how they can be resisted.

10.1.

Here is an argument that might be thought to strongly support REJECTION.[40]

- (P1) One should reject any contradiction.
- (P2) The liar sentence entails a contradiction.
- (P3) The negation of the liar sentence entails a contradiction.
- (P4) If ϕ entails ψ, then one should have at least as much confidence in ψ as in ϕ.
- (C) Therefore, one should reject both the liar sentence and its negation.

This argument, of course, extends to any sentence that, like the liar sentence, is such that both it and its negation entail a contradiction.

40. See Field 2003a, 467 for this argument.

Of course, we may want to extend the notion of indeterminacy to sentences that don't have this property, but in such cases considerations of uniformity could presumably be invoked.

The premise that I reject is (P1). Certain contradictions are, according to a paracomplete theorist, indeterminate, for example, $\lambda \wedge \neg\lambda$. In these cases, I hold that it should be indeterminate whether one rejects the contradiction in question.

The question is whether rejecting (P1) involves an unacceptable intuitive cost. Certainly (P1) is intuitive. It would, I think, be a significant drawback to the account I'm offering if there was nothing that I could say that could do justice to the intuitive pull of (P1). Ideally what we want is (a) an alternative principle that can capture at least some of the intuitive force of (P1) and (b) an error-theory that can account for our mistakenly taking (P1) to be correct. I think that both of these desiderata can be met.

While we cannot hold that one should always reject a contradiction, we can hold that one should never determinately fail to reject a contradiction. Using this latter fact, we can provide a plausible error-theory to account for our finding the former claim plausible. For, as noted earlier, the distinction between something being the case and its *determinately* being the case is not one to which our intuitions are sensitive, at least in advance of significant theoretical work. But if one ignores the distinction between something being the case and its determinately being the case, then the claim that one should never determinately fail to reject a contradiction will collapse into the claim that one should always reject a contradiction.[41] And so one who thinks that one should never determinately fail to reject a contradiction should not be surprised if this correct principle was commonly confused with the incorrect principle that one should always reject a contradiction.

I don't deny that many will find the above argument in favor of REJECTION convincing, at least at first sight. But I don't think that the costs of rejecting premise (P1) are all that significant. We can capture much of the intuitive force of this premise. And we can explain why one who was not attuned to the possibility of doxastic indeterminacy would, on this basis, find (P1) plausible. This seems to me to take much of the sting out of rejecting (P1).

41. To see that this is the case, note that if one ignores the distinction between the conditions for something to be the case and the conditions for it to *determinately* be the case, then ϕ and $D\phi$ will be equivalent. Given this, $\neg D \neg R\phi$ will be equivalent to $\neg\neg R\phi$ and so to $R\phi$.

10.2.

Here's another argument in favor of REJECTION. This argument appeals to degrees of belief. So far I've confined myself to talking about binary belief. This has merely been for the sake of simplicity. Ultimately I think that binary belief should be understood in terms of degrees of belief. For the purposes of this argument, let's make the following assumptions. To believe a proposition ϕ is to have a degree of belief above a certain threshold τ. To reject a proposition is to have a degree of belief below the co-threshold $1 - \tau$. Now we can argue as follows:

(P1) ϕ entails $D(\phi)$.
(P2) $D(\phi)$ entails ϕ.
(P3) If ϕ entails ψ, then one should have at least as much confidence in ψ as in ϕ.
(P4) One's degree of belief in ψ should be less than or equal to $1 -$ one's degree of belief in $\neg\psi$.

From (P1) – (P3), it follows that:

(C1) One should have the same degree of belief in ϕ as in $D(\phi)$.

Given (P4), it follows that:

(C2) If one has a degree of belief over the threshold for $I\phi$, that is, for $\neg D\phi \wedge \neg D \neg \phi$, then one should have a degree of belief below the co-threshold for $D\phi$ and for $D \neg \phi$.

So, given (C1):

(C3) If one has a degree of belief over the threshold for $I\phi$, then one should have a degree of belief below the co-threshold for ϕ and for $\neg\phi$.

To adequately assess this argument, we need to say a bit about the form of the models that Field employs and how entailment is defined in these models. The models involve an infinite set of semantic values that are partially ordered.[42] There is a top value 1. Given such a set of models,

42. These are the "fine-grained" semantic values. There is also another way one can describe the models in which there are only three values, with all the nonextremal values getting lumped together. To avoid confusion, I note that, for reasons of simplicity, it is the coarse-grained values that I employ in the technical appendix.

we can define at least two notions of entailment. Call these respectively "weak entailment" and "strong entailment."[43] We say that ϕ weakly entails ψ just in case, in every model in which ϕ gets semantic value 1, ψ gets semantic value 1. We say that ϕ strongly entails ψ just in case, in every model, ψ has a semantic value at least as great as the semantic value of ϕ.[44] As it turns out the claim that ϕ strongly entails ψ is equivalent to the claim that the conditional $\phi \rightarrow \psi$ is weakly valid, that is to say that it has semantic value 1 in every model.

Given the distinction between strong and weak entailment, there are two ways of understanding the above argument. We could understand (P1) − (P3) as involving weak entailment:

(P1w) ϕ *weakly* entails $D(\phi)$.
(P2w) $D(\phi)$ *weakly* entails ϕ.
(P3w) If ϕ *weakly* entails ψ, then one should have at least as much confidence in ψ as in ϕ.

Or we could understand these premises as involving strong entailment:

(P1s) ϕ *strongly* entails $D(\phi)$.
(P2s) $D(\phi)$ *strongly* entails ϕ.
(P1s) If ϕ *strongly* entails ψ, then one should have at least as much confidence in ψ as in ϕ.

Either way the argument can be resisted. If the argument is understood in terms of weak entailment, then ($P1_w$) and ($P2_w$) both hold. I claim, however, that we should reject ($P3_w$). Instead, we should accept only ($P3_s$). It is strong entailment, not weak entailment, that should be thought of as providing a normative constraint on our degrees of belief. However, accepting ($P3_s$) does not provide adequate materials for the argument in favor of REJECTION. For, while ($P1_w$) holds, ($P1_s$) does not. ϕ does not, in general, strongly entail $D\phi$. In fact, the only cases in which it does are ones in which ϕ is valid. So, the only case in which one is required by logic to have the same degree of belief in ϕ and $D\phi$ is when ϕ is a logical validity. Such cases don't provide any trouble since presumably one shouldn't believe that such sentences are indeterminate.

43. See Field 2008, 169 for a discussion of these two notions of entailment.

44. More generally, we can say that a set of sentences Γ strongly entails ψ iff in every model the semantic value of ψ is at least as great as the greatest lower bound of the semantic values of members of Γ.

It is certainly true that we want a notion of entailment that constrains our degrees of belief in the manner specified in (P3). However, adopting the model theory developed by Field for the treatment of indeterminacy does not force us to accept REJECTION in order to meet this desideratum. We can hold that the relevant notion is strong entailment. A minimal point, then, is that a defender of INDETERMINACY does have the resources available to resist this argument in favor of REJECTION.

But such a defender can say something stronger. For I think that there is good independent reason to think that it is strong entailment that should be thought of as having a normative role to play in constraining the degrees of belief of rational agents.

It is often said that belief aims at truth or has truth as its goal. A natural corollary to this thought is the following. The reason why our beliefs should be constrained in the manner described by (P3) is that truth is preserved under entailment. For, if our goal as believers is to believe the truth, then, given that whenever ϕ is true ψ is true, whatever confidence we have in ϕ should also be invested in ψ.

The important point is that if this is the justification for (P3), then it is strong entailment and not weak entailment that is the relevant notion. For while we can say that if ϕ strongly entails ψ, then if ϕ is true, then ψ is true (and of logical necessity), we cannot say the same thing of weak entailment. Let me explain. As noted above, the claim that ϕ strongly entails ψ is equivalent to the claim that $\phi \rightarrow \psi$ is valid. The latter is equivalent to the claim that $T\ulcorner\phi\urcorner \rightarrow T\ulcorner\psi\urcorner$ is valid. So given that ϕ strongly entails ψ, it follows of logical necessity that if ϕ is true, then ψ is true. We cannot, however, say the same for weak validity. It does not follow from the fact that an inference preserves semantic value 1 that it preserves truth; that is, we cannot infer from $\phi \vDash \psi$ to $\vDash T\ulcorner\phi\urcorner \rightarrow T\ulcorner\psi\urcorner$. Since ϕ and $T\ulcorner\phi\urcorner$ are intersubstitutable, the explanation for this is that the deduction theorem fails for weak validity. And the deduction theorem must fail, for otherwise the Curry paradox could not be given an adequate solution.[45]

It seems to me, then, that the most natural justification for (P3) motivates understanding this principle in terms of strong validity. Not only can the argument under consideration be resisted, but such resistance is independently motivated.

45. See Field 2008, chap. 19 for a discussion of the relationship between the deduction theorem for weak validity and the Curry paradox.

10.3.

Here's a final argument in favor of REJECTION.

The hope of providing an informative analysis of indeterminacy, at least as it applies to the liar paradox, is slim. Certainly standard analyses that have been thought promising in the case of vagueness are hopeless when we are dealing with semantic paradoxes. How, then, one might ask, are we to understand what it is for a proposition to be indeterminate?

A proponent of REJECTION can say the following. While we cannot provide an analysis of indeterminacy, we can come to understand the concept by seeing the role that it plays in our cognitive lives. Indeterminacy would, in this way, be like objective chance.[46]

I, however, can say no such thing. For I think that we need to *use* the concept of indeterminacy in order to characterize the distinctive cognitive role of indeterminacy.

REJECTION, then, gives us an independent grip on indeterminacy that INDETERMINACY does not. And this, so the argument goes, is a significant advantage of REJECTION over INDETERMINACY.

It must be admitted that it is a cost of my view that it deprives us of this independent grip on the concept of indeterminacy. Nonetheless, I suggest that on careful inspection the asymmetry between myself and the proponent of REJECTION is not that great.

To see this point, first recall that an adequate treatment of the liar paradox that avails itself of the notion of indeterminacy requires that there be a hierarchy of nonequivalent indeterminacy-type operators. The question arises, then, for each $n > 1$, what attitude should one have toward a proposition ϕ that one ought to take to be indeterminate^n? Let's focus on a simple case. Let ϕ be a proposition that one ought to believe is indeterminately indeterminate. What attitude should one have toward ϕ?

Here are the reasonable responses that I think are available to the proponent of REJECTION:

46. It is plausible to think, first, that objective chance cannot be analyzed in more basic terms and, second, that at least a large part of our understanding of objective chance consists in our knowing that whatever objective chance is it should play something like the following role in our cognitive lives: if one is rational and one believes that the chance of ϕ occurring is x, and one has no additional information about ϕ, then one will have credence x in ϕ. See Lewis 1986 for an argument that this provides the foundation for our understanding of objective chance. See Field 2003a, 479 for a comparison between indeterminacy and chance in this respect.

(i) The proponent of REJECTION may say that in every such case one should reject ϕ and its negation.

(ii) The proponent of REJECTION may say that, in at least some such cases, one should be such that it is indeterminately indeterminate whether one believes ϕ.

(i) can be motivated as follows. The proponent of REJECTION holds that it is a consequence of the claim that one ought to believe $\neg D\phi$ and $\neg D \neg \phi$ that one ought to reject ϕ and its negation. One who accepts this should, I think, be inclined to accept also the more general claim:

> **REJECTION*** For any proposition ϕ, it is a consequence of the claim that one ought to reject $D\phi$ and $D \neg \phi$ that one ought to reject ϕ and its negation.[47]

Using REJECTION*, we can provide an argument that $OBII\phi \vDash OR\phi \wedge OR \neg \phi$.

To argue for this, it will suffice to argue for the following claims:

(20) $OBII\phi \vDash OR \neg I\phi$

(21) $OR \neg I\phi \vDash OR\phi \wedge OR \neg \phi$

(20) is an obvious consequence of REJECTION.

(21) can be established by the following simple argument: First note that $\neg I\phi$ is equivalent to $D\phi \vee D \neg \phi$. We then have $OR \neg I\phi \vDash OR(D\phi \vee D \neg \phi)$. The following strikes me as a nonnegotiable norm governing rejection: $OR(\gamma \vee \psi) \vDash OR\gamma \wedge OR\psi$. In particular, then, we have $OR(D\phi \vee D \neg \phi) \vDash ORD\phi \wedge ORD\neg \phi$. So by transitivity of entailment, we have $OR\neg I \phi \vDash ORD\phi \wedge ORD \neg \phi$. By REJECTION*, we have $ORD\phi \wedge ORD\neg \phi \vDash OR\phi \wedge OR \neg \phi$. And so, finally, we have $OR \neg I\phi \vDash OR\phi \wedge OR \neg \phi$, which is (21).

The only reason that I can see for a proponent of REJECTION to not accept (i) (and so to try to find a way of resisting this argument) would be if she wanted to treat β, or some other similar proposition, as being indeterminately indeterminate, as a way of addressing the normative paradox developed in section 2. In this case, however, in order to avoid resurrecting the normative paradox, the proponent of REJECTION should hold that the correct response to the perceived higher-order in-

47. Field, for example, would accept this more general statement. For, as noted, he thinks that it is a rational requirement that one's degree of belief in ϕ be the same as one's degree of belief in $D\phi$.

determinacy is for it to be indeterminately indeterminate whether the agent believes the proposition in question. Thus if the proponent of REJECTION doesn't accept (i), she should accept (ii).

Now in either case we can argue that the the explanatory asymmetry between REJECTION and INDETERMINACY is not as great as it might at first appear.

If, on the one hand, the proponent of REJECTION decides to accept (i), then we can argue as follows. It is true that REJECTION gives us a grip on indeterminacy in terms that do not presuppose an understanding of this concept. Nonetheless REJECTION does not distinguish between indeterminacy and indeterminate indeterminacy, for there are equivalent principles governing this latter operator. While being indeterminate and being indeterminately indeterminate are not the same, we cannot understand the difference between them solely by appeal to the attitudes of rational agents toward propositions not involving indeterminacy.

If, on the other hand, the proponent of REJECTION decides to accept (ii), then we can argue as follows. While it is true that REJECTION does provide us with an independent grip on indeterminacy, the proponent of REJECTION must still allow that there are other indeterminacy-type statuses, in particular indeterminate indeterminacy, whose cognitive role cannot be specified in terms that don't presuppose an understanding of that very notion of higher-order indeterminacy.

In either case, the conclusion to be drawn is that the proponent of REJECTION can't provide us with an understanding of indeterminate indeterminacy and how it differs from indeterminacy by appeal solely to an independently specifiable cognitive role. In the first case, we can understand the cognitive role without having an independent grip on indeterminate indeterminacy. However, the cognitive role is insufficient to distinguish indeterminate indeterminacy from indeterminacy. In the second case, the cognitive role does distinguish the higher-order status from first-order indeterminacy. However, we can't understand this cognitive role without already having an understanding of indeterminate indeterminacy.

How, then, can a proponent of REJECTION account for our understanding of indeterminate indeterminacy and the difference between this status and simple indeterminacy? This is a difficult question. In outline, what I think the proponent of REJECTION will have to say is that our understanding of indeterminate indeterminacy and the way in which it differs from indeterminacy comes from our grasping (a) the different paradigm cases in which the higher-order operator applies and (b) the difference in its logical behavior from other indeterminacy-type operators. The difference between the proponent of REJECTION and myself is that I don't have, in addition to these resources, facts about cognitive significance to help explain our understanding of various indeterminacy-type operators. Now, it is certainly preferable ceteris paribus to have available more resources in order to explain the primitives of one's theory. But giving up one's right to appeal to facts about attitudes in one's explanation of indeterminacy does not strike me as intolerable given that one must, in any case, avail oneself of other resources in order to fully account for our understanding of related concepts. It is a cost of accepting INDETERMINACY that one can no longer appeal to attitudes in order to explain indeterminacy. But if, as I've argued, there are arguments that point strongly in favor of INDETERMINACY, then this is a cost, I think, worth incurring.

11. Conclusion

We started with the Normative Question: what attitude should a rational agent take toward a proposition that it takes to be indeterminate? The answer to this question is, I claimed, not at all obvious. Nonetheless, there has been a strong consensus that the correct answer is provided by REJECTION. I've argued, however, that attention to the normative paradox raised by Agent Alpha motivates, instead, INDETERMINACY as the answer to this question.

According to the picture that emerges, indeterminacy in the objects of a rational agent's doxastic states will filter up to the attitudes themselves. Such an agent's doxastic states will exhibit the same paradoxical features as the objects of its attitudes. If this is right, then the study of the semantic paradoxes has broader implications than has traditionally been thought. Attention to the semantic paradoxes can provide surprising insights into the nature of rationality and of the mental states of rational agents.

A. Technical Appendix

In this appendix, I outline the underlying technical machinery appealed to in the body of the article and prove certain results that are important for the arguments therein.

A.1. General Framework

The model theory that I'll sketch uses tools developed by Hartry Field and extends them to treat languages involving modal operators. For a full development of Field's model theory, see Field 2008, chaps.15–17.

The Field-style models that we'll be constructing make essential use of the fixed-point construction developed in Kripke 1975. Let me first give a brief sketch of how this construction works for a language involving modal operators. For a detailed exposition of how such constructions proceed (for languages not involving modal operators), see Kripke 1975.[48]

Kripke Models: Let us start with a standard classical model M for a language L containing a modal operator $\Box$ but not containing a truth predicate. M will be a quadruple $< D_m, \Delta_m, R_m, I_m >$. D_m is the domain of individuals. Δ_m is a set of points relative to which sentences are assigned truth-values. R_m is a relation of "accessibility" holding between members of Δ. I_m is an interpretation function that assigns classical values to the elements of L relative to members of Δ. Truth relative to a point and a sequence is defined in the standard way.

Let L^+ be the language that results by adding a truth predicate T to L. I'll now show how we can extend M to a model M^+ for the language L^+. M^+ will be a nonclassical model. M^+ will be identical to M except that I_m^+ will assign a semantic value to T relative to members of Δ. Unlike in a classical model, however, T is not assigned a single extension relative to a point δ. Instead it is assigned an extension $T^{\delta+}$ and an antiextension $T^{\delta-}$. $T^{\delta+}$ will be a set of sentences. $T^{\delta-}$ will be a set consisting of all nonsentences together with any sentence whose negation is in $T^{\delta+}$. Not every sentence will be in $T^{\delta+} \cup T^{\delta-}$. In this model, sentences can receive one of three semantic values relative to a point of evaluation: 1, 1/2, or 0. Sentences in the base language L will receive only values 1 and 0, but sentences involving the new predicate T may receive value 1/2.

M^+ is constructed by considering a series of models M_α^+. Each M_α^+ is, like M^+, identical to M except that I_α^+ assigns an extension and antiextension to T relative to each member of Δ. We let $T_\alpha^{\delta+}$ denote the extension of T under M_α^+. $T_\alpha^{\delta-}$ will denote the antiextension as defined above.

48. There are, in addition, a number of secondary source expositions. Two useful sources are Field 2008, chap. 3 and Soames 1999, chap. 6.

We first provide an inductive definition of what it is for a formula γ of L^+ to have a semantic value 1, 1/2, or 0 relative to a sequence s and point δ under M^+_α.

- If γ is a formula of the form $P(t_1, t_2 \ldots t_n)$, where P is a predicate of L, then:
  $[\![\gamma]\!]^{s,\delta,M^+_\alpha} = 1 \quad \text{iff} \quad < [\![t_1]\!]^{s,\delta,M^+_\alpha}, \ldots, [\![t_n]\!]^{s,\delta,M^+_\alpha} > \in [\![P]\!]^{s,\delta,M^+_\alpha}$
  Otherwise $[\![\gamma]\!]^{s,\delta,M^+_\alpha} = 0$
- If γ is a formula of the form $T(t)$, then:
  $[\![\gamma]\!]^{s,\delta,M^+_\alpha} = 1$ iff $[\![t]\!]^{s,\delta,M^+_\alpha} \in T^{\delta+}_\alpha$
  $[\![\gamma]\!]^{s,\delta,M^+_\alpha} = 0$ iff $[\![t]\!]^{s,\delta,M^+_\alpha} \in T^{\delta-}_\alpha$
  Otherwise $[\![\gamma]\!]^{s,\delta,M^+_\alpha} = 1/2$
- If γ is a formula of the form $\neg\phi$, then:
  $[\![\gamma]\!]^{s,\delta,M^+_\alpha}$ is equal to $1 - [\![\phi]\!]^{s,\delta,M^+_\alpha}$
- If γ is a formula of the form $\phi \wedge \psi$, then:
  $[\![\gamma]\!]^{s,\delta,M^+_\alpha}$ is equal to min $\{[\![\phi]\!]^{s,\delta,M^+_\alpha}, [\![\psi]\!]^{s,\delta,M^+_\alpha}\}$
- If γ is a formula of the form $\phi \vee \psi$, then:
  $[\![\gamma]\!]^{s,\delta,M^+_\alpha}$ is equal to max $\{[\![\phi]\!]^{s,\delta,M^+_\alpha}, [\![\psi]\!]^{s,\delta,M^+_\alpha}\}$
- If γ is a formula of the form $\forall x\phi$, then:
  $[\![\gamma]\!]^{s,\delta,M^+_\alpha}$ is equal to min $\{[\![\phi]\!]^{s',\delta,M^+_\alpha} : s' \in \Sigma^{s/x}\}$, where $\Sigma^{s/x}$ is the set of sequences differing from s at most with respect to x.
- If γ is a formula of the form $\Box\,\phi$, then:
  $[\![\gamma]\!]^{s,\delta,M^+_\alpha}$ is equal to min $\{[\![\phi]\!]^{s,\delta',M^+_\alpha} : \delta' \in \Delta^\delta\}$, where Δ^δ is the set of points δ' such that $\delta R_m \delta'$.

We say that a sentence ϕ has a semantic value x relative to a model M^+_α and a point δ, $[\![\phi]\!]^{\delta,M^+_\alpha} = x$, iff for all s $[\![\phi]\!]^{s,\delta,M^+_\alpha} = x$. Note that if ϕ is a sentence, then $[\![\phi]\!]^{s,\delta,M^+_\alpha} = [\![\phi]\!]^{s',\delta,M^+_\alpha}$ for all sequences s and s'.

We now construct a series of models M^+_α. We start by setting $T^{\delta+}_0 = \emptyset$ for all δ. At stage $\alpha + 1$, we let $T^{\delta+}_{\alpha+1}$ be the set of sentences ϕ such that $[\![\phi]\!]^{\delta,M^+_\alpha} = 1$. At a limit stage λ, we let $T^{\delta+}_\lambda$ be the set of sentences ϕ such that for some $\beta < \lambda \quad [\![\phi]\!]^{\delta,M^+_\beta} = 1$.

> **FIXED POINT THEOREM** There exists a least ordinal σ such that, for every δ, the set of sentences ϕ such that $[\![\phi]\!]^{\delta,M^+_\sigma} = 1$ is identical to $T^{\delta+}_\sigma$.

We let M^+ be M^+_σ.

Proof Sketch for the FIXED POINT THEOREM: Take some arbitrary well-ordering of the set Δ_m. This will associate with each $\delta \in$

Δ_m a corresponding ordinal up to some ordinal ξ.[49] Using this well-ordering, we can associate with any model M^+_α a sequence $< A_\beta : \beta \in \xi >^{M^+_\alpha}$, where A_β is the set of sentences that M^+_α assigns as an extension for T relative to the point associated with ordinal β.

Let $\leq$ be the following partial-order on such sequences: $< A_\beta : \beta \in \xi > \leq < B_\beta : \beta \in \xi >$ iff for every $\beta \in \xi$ $A_\beta \subseteq B_\beta$.

A simple inductive argument will show that the following property holds for every formula ϕ:

> (**) For every point δ and sequence s, if $< A_\beta : \beta \in \xi >^{M^+_\alpha} \leq < B_\beta : \beta \in \xi >^{M^+_\beta}$, then (a) if $[\![\phi]\!]^{s,\delta,M^+_\alpha} = 1$ then $[\![\phi]\!]^{s,\delta,M^+_\beta} = 1$ and (b) if $[\![\phi]\!]^{s,\delta,M^+_\alpha} = 0$ then $[\![\phi]\!]^{s,\delta,M^+_\beta} = 0$.[50]

It is a consequence of (**) that in the above construction if $\alpha \leq \beta$, then $< A_\beta : \beta \in \xi >^{M^+_\alpha} \leq < B_\beta : \beta \in \xi >^{M^+_\beta}$. As our construction proceeds, the extensions assigned to T relative to a point will either stay the same or grow. Growth, however, cannot continue indefinitely. Here's a way of seeing why this is the case. Imagine a construction in which we added at each stage a single sentence to the extension of T at a single point—the minimal amount of growth possible. Given such a construction, there will be some ordinal by which we will have run out of sentences to add. Let C^S be the cardinality of the set of sentences. We can think of ourselves as having a number of such sets, namely, one for every point in Δ_m. Let C^{Δ_m} be the cardinality of Δ_m. In total we have a set of sentences of cardinality $C^S \times C^{\Delta_m}$. There are, however, ordinals of greater cardinality than this. Once we have reached an ordinal of size greater than $C^S \times C^{\Delta_m}$, we will have run out of sentences to add to points. There will be some ordinal σ, then, such that $< A_\beta : \beta \in \xi >^{M^+_\sigma} = < B_\beta : \beta \in \xi >^{M^+_{\sigma+1}}$. And since the ordinals are well-ordered, there will be a least such ordinal.

Note that under M^+ a sentence ϕ and $T\ulcorner\phi\urcorner$ will have the same semantic value relative to a point δ and a sequence s, where $\ulcorner\phi\urcorner$ is any term that denotes ϕ under M^+ relative to δ and s. In general, then, substituting an occurrence of ϕ for $T\ulcorner\phi\urcorner$ in a sentence ψ will not change the semantic value of ψ under M^+ relative to δ and s.

49. By Zermelo's Theorem, any set can be well-ordered. Δ_m will be set-sized. It follows that there will be some well-ordering of this set.

50. For the sake of concision, I omit the proof of this here. The proof, however, is a fairly straightforward generalization of the proof, given in Kripke 1975, of the monotonicity of the function mapping extensions of T to formulas having semantic value 1 given that assignment.

Field Models: We will now show how to construct a model for a language L^{++}, which, like L^{+}, contains a truth predicate, but in addition contains a conditional $\rightarrow$. (Using the resources of this language, a determinacy operator D can be defined as $D\phi \leftrightarrow_{df} \phi \wedge \neg(\phi \rightarrow \neg\phi)$.) The construction developed by Field involves a transfinite sequence of Kripke constructions for a language not involving a modal operator. It is straightforward to apply the same sort of construction using instead a transfinite sequence of Kripke constructions for a language involving a modal operator.

The construction works as follows. At each stage in the sequence, we begin with a model M^{α}. M^{α} assigns to the elements of language L the assignments provided by M, it assigns to T the null set as the extension at each point δ, and, in addition, it assigns, relative to a sequence s and point δ, semantic values to formulas that have $\rightarrow$ as their main connective. Given such a starting model, we then construct a Kripke model M^{α}_{+} using the method described above.[51]

Consider an arbitrary formula with $\rightarrow$ as its main connective: $\phi \rightarrow \psi$. The assignment $[\![\phi \rightarrow \psi]\!]^{s,\delta,M^{\alpha}}$ is determined as follows:

- For all s and δ, $[\![\phi \rightarrow \psi]\!]^{s,\delta,M^{0}} = 1/2$.
- For all s and δ, $[\![\phi \rightarrow \psi]\!]^{s,\delta,M^{\alpha+1}} = 1$ iff $[\![\phi]\!]^{s,\delta,M^{\alpha}_{+}} \leq [\![\psi]\!]^{s,\delta,M^{\alpha}_{+}}$; otherwise, $[\![\phi \rightarrow \psi]\!]^{s,\delta,M^{\alpha+1}} = 0$.
- For limit ordinal λ, for all s and δ, $[\![\phi \rightarrow \psi]\!]^{s,\delta,M^{\lambda}} = 1$ iff there exists some stage $\beta < \lambda$ such that for all σ, $\beta \leq \sigma < \lambda$, $[\![\phi]\!]^{s,\delta,M^{\sigma}_{+}} \leq [\![\psi]\!]^{s,\delta,M^{\sigma}_{+}}$.
- For limit ordinal λ, for all s and δ, $[\![\phi \rightarrow \psi]\!]^{s,\delta,M^{\lambda}} = 0$ iff there exists some stage $\beta < \lambda$ such that for all σ, $\beta \leq \sigma < \lambda$, $[\![\phi]\!]^{s,\delta,M^{\sigma}_{+}} > [\![\psi]\!]^{s,\delta,M^{\sigma}_{+}}$; otherwise, $[\![\phi \rightarrow \psi]\!]^{s,\delta,M^{\lambda}} = 1/2$.

At each stage α, a formula ϕ will receive a semantic value relative to a sequence s and point δ, given the resulting Kripke model at that stage M^{α}_{+}. Certain formulas will at some point in this series stabilize at value 1 (relative to s and δ). We say that such formulas have ultimate value 1 (relative to s and δ). Other formulas will eventually stabilize at value 0 (relative to s and δ). In this case we say that such formulas have ultimate value 0 (relative to s and δ). Other formulas will either eventually stabilize at value 1/2 (relative to s and δ) or will never stabilize at any value. These

51. Typography note. Do not confuse M^{α}_{+} with M^{+}_{α}. The latter is the model at the α stage of the Kripke construction. The former, which we are now considering, is the Kripke model at the α stage of Field's construction.

formulas we say have ultimate value $1/2$ (relative to s and δ). Following Field we denote the ultimate value of a formula relative to a sequence s and point δ $|||\phi|||^{s,\delta}$.

This resulting assignment of values gives us our model. Ultimate value 1 is the designated value. We can then define validity within a class of modal Field-models $\mathcal{M}$ as follows. We say that for sentences ϕ and ψ, $\phi \vDash_{\mathcal{M}} \psi$ iff for every $M \in \mathcal{M}$, and every $\delta \in \Delta_m$, if $|||\phi|||^{\delta,M} = 1$, then $|||\psi|||^{\delta,M} = 1$.

I note that the following result holds:

FUNDAMENTAL THEOREM For any ordinal ρ, there exists an ordinal $\xi > \rho$ such that for every formula ϕ, sequence s, and point δ, if $|||\phi|||^{s,\delta} = x$, then $[\![\phi]\!]^{s,\delta,M_+^{\xi}} = x$.[52]

Ordinals such as ξ are called "acceptable." The existence of such ordinals ensures that, like the Kripke models, the logic induced by this class of models for the fragment L^+ is K_3.

A.2. Some Theorems

Having outlined the manner in which our models are constructed, I will now justify certain claims made in the paper about these models.

Let M be a classical possible-worlds model for which R_m is an equivalence relation, that is, a transitive, reflexive, and symmetric relation. Let B_α be treated in the model in the manner specified for $\Box$ above. For all s and δ, the following hold in the model generated by the Field construction.

(8) $|||B_\alpha T(\beta) \leftrightarrow B_\alpha B_\alpha T(\beta)|||^{s,\delta} = 1$

For each stage α in the construction $[\![B_\alpha T(\beta)]\!]^{s,\delta,M_+^{\alpha}} = [\![B_\alpha B_\alpha T(\beta)]\!]^{s,\delta,M_+^{\alpha}}$. So at each stage $\alpha > 0$, $[\![B_\alpha T(\beta) \leftrightarrow B_\alpha B_\alpha T(\beta)]\!]^{s,\delta,M_+^{\alpha}} = 1$.

(9) $||| \neg B_\alpha T(\beta) \leftrightarrow B_\alpha \neg B_\alpha T(\beta)|||^{s,\delta} = 1$

As above, for each stage α in the construction $[\![\neg B_\alpha T(\beta)]\!]^{s,\delta,M_+^{\alpha}} = [\![B_\alpha \neg B_\alpha T(\beta)]\!]^{s,\delta,M_+^{\alpha}}$. So at each stage $\alpha > 0$, $[\![\neg B_\alpha T(\beta) \leftrightarrow B_\alpha \neg B_\alpha T(\beta)]\!]^{s,\delta,M_+^{\alpha}} = 1$.

(10) $|||B_\alpha(T(\beta) \leftrightarrow \neg B_\alpha T(\beta))|||^{s,\delta} = 1$

52. Again, for the sake of concision, this proof is omitted here. The proof is a fairly straightforward generalization of the proof of the Fundamental Theorem given in Field 2008.

For each stage $\alpha > 0$ in the construction and each point δ', $[\![T(\beta)\leftrightarrow \neg B_\alpha T(\beta)]\!]^{s,\delta',M_+^\alpha} = 1$. It follows that for every stage $\alpha > 0$, $[\![B_\alpha(T(\beta)\leftrightarrow \neg B_\alpha T(\beta))]\!]^{s,\delta,M_+^\alpha} = 1$.

(11) $|||B_\alpha\phi \rightarrow \neg B_\alpha \neg\phi|||^{s,\delta} = 1$

Assume $[\![B_\alpha\phi]\!]^{s,\delta,M_+^\alpha} = 1$. It follows that for every point δ', $[\![\phi]\!]^{s,\delta',M_+^\alpha} = 1$, and so $[\![\neg\phi]\!]^{s,\delta',M_+^\alpha} = 0$. It follows that $[\![B_\alpha \neg\phi]\!]^{s,\delta,M_+^\alpha} = 0$ and so $[\![\neg B_\alpha \neg\phi]\!]^{s,\delta,M_+^\alpha} = 1$.

Assume $[\![B_\alpha\phi]\!]^{s,\delta,M_+^\alpha} = 1/2$. It follows that for no point δ' is it the case that $[\![\phi]\!]^{s,\delta',M_+^\alpha} = 0$, and for some point δ' it is the case that $[\![\phi]\!]^{s,\delta',M_+^\alpha} = 1/2$. From this it follows that every point δ' is such that either $[\![\neg\phi]\!]^{s,\delta',M_+^\alpha} = 0$ or $[\![\neg\phi]\!]^{s,\delta',M_+^\alpha} = 1/2$. Thus, $[\![B_\alpha \neg\phi]\!]^{s,\delta,M_+^\alpha} \leq 1/2$, and so $[\![\neg B_\alpha \neg\phi]\!]^{s,\delta,M_+^\alpha} \geq 1/2$.

For every stage α, then, $[\![B_\alpha\phi]\!]^{s,\delta,M_+^\alpha} \leq [\![\neg B_\alpha \neg\phi]\!]^{s,\delta,M_+^\alpha}$. And so for every stage $\alpha > 0$, $[\![B_\alpha\phi \rightarrow \neg B_\alpha \neg\phi]\!]^{s,\delta,M_+^\alpha} = 1$.

(12) Let Γ be a set of sentences. Let $B_\alpha\Gamma$ stand for the set of sentences that result by appending B_α to each member of Γ. For any $\mathcal{M}$, $\Gamma \vDash_{\mathcal{M}} \psi \Rightarrow B_\alpha\Gamma \vDash_{\mathcal{M}} B_\alpha\psi$.

Assume that $\Gamma \vDash_{\mathcal{M}} \psi$. Take an arbirary $M \in \mathcal{M}$ and assume that in this model $|||B_\alpha\gamma|||^{\delta} = 1$ for every $B_\alpha\gamma \in B_\alpha\Gamma$. We'll argue that $|||B_\alpha\psi|||^{\delta} = 1$. At every point δ' such that $\delta R\delta'$, $|||\gamma|||^{\delta'} = 1$, for every $\gamma \in \Gamma$. Given that $\Gamma \vDash_{\mathcal{M}} \psi$, it follows that at every such δ', $|||\psi|||^{\delta'} = 1$. And so $|||B_\alpha\psi|||^{\delta} = 1$. Thus $B_\alpha\Gamma \vDash_{\mathcal{M}} B_\alpha\psi$.

(13) $|||I \neg B_\alpha T(\beta) \leftrightarrow B_\alpha I \neg B_\alpha T(\beta)|||^{s,\delta} = 1$

Let α be an arbitrary ordinal. I claim that $[\![I \neg B_\alpha T(\beta)]\!]^{s,\delta,M_+^\alpha} = [\![B_\alpha I \neg B_\alpha T(\beta)]\!]^{s,\delta,M_+^\alpha}$. To see this, note that $I \neg B_\alpha T(\beta)$ has the same semantic value relative to each point in Δ_m, given M_+^α. For we have $I \neg B_\alpha T(\beta) \leftrightarrow (\neg DB_\alpha T(\beta)\wedge \neg D \neg B_\alpha T(\beta)) \leftrightarrow [\neg(B_\alpha T(\beta) \wedge \neg(B_\alpha T(\beta) \rightarrow \neg B_\alpha T(\beta)))\wedge \neg(\neg B_\alpha T(\beta)\wedge \neg(\neg B_\alpha T(\beta) \rightarrow B_\alpha T(\beta)))]$. Given that R_m is an equivalence relation, it follows that each of $B_\alpha T(\beta)$, $B_\alpha T(\beta) \rightarrow \neg B_\alpha T(\beta)$ and $\neg B_\alpha T(\beta) \rightarrow B_\alpha T(\beta)$ have the same semantic value relative to each point of evaluation, given M_+^α. This is sufficient to guarantee that $I \neg B_\alpha T(\beta)$ has the same semantic value x relative to each point of evaluation. And whatever value x is, that will be the same value that $B_\alpha I \neg B_\alpha T(\beta)$ has at every point of evaluation. In particular, $[\![I \neg B_\alpha T(\beta)]\!]^{s,\delta,M_+^\alpha} = [\![B_\alpha I \neg B_\alpha T(\beta)]\!]^{s,\delta,M_+^\alpha}$. It follows that for every $\alpha > 0$, $[\![I \neg B_\alpha T(\beta) \leftrightarrow B_\alpha I \neg B_\alpha T(\beta)]\!]^{s,\delta,M_+^\alpha} = 1$.

(18) $\neg D\phi \wedge \neg D\psi \wedge \neg D\xi \vDash \neg D(\phi \vee \psi \vee \xi)$

Instead of proving (18), I'll sketch the proof for the simpler $\neg D\phi \wedge \neg D\psi \vDash \neg D(\phi \vee \psi)$. The proof of (18) is a straightforward generalization of this. The cases one needs to deal with are greater but add no further complexities.

Let $|||\neg D\phi \wedge \neg D\psi||| = 1$. We want to show $|||\neg D(\phi \vee \psi)||| = 1$. By the FUNDAMENTAL THEOREM, we know that there is a class of acceptable ordinals. Let's denote the least such ordinal σ. To show that $|||\neg D(\phi \vee \psi)||| = 1$ follows, on the assumption that $|||\neg D\phi \wedge \neg D\psi||| = 1$, it suffices to show that $[\![\neg D(\phi \vee \psi)]\!]^{M^\sigma_+} = 1$ follows, on the assumption that $[\![\neg D\phi \wedge \neg D\psi]\!]^{M^\sigma_+} = 1$.

Now, $[\![\neg D\phi \wedge \neg D\psi]\!]^{M^\sigma_+} = 1$ just in case $[\![\neg(\phi \wedge \neg(\phi \rightarrow \neg\phi)) \wedge \neg(\psi \wedge \neg(\psi \rightarrow \neg\psi))]\!]^{M^\sigma_+} = 1$. And this holds just in case either (i) $[\![\phi]\!]^{M^\sigma_+} = [\![\psi]\!]^{M^\sigma_+} = 0$, or (ii) $[\![\phi]\!]^{M^\sigma_+} = 0$ and $[\![\psi \rightarrow \neg\psi]\!]^{M^\sigma_+} = 1$, or (iii) $[\![\psi]\!]^{M^\sigma_+} = 0$ and $[\![\phi \rightarrow \neg\phi]\!]^{M^\sigma_+} = 1$, or (iv) $[\![\phi \rightarrow \neg\phi]\!]^{M^\sigma_+} = [\![\psi \rightarrow \neg\psi]\!]^{M^\sigma_+} = 1$.

$[\![\neg D(\phi \vee \psi)]\!]^{M^\sigma_+} = 1$ just in case $[\![(\phi \vee \psi) \wedge \neg((\phi \vee \psi) \rightarrow \neg(\phi \vee \psi))]\!]^{M^\sigma_+} = 0$. We'll show that the latter holds whichever of (i) – (iv) obtains.

If (i) holds, then $[\![\phi \vee \psi]\!]^{M^\sigma_+} = 0$, and so clearly $[\![(\phi \vee \psi) \wedge ((\phi \vee \psi) \rightarrow \neg(\phi \vee \psi))]\!]^{M^\sigma_+} = 0$.

If (ii) holds, then we have that there exists a $\beta < \sigma$ such that for all γ such that $\beta \leq \gamma < \sigma$, $[\![\psi]\!]^{M^\gamma_+} \leq 1/2$. Since σ is an acceptable ordinal, we can be assured that given that $[\![\phi]\!]^{M^\sigma_+} = 0$, it follows that for all γ such that $\beta \leq \gamma < \sigma$, $[\![\phi]\!]^{M^\gamma_+} = 0$. Thus, for all γ such that $a \leq \gamma < \sigma$, $[\![\phi \vee \psi]\!]^{M^\gamma_+} \leq 1/2$. This assures us that $[\![((\phi \vee \psi) \rightarrow \neg(\phi \vee \psi))]\!]^{M^\sigma_+} = 1$, and so $[\![(\phi \vee \psi) \wedge \neg((\phi \vee \psi) \rightarrow \neg(\phi \vee \psi))]\!]^{M^\sigma_+} = 0$.

Proving the cases for (iii) and (iv) simply involves applying the reasoning from the previous two cases in an obvious way.

(19) For any class of modal Field-models $\mathcal{M}$, $B_\alpha I\phi \vDash_{\mathcal{M}} IB_\alpha\phi$.

Let M be some model in $\mathcal{M}$. Assume that in this model $|||B_\alpha I\phi|||^\delta = 1$. We'll now show that given this assumption, $|||IB_\alpha\phi|||^\delta = 1$.

By the FUNDAMENTAL THEOREM, we know that there is a class of acceptable ordinals. Again denote the least such ordinal "σ." To show that $|||IB_\alpha\phi|||^\delta = 1$ follows, on the assumption that $|||B_\alpha I\phi|||^\delta = 1$, it will suffice to establish:

(D1) If $|||B_\alpha I\phi|||^\delta = 1$, then for all δ' such that $\delta R_m \delta'$, for all $\xi > \sigma$, $[\![\phi]\!]^{\delta', M^\xi_+} = 1/2$.

For if for all δ' such that $\delta R_m \delta'$, for all $\xi > \sigma$, $[\![\phi]\!]^{\delta', M^{\xi}_{+}} = 1/2$, then it follows that for all $\xi > \sigma$, $[\![B_\alpha \phi]\!]^{\delta, M^{\xi}_{+}} = 1/2$. And from this it follows that for all $\xi > \sigma + 2$, $[\![IB_\alpha \phi]\!]^{\delta, M^{\xi}_{+}} = 1$. This can be verified by noting that for any model M^{ξ}_{+} and point δ', $I\phi$ will receive the same semantic value as $(\neg\phi \vee \phi \rightarrow \neg\phi) \wedge (\phi \vee \neg\phi \rightarrow \phi)$.

To show that (D1) holds, we will establish:

(D2) If $|||B_\alpha I\phi|||^{\delta} = 1$, then for all δ' such that $\delta R_m \delta'$, $|||I\phi|||^{\delta'} = 1$.
(D3) If $|||I\phi|||^{\delta'} = 1$, then for all $\xi > \sigma$, $[\![\phi]\!]^{\delta', M^{\xi}_{+}} = 1/2$.

Justification for (D2): (D2) is a straightforward consequence of our treating B_α as a universal quantifier over the set of δ' such that $\delta R_m \delta'$.

Justification for (D3): Assume $|||I\phi|||^{\delta'} = 1$. We'll show that on this assumption it follows that for all $\xi > \sigma$, $[\![\phi]\!]^{\delta', M^{\xi}_{+}} = 1/2$. We do so by reductio. We'll assume that there is some $\xi > \sigma$, $[\![\phi]\!]^{\delta', M^{\xi}_{+}} \neq 1/2$ and show that from this, together with our other assumption, a contradiction can be derived. This is sufficient to establish (D3).

To argue for this we first establish the following:

(D4) If $|||I\phi|||^{\delta'} = 1$, then for all $\xi > \sigma$, if $\exists \sigma' \forall \sigma''\ \sigma' \leq \sigma'' < \xi$, $[\![\phi]\!]^{\delta', M^{\sigma''}_{+}} \neq 1/2$, then $[\![\phi]\!]^{\delta', M^{\xi}_{+}} \neq 1/2$.

Justification for (D4): Assume that $|||I\phi|||^{\delta'} = 1$ and that $\exists \sigma' \forall \sigma''\ \sigma' \leq \sigma'' < \xi\ [\![\phi]\!]^{\delta', M^{\sigma''}_{+}} \neq 1/2$. If $|||I\phi|||^{\delta'} = 1$, then for all $\xi > \sigma$, $[\![I\phi]\!]^{\delta', M^{\xi}_{+}} = 1$.[53] Our first assumption, then, guarantees that $[\![I\phi]\!]^{\delta', M^{\xi}_{+}} = 1$. Given this and our second assumption, it follows that $[\![\phi]\!]^{\delta', M^{\xi}_{+}} \neq 1/2$. To see this, note that for any model M^{ξ}_{+} and point δ', $I\phi$ will receive the same semantic value as $(\neg\phi \vee \phi \rightarrow \neg\phi) \wedge (\phi \vee \neg\phi \rightarrow \phi)$. So we have $[\![(\neg\phi \vee \phi \rightarrow \neg\phi) \wedge (\phi \vee \neg\phi \rightarrow \phi)]\!]^{\delta', M^{\xi}_{+}} = 1$. But it is clear that, given that $\exists \sigma' \forall \sigma''\ \sigma' \leq \sigma'' < \xi\ [\![\phi]\!]^{\delta', M^{\sigma''}_{+}} \neq 1/2$, then in order for the above formula to have value 1, either $[\![\phi]\!]^{\delta', M^{\xi}_{+}} = 1$ or $[\![\phi]\!]^{\delta', M^{\xi}_{+}} = 0$. For given that $\exists \sigma' \forall \sigma''\ \sigma' \leq \sigma'' < \xi\ [\![\phi]\!]^{\delta', M^{\sigma''}_{+}} \neq 1/2$, we cannot have both $[\![\phi \rightarrow \neg\phi]\!]^{\delta', M^{\xi}_{+}} = 1$ and $[\![\neg\phi \rightarrow \phi]\!]^{\delta', M^{\xi}_{+}} = 1$. In order, then, for both disjuncts to have value 1, ϕ must have either value 1 or value 0. This suffices to establish (D4).

53. To see this, note that the sequence of Kripke constructions will eventually fall into a cyclical pattern. This is a consequence of the fact that there are ordinals of greater cardinality than the cardinality of the set of possible functions from formula, sequence, point triples to values 1, 0, 1/2. At some point then, a valuation will reoccur, and this will institute a cyclical pattern. And such a pattern will have been instituted by the time the first acceptable ordinal occurs since this is one of the reoccurring valuations.

From (D4), together with our assumption that $|||I\phi|||^{\delta'} = 1$ and our assumption that there is some $\xi > \sigma$, $[\![\phi]\!]^{\delta', M_+^{\xi}} \neq 1/2$, it follows that for all $\xi' \geq \xi$, $[\![\phi]\!]^{\delta', M_+^{\xi'}} \neq 1/2$. By the FUNDAMENTAL THEOREM, there exists an acceptable ordinal σ' such that $\sigma' > \sigma$. We have that either $[\![\phi]\!]^{\delta', M_+^{\sigma'}} = 1$ or $[\![\phi]\!]^{\delta', M_+^{\sigma'}} = 0$. Both of these, however, conflict with our assumption that $|||I\phi|||^{\delta'} = 1$. Given that $|||I\phi|||^{\delta'} = 1$, it follows that $[\![I\phi]\!]^{\delta', M_+^{\sigma'}} = 1$, and so $[\![(\neg\phi \vee \phi \rightarrow \neg\phi) \wedge (\phi \vee \neg\phi \rightarrow \phi)]\!]^{\delta', M_+^{\sigma'}} = 1$. But this cannot be the case if either $[\![\phi]\!]^{\delta', M_+^{\sigma'}} = 1$ or $[\![\phi]\!]^{\delta', M_+^{\sigma'}} = 0$. Assume $[\![\phi]\!]^{\delta', M_+^{\sigma'}} = 1$. Then in order for $[\![I\phi]\!]^{\delta', M_+^{\sigma'}} = 1$, it must be that $[\![\phi \rightarrow \neg\phi]\!]^{\delta', M_+^{\sigma'}} = 1$. However, since σ' is an acceptable ordinal, it follows that $|||\phi|||^{\delta'} = 1$. And so $\exists\sigma''\forall\sigma'''\ \sigma'' \leq \sigma''' < \sigma\ [\![\phi]\!]^{\delta', M_+^{\sigma'''}} = 1$. It follows that $[\![\phi \rightarrow \neg\phi]\!]^{\delta', M_+^{\sigma'}} = 0$, and so $[\![I\phi]\!]^{\delta', M_+^{\sigma'}} \neq 1$. A similar argument will show that on the assumption $[\![\phi]\!]^{\delta', M_+^{\sigma'}} = 0$, $[\![I\phi]\!]^{\delta', M_+^{\sigma'}} \neq 1$.

So on the assumption that $|||I\phi|||^{\delta'} = 1$ and that there is some $\xi > \sigma$, $[\![\phi]\!]^{\delta', M_+^{\xi}} \neq 1/2$, it follows that $[\![I\phi]\!]^{\delta', M_+^{\sigma'}} = 1$ and $[\![I\phi]\!]^{\delta', M_+^{\sigma'}} \neq 1$. By reductio, it follows, on the assumption that $|||I\phi|||^{\delta'} = 1$, that for all $\xi > \sigma$, $[\![\phi]\!]^{\delta', M_+^{\xi}} = 1/2$. This gives us (D3).

References

Arntzenius, F. 2008. "No Regrets, or: Edith Piaf Revamps Decision Theory." *Erkenntnis* 68: 277–97.

Burge, T. 1978. "Buridan and Epistemic Paradox." *Philosophical Studies* 34: 21–35.

———. 1984. "Epistemic Paradox." *Journal of Philosophy* 81, no. 1: 5–29.

Conee, E. 1982. "Against Moral Dilemmas." *Philosophical Review* 91, no. 1: 87–97.

———. 1987. "Evident, but Rationally Unacceptable." *Australasian Journal of Philosophy* 65: 316–26.

Davidson, D. 1980a. "Actions, Reasons, and Causes." In *Essays on Actions and Events*, 3–19. New York: Oxford University Press.

———. 1980b. "Mental Events." In *Essays on Actions and Events*, 207–27. New York: Oxford University Press.

Feldman, R. 2004. "Having Evidence." In Earl Conee and Richard Feldman, *Evidentialism*, 219–41. New York: Oxford University Press.

Field, H. 2003a. "No Fact of the Matter." *Australasian Journal of Philosophy* 81, no. 4: 457–80.

———. 2003b. "The Semantic Paradoxes and the Paradoxes of Vagueness." In *Liars and Heaps*, ed. J. C. Beall, 262–311. New York: Oxford University Press.

———. 2007. "Solving the Paradoxes, Escaping Revenge." In *Revenge of the Liar*, ed. J. C. Beall, 78–144. New York: Oxford University Press.

———. 2008. *Saving Truth from Paradox*. New York: Oxford University Press.

Gibbard, A., and W. Harper 1978. "Counterfactuals and Two Kinds of Expected Utility." In *Foundations and Applications of Decision Theory*, ed. C. Leach, E. McClennen, and C. Hooker, 25–162. Dordrecht: D. Reidel.
Kripke, S. 1975. "Outline of a Theory of Truth." *Journal of Philosophy* 72, no. 19: 690–716.
Lemmon, E. J. 1962. "Moral Dilemmas." *Philosophical Review* 71, no. 2:139–58.
Lewis, D. 1974. "Radical Interpretation." *Synthese* 23: 331–44.
———. 1986. "A Subjectivist's Guide to Objective Chance." In *Philosophical Papers*, vol. 2, 83–132. New York: Oxford University Press.
———. 1999. "Reduction of Mind." In *Papers in Metaphysics and Epistemology*, 291–324. New York: Cambridge University Press.
Marcus, R. B. 1980. "Moral Dilemmas and Consistency." *Journal of Philosophy* 77, no. 3: 121–36.
McGee, V. 1991. *Truth, Vagueness, and Paradox*. Indianapolis: Hackett.
Osborne, M., and A. Rubenstein. 1994. *A Course in Game Theory*. Cambridge, MA: MIT Press.
Parsons, T. 1984. "Assertion, Denial, and the Liar Paradox." *Journal of Philosophical Logic* 13: 137–52.
Priest, G. 2002. "Rational Dilemmas." *Analysis* 62, no. 1: 11–16.
Richard, M. 2008. *When Truth Gives Out*. New York: Oxford University Press.
Soames, S. 1999. *Understanding Truth*. New York: Oxford University Press.
Sorensen, R. 1988. *Blindspots*. New York: Oxford University Press.
Stalnaker, R. 1984. *Inquiry*. Cambridge, MA: MIT Press.
Van Fraassen, B. C. 1973. "Values and the Heart's Command." *Journal of Philosophy* 70, no. 1: 5–19.
Williams, J. R. G. 2008. "Aristotelian Indeterminacy and the Open Future." Unpublished manuscript.
———. 2011. "Decision Making under Indeterminacy." Unpublished manuscript.
Williamson, T. 2000. *Knowledge and Its Limits*. New York: Oxford University Press.
Yablo, S. 2003. "New Grounds for Naive Truth." In *Liars and Heaps*, ed. J. C. Beall, 312–30. New York: Oxford University Press.