Notre Dame Journal of Formal Logic
Volume 38, Number 4, Fall 1997

Impossible Worlds: A Modest Approach

DANIEL NOLAN

Abstract Reasoning about situations we take to be impossible is useful for a variety of theoretical purposes. Furthermore, using a device of impossible worlds when reasoning about the impossible is useful in the same sorts of ways that the device of possible worlds is useful when reasoning about the possible. This paper discusses some of the uses of impossible worlds and argues that commitment to them can and should be had without great metaphysical or logical cost. The paper then provides an account of reasoning with impossible worlds, by treating such reasoning as reasoning employing counterpossible conditionals, and provides a semantics for the proposed treatment.

1 Introduction Some things just can't happen. Reasoning about such impossibilities, however, seems perfectly possible, and indeed important: I find myself doing it often. Some people have taken the legitimacy of reasoning about impossibilities to show that our logic must be weakened—it is thought that logic must not only cater for relatively well behaved worlds like ours, but must be suitable for dealing with the more logically unruly cases we nevertheless have to consider.[1] I think such reasons are not good reasons to tamper with our logic—we can, for example, keep even classical logic while making adequate room for thinking about impossibilities. In this paper I will outline what I take to be a modest approach to impossibilities, and impossible worlds. Of course, one person's modesty is another's extravagance and often a third's cowardice, so I do not expect this approach will appeal to everyone. It should however serve as a challenge to both those who are suspicious of impossible worlds and those who embrace impossible worlds more thoroughly than I do, by revising their logics to welcome them: since impossible worlds can be had comparatively cheaply, why not accept them? On the other hand, why pay more? In this paper, I will begin by discussing some of the uses impossible worlds have, and why I find reasoning about cases which are not even possible so important. Then I will present my modest proposal and argue that the advantages of impossible worlds can be had rather cheaply. Finally, I will explore some of the important difficulties which remain.

Received November 20, 1997; revised December 16, 1997

2 Why we need impossible worlds There are a variety of areas in which it is useful to be able to reason about impossible situations and to do so in a nontrivial way (so that it is not good enough to just throw up one's hands and say that everything follows). The mere fact that we can think about what is impossible does not commit us to impossible worlds, any more than the mere fact that we claim that some claims are necessary or possible commit us to possible worlds. But just as it is a natural way to cash out our talk of necessity and possibility in terms of possible worlds, it is tempting to talk about impossible worlds, or situations, or ways things couldn't be. My defense of impossible worlds, then, will mostly be a defense of the need to nontrivially reason about claims or theories which we think cannot possibly be correct. If this is established, then the same sorts of reasons that encourage people to move from talking of possibility and necessity to possible worlds can be applied, *mutatis mutandis*, to move from talking of impossibility to impossible worlds.[2]

One of the first reasons which springs to mind for trafficking in the impossible is the understanding of logics which one takes to be incorrect. In one sense, to understand a logic different from the one(s) that one prefers does not require any special reasoning about the impossible: treating a logic as a set of symbols, or as merely representations of a set-theoretic semantics, does not require any suppositions about the impossible (especially when the metalanguages of such logics, and the logics governing the semantics, are not at all unusual: a classical logician can happily experiment with rival "logics" whose metalanguages are classical, and for which the set theory governing their models is classical). However, when we take rival logics to be claims about what really follows from what, or what inferences are licensed using connectives much like the English 'and', 'or', and 'not' (and other connectives), then we do get into the realm of serious disagreement. Nobody denies that there is a formal system where $\sim\sim p \rightarrow p$ fails to be a theorem. But most nonintuitionists balk at the suggestion that there is (or even possibly is) a proposition p such that it is the case that not-not-p but it is not the case that p.

One strategy which has some popularity is to suppose that logicians who disagree with one are really talking about something else (they have "changed the subject," as Quine would say ([26], p. 70): so intuitionists are really talking about provability, or some such; quantum logicians are really only talking about set-theoretic operations of various sorts in Hilbert spaces; "truth-value gap" logicians are talking about sentences rather than propositions, or whatever. And some of the time this strategy is the correct one: some people exploring or applying a "logic" may not take it to tell us anything about what really does follow from what, or how the logical connectives of natural language work (or work, given suitable idealizations). But sometimes this strategy seems to me to be clearly inapplicable: there are genuine disputes in logic, and people do often mean what they appear to mean: some intuitionists, dialetheists, and classical logicians really do disagree with each other about negation (among other things), Aristotelians really do disagree about nonsyllogistic inferences, **S4**-ers and **S5**-ers really do disagree about which modal inferences are acceptable, and so on.

Another strategy is, of course, to take one's logical opponents to be talking incoherent nonsense, and so there is nothing useful to understand. Hardline intuitionists who claim that classical logic, insofar as it goes further than intuitionistic logic, is

incoherent and nonsensical, or makes meaningless assertions, are an example, as are those hardline classical logicians who dismiss deviance as gibberish. These strategies are much less common these days, as far as I can tell, and rightly so: the theory that rival logical schools are saying things that are utterly meaningless, only seem to communicate between themselves, and make no claims which are even assessable for plausibility, seems to me to border on a conspiracy theory. Surely the simplest explanation is that they have a theory: perhaps one which could not even possibly be correct, but one which is at least intelligible.

Finally, there is one more approach to rival logical systems which would allow them to be understood and evaluated without entertaining things one takes to be impossible. Some logical systems have only a proper subset of the axioms or rules of inference of others. When one accepts a strictly stronger logic, it is comparatively easy to understand weaker logical systems, since all of the inference rules of such systems are acceptable: they simply do not exhaust the acceptable inference rules. So, for example, propositional logic does not capture important quantification-involving inferences, but nobody finds the propositional fragment of their preferred logic mysterious or incoherent on that account (even though, for example, there will be models where 'All men are mortal' and 'Socrates is a man' will hold, but 'Socrates is mortal' will not). It is possible to consider weaker logics as merely leaving out some of the principles which could have been added. This approach to weaker logics is not entirely satisfactory in all cases either—defenders of some relevant logic as the one true logic, for example, are genuinely disagreeing with a classical logician in a way that someone who was primarily interested in the propositional fragment of classical logic would not be.

If these are genuine, meaningful disagreements, and at most one of the parties to the various disputes can be correct, then it seems that the other parties are reasoning about, and even believing in, impossibilities (albeit without knowing it). Furthermore, even those who have the good fortune to be correct, if they understand and can draw out the implications of their rivals' views, will need to be able to consider and usefully reason about logical impossibilities. This nontrivial reasoning about impossible worlds is a central feature of much logical debate: and we seem to be quite proficient at it, even those who, in their philosophical moments, claim that such reason and understanding is impossible, or very limited, when it comes to situations or theories which could not even possibly obtain or be correct.

A second area where people reason about the impossible, and where we may have to examine the commitments of impossible theories, is mathematics. One aspect of this reasoning about impossible situations has received a lot of notice: the apparent nontriviality of reductio proofs. When I start from supposing some premises which are in fact inconsistent, there seems all the difference in the world between a proof to the negation of one of the premises via derivation of a reductio, and a proof with the same premises and conclusion, but without the intervening steps (or with intervening steps which seem totally irrelevant: 'p, q, r, so the moon is made of green cheese, so not-r', or some such). I do not think that these cases need the invocation of a distinction between trivial and nontrivial reasoning in the presence of inconsistency in order to distinguish them, however. The difference between an acceptable reductio proof and an unacceptable proof from the same premises can be found elsewhere. Besides

having a proof which is formally valid, we typically require that a proof is obviously formally valid, or at least is such that the rule used at each step is reasonably obviously formally valid. If a proof lacks this feature, then it is not as useful for the purposes to which we put proofs: convincing ourselves of the conclusions, or making sure of our results, or offering pieces of reasoning which we hope will convince, or any of these purposes. A proper reductio, as opposed to a chaotic jumble with steps like 'the moon is made of green cheese', is one where the steps from the supposition to the explicit contradiction (or whatever other *absurdum* which is employed) are obviously (or reasonably obviously) formally valid. So I think that reductio proofs do not show the need for nontrivial reasoning about impossibilities which many take them to, since the reason why some steps of derivation are acceptable and others are not can be explained in terms of the *obviousness* of the validity of some deductions but not others, rather than a distinction between the validity of different possible deductive steps per se.

There is another sort of case which worries me more. In set theory, for example, there are axioms which some accept and others reject: the continuum hypothesis and the axiom of choice are the two most famous. There are also debates about areas of set- and setlike theory: whether there are non-well-founded sets, whether there are proper classes, whether category theory is about a new domain of mathematical objects, or whether its real interpretation is to be found in set theory with extra large-cardinal axioms, and so on. Positions in such debates are often inconsistent with each other (obvious enough when we examine systems, one of which has the continuum hypothesis (or generalized continuum hypothesis) as an axiom, and the other the negation of the continuum hypothesis as an axiom)—yet it is not obvious that either side of such disputes is incoherent: in fact, with many of these issues it has been proved that such systems are consistent if standard set-theoretic systems are (both the axiom of choice and its negation are relatively consistent with Zermelo-Fraenkel set theory, for example—that is, if ZF is consistent with itself, then it is consistent with the axiom of choice, and it is consistent with the negation of the axiom of choice). However, a very common view, and one which I share, is that the truths of mathematics have the same necessary status as the truths of logic:[3] so it seems that at least one of the sides in these debates is committed to a theory which could not even possibly be true.

Some people at least think that there is a sensible question about what mathematical system is correct: for example, many people do not believe in proper classes, whereas I do. My opponents will typically admit that there are interpretations of the axioms of a theory of proper classes which have set-theoretic models which they do believe in, and so in that sense at least are as consistent as the theory in which they have models (see e.g., Fraenkel, Bar-Hillel, and Levy [9], p. 141). If the realm of mathematics is a realm of necessary truths, then it seems, prima facie, one of us holds a belief which is necessarily incorrect. But both sides in this disagreement are perfectly adept at reasoning about what would be true if our opponent's view was correct (as well as reasoning about what would be true if our own view was correct).

Others might think that the debate as set up here is misguided. Perhaps the appearance of dispute is caused by an illegitimate Platonism, and that what is going on is not a dispute about the existence of sorts of objects (or the objective truth of

mathematical propositions) at all. Or perhaps our Platonism is not generous enough: perhaps there are not only Zermelo-Fraenkel sets, but New Foundations sets, and non-well-founded sets, and Gödel-Bernays-von Neumann classes, and Kelly-Morse classes, and categories, and systems of sets with large cardinals, and systems without large cardinals, and some systems with a generalized continuum hypothesis, and some systems without . . . . if all of the mathematical systems happily lived together in Platonic heaven our dispute would not be as it seemed either. Or perhaps the disputes are not as they seem because we are misled into thinking that it is a once-and-for-all issue what sets are like, or how sets and setlike objects (like classes) really are: perhaps the true picture is some sort of modal story, or a structuralist story, or both, or something other again, so that it turns out that there are not real disagreements between rival systems, since they each capture in an equally valid way one of the systems which mathematics could appropriately describe.

Maybe the battle lines in these disputes are drawn in an unobvious way—perhaps the simple picture I put forward of partisans of different axiom systems being in the uncomfortable position of at most one being right is not the right picture. Nevertheless, when partisans of rival systems do disagree—disagree about what system is right, rather than merely which one is more convenient, or which one is the most interesting, or which lends itself to interesting applications—then still, at most one can be right. For even if, say a partisan of Zermelo-Fraenkel and a partisan of New Foundations could both be right, when each condemns the other (perhaps by saying "your axioms, taken together, are false") still at most one can be correct—indeed, if they are both correct as far as their positive views go (say, because each system is present separately in Gödelian heaven, or because there are appropriately interpretable structures for each), then both will be *incorrect* in their rejection of the truth of their rival's system. Indeed, regardless of the exact truth of the matter, provided only that people are coherently disagreeing with each other, then some will be coherently operating with a mathematical framework which is necessarily incorrect, should mathematics be necessary.[4]

A third area which I find it important to have the ability to coherently discuss impossibilities is in metaphysics. Many metaphysical views seem to be such that if they are true at all, they are necessarily true, and if false, necessarily so: yet rivals understand each other, and we metaphysicians flatter ourselves that we are engaging in real debates, where argument and invocation of considerations are important: we are not babbling mere nonsense, even when some of our number (or many of our number) fall into necessary falsehood. The metaphysics of modality and possible worlds is only the most obvious example: when a metaphysical picture commits one to claims about the nature of possible worlds, and modal claims as a result about what is and is not possible (like Lewis's denial that there could be several disconnected spacetimes not otherwise connected by some special natural external relations[5]), it is often involved in commitments that are necessarily false if certain of its rivals turn out to be true instead. Nevertheless, debate over modal questions continues, and exploration of systems of modal metaphysics other than the sort one accepts is a standard part of such investigation, both to see if one can put one's finger on what one finds unattractive, and to see whether one should switch one's views to a rival which proves more plausible.

As well as the debate about modality and possible worlds themselves, there are closely related debates with modal implications. Consider the debate about identity. There are views according to which I am not identical to the sum of my parts, but am only constituted by it; views according to which I am identical to the aggregate of my parts, but only contingently; views according to which I am essentially identical to the aggregate of my parts; and more besides. It seems plausible that if one of these alternatives is true, the others are necessarily false.[6]

Yet we need to reason about what would follow from these various theories if we are to hope to discover which one is the most plausible. There are other debates which seem to me to have modal implications: the debate between a realist about properties who claims that there must be a property of redness if a rose is to be red (and indeed that properties and relations are needed to underwrite all, or at any rate virtually all, predication), and a nominalist who claims that roses are red without there being any such property (and indeed that predication never needs to be underwritten by properties or relations), does not seem to be a debate about a matter which may be true in some worlds and false in others. Similarly the debate about whether normativity could reduce to dispositions seems to be one where many positions, if they are right at all, are necessarily so. I find myself, if I am to seriously and sympathetically understand and evaluate rival positions, forced to consider a range of options, many of which are impossible (though we will argue about which are the impossible ones, of course). It seems I must think about, and distinguish between, ways the world could not have turned out, as well as ways that it could if I am to best work out which are the impossible ones and which one is the actual one, or which of a small handful of the alternatives are genuinely possible.

In all of these areas, discussing alternatives to the ways things could possibly be seems important. Of course, one could balk at moving from discussing different impossibilities to acceptance of the existence of different impossible worlds: just as one could balk at the equivalent move of going from talk of what is possible to talk of different possible worlds. But it certainly simplifies matters to be able to talk, for instance, about a world where Leibniz's metaphysics is correct—and it seems proper to do so even while the question of whether such an account is even possibly correct is still undecided (I happen to think that Leibniz's metaphysical picture is necessarily false, as, I believe, is Spinoza's—but I think it would be an absurd historian of philosophy who could not agree that the two systems are different, and different things are true according to each of them). In any case, I find myself inextricably committed to reasoning about systems, not all of which seem to be possible: thus, I find I need to be able to reason in a nontrivial way about impossible cases. This is by no means the only sort of consideration which drives people to want to draw a distinction between good and bad reasoning when thinking about impossible cases: the need to reason with inconsistent information, or inconsistent beliefs, is a pressing one for many concerned with reasoning or with modeling rational belief revision. We have good reason, then, to deny that just any old thing would be the case were something which is in fact impossible to be the case: and it would certainly be nice heuristically and formally if we could employ talk of different impossible worlds to represent cases where impossible things are the case. This is a prima facie case for impossible worlds, and their usefulness: next I must address the issue of what the theoretical cost of their

admission is, and how they are to be used—for they will look much less appealing if an attempt to use them lands us in logical trouble.

3 Counting the cost The first issue which needs to be considered is how much of an extravagance impossible worlds would be. On some accounts, the cost would be too extreme. Lewis emphatically rejects impossible worlds, (e.g., [17], p. 7, n. 3). But much of this flows from his conception of what possible (or impossible) objects are like. On his conception, possibilia do in fact have the features which we associate with them: the merely possible blue swans are literally blue and literally swans, for example. Possible worlds for Lewis, notoriously, are just large objects much like our own cosmos—so the worlds where there are blue swans are just cosmoi with blue swans (among other things) in them. Extending this approach to impossible objects produces literal impossibilities, it seems: if the impossibilium corresponding to the blue swan-and-not-a-swan is literally a swan and is literally not a swan, then a contradiction is literally true. The problem does not just arise either for the nondialetheists among us: there are other things which cannot possibly exist which would cause trouble. The existing-at-all-possible-worlds God of Anselm's imagination does not exist at every world—and it simply fails to exist at this world, full stop—it is not that it both literally exists in this world and literally does not exist in this world.[7] There could not be a thing which made all disjunctions false by its mere existence: but if we are to infer from this that there is an *impossibilium* which literally makes all disjunctions false by its mere existence, then we are in deep trouble.

However, this problem for construing impossible objects and worlds along the lines of possible objects and worlds only arises in systems which hold that there are merely possible objects, and they do literally have the features associated with them. Most believers in possible worlds do not accept that there are possibilia which literally have the features associated with them: most believers in possible worlds reject the existence of talking donkeys, phlogiston, crystal spheres spinning around the center of their universe, and so on, even though they admit that such things are possible. Most believers in possible worlds either do not believe in (mere) possibilia at all, or at the very least they deny that possibilia literally have many of the features associated with them. Those who deny that there literally are mere possibilia (like blue swans, to stay with the standard example) may admit that there are possible worlds *according to which* there are such things, or that it is *true in* some possible worlds that there are blue swans, or that some possible worlds *represent that* there are blue swans: but this is all. Such people should probably not take possibilist quantification literally, which could be an unfortunate limitation for some purposes to which possible worlds are often put—but they may have (and I think they should develop) a way of understanding such possibilist quantification so that it is respectable after all. Those that accept the existence of possibilia do not need to say that any of them are literally blue swans—they may say instead only that some possibilia *represent that there* is a blue swan, or that some possibilia are such that *were they actualized, they would be blue swans*, or something of this sort. Such abstract possibilia seem to serve admirably for most purposes. Abstract impossibilia of this sort would not pose the same risk of incoherence as impossibilia which literally had the features associated with them: for an object that represented that there was a round square table, or an object which is such that

were it (*per impossibile*) to be actualized, it would be a round square table, are not nearly so logically ill-behaved.

For most abstractionists, in fact, it would seem that accepting impossible worlds, and even impossibilia, would be only accepting ontology of a sort which they are already committed to. In some cases, they would not even need to accept anything new: someone who took possible worlds to be sets of propositions, or sets of sentence-like representations, is probably already committed to sets of sentences which are not maximal (in the sense of containing each proposition/sentence or its negation), or consistent (containing at most one of each contradictory pair of sentences or propositions), or either. These other sets may well represent perfectly adequately ways the world could not turn out. And those, like van Inwagen [36] and apparently Stalnaker [34], who postulate *sui generis* abstract possible worlds, need postulate nothing more strange when they postulate the existence of *sui generis* impossible worlds also.

Furthermore, there are those who do not admit the literal existence of possible worlds, but engage in talk of them all the same: fictionalists (Hinckfuss [12], Armstrong [2], Rosen [29], to name a few), instrumentalists (van Fraassen [35] is either an instrumentalist or a fictionalist about possible worlds, Merrill [20] may count as one, and indeed counts himself as one, though he might also with justice be classified as a Meinongian or a fictionalist), Meinongians (e.g., Routley [30]) and others (such as Forbes's [10] "instrumentalist" strategy, which seems to me more a contextualist paraphrase strategy about talk of possible worlds). In principle, these theorists have an easy time with impossible worlds, since after all these theorists already did not admit the existence of such things! There are various accounts around which explain the intelligibility of impossible fictions and inconsistent instrumental machinery, and indeed I will make some suggestions about how to go about reasoning about impossible contexts—suffice it to say that it seems very plausible that impossible worlds will not cause serious ontological problems for theorists who did not propose to literally believe in them in the first place!

So on many standard accounts of the ontological commitments of a theory of possible worlds, adding belief in impossible worlds will not be a notably more extravagant gesture. Given the prima facie desirability of them for a range of theoretical purposes, then, the case for believing in impossible worlds (or, in the case of instrumentalists and fictionalists and suchlike, for talking as if we believed in impossible worlds) is in good shape.

The question of the extent of impossible worlds can receive, I believe, a very generous answer. I think the most plausible comprehension principle for impossible worlds is that for every proposition which cannot be true, there is an impossible world where that proposition is true. This comprehension principle, while natural, will be inconsistent with most accounts of impossible worlds, according to which impossible worlds obey some constraints, but not as many as possible worlds. This comprehension principle is at least a good working hypothesis, and does accord with our normal practice of apparently quantifying over "ways": when faced with an impossible description, or specification, we are pretheoretically tempted to say that we have been presented with a way things can't happen, or a way things cannot be: we do not, it seems to me, require that the specifications of ways things cannot happen meet any particular requirement, except that they not be ways things could happen.

Worries about the nature of impossible worlds, then, should not be worries which concern many, and the question of their extent can be given a simple and plausible answer. There is another sort of worry often evinced about impossible worlds, and nontrivial reasoning involving the impossible in general: the worry is that allowing such things respectability will bring the evils of nonclassical logic in their train. While I am not filled with dread by such a prospect (and it may well turn out to be something that we should embrace on other grounds in any case), I do not think that the classical notion of logical consequence needs to be tampered with to give impossible worlds their place in the sun. Indeed, I think there is a good general reason why modifications of one's theory of logical consequence should not be the resource employed to accommodate impossible worlds for the tasks I outlined in Section 1. To explain why, let me outline my "modest proposal" for accommodating impossible worlds.

4 A modest proposal After use as models of logical consequence, one of the most technically useful tasks for which possible worlds have been employed is in providing the semantics for conditionals. While there is a great deal of debate about their applicability to "indicative" conditionals,[8] it is perhaps the majority view in the debate that they shed light on the best model of the "subjunctive" or "counterfactual" conditional[9] (though it is by no means an uncontroversial contention). The idea, in a nutshell, is this: when evaluating conditionals of the appropriate sort, one checks the "nearest" possible worlds where the antecedent is true, and if in all of those worlds the consequent is true as well, then the conditional itself is true. More strictly speaking, a conditional (of the appropriate sort) is true if and only if in the nearest possible worlds where the antecedent is true, the consequent is true as well.[10] Much of the weight of the analysis, and much of the controversy within this tradition, concerns what exactly "nearness" amounts to. Lewis [14] suggested that "nearness" is a matter of similarity in appropriate respects, and of course a lot of effort has gone into locating the right respects of similarity ([15] and [18] show clearly that similarity in the right respects can come apart dramatically from overall similarity of worlds). Nevertheless I think that the basic idea of treating "nearness" as similarity in appropriate respects is right—though I will not defend it here, and many of my remarks can be reinterpreted in the light of another account of "nearness," if desired. Other debates have centered around how nearness functions: Stalnaker claimed in [33] that there was always a unique nearest world in which the antecedent was true (provided the antecedent was possible); Lewis wanted to allow that worlds could tie. Lewis and Stalnaker, and many besides, wanted to claim that the actual world was more near to itself than any other world, in any respect: others (e.g., Read [27], p. 94) have wanted to claim at least that nearness in relevant respects could be tied between the actual world and nonactual ones, thus providing the mechanism for denying that 'if *A* then *B*' is true whenever *A* and *B* happen to be, regardless of the lack of connection: 'if I go home this evening, horrible atrocities will be carried out on children tomorrow morning' does not seem right—I have nothing to do with such atrocities, and no tormentor of children cares particularly how I spend my evening, but surely in a sorry world as big as ours (and given that it's not going to be a very eventful evening for me) both the antecedent and consequent are true. Some might even be prepared to say that the actual world is not one of the "closest" for the evaluation of some condi-

tionals, though this would indicate that similarity to the actual world should not play quite the usual role assigned to it, since surely there could not be something more similar to the actual world than the actual world is to itself, in any respect. Furthermore, counterexamples to modus ponens will appear unless one is very careful. There are no doubt other points of contention within the closest-world approach to conditionals as well as the ones I have mentioned, but I do not need to buy into any of these controversies here.

It is clear that if the notion of similarity-in-relevant-respects between worlds is a useful device for thinking about the truth-conditions of some class of conditionals, then we must have a better implicit grasp of the relevant respects to take as similar to our typical explicit grasp: for we manage the context-sensitivity and judgment required to evaluate conditionals effortlessly in normal conversation, but philosophers have difficult wrangles over how best to explicate what grounds are relevant for determining similarity relations explicitly. I do not hope to add much to our explicit understanding of this notion here. However, I hope to rely on our understanding of the notion of similarity-in-relevant-respects, and more generally our use of worlds in evaluating conditionals, to find useful work for impossible worlds.

The thought is a simple one: that some impossible worlds are more similar, in relevant respects, to our actual world than others. The "explosion" world—the impossible world where every proposition is true—is very dissimilar from our own. Indeed, it seems to be one of the most absurd situations conceivable.[11] On the other hand, the world which is otherwise exactly like ours, except that Hobbes succeeded in his ambition in squaring the circle (but kept it a secret), is far less dissimilar. Of course, that world is still a strange one—and perhaps a world where Hobbes succeeded and later mathematicians failed to provide a proof that it is impossible to so succeed would be less dissimilar still. If this is right, then we have room for distinguishing true from false counterpossibles: we can say that 'if Hobbes had squared the circle, sick children in the mountains of South America at the time would not have cared' is true, but 'if Hobbes had squared the circle, then everything would have been the case' is false, if the nearest (most relevantly similar) Hobbes-squaring-the-circle impossible worlds contain no interested sick South American children, but that the explosion world is not as near as other Hobbes-squaring-the-circle impossible worlds.

The suggestion that impossible worlds might be used to allow for nontrivial counterpossible conditionals is not a new one: Read ([27], pp. 90–91) is one who advocates it, and Routley ([31], pp. 294–301) provides a formal apparatus designed to model counterlogical conditionals using a selection function on worlds analogous to the usual selection function of conditional logics. (As will become clear, however, I disagree with Routley about how to treat counterpossible conditionals, even if our general framework has important similarities). It can be extended to allow us to capture much of what we might want to express about impossible cases. Let us suppose that the debate between defenders of the axiom of foundation and non-well-founded-set theorists is as it seems to some to be: a debate about the one true system of sets in Platonic heaven, and whether the Axiom of Foundation is true for them (where the axiom is interpreted straightforwardly). Further, let us suppose for the sake of the example that the Foundation defenders are right, and their opponents wrong. Still, we know that were the axiom of foundation to be false, there would be a non-well-

founded set: but were the axiom of foundation to be false, it would still not be that 1 equaled 2. In such a case we have a pretty good idea of how things would be if the axiom of foundation was false: set theory would be a little different in well-explored ways, and not much would be different here in the world of change and decay.

Similarly, as historians of philosophy we often have a reasonable grasp of what the world would be like if, for instance, Kant or Aristotle were correct. When we puzzle about exactly how Plato took the world to be, we should not take it that anything we please would be the case were Plato to be correct, even if we do suspect that he could not even possibly have been right. A world with Platonic Forms, which objects rely on for their being and character (if indeed this was Plato's view) is, even if impossible, not so distant as the explosion world, or even Plato's world with the addition that the appearances are that everyone is a ten-foot-tall scaly lightbulb. Metaphysical disputants can often work out how the world would be (or the worlds would be) if their opponents are correct, and indeed by working out how that world would be, they can try to find particularly salient absurdities to report to their opponents and to interested onlookers. This practice seems to display at least a grasp of relative nearness of positions they take to be impossible: I may be suspicious of the view that necessarily, everything that exists belongs to a substance-sortal which determines its essence (indeed, I believe this doctrine to be necessarily false): but I do not believe that were this to be the case, it would be the case that two plus two equaled five. A world where objects have substance-sortals which determine their essence (even a world where this happens of necessity) is more similar to our own than one where this happens, and in addition two plus two make five. I would have a hard time proving this judgment of relative similarity to a skeptic—but I am confident enough of it to resist charging a believer in the above view that were their view to be right, two plus two would have to equal five.[12] I suspect I am far from alone in this relative similarity judgment, even if some resist this intuition for theoretical reasons.

This notion that not all impossible worlds are equally distant can even permit us to discuss what might be the case where the logical laws are different: we have a quite extensive literature on what the theorems of logic would be if intuitionistic logic were the One True Logic, or if any of the relevant family were correct. We also have an almost exhaustive investigation of what laws would hold and what results could be obtained if classical logic were correct. We are often able to say quite exactly what would be the case, logically speaking at least, in the closest impossible worlds where an actually false logic is true. There may be areas of uncertainty: how much of classical mathematics is the case in the nearest intuitionistic world? Indeed, there are likely to be areas of indeterminacy: exactly how nonprime would the world be (if at all) if disjunctive syllogism were not valid? It seems very plausible that many things would be pretty much the same at most of the closest logically impossible worlds: there does not seem to be much reason to suppose the price of eggs would be terribly different were one of the actually incorrect but plausible rivals to be correct instead.

With this capacity to usefully employ counterpossible conditionals, we allow ourselves to talk of what cannot be, in a way which allows us to nontrivially make claims about how things would be if various impossibilities were the case, and which allows us to coherently have disagreements about how things would be were certain things to be impossible. Note that this ability is purchased with no alteration to our

logic except an expansion of the semantics of the conditional—our notion of logical consequence, and the rules governing this logical consequence for the other connectives, can remain unchanged. So, for example, even the classical logician can allow for a sort of nontrivial reasoning involving impossible situations and ways things could not be. Some of the details of how this might be used will be explored more in Section 5.

This is fine and dandy for logical conservatives, you might think. But why, one might ask, is it any advantage to be able to keep a logic as unfriendly to exotic possibilities as classical logic in the first place? Even if one can allow oneself to talk of impossibilities while being classical, why beat around in the bush? Why not adopt a logic with a consequence relation less likely to produce explosion when impossibilities are in play? One answer would be a defense of the attractiveness of classical logic—this is well-tilled ground, and would be unlikely to convince one inclined to ask the question in the first place. In any case, this is not the sort of answer I wish to offer here. One of the purposes of discussing the behavior of this counterpossible conditional in classical logic is to demonstrate that the admission and use of impossible worlds need not have much in the way of logical ramifications. I do, however, think that there is something to be said for dealing with impossibilities by means of the conditional rather than modifying one's notion of logical consequence—there may be other reasons for adopting a nonclassical account of logical consequence but impossibilities should not be one of them.

Modifying logical consequence to deal with impossibilities is a popular move—though pointing to specific examples can be controversial. Many paraconsistent logicians, for example, wish to do away with the *ex falso quodlibet* rule and *disjunctive syllogism*, not because they believe that there are actual cases where a contradiction is true, or which are nonprime (have a disjunction being true with neither of the disjuncts being true), but because there are nonactual nontrivial situations where such things happen. These paraconsistentists are by no means the only paraconsistentists—a more moderate paraconsistent line is to not be an alethic paraconsistentist at all, but to examine paraconsistent "logic" for modeling belief revision, or some such (see, e.g., Anderson, Belnap, and Dunn [1], pp. 506–63). On the other hand, the more robust paraconsistentist position is to be dialetheist as well, and accept actual contradictions—the *ex falso quodlibet* rule and *disjunctive syllogism* should be rejected, for a dialetheist, at least partly because they have actual counterexamples (Priest [22], [30])! I am not concerned here to take issue per se either with the nonalethic paraconsistentists (who might not disagree with me at all about alethic logic, which is what I am primarily interested in here), nor with the full-blown dialetheists—if there are actual contradictions, then a logic which takes truth-preservation to be even a necessary condition of logical consequence will need to take that into account. However, paraconsistentists (both the middle-of-the-road variety and ones who happen to be dialetheists as well) do sometimes talk as if there is a quick argument from the need to reason about impossibilities to a paraconsistent logic. For instance, Routley [32] claims that "it suffices for the falsification of Disjunctive Syllogism that there are nontrivial but inconsistent deductive situations and theories" (p. 156).[13] Again, Priest and Routley ([25], pp. 483–583) argue at length that there are cases of interesting or important positions which are inconsistent, and

so a paraconsistent logic should be adopted if we are to reason usefully about the commitments of these theories.[14] Of course Priest and Routley may well not accept that these inconsistent theories are necessarily false: but since they present this argument as being independent of their argument that some contradictions are true, it is reasonable to see them as recommending the move to a paraconsistent logic even to those not prepared to accept actually true contradictions.

I think, however, that modifying one's account of logical consequence in order to accommodate impossible situations is a mistake. For if there is an impossible situation for every way we say that things cannot be, there will be impossible situations where even the principles of subclassical logics fail. Were conjunctions to behave like logically atomic sentences, there would be no guarantee that a conjunct was true whenever a conjunction was: and if this always failed, then in that impossible situation *A* and *B* would be true, but *A* would not be, and *B* would not be either. Most subclassical logics keep the rules '*A* and *B*, therefore *A*' and '*A* and *B*, therefore *B*', however.[15] If the motivation is to provide a logic which applies to every situation, possible or not, then that logic will have few principles indeed. And indeed for just about any cherished logical principle there are logics available where that principle fails: one favorite is the system *S*, where '*A*, therefore *A*' is an invalid rule of inference. This logic nevertheless puts systematic restrictions on what follows from what (see Martin and Meyer [21] for details), and even champions of having a notion of consequence which can apply to logically exotic situations tend to rule out this as an acceptable notion of logical consequence. Yet we are able to nontrivially evaluate how things would be were *S*, or one of its extensions, to be correct.

The proponent of modifying our notion of logical consequence to handle impossible situations as well as possible ones has two alternatives: they can deny that there really are such things as the impossible worlds where their favorite logic fails to preserve truth—they can put their foot down and insist that there is no impossible world where, for some *A* and *B*, '*A* and *B*' is true, but both *A* and *B* fail to be true. Presumably, though, they think it is impossible for that to happen, so they will have to distinguish between the impossibilities which obtain in some impossible worlds and those impossibilities which obtain in no worlds, even impossible ones: and this seems a distinctly uncomfortable halfway house between those who deny that there are impossible worlds (perhaps the standard position), and on the other hand my position, which maintains that for every impossibility, there is some impossible world where it holds.[16]

The other alternative they can face is to say that their notion of logical consequence is not to be associated with truth-preservation in all worlds, but only truth preservation in all possible worlds and some impossible worlds (that is, they allow that there are some impossible worlds where the premises of an argument which is valid in their favorite logic hold, but the conclusion of that argument fails to hold).[17] Those that pursue this alternative need some other mechanism to talk about and reason about worlds where their favorite logic fails—I recommend that we discuss these cases though talking about what *would* have been the case, were things to have gone that way—to use conditionals. For example, if conjunction elimination sometimes failed to hold, then there would be propositions *A* and *B* such that '*A* and *B*' was true, but *A* would fail to be true. And if they both modify their logic to accommodate

some possibilities, and then follow some other strategy (such as allowing for the nontriviality of counterlogical conditionals, as I propose), then they are using two stones to kill their bird: and since the second stone kills the bird anyway, the use of the first one (modifying their logic to accommodate impossibilities) is unmotivated.

My point is that to attempt to produce an ultralogic which is truth-preserving in any world that we can reason about is a misguided venture, since we can reason in the absence of any given principle of logic (as the case of reasoning about what would be the case were *S* the one true logic should demonstrate[18]). Better, instead, to only worry about possibilities when considering what notion of logical consequence our logic should capture, and use the device of conditionals to carry out our discussion of impossible cases. As well as a recommendation, it seems plausible to me that this is what actually happens—people are more prepared to consider hypothetically (or entertain a "what if" about) what things might be like if some other system of logic besides the ones they favored were correct, than they are to reject principles of their favorite logical system on the basis that we can reasonably consider what would be the case were those principles not to hold. They are more hospitable to counterlogical conditionals than they are to revising their logic itself.

Let me reiterate that there may be other reasons for modifying our logic—my example of a basic logic has been classical logic, and as a matter of fact I think that logical possibility goes along with a more or less classical conception, but my remarks would have been equally applicable had I been assuming that some other logic was the default assumption. No matter what one's logic is, one should not be tempted to weaken the notion of logical consequence to handle impossible situations: for no such weakening will handle every impossible situation it might be worth considering or reasoning about, unless one makes one's logic so weak that there are hardly any principles governing logical consequence at all. This should be kept in mind for the rest of the paper—from now on, I'll often be talking as if our base logic is classical, but many of my remarks apply just as well (perhaps with altered examples) if one's base logic is nonclassical.

Employing conditionals to make nontrivial claims about impossible worlds is a way of making claims about impossibilities with which we are relatively familiar. If this proposal is to be adequate for our purposes in evaluating and arguing about impossible situations, however, it must provide an account of how we are to reason about impossibilities too—how these conditional claims are to be handled in inferences. Another aspect of our talk about impossibilities which it would be useful to capture is that section of our talk which does not have the surface structure of conditional utterances—proving theorems in a logic we do not believe, or in discussing a metaphysics which we do not accept, we obviously do not begin every sentence uttered "Were the assumptions I am hypothetically endorsing to be true . . . ," or any such thing. In accommodating this, I introduce an account of presuppositions slightly different from many standard accounts.

5 Counterpossibles in action: inference and presupposition In many respects, counterpossible conditionals obey the fundamental principles governing conditionals (and just as well . . .): modus ponens is satisfied (from 'if Hobbes squared the circle, π is rational' and 'Hobbes squared the circle', we can safely infer that π is ra-

tional), as is modus tollens (π is not rational, so Hobbes didn't square the circle after all). Modus ponens is satisfied on the assumption of *weak centering*: that the actual world is at least as similar to itself as any other world (in particular, more similar to itself than any impossible world is): so the conditional interpreted as "at the closest antecedent worlds, the consequent holds as well," added to the information that the antecedent holds at the actual world, is sufficient to ensure the consequent holds here as well. As for modus tollens: of course the antecedent is false! It's a counterpossible, remember? The antecedent is not even possibly true! A fortiori, the antecedent is false when the consequent is.

Conditional proof, on the other hand, does not always hold: from the fact that a set of premises Σ, including A, have as a consequence B, it does not automatically follow that a consequence of Σ minus A is 'if A then B' (in the sense of 'if' captured by the counterpossible conditional). This should not be surprising: conditional proof fails for the ordinary nearest-world counterfactual.[19] What is slightly more surprising is that 'if A then B' is not a theorem when B follows from A alone: this restricted form of conditional proof is satisfied by the counterfactual conditional, but since there will be many impossible worlds where logic fails to be even truth-preserving[20] models where not everything is the case in the nearest impossible world where, for instance, a given contradiction is true (for instance the aforementioned contradiction concerning a far-distant electron), there will be cases where (classically)

$$p \,\&\, {\sim}p \vdash q$$

but where it will not be that

$$\vdash (p \,\&\, {\sim}p) \rightarrow q.$$

For instance, it follows classically from there being an electron that is negatively charged and is not negatively charged that Australians all vote for the Greens, but a model where the closest impossible world where there is such an electron but Australians do not all vote for the Greens is a perfectly respectable model. Indeed, I believe that it is in fact false that were there to be such an electron, Australians would all vote for the Greens—we would be utterly unaware of such an electron, and even its discovery would not raise environmental awareness very much (though it may well cause consternation to physicists and logicians). I do not think that the failure of this restricted conditional proof should worry even a classical logician: logic is modeled by what happens at all possible worlds, and the conditional is concerned also with the behavior of close impossible worlds. One way of looking at it is to see the failure of this restricted conditional proof as analogous to the failure of full-blown conditional proof in the case of the ordinary counterfactual conditional. The mere fact that some things happen to be true which, in conjunction with the antecedent, necessitate the consequent, does not mean that the conditional is true, since the truth of the antecedent would not come about (or would not come about in any close world) with those particular facts: for example, my being in this room, and the room's not moving, together ensure that I will not be falling to the ground outside the window. Nevertheless, I cannot infer from my being in the room and the room's not moving that were I to have jumped out the window, I would not have fallen to the ground, since it is precisely those facts which would not have obtained had I jumped out of the window.

Similarly, it follows classically from the electron both having and failing to have negative charge that Australians vote for the Greens: but it is precisely the sort of logical principle that ensures this that would have to be suspended, or have had exceptions, were it to have turned out that there was such an electron. This analogy is only a loose one: there is an important distinction between logical rules or principles on the one hand, and premises of arguments on the other, and one ignores that distinction at one's peril. Nonetheless, I think there is something to the analogy.

Conditionals with impossible antecedents then may not hold even when the corresponding consequence relation does. However, this difference between my account and a Lewis-Stalnaker account which restricts worlds relevant to the conditional to possible worlds does not show up very much in practice: for many everyday conditionals, impossible worlds do not seem called for—which is why Lewis's or Stalnaker's restriction of the worlds relevant to evaluation of conditionals is adequate in most everyday circumstances. In general, I am tempted to impose the following restriction on the relevant similarity relations which might be used to evaluate these counterpossible conditionals:

Strangeness of Impossibility Condition: any possible world is more similar (nearer) to the actual world than any impossible world.

I think this has a fair bit of intuitive support—the heavens will fall before (correct) logic fails us. When I utter conditionals normally, and say that if I have no legs and feet I cannot walk, I presumably ignore those worlds where I have no legs, but can nevertheless walk—perhaps because I am an inconsistent object, which has legs despite having no legs, or where walking is (impossibly!) something which does not involve legs or feet at all—for instance the world where walking is a special sort of mathematical operation (absurd, yes—and so surely something which just can't be!). By and large, we do not consider impossible situations when working out how things would be, were things to be otherwise in certain respects. We do sometimes, however—when explicitly asked to consider something we consider impossible, we seem often to be able to have some ability to do so.

The Strangeness of Impossibility Condition (SIC) is offered as a conjecture about how we treat relative similarity—we could easily take impossible worlds to be more relevantly similar than some possible worlds if we chose. I am not even confident that the SIC is always adhered to. Consider the following conditional.

> If intuitionistic logic came to be thought a much more satisfactory basis for mathematics by the experts, and if intuitionistic investigations led to breakthroughs in many areas of inquiry, and if important technological advances were made by the best minds in the field, which they would not have come to if they had been stuck in the rut of nonintuitionistic logic, then intuitionistic logic would turn out to be correct after all.

The above conditional might seem like an appropriate one to utter—we are not dogmatic, and we (or I at least) think that total inquiry may have some evidential bearing on the correct logic. If the SIC is enforced strictly then the statement can be dismissed out of hand by a believer in classical logic, since the closest *possible* world where the experts and human inquiry behave in the way the antecedent states is still a world

where intuitionistic logic is incorrect. Perhaps the example is defective—perhaps the conditional employed is not a "counterfactual" or "subjunctive" conditional of the right sort, but merely some sort of evidential conditional instead. Or perhaps some readers might find it obvious that it is false. There are enough examples like this, though, for me to suspect that on some occasions the SIC fails, and well behaved, only slightly impossible worlds, are to be preferred to some particularly bizarre possible worlds.[21] Whether or not this is true or not is a sociological matter—by and large we take the possible to be more relevantly similar to the actual than the impossible (and things would be much less convenient if we did not pay attention to the difference between the possible and the impossible), but we are not forced to do so always.

But the introduction and the elimination of the conditional are by no means all of the inferences that are important concerning the conditional.[22] We can employ the information conveyed by such conditionals to reason to further conclusions, or put the information of the conditionals together. If I find out that 'if I proved π was rational, I would become rich', I can infer that 'if I proved π was rational, I would own more than five dollars worth of property'. I might reasonably come to believe, on the same basis, that 'if I proved π was rational, I would be approached to give money to worthy causes more often', and if I came to believe that 'if I proved π rational, I would annoy all of the mathematicians', I might conclude that 'if the cup in front of me is a mathematician in disguise, then if I prove π rational, I would annoy the cup in front of me'. These sorts of inferences seem okay to me, whereas if, on the basis of the above beliefs, I came to believe 'if I proved π rational, and if the cup in front of me is a mathematician in disguise, then my troops will conquer Paris', it would seem that I was deluded about my military power.[23]

These inferences consist of making inferences in the scope of conditionals: one of the simplest cases is an inference from 'if *A* then *B*' to 'if *A* then *C*', where *C* is a logical consequence of *B*. Such inferences are valid on the standard Lewis-Stalnaker style systems: if the nearest *A* world is a *B* world, it must also be a *C* world if *C* is a consequence of *B* (since all of the logical consequences of *B* hold in every *B*-world). The inference may not be formally valid when the worlds being considered include impossible worlds, since with impossible worlds there is no guarantee that all of the logical consequences of a proposition will hold at a world if that proposition does. If one considers classical consequence, then counterexamples will abound: take the case mentioned above at p. 549, of the world with one inconsistent electron. It classically follows from that electron having negative charge and lacking negative charge that I am a tap-dancing squid: but the closest impossible world which contains such an electron is not one where I am a tap-dancing squid. So it does not follow from 'if there were an electron in a distant galaxy which both has and fails to have negative charge, then there would be an electron in a distant galaxy which both has and fails to have negative charge' that 'if there were an electron in a distant galaxy which both has and fails to have negative charge, then I would be a tap-dancing squid'. Nevertheless the inference will often be rationally justified, since the impossible worlds most similar to ours will by and large have the consequences of most of the propositions which are true at them also true at them. So, to use an example mentioned previously, it is safe to infer from 'if I proved π was rational, I would become rich', that 'if I proved

π was rational, I would own more than five dollars worth of property'.

A principle which plays a similar inferential role is an inference of the form 'if *A* then *B*', 'if *A* then *C*', therefore 'if *A* then *D*', where *D* is a logical consequence of *B* and *C* together. Inferences of this simple sort again are (almost always)[24] valid on standard closest-world semantics for conditionals (if the closest *A*-world is a *B*-world, and the closest *A*-world is also a *C*-world, then both *B* and *C* hold in that world, so any logical consequence of *B* and *C* hold there too). Let me call this principle the *Conjoining Consequents* principle. Again it will not be unreservedly true on a semantics which includes impossible worlds: for then we lose the guarantee that when *B* is true at a world, and *C* is true at a world, then any consequence of *B* and *C* together is true at that world. Consider the following argument:

> if the only two things that were the case were that there existed a red rose and that there existed a red apple, then there would exist a red rose;
>
> if the only two things that were the case were that there existed a red rose and that there existed a red apple, then there would exist a red apple;

therefore,

> if the only two things that were the case were that there existed a red rose and that there existed a red apple, then some flower would be the same color as some fruit.

I think both premises are true, but the conclusion is false—the closest world where the antecedent is true is an impossible one (since if there were a red apple and a red rose, then many other things would be the case too), and it is one where only two propositions are the case: so the propositions which are actually strictly implied by those two things being true together will not necessarily hold in that world—and indeed all but two of them will not. This sort of case serves as a useful reminder that some antecedents—antecedents which are true only at rather distant impossible worlds—will cause conditionals to fail to obey many of the conditions which Lewis-Stalnaker conditionals restricted to possible worlds do.

However, as before, many of the less bizarre impossibilities will not be so badly misbehaved. Most of the closest impossible worlds will, I take it, be relatively well behaved logically, with the odd isolated glitch, rather than logical impossibilities being completely pervasive. In situations where we have no general reason to suspect that the consequences of *B* and *C* will hold if *B* holds and *C* holds (though we may have good reason to suppose that there are specific isolated instances where this will not), then such inferences will work. In the case of my proving the rationality of π, for example, in the nearest world as I judge it the impossibility is mostly confined to some areas of mathematics and perhaps some areas of our mathematical activity: it will not spread to quotidian facts of finance, like the fact that being rich involves having more than five dollars. Indeed, in this respect, the closest such impossible worlds will be the same as the actual world. Thus the inference

> if I proved π was rational, I would become rich;

therefore,

> if I proved π was rational, I would own more than five dollars worth of property,

is rationally justified: the nearest impossible worlds where I prove π rational are worlds where the basic facts of finance are not very different.

I say the inference is rationally justified: but I do not know whether it is fruitful to see this inference as a matter of deduction. Putting the inference in a form which makes it formally valid is difficult, especially when it is taken into account that the argument is valid even if the premise is false—take some hermit, H, who, as a matter of fact, would not become a commercial success regardless of what mathematical results she proved. The analogous argument still works, it seems to me:

> if H proved π was rational, H would become rich;

therefore,

> if H proved π was rational, H would own more than five dollars worth of property.

One could attempt this through enthymemes—the facts of finance, some premise to capture the fact that H's coming to be in a position to be a commercial success would not involve massive deflation, and so on—and while any rational inference can be transformed into a deductive one with sufficient enthymematic machinery, I am not sure there would be much advantage to it. Having reasoning within the scope of conditionals which is difficult to represent formally is a pervasive feature of nearest-world conditionals in any case: it is very hard to formally represent the following inference as *logically* valid, even though I think it is a perfectly reasonable one to make:

> if Caesar were in the Korean war, he would have used catapults;

therefore,

> if Caesar were in the Korean war, more large rocks would have been hurled than in fact were.

Notoriously, there is disagreement about the truth of the premise, and even whether the premise has a determinate truth value. Furthermore, it is not true that in every world in which the premise is true, the conclusion is also—for there are possible worlds where in fact, in that world, so many boulders are needlessly tossed in Korea that the nearest world to that world where Caesar is involved, catapults and all, is one where fewer large rocks altogether are hurled. As well as those worlds, there are the worlds where Caesar's catapults do not throw large rocks, but throw pitch or steel spheres or mustard gas shells instead—in the worlds sufficiently close to them, were Caesar to be involved in the Korean war with his catapults, it would not affect the number of rocks hurled. Nevertheless the inference seems reasonable to make—however I at least can see no way to systematically supply a convenient enthymeme to make such inferences formally valid.

So I hope I may be excused for proposing only to demonstrate why some inferences involving conditionals are rationally justified, rather than proving very many

theorems which hold regardless of the content of the antecedent or consequent of the conditionals. This will no doubt disappoint some, who hope for a set of axioms, an interesting proof theory, with perhaps some surprising (or at least complex) theorems. I fear, however, that I think impossible worlds are too badly behaved for there to be many principles which hold distinctively of the conditional regardless of the content of the antecedent and the consequent: though this is not to say that there is anything wrong with more restricted conditionals which ignore the most misbehaved impossible worlds, or even which ignore the impossible worlds at all: there may well be contexts where one should ignore them, as never being relevantly similar enough to count as being in even a distant sphere of similarity (just as Lewis allows that the system of spheres may not be universal, to permit the modeling of the idea, if desired that some *possible* worlds are so bizarre as to be left out of consideration—see [14], p. 16). The context which ignores all of the impossible worlds is useful for some purposes—I do not know if ordinary folk ever employ such a context, but philosophers and logicians have been in the habit of doing so from time to time (though in their discussions of rival logical, mathematical, or metaphysical systems they often slip into a less senatorian usage without noticing): the Lewis-Stalnaker formal systems can be seen from my broader perspective as attempts to capture which principles always hold of conditionals in this limited but often convenient context.

Since most principles concerning the conditional have counterexamples, when sufficiently strange antecedents are employed, I think that there will be almost no distinctive theorems which hold of conditionals regardless of what propositions make up their antecedents and the consequents. One position which I find appealing is to say that there are none whatsoever: even ones usually accepted, such as 'A and B, therefore A' will fail for suitably inhospitable A or B (for a counterexample, see p. 552). Statement modus ponens will fail as well, I take it: the nearest worlds where, for example, modus ponens fails might well be worlds where modus ponens fails, and so a conditional such as 'if modus ponens failed and it would be the case that logic would be intuitionistic if modus ponens failed, then logic would be intuitionistic' might not come out as true: though I have to admit that my pretheoretic intuitions about this example deliver no clear verdict. The failure of the statement form of such a principle should not worry us, however, since the rule form is not invalidated, as I argued above. (Incidentally, since 'statement' *ex falso quodlibet* fails, this system has some claim to be counted as paraconsistent, if paraconsistency is defined in terms of rejection of *ex falso quodlibet*. It would have to be one of the weakest forms of paraconsistency on the market, in that case. And certainly if the counterpossible conditional is grafted onto a classical base, rule *ex falso quodlibet* will still be valid, as will claims like the material conditional reading of statement of *ex falso quodlibet*, and the strict conditional reading.)

However, one may not be able to say that there are no theorems employing the conditional as the main operator. Identity ($A \rightarrow A$) seems to be preserved (depending on the exact details of the construction of the system): the closest worlds where A is true must, it seems, be worlds where A is true. (It might be false as well at such a world, but that will not serve to make the conditional false as well: at most, that will simply ensure the truth of $A \rightarrow \sim A$ for that A). Whether or not one can find plausible counterexamples to identity in this way of modeling counterpossibles depends on

how one understands the selection function that gives one a world (or set of worlds), given an antecedent and a context. If one goes to the nearest world where the antecedent holds, then it seems one must allow that conditionals of the form '$A \rightarrow A$' are theorems. On the other hand, some other understanding of what the selection function[25] does may permit one to hold that there are no theorems employing the conditional (or at least which have the conditional as their main operators: theorems such as '$A \rightarrow B \supset ((A \rightarrow B) \vee C)$' or '$A \rightarrow B \vee \sim(A \rightarrow B)$' will still be theorems, if '$A \supset (A \vee C)$' and '$A \vee \sim A$' were theorems to begin with. But these are not distinctive theorems involving the conditional, since they are just substitution instances of theorem schemata which need not mention the conditional). For instance, it might be thought that the selection function does something more like taking the world which the antecedent is truly *about* in some sense. This distinction sounds mysterious, let me attempt to motivate it with an example. Consider the antecedent 'if nothing were to be true . . . '. If identity holds, to evaluate such a conditional we must go to a world which has true according to it 'nothing is true'. However, at least as good a candidate to be selected, intuitively, would be an impossible world which did not have anything true at it, even the claim that 'nothing is true'. However, such a world (let me call it the 'null world' for convenience)[26] is not selected by the selection function as it is normally understood, since even the antecedent of the conditional in question is not true there. However, there is clearly some connection between the antecedent in question and the null world—one which one might attempt to capture with a different account of the selection function. For the purposes of this paper, I am happy to count instances of identity as theorems—I merely note the possibility of fiddling with the selection function if one believed, for whatever reason, that there should be counterexamples to it.

There are few exceptionless principles, but restrictions of context mean that many inference moves which are not formally valid will be acceptable in a wide range of circumstances. In this regard, the principles employing the conditional as a main operator which are theorems in Lewis-Stalnaker systems are like principles such as *strengthening the antecedent*, *contraposition*, and *transitivity* ($A \rightarrow B, \vdash A \,\&\, C \rightarrow B$; $A \rightarrow B, \vdash \sim B \rightarrow \sim A$; and $A \rightarrow B, B \rightarrow C, \vdash A \rightarrow C$, respectively). These three principles fail even in Lewis-Stalnaker conditional logics, though they are principles which appear to sometimes be employed in reasoning. However, they are not formally valid, since there are some slightly unusual cases where they have counterexamples. When the nearest world where A holds is as close as the nearest world where A and C both hold, for example, strengthening the antecedent will be an allowable step. And we can often tell that strengthening the antecedent in this way is harmless—when we have good reason to think that C is fairly orthogonal to A. When put together with the conjoining consequents, strengthening the antecedent, allows us to come to conclusions about more complicated situations much more easily: from 'if kangaroos had no tails, they would topple over', and 'if bicycles had three wheels, etymologists would think they were badly named', we can safely infer 'if kangaroos had no tails and bicycles had three wheels, then kangaroos would topple over and etymologists would think bicycles were badly named', since we know the tails of kangaroos have no effect on the ruminations of etymologists about bicycles, or vice versa. The fact that this inference is not formally valid does not prevent us from relying on it

in this sort of harmless case. Similarly, even for possible cases transitivity ($A \rightarrow B$, $B \rightarrow C, \vdash A \rightarrow C$) fails, but it is a useful principle in a wide range of cases where moving to the closest B world does not affect the truth of either conditional. Again, this will often be known to be true: in the case 'if it rains, I'll take an umbrella' and 'if I take an umbrella, I will not lose it' it will normally be quite plain that whether or not it rains the truth of the second will not be affected (there would not be much point to saying the second in many ordinary circumstances if it was liable to be false if rain should occur). Whether or not in a given circumstance the conditional 'if it rains, I will not lose my umbrella' (or 'even if it rains, I will not lose my umbrella') would be appropriate to infer is a matter of judgment, not of algorithm, but nevertheless we often correctly draw inferences in these patterns even though we cannot always do so.

The inferences captured by the theorems of Lewis's logic of counterfactuals in [14] are rarely formally valid when counterpossibles are in play, but they are often safe enough when the context is not too extraordinary. Reasoning involving counterpossibles will also often safely involve the use of principles such as strengthening the antecedent or transitivity which are not valid even if we restrict ourselves to the domain of noncounterpossibles (though a fortiori such principles are not formally valid arguments when employing the proposed conditional). So from 'if Ms. A is correct in her views about logic, she will convince people eventually' and 'if Mr. B disproves Fermat's Last Theorem (Fermat's Last Conjecture?), he will upset many mathematicians', we can conclude that 'if Ms. A is correct in her views about logic, and Mr. B disproves Fermat's Last Theorem, Ms. A will convince people eventually and Mr. B will upset many mathematicians', provided we are confident enough that the one would not interfere with the other: and when, for instance, speculating about the future of the interaction between logicians and mathematicians, we might rightly perform such an inference, even if we happen to think Ms. A probably wrong, and Mr. B's task probably futile (and let us suppose for the purposes of the example that we are in fact right in our assessments of Ms. A's logic and Mr. B's chances). Provided we keep an eye on the possibility that such inferences might let us down (which only happens in unusual, often easily recognizable ways), we can engage in quite complicated chains of reasoning to yield results of various sorts: that X is a theorem if Y is the one true logic, for example, might take a lot of reasoning to show (at least as much as the believers in Y engaged in before they were prepared to assert that X was a theorem in the first place). Or take a case from the literature: some relevant logicians have argued that a package including their preferred logic may be simpler overall than a classical package, because they can have a smoother set theory (indeed, naïve set theory can be formulated so as to be absolutely consistent in many relevant frameworks). To do so, they need to be able to show what mathematics would be like were their favorite logic plus a suitably naïve set theory both correct (e.g., that the package was absolutely consistent, that the contradictions would not automatically start to spread to nonmathematical/logical areas, etc.). They do not necessarily need to employ conditionals which *they* take to be counterpossibles, of course. But when someone like me talks about what the theorems of set theory would be were one to adopt various paraconsistent alternatives, and in speaking about things being theorems or not in these situations, I speak hypothetically: what contradictions would

be theorems were a naïve abstraction correct and, say, the logic B to be correct. I can mix and match axioms and logical systems hypothetically, building on conditionals already established (I know that *X* would be a theorem in system *Q*, and that *Y* follows (i.e., would follow) from *X* by a rule of *Q*, so I can infer that if *Q* were the correct system, *Y* would be a theorem). This is not just hypothetical reasoning for the fun of it either: naïve set theory has a great deal of intuitive attraction, and holds out the prospect of a simpler and more intuitive foundation for mathematics. Accepting a paraconsistent logic is a cost, I think, and as a matter of fact I am not convinced it is worth it: but I am convinced that it is an option worthy of investigation. For *me* to investigate that option, the most convenient way is to investigate it hypothetically: by thinking about how different options *would* work, *if* naïve set theory plus some paraconsistent logic were correct. The situations I consider when I do so, which I still take to be impossible, are logically fairly well behaved—some have inconsistencies, and some might have gaps (as when there is no fact of the matter whether or not two specifications of sets specify the same set). But conjoining consequents does not let me down, and within wide limits transitivity and strengthening of the antecedent do no harm either. This example may be an ill-chosen one if I am wrong and naïve set theory turns out to be true after all (and so not impossible): but it serves to illustrate a situation where it may be useful to reason 'in the scope of conditionals', yielding interesting conditional conclusions from starting points which are also conditionals.

This practice of reasoning "within the scope of conditionals"—coming to accept new conditionals on the basis of accepting old conditionals whose antecedents and consequents are related in various ways—gives us a new alternative for representing hypothetical reasoning. Hypothetical reasoning has not, to my knowledge, received a great deal of independent attention by contemporary philosophers: the standard treatment is to treat hypothetical reasoning just like nonhypothetical, categorical reasoning: one's hypotheses are the premises, and one applies one's deductive and nondeductive norms of reasoning to work out the consequences of the hypothesis. It is just like nonhypothetical reasoning, except that one need not believe the premises, and soundness (as opposed to validity) is not a priority in the same way (*mutatis mutandis* for nondeductive inferences: they should be rational, or secure, or inductively strong, or whatever, but they need not be based on true premises).

This orthodoxy is so pervasive that it is not very often stated, let alone the sort of thing which one might think is in need of defense (apart, perhaps, from fringe views such as the view that logic should not be used in the real world, or the boneheaded actualism that refuses to engage in hypothetical reasoning at all). Nevertheless, I am going to deny it (question it might be more accurate, but I've always wanted to deny a dogma so entrenched that it's hardly been noticed). Instead, I want to propose that a better way of representing hypothetical reasoning is to take the hypotheses to be antecedents of conditionals, and one's hypothetical conclusions to be consequents of conditionals with the hypotheses as antecedents, where those conditionals are nonhypothetically accepted, or believed to be correct.

This has obvious advantages when one considers impossible hypothetical situations: if one is classical, adding a logical falsehood to one's premises produces absurdity, and even if one is not classical, things remain logically well behaved even when one of the premises is to the effect that they are not. (Hypothetically reasoning

using the relevant logic R, for example, from a bunch of premises which support '*A* and *B*', you can deduce from that bunch of premises that *A*: even when one of the premises states that conjunction elimination fails for '*A* and *B*'. But I think it is plausible that were we to suppose that conjunction elimination fails for '*A* and *B*', and furthermore were we to suppose '*A* and *B*' was true, we should not conclude that on those hypotheses, *A* is true: though there are other things which we could conclude on the basis of those hypotheses: that things were very strange, or that conjunction would not function in the way we ordinarily think that it does.

I think the *hypothetical as conditional* model has something going for it apart from it being more accommodating to dealing with impossibilities. (This is fortunate, since otherwise the suggestion looks a little ad hoc). Hypothetical reasoning has other features which distinguish it from straightforward categorical reasoning. One difference is what premises, or antecedents (or "starting points," to use an unloaded phrase) may be introduced. In categorical reasoning, a new starting point can nearly always be introduced if it is known to be true. Arguments can easily be enthymematic (and usually are), appeal can be made to all sorts of well-known facts, and by and large people do not mind—categorical reasoning often has as a goal the aim of coming to interesting new truths which had not been previously explicitly believed, and any known truths dragged in to serve the goal will help. Hypothetical reasoning, on the other hand, is at once less strict and more strict with regard to starting points: less strict, in that propositions can be entertained when they are not thought to be true, or even thought to be false, but more strict in that unrestricted importation of things known to be true can disrupt the exploration of the hypothesis, or even change the subject. So if I am attempting to work out how things would have gone if Japan had attempted to invade Australia during World War II, I can include among my "starting points" claims about Japanese intentions which I do not in fact believe (since I do not believe that Japan ever in fact intended to invade Australia, but I do think that if Japan had attempted the invasion, they would have intended it—I don't think a full-scale attempted invasion would be the result of a navigation error, for example). Other starting points would be acceptable too—it would be permissible for me to suppose that Japan built more troop transports than they in fact did, or even that they might have diverted more troops and resources into the South Pacific theater. I could then happily work away at my "what if"—the Japanese land forces when they arrived in Australia would have started near the coasts, of course, and they would be unlikely to make Melbourne the beachhead for an invasion . . . and off I go. Some claims which I take to in fact be true would be inadmissible into the "what if": the claim that no more than a small handful of Japanese military ever voluntarily landed in Australia, for example, or the claim that the Prime Minister never came to believe that Australia was being invaded in force. So some known facts are inadmissible, and some claims known to be false are admissible as "starting points," or even as not explicitly stated information introduced during the course of the reasoning. How do I tell the difference when I am reasoning on the basis of a hypothesis which I do not accept? How can we tell, for example, to include the assumption that Melbourne is on the southern coast of Australia, but not the assumption that the Australian military never detected a landing in force? The answer should be obvious—we see what is true at the nearest world (or most relevantly similar world) where the explicitly given

"starting points" are true, and that is what is admissible—or at least the portion of it which is known is admissible. This gives us a different guide to which other starting points are admissible without changing the subject: the question of what would have happened if the Japanese had attempted an invasion but no Australian ever became aware of it is potentially a very different question from the question of what would have happened if the Japanese had attempted an invasion, so to insist on introducing the claim that no Australian ever became aware of a Japanese invasion is to begin to reason about a different hypothesis (and so not admissible when reasoning about the original hypothesis), even though it is the addition of something which is in fact true.

What can be reasoned to be a consequence of a hypothesis outruns the mere logical consequences of that hypothesis, even for perfectly everyday hypotheses. In working out those consequences, it is appropriate to employ other pieces of information—some, but only some, of our knowledge, plus some other claims. I think it is very plausible that the sort of information we can help ourselves to is just that information which holds in the nearest (most relevantly similar) worlds where the "starting points" of the hypothesis are true. This explains why we can legitimately suppose "in the scope of the hypothesis" about a Japanese invasion that Melbourne is on the southern coast of Australia, but we cannot legitimately suppose in the scope of the hypothesis that the Australian military never became aware of a Japanese invasion. At least we have no immediate warrant to: perhaps if it came to light that the Americans had many submarines in the Torres Strait which were capable of sinking transports and landing craft, but for some reason kept this fact from the Australians, we might then legitimately suppose that it could be that were the Japanese to have attempted an invasion, the Australian military would have been unaware of it. But notice that even in this case the sort of consideration that licenses the claim 'the Australian military were not aware of any attempted invasion of Australia by Japan' in the scope of the hypothesis of an attempted Japanese invasion is the sort of consideration that makes us accept the corresponding conditional: 'were Japan to have attempted to invade Australia, the Australian military would have been unaware of it'. There is a close link between reasoning hypothetically, or about what holds in hypothetical situations, and the conditions for accepting subjunctive conditionals.

On the other hand, reasoning categorically we would not normally have license to ignore any relevant knowledge. If I shifted to consider whether the Japanese did in fact attempt an invasion of Australia, then I cannot ignore the information that the Australian military never became aware of such an invasion. This information is clearly relevant to the question: for while it is not formally inconsistent with their being such an invasion attempt, it provides good evidence that there was no such attempt, since the Australian military were in a good position to notice any such attempt.

If I am right that hypothetical reasoning is quite similar to accepting or rejecting conditionals: if, to put it straightforwardly, one should accept B under the hypothesis A if and only if one is prepared to accept the subjunctive conditional if A then B: then when entertaining a theory which I do not currently accept as even possible, I should take B to be "part of" the theory T, or "flow from" the theory, just in case I accept the (counterpossible but nontrivial) conditional 'if T then B'. This allows even a classical logician to reason nontrivially about, for instance, an inconsistent theory: since

the theory will almost certainly be true at impossible worlds closer than the explosion world, and provided the theory is reasonably precise and not too absurd, it should even be possible to tell what sorts of things are relevant in determining which impossible worlds are closest. At least, this is so when T is taken to be a set of sentences or propositions given more or less explicitly: if one were to insist on taking T to be not only the explicit content but all of the (classical) logical consequences of the sentences or propositions given explicitly, then T would just describe the explosion world. Taking theories to consist of propositions or sentences given explicitly plus all of their consequences is one perfectly respectable conception of theories: it is the orthodox one, and is convenient for many purposes. But the notion of the "commitments" of a theory, or what "flows from" a theory (or whatever piece of jargon you want to reserve for these propositions), as consisting of the explicit content of a statement of a theory, plus those propositions that are consequents of true conditionals which take the explicit content as antecedents (let us call this 'T^{*}') is one that even classical logicians can help themselves to, and it is what we are often interested in when we examine or criticize views which we take to be impossible. It is a *commitment* of naïve set theory that there is a set of all non-self-membered sets: but it is not a *commitment* of the theory (though it is a classical consequence), that I am a radioactive motorcycle. And extensive reasoning in the scope of the hypothesis that naïve set theory plus some suitably weak logic are correct is something which is possible: Restall [28] is a good example. Such reasoning is even useful for those wishing to demolish theories they take to be impossible: for it is often easier (despite it being formally much more messy) to agree with one's opponents about what the *commitments* of their theories are than what the logical consequences are: and if it can be shown that a *commitment* of a theory is unacceptable, this will serve as a better refutation of the theory than showing that a logical consequence is unacceptable. To take a well-known example: suppose that bivalence is in fact correct, so that a proposition's not being true is sufficient for its being false. And consider the view of someone who takes a basic liar sentence, and indeed all such sentences ('this sentence is false', for example) to express a proposition—roughly, the proposition it appears to express—and accepts most of the usual naïve semantic machinery (T-schema, semantic closure, excluded middle, etc.) but rejects bivalence, and who attempts to avoid contradiction (which they regard as unacceptable) by supposing that there is a third, nondesignated truth value (*Other*). This truth value is defined to be such that a negation of a proposition with truth value O itself has truth value O, and that the liar sentence previously mentioned takes this truth value. There is a quick way of showing that this view has unacceptable logical consequences: the classic liar argument will show it to be inconsistent. The classical liar argument relies on bivalence, however, so will seem question-begging, and not convince the O-proponent. However, a slightly longer detour will serve better. It is not merely a consequence, but a *commitment* of this theory, that an extended liar sentence such as 'this sentence is not true' receives a truth value: and an argument from naïve semantic premises which does not appeal to bivalence will show that this leads to a contradiction even in a system with three truth values. The O-defender will be more likely to give up his/her theory when it is demonstrated that it has inconsistent commitments, since that makes it inconsistent by his/her own lights, and not merely by the lights of a defender of bivalence. When people make the well-known point

that adding a truth value doesn't automatically help much (or truth value gaps, for that matter), they typically use the extended liar to show this. And this is so even for those who think that the third-truth-value solution is not even logically possible (e.g., they may accept as a theorem all instances of '$Ta \vee T{\sim}a$'). Why? Just because it is more widely accepted that the extended liar is inconsistent given the commitments of the theory, whereas the classic argument relies on a principle which is very much at issue: a principle which a classical logician may be prepared to claim is a logical consequence of the 3-valuist view (if the principle is a theorem or axiom of the logic of truth), but should be wary of claiming that it is a *commitment* of the view.

The commitments of a theory, in this sense, will not contain every proposition, regardless of how strong one's favored logical consequence is. They provide a more useful handle on impossible rival theories, for the most part, and it is plausible that it is these we are considering even when we hypothetically consider a theory which is possible, but which we do not necessarily believe. In cases where the theory is consistent, of course, there need not be such a gap between its commitments and its logical consequences: for many such theories, there may be no gap at all between its commitments and the logical consequences of it and some common background assumptions. When a theory's possibility is in doubt, on the other hand, it is often important (and sometimes vital if one's base logic is classical) to distinguish them: and in such cases, it seems plausible, it is the commitments, not the consequents, which are important for exploring, evaluating, and criticizing the theory.

6 Conclusion Here then is an outline of a new way of approaching the task of dreaming the impossible dreams—or at least of engaging constructively with people who do so. Reasoning about impossibilities and impossible worlds is important, and the metaphysics of impossible worlds itself need not be terribly costly—certainly no more costly than possible worlds, on many accounts of what they are. The theory of counterpossible conditionals outlined goes a long way toward solving the logical problems of seriously assessing and reasoning about impossible situations, as well as being a natural extension of the standard Lewis-Stalnaker style accounts of the semantics of subjunctive conditionals more generally. Hypothetical reasoning, which can be represented as reasoning using subjunctive conditionals and which is plausibly equivalent to such reasoning, provides a way for even logically conservative classical logicians to assess the commitments of impossible theories. Impossible worlds can be had without extra suspicious metaphysics and without interfering with your favorite notion of logical consequence (whatever that might be). Since they are to be had, and serve useful functions, why not accept them? The price is low, and abundantly worth paying.

7 Appendix Let me briefly outline a model theory for the proposed closest-world conditional discussed. Since there are few constraints on impossible worlds, and few formal constraints on relative similarity, I do not take the conditional to have too many interesting formal properties. Nevertheless, a brief formal treatment is a concise way to bring together features argued for in the text. The semantics offered is modeled on the semantics offered by Lewis in [14] with a few variations. I will assume that the base logic is classical—it need not be so, of course, but I make this assumption for

simplicity (some of the features of the model will need to be altered if certain non-classical logics are adopted, and the reliance on the material conditional in the specifications of the semantics will need to be adapted, for example, for logics for which material detachment is not valid).

Let each model consist of a 5-tuple $\langle W, I, \pi, \$, \nu \rangle$: W is the set of possible worlds, I is the set of impossible worlds, π is the set of propositions, \$ is a function from worlds to systems of spheres of worlds (both possible and impossible), and ν is a function from a pair of a world (possible or impossible) and a proposition to a truth value (the function is fully defined for values of possible-world/proposition pairs), which represents "truth at a world." The function will be written, for a given world w and proposition p, $\nu_w(p) = x$, where x is a subset of $\{1, 0\}$. For convenience, I will take there to be four "truth values," the truth values being the subsets of the set $\{1, 0\}$, with $\{1\}$ being "true," $\{0\}$ being "false," $\{1, 0\}$ being "true and false," and $\varnothing$ as "neither true nor false." This slightly unusual machinery is to allow us to represent that some impossible worlds have propositions being both true and false according to them (when a proposition p takes the truth value $\{1, 0\}$), and that other impossible worlds might have "gaps," and have some proposition being neither true nor false at such a world (where p takes the truth value $\varnothing$). An alternative method of proceeding is to take n to be a partial relation, with 1 and 0 as the only truth values.[27]

One could also add an accessibility relation on worlds in the usual manner: I have not done so here for simplicity. If I were to do so, the accessibility relation would be standardly defined on possible worlds, and for normal modal logics at least I would stipulate that no impossible world was accessible from any possible world. One could also introduce a distinguished actual world, @, and define truth or falsehood as follows: a proposition p is true just in case $\nu_@(p) = \{1\}$ and false just in case $\nu_@(p) = \{0\}$ (@ is, of course, a possible world, and as we shall see these are the only two values propositions can take at possible worlds).

Possible worlds will behave in the usual way in the model: to begin with, the only truth values that propositions can take at possible worlds are $\{1\}$ and $\{0\}$. The truth value of truth functional compounds of propositions will be determined in the usual way at possible worlds. Take conjunction and negation as primitive, and define the other truth-functional connectives in the usual way. At a given world $w \in W$,

$$\nu_w(\sim p) = \{1\} \text{ just in case } \nu_w(p) = \{0\}$$

and

$$\nu_w(\sim p) = \{0\} \text{ otherwise.}$$

$$\nu_w(p \,\&\, q) = \{1\} \text{ just in case } \nu_w(p) = \{1\} \text{ and } \nu_w(q) = \{1\}$$

and

$$\nu_w(p \,\&\, q) = \{0\} \text{ otherwise.}$$

However, no such constraints can be assumed for the impossible worlds. As I say on p. 542, I think that a very generous comprehension principle for impossibilities is called for: and I model this by not putting any constraints on assignment of truth values to propositions at impossible worlds. So when $w \in I$, I will allow that $\nu_w(p \,\&\, q) = \{1\}$ might be the case even though it might not be the case that $\{1\} \in \nu_w(p)$ or that $\{1\} \in \nu_w(q)$, or even that they may both fail. Of course, there

will be many $w \in I$ which do obey the constraints which possible worlds obey for most propositions and truth-functional compounds of propositions (and in the intended model they tend to be "closer"), but there are less well-behaved worlds in I as well. Theoremhood, then, is truth at all possible worlds in the model (since there would be no theorems if it was defined in terms of truth at all worlds). Similarly for validity: an argument is formally valid just in case in every model the conclusion is true at every possible world in which the premises are true.

Other logical operations are treated analogously to the truth-functional operations: when $w \in W$, $\nu_w(Lp) = \{1\}$ just in case, for all $v \in W$, $\nu_v(p) = \{1\}$, and $\nu_w(Lp) = \{0\}$ otherwise. (Alternatively, $\nu_w(Lp) = \{1\}$ just in case for all worlds v accessible from w, $\nu_v(p) = \{1\}$, if an accessibility relation is used. Or it may even be that it should be that for all worlds v accessible from w, $\{1\} = \nu_v(p)$, if a nonnormal modal logic utilizing impossible worlds as the nonnormal worlds is being employed.) Were quantificational logic to be introduced, the quantifier rules would similarly be restricted to possible worlds, with an "anything goes" approach to the behavior of quantifier propositions at impossible worlds. One may wish to add as a constraint on the impossible worlds that each fails to obey at least one constraint which the possible worlds obey (that's what makes them impossible, after all): for convenience, I will not bother to add this constraint here.

Because of this laxity about impossible worlds, it is best not to identify propositions with arbitrary sets of worlds. It is an advantage of recognizing impossible worlds that a treatment of propositions as sets of worlds gains the ability to distinguish between logically equivalent propositions (they will have the same possible worlds in them, but differ with respect to which impossible worlds are found in them). However, not every arbitrary set of worlds should count as a proposition once enough impossible worlds are admitted—since an impossible world w which is obtained by adding further things true to all of the things true at a possible world v (and surely there will be such worlds) will be such that v will occur in every proposition in which w occurs (on pain of a proposition being true at v which is not true at w): so those sets containing w but not v should not count as propositions. A propositions-as-sets-of-worlds account is still viable, and indeed such an account will provide useful models in the same sorts of cases as before, but it should I think be an account where only specified subsets of the total domain of worlds count as propositions. So for simplicity in this model I have not pursued such a definition, but treated propositions separately.

Thus far I have discussed W, I, π, and ν. It remains only to discuss \$, and to define the nearest-world conditional in terms of these resources. \$, as is standard, is a function from worlds to sets of spheres of worlds, where a sphere of worlds is a set of worlds (either possible or impossible) which obeys some constraints.[28] $\$_i$ is the set of spheres associated with the world i. Three constraints on systems of spheres are completely standard ([14], p. 14):

1. Each of the spheres is *nested*: when two spheres S and T are members of any given $\$_i$, then either S is a subset of T or T is a subset of S.
2. The spheres are closed under unions: whenever S is a subset of $\$_i$ and $\bigcup S$ is the set of all worlds which are members of members of S, then $\bigcup S$ is a member

of $\$_i$.

3. The spheres are closed under intersection: whenever S is a nonempty subset of $\$_i$ and $\bigcap S$ is the set of all worlds which belong to all of the members of S, then $\bigcap S$ is a member of $\$_i$.

There will be more conditions put on spheres in a moment—these are just a bare minimum. Once the spheres are properly defined, a conditional of the sort I am discussing, of the form $\varphi \rightarrow \psi$ will then be true at a possible world in the model just in case, to use Lewis's definition, either there is no world where φ is true which is a member of any of the spheres associated with that world, or that there is a sphere associated with that world in which $\varphi \supset \psi$ holds in every member of that sphere ([14], p. 16). Intuitively, this means that either there are no φ worlds available, or that there is a φ & ψ world closer than any φ & $\sim\psi$ world. (If one employs the Limit Assumption (see [14], pp. 19–21) one can define the conditions in other ways which are in some respects more initially intuitive—but I will not rely on a Limit Assumption here). In symbols:

$$\text{for all } w \in W, \nu_w(\varphi \rightarrow \psi) = \{1\}$$

if and only if

$$(\forall x)(Wx \,\&\, \nu_x(\varphi) = \{1\}) \supset ((\forall S)S \in \$_w \supset x \notin S)$$

or

$$(\exists T)(\exists x)(Wx \,\&\, \nu_x(\varphi) = 1 \,\&\, x \in T \,\&\, T \in \$_w \,\&\, (\forall y)((Wy \,\&\, y \in T) \supset \{1\} = \nu_y(\varphi \supset \psi))))$$

and

$$\nu_w(\varphi \rightarrow \psi) = \{0\} \text{ otherwise.}$$

'W' here is used as a predicate for worlds, both possible or impossible: it covers any member of W or I.

Again, notice that the definition is restricted to saying when conditionals are true in possible worlds. I am inclined to think that some impossible worlds do not obey this constraint (just as they disobey many other constraints of possible worlds): which conditionals are true at which impossible worlds has no formal constraints on it in the system as it stands (and I think no formal constraints will hold of all impossible worlds, though distinguished "better behaved" subclasses of them may behave much as possible worlds do in this regard).

It remains only to specify what other conditions hold for the spheres. There will be a certain amount of controversy about what conditions should be met, but the following three from [14], p. 120 are standard.

Normality (N):	$\$$ is *normal* iff, for each i in $(W \cup I)$, $\bigcup \$_i$ is nonempty.
Total Reflexivity (T):	$\$$ is *totally reflexive* iff, for each i in $(W \cup I)$, i belongs to $\bigcup \$_i$.
Weak Centering (WC):	$\$$ is *weakly centered* iff, for each i in $(W \cup I)$, i belongs to every nonempty member of $\$_i$, and there is at least one nonempty member of $\$_i$.

(As Lewis notes, WC implies T, and T implies N, but not conversely). WC ensures the validity of modus ponens: when $\varphi \rightarrow \psi$ is true at a possible world w, and φ is true at w, there will be a sphere in $\$_w$ which contains a φ world (since w is a φ world and w belongs to every sphere in $\$_w$), and so, since $\varphi \rightarrow \psi$ is true at w, there will be a sphere in $\$_w$ such that all of its members will be $\varphi \supset \psi$ worlds: since w is a member of that sphere, $\varphi \supset \psi$ will hold at w: and since φ and $\varphi \supset \psi$ both hold at w, ψ will do so too.

The conditions on \$, as they are currently stated, do not guarantee the theoremhood of *identity* $(\varphi \rightarrow \varphi)$. A model of this failing is a model where the smallest sphere around a possible world i contains two worlds: i, and an impossible world where $\varphi \supset \varphi$ fails (either through receiving the truth value $\{0\}$ or through receiving the truth value $\varnothing$). $\varphi \rightarrow \varphi$ will fail to be true at that world (since the spheres are nested, if the smallest sphere in $\$_i$ contains a world where $\varphi \supset \varphi$ fails, all the spheres in $\$_i$ will contain that world). This may be seen as an infelicity of the current definition: perhaps the second clause of the definition should say, of the relevant sphere(s) which contains a φ world that for each of the worlds w in that sphere, either φ fails to be true at w or both that φ is true at w and ψ is true at w, where 'φ fails to be true' means that $1 \notin \nu_w(\varphi)$. With this amended definition, identity does become a theorem (since every world, even impossible ones, where φ is true, φ will be true—even though it may be false as well, and even though $\varphi \supset \varphi$ might be false at that world). And as I mentioned in the text, one might try other ways of defining the truth-conditions for the conditional if one wished identity to fail to be theorem (perhaps if one thought there were counterexamples with sufficiently bizarre antecedents).

There are assorted other principles which one may or may not add to one's conditional, according to philosophical taste (see [14], p. 120 for many more): I think it is useful to add

Universality (UT):	\$ is universal if and only if, for each i in $(W \cup I)$, $\bigcup \$_i$ is $(W \cup I)$ (i.e, all of the worlds appear in the system of spheres).

There are three other common additions that I myself would be wary of adding but that often are included (all of which can be found in [14], p. 120).

Centering (C):	\$ is centered if and only if, for each i in I, $\{i\}$ belongs to $\$_i$.
Limit Assumption (L):	\$ satisfies the *Limit Assumption* if and only if, for any proposition φ, if it is true at any members of a given $\bigcup \$_i$, then there is some smallest member of S_i which has as a member a world where φ is true.
Stalnaker's Assumption (S):	\$ satisfies *Stalnaker's Assumption* if and only if, for any φ, if it is true at any members of a given $\bigcup \$_i$, then there is a member of $\$_i$ which has exactly one φ-world among its members (where a φ world is just a world i such that $1 \in \nu_i(\varphi)$).

Lewis himself defends C, while L and S are features of Stalnaker's preferred system. C delivers the inference $A \,\&\, B \vdash A \rightarrow B$. However, Stalnaker's Assumption is not

so well named in the current system, since it does not deliver conditional excluded middle $((A \rightarrow B) \vee (A \rightarrow \sim B))$, the principal theorem which rests on S in standard Stalnaker and Lewis style systems. To see why conditional excluded middle fails, consider where, for a $\$_i$ for a given i, the closest sphere containing a φ-world contains exactly one φ-world (though it may contain many non-φ-worlds), and that world is an impossible world where neither ψ nor $\sim\psi$ holds. Neither $\varphi \rightarrow \psi$ nor $\varphi \rightarrow \sim\psi$ holds in i in such a case, but S does hold.

The only substantially new condition I suggest (though I am hesitant to endorse) for this conditional is the strangeness of impossibility condition, discussed in the text:

Strangeness of Impossibility Condition (SIC):	For any $i \in W$, any $v \in W$ and any $w \in I$, if $v \in \bigcup \$_i$ and $w \in \bigcup \$_i$, then there is a sphere S such that $S \in \$_i$ and which is such that it has v as a member and does not have w as a member.

This condition ensures that any possible world which is a member of any sphere which is a member of a given $\$_i$ is "closer" to i than any of the impossible worlds which are members of the spheres in $\$_i$, where i itself is a possible world.

There are several variants of this base. The clause stipulating that w and v must both be members of $\bigcup \$_i$ can safely be dropped if UT is already present in the system, since UT will ensure that in any case. The restriction that i must be a possible world can also be dropped—let us call this the *Extended Strangeness of Impossibility Condition*, or ESIC—but caution must be taken. Dropping this restriction allows that possible worlds are all "closer" to impossible worlds than any impossible world is: a proposition which will not be thought attractive by many, and which is inconsistent with WC and C unless no possible worlds are to be found in $\bigcup \$_i$, or there are no impossible worlds at all in the model. Obviously, the triad of ESIC, UT, and WC is inconsistent with the assumption that there are impossible worlds, as is, a fortiori, the triad ESIC, UT, and C.

One could also relax SIC, while still retaining some of the thought behind it, and claim only that no impossible world was less distant than any possible world. This Lesser Strangeness of Impossibility Condition might be stated as follows:

Lesser Strangeness of Impossibility Condition (LSIC):	For any $i \in W$, any $v \in W$ and any $w \in I$, if $v \in \bigcup \$_i$ and $w \in \bigcup \$_i$, then there is no sphere S such that $S \in \$_i$ and which is also such that it has w as a member and does not have v as a member.

Of course LSIC could also be extended by dropping the constraint that i must be a possible world. Call this extension ELSIC. This extension is slightly more attractive than ESIC, since it no longer conflicts with WC. This is because the innermost sphere around an impossible world j could be composed of all of the possible worlds, as well as j (as well as perhaps other impossible worlds). Intuitively this will probably still not seem terribly attractive, however. Furthermore, ELSIC, UT plus C will still be inconsistent with the assumption that there are impossible worlds, as will the combination of ELSIC, UT, and S, since S implies C.

The point of having some form of Strangeness of Impossibility Condition is that it ensures much more formal predictability for conditionals with possible antecedents.

SIC permits conditionals with possible antecedents to be treated with a more standard Lewis or Stalnaker conditional logic: when an inference is formally valid in the appropriate possible-worlds-only logic, the conclusion will be true when the premises are, when the antecedent is possible. Furthermore, one is guaranteed that a conditional which is a tautology in the appropriate possible-worlds-only system will be true (though probably not a theorem) when the antecedent is possible, given SIC. LSIC provides much less reassurance in this regard, since impossible worlds can be among the equally closest worlds for the purpose of an evaluation of a conditional.

Not many interesting theorems or inference patterns emerge from this system, but I do not think that this is a drawback, but is a reflection of how ill-behaved impossible worlds are, and so how bizarre the results might be when sufficiently impossible antecedents are employed. However, while this may limit the algorithmic usefulness of the theory to an extent, the system modeled may still be powerful enough to carry on our hypothetical consideration of impossibilities. The intended model(s), after all, are those in which much more information about which impossible worlds are nearer, or more similar, than the others: and it is primarily facts about the relative closeness of spheres which determine which conditionals are true and which inferences are acceptable, as opposed to which are theorems or which inferences are valid in virtue of their logical form.

Acknowledgments Thanks to James Chase and Graham Priest, and special thanks to Greg Restall for discussion of the topics of this paper: though they should not be held responsible for the opinions expressed!

NOTES

1. See e.g., Routley et al. [32], p. 58, Priest and Routley [24], pp. 151–3.
2. Some prefer to talk of situations rather than worlds. Myself, I do not see an important difference for many purposes between constructing situations from worlds, or worlds from situations. (There would be a more important difference if the concept of worlds as totalities was incoherent as some have claimed (e.g., Grim [11]). But I, for one, do not think it is.) In any case, the game is very similar—the provision of ontological correlates for modal claims which provide advantages such as easy generalizations (there are several ways things could be, a convenient generalization which is difficult to express in a language which has modal operators as its only modal resources). In what follows, I have no real objection if people want to reconstrue my arguments for impossible worlds merely as arguments for impossible situations.
3. This view is not unchallenged: famously, Field takes the claims of mathematics to be literally false and so certainly not necessarily true (Field [7]). Furthermore, he defends the contingency of his claim about mathematical objects (Field [8]). He also seems prepared to admit that two conflicting mathematical systems can sometimes both be logically possible (see [7], pp. 240–2, where he claims that various of the rival systems are conservative, together with pp. 250–2, which provides a definition of conservativeness in terms of logical possibility). However, the worries of this section will have force for the many who do take the truths of mathematics to hold in all possible worlds. (Furthermore, see the next note for a similar problem, even if Field is right).
4. Indeed, we may have coherent and interesting claims being made even should the truths of mathematics be contingent since then, all who claim their preferred mathematical sys-

tem is necessarily true claim something which is necessarily false, should **S5** be the correct modal system.

5. The claim that Lewis denies this may sound strange to some since he indeed thinks that there exist many disconnected spacetimes since these are the possible worlds (or some of the possible worlds). But to assess a modal claim involving a modal operator such as 'could', the question becomes whether any world contains several disconnected spacetime regions not otherwise linked by some external relation: and there is no such world, according to Lewis, since disconnected spacetimes (not otherwise connected . . .) are ipso facto parts of different worlds (see Lewis [17], pp. 69–78, especially pp. 71–72). At the very least, this means that there is one sensible interpretation of the claim that there could not be several disconnected spacetimes on which Lewis should agree. There *may* be another sense in which he should disagree: the sense in which there are several disconnected spacetimes (when I employ a less restrictive quantifier) may be a sense in which there could be several disconnected spacetimes, too.

6. I intend to necessitate the consequent and not merely the conditional here: I mean not only that it is impossible that more than one alternative be true at once, but also that the alternatives which are not true are necessarily false.

7. As always, it is difficult to signal simple falsity when one is talking in a context where dialetheism is in the air—but perhaps I can put it this way: I deny that Anselm's God both literally exists in this worls and literally does not exist in this world.

8. Among others, Stalnaker [33], Davis [4], and Nolan (in a forthcoming paper) argue that indicative conditionals are to be handled with possible worlds as well (or at least I argue that this position is more attractive than many have thought), but perhaps the majority of writers in the area do not: Lewis [14] and [19], Jackson [13], Edgington [5] and [6], Bennett [3], etc.

9. All of the writers mentioned in the previous note believe this, with the exception of Edgington [6].

10. Even this way of putting it might sound as if it is assuming the limit assumption (see [14], p. 19). For an official statement of the truth-conditions of the conditional, see the appendix to this paper.

11. Those who tie possibility to conceivability (and some who do not) will, of course, claim that it is not conceivable. It is odd that we can communicate intelligibly about it if it is so inconceivable, and myself, I seem to be able to conceive what is involved with a world being the explosion world. But for those who insist that it is not conceivable, they can read me as saying that it is one of the most absurd situations, even of the inconceivable ones.

12. There are other explanations of my reluctance to lay this charge, of course—even were I to think that the charge was, strictly speaking, correct, I might resist making it as being unhelpful or question-begging (since I have no argument for it from premises which my opponent would accept, or even from premises which my opponent might find more plausible than arguments directly leveled against the impossibility of her view). So the intuition that the charge would be inappropriate needs to be handled with care. Even with this care taken, however, I suspect that my intuition that the charge is incorrect (and not just unhelpful or ill-advised) is one which is shared.

13. Given a technical sense of 'theory', according to which a theory must be closed under logical consequence (and a similarly technical sense of 'deductive situation'), Routley et al. are right to point out a tension between accepting disjunctive syllogism and accepting nontrivial but inconsistent "theories" or "deductive situations." But on a more intuitive

understanding of what a theory or a deductive situation might be, nontrivial reasoning about impossibilities need not jettison any classical principles. See Section 5 for details of how such reasoning might proceed.

14. Those who, in fact, accept the inconsistent theories will need a paraconsistent logic, of course, but my argument is that those of us considering the theories and reasoning about them need not.

15. By 'would not be true', I mean to rule out its being true or reject the claim that it is true: dialethetic compromises, where the statement is not true, but is true as well are not good enough for the sense which I intend to be using here.

16. Of course, my statement of my position might not quite capture my difference from my imagined opponent: for if they think it impossible that for some A and B, 'A and B' holds but A fails to hold and B fails to hold, then they may deny that it is an impossibility that this is the case (since quantifying over impossibilities already seems committed to impossible situations). One of the advantages of a commitment to impossible worlds is an ability to quantify over ways things can't be: my opponent lacks such a general quantificational device if they think that some propositions are impossibly true but there is no impossible situation (impossible world, way things can't be, or whatever) such that those propositions describe that situation.

17. Again, by 'fails to hold' I mean to rule out its holding or reject the claim that it holds: it is not enough that it fails to hold, if it holds as well (as some dialetheists might allow).

18. Partisans of **S** as the correct alethic logic (if there are any) are invited to substitute another example.

19. It should be clear that what I am calling conditional proof is not the only inference that deserves the name: the inferences from $\Sigma(A) \vdash B$ to $\Sigma - A \vdash A \supset B$ and $\Sigma(A) \vdash B$ to $\Sigma - A \vdash L(A \supset B)$, among others have some claim to the title, and these inferences, like all classical inferences not concerning the conditional, are unaffected by the addition of the proposed counterpossible conditional. The conditional proof I am concerned with is only that conditional proof which results in a statement involving the conditional: and for this conditional, both $\Sigma(A) \vdash B$ to $\Sigma - A \vdash A \rightarrow B$; and $A \vdash B$ to $\vdash A \rightarrow B$ fail.

20. Priest claims that the logical laws are different at nonnormal worlds rather than that they are the same but are broken (see [23]). I do not think this difference makes a difference in this context: what is not disputed is that our laws of logic fail to be truth-preserving there, but whether our laws of logic are *laws of logic* in impossible worlds, or whether it is rather that there are other laws (or no laws at all), seems to me to be largely a matter of terminological decision. Myself, I am tempted to think that in some impossible worlds it is our laws of logic which hold, but which are broken, and in others it is different laws of logic which hold (and some where the laws are different and are broken too).

21. Another example: take someone, in awe of Gödel's ability and believing that Gödel has one of the best mathematical intuitions, who is prepared to assert 'for any mathematical proposition, if Gödel had come to believe it, it would have been true'. He/she might come to accept 'if Gödel had believed Fermat's last theorem to be false, it would have been', and accept that even though he/she believes Fermat's last theorem to be true, and necessarily so. (Of course, when you remind him/her of this, he/she will remind you that he/she does not believe that Gödel would have believed it). The point is not that it is rational of him/her to think this of Gödel or that any of his/her beliefs are correct (even his/her conditional belief)—it is just that if the known-to-be counterpossible conditional he/she asserts is an appropriate one for him/her to assert given what he/she believes, then he/she is treating SIC as having exceptions—in this context, it seems appropriate for him/her to take it that the world where Fermat's last theorem is false, strange though it would be, is not as bizarre as the world where Gödel makes that sort of mistake.

22. There is another introduction rule which some will wish to add: those that believe that the nearness-conditional should be governed by centering will hold that one can infer $A \rightarrow B$ from A & B. I am not myself a fan of centering, though I do endorse the condition known as weak centering (that the actual world is at least as near to itself as any other world or that the actual world is at least as relevantly similar to itself as any other world).

23. Or that I was in the grip of a philosophical theory—people are so much more understanding of strange outbursts and conditional utterances when they've been exposed to philosophers and philosophical theories.

24. An exception might be if consequents of conditionals serve as one of the determinants of context: pragmatic rules of accommodation will often tweak the relevant similarity relation to ensure the utterance is correct (see Lewis [16] for a general story), and so different consequents may produce a difference in context which causes a different antecedent-world to be selected: 'if he sells drugs, he'll be wealthy' and 'if he sells drugs, he'll be caught and have his drug money confiscated plus face a massive tax bill' might both be claims that are taken to be acceptable when uttered: the first because there is a lot of money to be made selling drugs (and this is made especially salient and relevant in determining the similarity relation), the second because he's very likely to get caught (and this fact is made especially salient by its being mentioned, and is made much more important in determining the similarity relation in the second case). But it may not be appropriate to infer that 'if he sells drugs, he'll be wealthy and have his drug money confiscated and face a massive tax bill', since the confiscation and tax bill would serve to prevent his becoming wealthy. This example might not spark everyone's intuitions: but if consequents are a factor in determining relevant similarity, there will most likely be room for some such failure of implication. Another example: take the conditionals 'if wishes were horses, beggars would ride' and 'if wishes were horses, then beggars wouldn't be able to afford to have any wishes anymore'.

25. It is more convenient to talk about this system in terms of selection functions, and picking out a world (rather than worlds or spheres): formally, this may not be ideal, if modeling nearest-world conditionals with a selection function, since it seems to presuppose the dubious limit assumption (see [14], pp. 57–60). I thus do not employ a sematics of selection functions in the appendix to this paper—but I will employ this language in the text, as once the points are grasped, it is easy enough to see how to extend them to a system without selection functions or the limit assumption.

26. It will be controversial whether there is such a null world, even among the impossible worlds. A generous comprehension principle may well allow it, though—it forms a nicely symmetric dual with the explosion world where everything is true. In any case, I will treat it as if it is an impossible world for exploring this option. However, I do not think that a great deal more than this example hangs on the question.

27. Restall has pointed out that given my lack of constraints on negation in impossible worlds, the system can also be formulated with just the two truth-values 1 and 0, while retaining the functionality of ν. A world where p would be assigned "both" is, in the alternative formulation, a world where p and $\sim p$ are both assigned 1, and a world where p would be assigned "neither" is, in the alternative formulation, a world where p and $\sim p$ are both assigned 0.

28. A more sophisticated approach would be to employ, instead of a function from worlds to sets of spheres, a function from worlds and contexts to sets of spheres. This may deal better with representing our use of relevant similarity in determining the spheres, since what is more relevantly similar than what is often (always?) a matter of context. Contexts themselves are not monolithic, of course, and there is a potential to develop a quite

sophisticated formal mechanism for modeling the selection of sets of spheres. I ignore these subtleties here for simplicity and convenience.

REFERENCES

[1] Anderson, A. R., Belnap, N. D., and Dunn, J. M., *Entailment*, vol. 2, Princeton University Press, Princeton, 1992. Zbl 0921.03025 MR 94b:03042 4

[2] Armstrong, D. M., *A Combinatorial Theory of Possibility*, Cambridge University Press, Cambridge, 1989. 3

[3] Bennett, J., "Classifying conditionals: the traditional way is right," *Mind*, vol. 104 (1995), pp. 331–54. MR 1337607 7

[4] Davis, W. A., "Indicative and subjunctive conditionals," *Philosophical Review*, vol. 88 (1979), pp. 544–64. 7

[5] Edgington, D., "Do conditionals have truth-conditions?" *Critica*, vol. 18, 52 (1986), pp. 3–30. Reprinted in *Conditionals*, edited by F. Jackson, Oxford University Press, Oxford, 1991. MR 1260670 7

[6] Edgington, D., "On conditionals," *Mind*, vol. 104 (1995), pp. 235–329. MR 97a:03026 7, 7

[7] Field, H., *Realism, Mathematics and Modality*, Basil Blackwell, Oxford, 1989. MR 92b:03003 7, 7

[8] Field, H., "The conceptual contingency of mathematical objects," *Mind*, vol. 102 (1993), pp. 285–99. MR 1219711 7

[9] Fraenkel, A. A., Y. Bar-Hillel, and A. Levy, *Foundations of Set Theory*, 2d revised edition, North-Holland, Amsterdam, 1973. Zbl 0248.02071 MR 49:10546 2

[10] Forbes, G., "Physicalism, instrumentalism and the semantics of modal logic," *Journal of Philosophical Logic*, vol. 12 (1983), pp. 271–98. Zbl 0513.03004 MR 84k:03013 3

[11] Grim, P., *The Incomplete Universe: Totality, Knowledge and Truth*, The MIT Press, Cambridge, 1991. 7

[12] Hinckfuss, I., "Suppositions, presuppositions, and ontology," *Canadian Journal of Philosophy*, vol. 23 (1983), pp. 595–617. 3

[13] Jackson, F., *Conditionals*, Basil Blackwell, Oxford, 1987. 7

[14] Lewis, D. K., *Counterfactuals*, Basil Blackwell, Oxford, 1973. MR 54:9979 4, 5, 5, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7

[15] Lewis, D. K., "Counterfactuals and comparative possibility," *Noûs*, vol. 13 (1979), pp. 455–76; reprinted in *Philosophical Papers*, vol. 2, Oxford University Press, Oxford, 1986. 4

[16] Lewis, D. K., "Scorekeeping in a language game," *Journal of Philosophical Logic*, vol. 8 (1979), pp. 339–59. 7

[17] Lewis, D. K., *On the Plurality of Worlds*, Basil Blackwell, Oxford, 1986. 3, 7

[18] Lewis, D. K., "Postscript to 'Counterfactual dependence and time's arrow'," pp. 52–66 in *Philosophical Papers*, vol. 2, Oxford University Press, Oxford, 1986. 4

[19] Lewis, D. K., "Postscript to 'Probabilities of conditionals and conditional probabilities'," in *Philosophical Papers*, vol. 2, Oxford University Press, Oxford, 1986. 7

[20] Merrill, G. H., "Formalization, possible worlds and the foundations of modal logic," *Erkenntis*, vol. 12 (1978), pp. 305–27. 3

[21] Martin, E. P., and R. K. Meyer, "Solution to the P-W problem," *The Journal of Symbolic Logic*, vol. 47 (1982), pp. 869–86. Zbl 0498.03011 4

[22] Priest, G., *In Contradiction*, Martinus Nijhoff, The Hague, 1987. Zbl 0682.03002 MR 90f:03007 4

[23] Priest, G., "What is a non-normal world?," *Logique et Analyse*, vol. 35 (1992), pp. 291–302. Zbl 0834.03002 MR 98k:03056 7

[24] Priest, G., and R. Routley, "Systems of paraconsistent logic," pp. 151–86 in *Paraconsistent Logic: Essays on the Inconsistent*, edited by G. Priest, R. Routley, and J. Norman, Philosophia Verlag, Munich, 1989. Zbl 0697.03015 7

[25] Priest, G. and R. Routley, "The philosophical significance and inevitability of paraconsistency," pp. 483–539 in *Paraconsistent Logic: Essays on the Inconsistent*, edited by G. Priest, R. Routley, and J. Norman, Philosophia Verlag, Munich, 1989. Zbl 0688.03005 4

[26] Quine, W. V., *Philosophy of Logic*, Prentice Hall, Englewood Cliffs, 1970. MR 57:9465 2

[27] Read, S., *Thinking About Logic*, Oxford University Press, Oxford, 1995. 4, 4

[28] Restall, G., "A note on naïve set theory in **LP**," *Notre Dame Journal of Formal Logic*, vol. 33 (1992), pp. 422–32. Zbl 0768.03033 MR 93k:03051 5

[29] Rosen, G. "Modal fictionalism," *Mind*, vol. 99 (1990), pp. 327–54. MR 1070667 3

[30] Routley, R., *Exploring Meinong's Jungle*, Australian National University, Departmental Monograph No. 3, Canberra, 1980. 3, 4

[31] Routley, R., "Philosophical and linguistic inroads: multiply intensional relevant logics," pp. 269–304 in *Directions in Relevant Logic*, edited by J. Norman and R. Sylvan, Kluwer, Dordrecht, 1989. 4

[32] Routley, R., R. Meyer, V. Plumwood, and R. Brady, *Relevant Logics and Their Rivals*, Ridgeview, Atascadero, 1982. Zbl 0579.03011 MR 85k:03013 4, 7

[33] Stalnaker, R., "Indicative conditionals," *Philosophia*, vol. 5 (1975), pp. 269–86. 4, 7

[34] Stalnaker, R., *Inquiry*, The MIT Press, Cambridge, 1984. 3

[35] van Fraassen, B., "Probabilities of conditionals," pp. 261–308 in *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, vol. 1, edited by W. L. Harper and C. A. Hooker, D. Reidel, Dordrecht, 1976. Zbl 0339.02025 MR 58:16148 3

[36] van Inwagen, P., "Two concepts of possible worlds," *Midwest Studies in Philosophy*, vol. 11 (1986), pp. 185–213. 3

School of History, Philosophy, and Politics
Macquarie University
Sydney NSW 2109
AUSTRALIA
email: dnolan@laurel.ocs.mq.edu.au