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

EDITOR'S INTRODUCTION

GRAHAM PRIEST

1 Background to the issue The origin of this special issue lay in a visit that Mic Detlefsen, Editor of the *Notre Dame Journal of Formal Logic*, made to Australia in 1995 as a guest of the Australasian Association of Logic and the Department of Philosophy at the University of Queensland. The visit allowed many enjoyable discussions concerning logic and philosophy. On one occasion, discussion turned to the fact that, though logicians in Australasia and Northern America work in much the same areas and publish in the same journals, there are, nonetheless, subjects on which they tend to have rather different perspectives. On why this is, the discussion was somewhat inconclusive. But after Mic's return to the U.S., it was decided to produce an issue of the *Journal* in which one of these differences would be explored in more detail. After some deliberation, we decided to make the topic that of impossible worlds. This bears on a number of different, but substantial, issues, both formal and philosophical, many of which are represented in the papers in this issue. Moreover, we thought, Australasian logicians are much more sympathetic to the notion of such worlds than North American logicians. Hence, the topic seemed a good one. Accordingly, we wrote to a number of philosophers and logicians on both continents. Essentially, this issue contains the product. The result is, perhaps, a rather surprising one. (At least, it surprised me.) All the papers, whether written by Australasians or North Americans, are sympathetic to the notion of an impossible world. The difference between the two groups emerged elsewhere. I will return to this matter later in this introduction.

I would like to thank Mic warmly for everything that he has done to make this issue itself not impossible. In the rest of this introduction, I want to put the papers in the collection, and the topics with which they deal, into some sort of perspective.

2 What is an impossible world? The first question that needs to be addressed is "what is meant by 'impossible world'?". Let us take the notion of a world itself for granted for the nonce. There are still many different kinds of impossibility: epistemic, physical, metaphysical, logical. Though some of the contributors (notably Barwise) cast their nets wider, it is primarily logical impossibility that is the focus of the issue. But what is meant by 'a logically impossible world'?

Received March 5, 1998; revised March 31, 1998

There are several things that might be caught by this rubric. One thing, probably the most obvious, is simply a world where some contradictory sentences, of the form α and $\neg\alpha$, hold. A rationale for calling such a world impossible is simply that such a pair of sentences can be true in no classical interpretation. But if this is the rationale, it is natural to extend the epithet 'impossible world' to any world at which the set of things that hold (at least as far as the connectives $\wedge$, $\vee$, $\neg$, and $\rightarrow$ go) is not the set of things that hold in any classical interpretation. This obviously generalizes the first use; not just because there might be things other than explicit contradictions that hold in the world, though in no classical interpretation (such as something of the form $\neg(\alpha \rightarrow \alpha)$); but because there might be things that *fail to hold* in the world but that hold in every classical interpretation (such as something of the form $\alpha \vee \neg\alpha$). In this sense, a world where the set of things that hold is the set of truths of first-order intuitionist arithmetic, is just as much an impossible world.

There is yet a third thing that one might mean by 'impossible world'. The rationale for the first two definitions is that classical logic is the correct logic. But one might, of course, disagree with this, endorsing instead some other logic, L. Thus, more generally, one might define an impossible world to be one where the set of truths is not one that holds in any interpretation of L. Thus, if one takes intuitionist logic to be correct, then a world that verifies exactly the truths of intuitionist arithmetic is a possible world; but a world where a contradiction holds is still impossible.

It should be noted, though, that an impossible world in the first sense (where a contradiction holds) is not necessarily an impossible world in the third sense. If one takes the correct logic to be a paraconsistent or relevant logic, there are interpretations in which a contradiction may be true. Indeed, one might even suppose that the set of logical truths is itself inconsistent. (This is less bizarre than might at first be thought if one takes the T-schema to be a logical truth; the schema may give rise to familiar contradictions.)

It might be thought that once one goes this far, there is nothing that will count as a logically impossible world, but this is not so. In a relevant logic, for example, all the logical truths still hold in each interpretation—that is, at its base world or, more generally, at any normal (regular) world in the interpretation. But there will, in general, be other worlds in it where logical truths fail. These are therefore impossible worlds in the third sense.

It should be noted that the question of whether there are impossible worlds is not the same as that of whether contradictions can be true (dialetheism). Even where the connection is closest, in the case of impossible worlds of the first kind, the connection is only partial. The dialetheist must hold that there are impossible worlds in this sense: the actual world is one. But one may hold that there are such worlds for various reasons (which we will come to in a moment), without supposing that the actual world may be one of them, though this at least forces us to ask the question 'why not?'.

3 Applications of impossible worlds Having clarified the notion of impossible world, let us turn to its applications. In any paraconsistent logic in which logical consequence may be defined semantically, there will be interpretations where contradictions hold (but where not everything does). Now, interpretations may not be the same as worlds—even when we have a "one world" interpretation, as in classical logic; but it is at least natural to think of an interpretation as representing a world or worlds in a certain way. In this case, impossible worlds of the first (and sometimes second) kind are playing an essential role in these semantics. Close encounters with impossible worlds of the third kind occur specifically in the world semantics of relevant logics. Indeed, it was one of the great insights leading to the production of these semantics by the Routleys and Meyer that there should be worlds where logical truths fail. This is the key idea behind getting "irrelevances" such as $\alpha \to (\beta \to \beta)$ to fail in an interpretation, even though $\beta \to \beta$ is a logical truth.

The application of impossible worlds in the semantics of relevant and other paraconsistent logics is a cardinal one. It is certainly not the only one. Another major one concerns the notion of propositional content (as discussed, e.g., by Vander Laan). Suppose that one analyzes this notion, in a familiar way, in terms of worlds. It is most implausible to suppose that two logically false propositions such as 'the sun is shining and the sun is not shining' and 'Fermat's Last Theorem (FLT) is false', express the same proposition. Yet they both hold in the same set of possible worlds, the null set. In *some* sense, it seems, one of these could be true without the others being so. Thus if this modality is cashed out in terms of worlds, some of these must be impossible worlds.

Propositional content plays a central role in the analysis of intensional states. It has long been observed that the contents of such states may be inconsistent. It is therefore natural to employ impossible worlds in an analysis of such states. Take belief, for example. (This example is deployed by Restall.) One may analyze the content of a belief-state as the set of worlds where things are as they are believed to be—which may be impossible.

It might be thought that veridical intensional notions, like knowledge, being veridical, would find no place for impossible situations; but this would be a mistake (as Barwise notes). For in real life, logically impossible things (such as the falsity of FLT) may well be *epistemic* possibilities. Hence, the content of our epistemic states may include impossible worlds, for example, one where FLT fails. One important intensional state that is less frequently discussed is perception. But as Mortensen notes, the contents of such a state may be impossible too, if we are viewing an impossible object (as depicted, say, in an Escher drawing), or are having a visual illusion, such as the waterfall illusion.

The notion of an impossible world is important in related, but rather different, kinds of cases. In particular, it is important for an analysis of information (as stressed by Barwise), quite independently of an analysis of intensional states. The most obvious case of this is where the information concerned is that in some computational data base. This may be inconsistent, despite our best efforts (there being no algorithm for inconsistency). But it is bizarre to suppose that the data 'the next flight to Sydney is at 9:45 and at 10:00—and so not at 9:45' carries the information that the next flight to Melbourne is at 11:00. The situation, as characterized by the information in the data base, is an impossible, but nontrivial, one (and hence, a paraconsistent logic is required to determine it).

We deal with inconsistent information in quite different contexts too, for example, with certain works of fiction. These may well be inconsistent, either wittingly or unwittingly (see Priest and Varzi, respectively). Time-travel stories often fall into this category. If we take the informational content of the work of fiction to be individuated by the collection of situations (note the plural; in general, fiction underdetermines) which verify what is explicitly given by the author, together with appropriate background assumptions, then these will be impossible situations.

Finally, there is the closely related issue of certain kinds of counterfactuals, as noted especially by Mares, but also by Zalta. We often consider what would be the case in various impossible situations. We may (to use examples given by Nolan) consider what would be the case if some incorrect logic were, in fact, correct; or what would be the case if I were you. What would obtain, in each case, is what follows from the information explicitly supposed, together with appropriate background information. A natural way of cashing out this idea is in the familiar Stalnaker/Lewis way. What obtains is whatever is the case in the worlds most similar to ours—that is, where the background information holds—where the supposition also holds. *Ex hypothesi*, such situations are impossible.

4 The structure of impossible worlds Most of the contributors would be fairly sympathetic to most of the applications of impossible worlds just mentioned. Though this is certainly not uniformly the case. (For example, Zalta rejects the use of impossible worlds in an analysis of propositional content, especially in connection with intensional states). There is, however, much less agreement about how impossible worlds are to be analyzed. Notoriously, philosophers and logicians disagree about the nature of possible worlds: are they abstract objects, nonactual concrete objects, sets of sentences? A similar variety of positions concerning the nature of impossible worlds is also possible, as the papers in this collection show. But the striking difference between the authors in this collection, and where they do tend to diverge along hemispherical lines, is not on this question. It is on the question of the *structure* of impossible worlds.

For most of the Australasians (Mares, Mortensen, Restall), worlds are represented by the interpretations of some paraconsistent logic, which therefore, gives them a determinate structure. Moreover, and consequently, the content of each world is closed under a paraconsistent consequence relation, one which includes the rule of adjunction.

In his article, (the North American) Barwise takes impossible worlds to be situations of a certain kind; and the structure of situations needs no further specification (for the project at hand). However, it is worth noting that in standard Barwisean situational semantics, a situation is a set of infons. And such a set is, in effect, just an interpretation of the relevant logic of first-degree entailment. We may therefore think of Barwise as an honorary Australasian.

The North Americans Vander Laan and Zalta, whilst disagreeing about what impossible worlds are, agree on the fact that they have no significant logical structure (other than being large in a certain sense, namely verifying either α or $\neg\alpha$ for every α). In particular, they are not structured as are the interpretations of any particular logic, and so not closed under any interesting notion of logical consequence. Nolan (an Australasian), too, endorses this view. He is, therefore, an honorary North American.

Finally, Varzi is a border-line case. (He is, after all, Italian.) For him, worlds are structured as the interpretations of a paraconsistent logic, but a paraconsistent logic of the nonadjunctive variety. This arises because he analyzes impossible situations, combinatorially, into classical worlds. In particular, impossible worlds have a limited amount of logical structure: they are closed under single-premise classical inference; they are not closed under any essentially multiple-premise inferences, however—at this level, they are anarchic.

The fact that the worlds of the nonparaconsistent logicians are, logically, completely anarchic might be thought to be a significant advantage for them. After all, if a situation is impossible, why shouldn't anything go? On the other hand, in the case of the worlds of relevant logic (though not nonadjunctive paraconsistent logics), there are still, for any sentence, worlds where it holds and worlds where it fails. And this trick is turned without the cost of surrendering logical control over worlds.

5 A critique of impossible worlds Although all of the contributors to this issue support the importance of the notion of impossible worlds in some form or other, not all philosophers do. One of the papers that I would like to have had in this collection is a critique of the notion by (the North American) Stalnaker [1]: our invitation to him to contribute an essay arrived just too late. But since his paper is not in this issue, let me take the liberty of saying a few words about it.

Stalnaker is prepared to agree that if by 'impossible world' one means just an inconsistent (or arbitrary) set of propositions, then the notion is unproblematic enough. He is just skeptical that it can do any interesting philosophical work ([1], p. 200). In particular, if one analyzes worlds in terms of propositions, one cannot then analyze propositions and their modal properties in terms of worlds. One cannot but agree with this.

Stalnaker's paper is in dialogue form, which means that his arguments against a nonreductive account of impossible worlds are somewhat fluid. But if I understand him correctly, he has essentially three such arguments. The first is a recurrent complaint that the notion of an impossible world is just too unclear. Now the notion of an impossible world certainly requires clarification, but this is what many of the papers in this collection try to do in their different ways. Moreover, the semantics of relevant and other paraconsistent logics, give us quite precise characterizations. So I will say no more about this argument.

Stalnaker's major argument against impossible worlds of the first kind is taken from Lewis. It starts by assuming that $\neg\alpha$ is true at a world, w, if and only if α is not true at w. Now suppose that α and $\neg\alpha$ are both true at w. Then α both is and is not true at w. Thus, an impossible world gives rise to a violation of the Law of Noncontradiction. Most who subscribe to the notion of impossible worlds will not like this conclusion. Even those who think that violations of the law may be possible are unlikely to take it that a contradiction at an impossible world should automatically spill over into an inconsistency in this one.

The problem with this argument, as the other party in the dialogue points out, is simply the truth conditions for negation. As soon as one supposes that there are worlds where things may be both true and false—or neither true nor false—these conditions are manifestly incorrect. If there are gaps, α can fail to be true at w, without

$\neg\alpha$ being true there; and if there are gluts, $\neg\alpha$ can be true at w without α failing to be true there. I quote Stalnaker's reply:

> . . . I thought that [the negation sign] had a pretty clear and uncontroversial semantics, at least for those of us who accept classical logic. My assumption about the meaning of "$\sim$" is this: $\sim P$ is true if and only if P is false. Or in other worlds, the set of worlds in which $\sim P$ is true is the complement of the set of worlds in which P is true. I learned this rule in my first logic class years ago. I suppose that one might use the symbol differently, but it is hard to see how any metaphysical question could turn on whether we stick with the traditional truth-table account of the negation symbol . . . . ([1], p. 196 f.)

There are a number of comments to be made about Stalnaker's words. First, even one who supposes that contradictions may be true (or true at a world) may agree that the negation of P is true if and only if P is false. This holds in the four-valued semantics of first-degree entailment, for example. More to the point is whether the negation of P is true if and only if P is *not* true. But even this is not quite right. Someone who subscribes to impossible worlds may well agree that this is, in fact, so. For it concerns truth *simpliciter*, that is, what holds at the actual world. It is what happens at other worlds that is at issue. (Note how Stalnaker slides from truth to truth in a world.) What needs to be defended is that for *every* world w, the negation of α is true at w if and only if α is not true there. It is not that Stalnaker is confused about what is really at issue here, but this claim should not be allowed to draw support from other, less contentious, things.

Stalnaker's defense of this claim? He says that he learned it in his first logic class. He probably did, but not much weight can be put on this fact. He probably learned Newtonian dynamics in his first mechanics class too. But this is wrong. A logical semantics provides an account of the meanings of the logical particles. I think that the best way of understanding such a semantics is as a *theory* about how certain vernacular notions function. That is, a semantics for negation is an account of the meaning of negation, as expressed in the vernacular in claims such as 'Socrates is not mortal', 'FLT holds or it doesn't', 'This sentence is not true'. And it is exactly the correctness of the "classical" theory of negation that someone who subscribes to impossible worlds will question—especially if (but not only if) they do not subscribe to classical logic, but to some relevant or paraconsistent logic. Quine's famous argument, that to change the logic is to change the subject, may be right to this extent: classical negation and nonclassical negations have different meanings. But the substantial issue that Quine never addressed is why we should suppose that the meaning of the vernacular negation is classical.

One might also think of a semantics as simply stipulative, as Stalnaker's last sentence tends to suggest. In this case, we may have a symbol, say $\neg$, whose meaning is given by one set of truth(-at-a-world) conditions, and which allow for α and $\neg\alpha$ both to be true at a world. (Stalnaker calls such a notion, somewhat question-beggingly, 'quasi-negation'). But we also have another symbol, $\sim$, whose meaning is given by Stalnaker's truth conditions. For such a notion, at least, no things of the form α and $\sim\alpha$ can ever hold at a world—or at least, if they do, contradictions spill over into the actual world.

This brings us to Stalnaker's third argument. He seems to suggests ([1], p. 201) that the fact that there are sentences that hold at no world makes trouble for someone who supposes there to be impossible worlds. Why, exactly, is not so clear; but I presume that the reason is something like this. If we are going to have to say that there are some things that hold in no world, why not just suppose that all impossible statements are like this? An obvious answer is that if we do so, then all sorts of counterintuitive things seem to happen, with propositional content, intensional contexts, inference, information content, counterfactuals, and so on. But is this then not also going to be the case for sentences containing $\sim$? No. If the correct semantics for (vernacular) negation are those for $\neg$, then $\sim$ is a connective which, though it may have a technical sense, does not correspond to any familiar vernacular notion. We have no linguistic intuitions about how it ought to behave. *A fortiori*, it cannot behave counterintuitively. We may simply suppose that its behavior is as given by the theory that is best in other respects.

6 Do impossible worlds have a future? Whatever one makes of all the above matters, the breadth and depth of the topics covered in this issue indicate that the notion of an impossible world is coming to play a role in the theorization and unification of a number of issues in philosophical logic similar to that which the notion of a possible world itself did some twenty-five years ago. Of course, the notion of a possible world was highly contentious twenty-five years ago. In some ways, it still is, though not in the same way. Whereas then, a good part of the debate was about whether one could make any sense of the notion at all, now its technical applications are so well entrenched that the debate is only about how best to understand the notion. Impossible worlds appear presently to be in a situation similar to that in which possible worlds found themselves twenty-five years ago. My prediction, for what it is worth, is that the debate concerning them will go the same way and for exactly the same reasons—but then, I'm an Australasian.

REFERENCES

[1] Stalnaker, R.,"Impossibilities," *Philosophical Topics*, vol. 24 (1996), pp. 193–204. 5, 5, 6, 6

Department of Philosophy
The University of Queensland
Brisbane QLD 4072
AUSTRALIA
email: ggp@lingua.cltr.uq.oz.au