Logique & Analyse **139-140** (1992), 291-302

WHAT IS A NON-NORMAL WORLD?

Graham PRIEST

1. *The History of Non-Normal Worlds*

Within the panoply of techniques that constitute possible-word semantics, the use of non-normal worlds is a singularly useful one. These were introduced by Kripke [2] in order to model the Lewis systems weaker than *S4*. In such systems the rule of *Necessitation* (from $\vdash \alpha$ infer $\vdash L\alpha$) fails. This requires there to be worlds where the necessitation of a theorem may fail. To achieve this, Kripke suggested that we let there be worlds in which $L\alpha$ is false for *all* α. These are the non-normal worlds. The idea was generalised by Routley and Meyer in their semantics for relevant logics (to be found, e.g., in [8]). If β is a theorem of some logic with standard possible-world semantics, β is true in all worlds. Given the usual understanding of strict implication, this means that $\alpha \rightarrow \beta$ will also be true in all worlds, whatever α, and so logically valid. Hence we have "paradoxes of implication". To get around these, Routley and Meyer suggested employing a ternary relation to state the truth conditions of $\rightarrow$. The important upshot of this, for present purposes, is the existence of worlds where any formula of the form $\alpha \rightarrow \beta$ may fail. Given that in these semantics arrow formulas are essentially necessitives (their truth conditions quantify over all worlds) this is a clear generalisation of the Kripkean technique. This comes out most clearly in the simplified semantics for relevant logics [6]. In these, there are normal worlds, where $\rightarrow$ formulas receive the usual *S5* truth conditions, and non-normal worlds where the ternary relation pokes its nose into their truth conditions. Despite the usefulness of the notion of non-normal worlds, no one has been able to say much about what they are. There are some brief suggestions in Creswell [1], and the idea that there are situations in which an arbitray formula may fail is part of the folklore of relevant logic. Yet little of philosophical substance has ever been made of the idea. Many have therefore felt that non-normal worlds are a mere technical device, with no real significance. This justifies a general attitude of dismissal to both non-normal modal systems and relevant logics. The point of this paper is to show that good philosophical sense *can* be made out of the notion of non-normal worlds, with appropriate consequences for an analysis of entailment.

2. *Worlds Where Logic Fails*

As should be clear from the discussion in section one, non-normal worlds are essentially those where theorems, that is, semantically, logical truths, may fail. This is the key observation. The rest of the paper merely develops this idea. We all suppose that things could (have) be(en) otherwise, and that there are, therefore, in some sense, possibilities other than the actual, or "possible worlds". Quite how one should understand these metaphysically is, of course, a contentious question, and I want nothing I say here to prejudice that issue. For the rest of this article, you (the reader) are at liberty to read in your favourite story about the nature of possible worlds. Whatever worlds are, it is clear that there are possible worlds where there are relatively minor differences from the actual; where, for example, everything is the same as the actual world except that the next coin you toss comes down differently. There are, however, worlds where there are differences of a much more profound sort, where, for example, the laws of nature are different; where, e.g., things can travel faster than the speed of light. We might call these nomologically impossible worlds. That there should be nomologically impossible worlds does not strain the imagination too much. Science fiction writers delight in describing them to us.

But just as there are possible worlds where the laws of physics are different, so there are possible worlds where the laws of logic are different. Anyone who understands intuitionist logic or quantum logic, for example, has some idea of what things would be like if these were correct (assuming, for the sake of argument, that they are not). Few novelists have (yet) explored the genre of stories about worlds where the laws of logic are different (Logic Fiction?). But it shouldn't be too difficult to write interesting such stories. Quantum logic, for example, is quite intelligible, and it is clear that one could write a coherent story within its framework. (It might just be a story about a world where Planck's constant equals 3.) Such stories may bend the mind, but no more so than stories set in worlds with strongly non-Euclidean geometries. (A genre of science fiction that already exists.) By analogy with the case where the laws of physics are different, we might call worlds where the laws of logic are different logically impossible worlds. The phrase 'logically impossible possible world' sits ill on the ear. If you find it unbearably so, just use the phrase 'logically impossible situation'.

Now, the prime notion of logic is inference; and valid (deductive) inferences are expressed by statements of entailment, $\alpha \rightarrow \beta$, (that α entails that β). Hence, in a logically impossible world we should expect statements of this form to take values other than the correct ones. Is there any limit to the value that such a conditional might take? I do not see why. Just as we can imagine a world where the laws of physics are arbitrarily different, indeed, an anomalous world where there are no such laws; so we can imagine worlds where the laws of logic are arbitrarily different, indeed, an anomalous world where there are no such laws. (This theme is pursued by Mortensen in [3].) Whatever the case, the worlds where statements of entailment may take on values other than the actual are exactly non-normal worlds.

3. *Formal Semantics*

Let us see how these observations cash out in terms of a formal semantics. Let us suppose, for a start, that we are dealing with a propositional logic whose connectives are $\vee$, $\wedge$ and $\rightarrow$. The addition of quantifiers causes no problems (or at lest, none that are germane here). Negation we will come to in due course.

An interpretation for the language is a structure $\langle W, L, [.], f\rangle$. W is a non-empty set (of possible worlds); L is a non-empty subset of W (the normal worlds), $[.]$ is a function whose arguments are propositional parameters and whose values are subsets of W; f we will come back to in a moment. The function $[.]$ assigns a semantic value (proposition) to each propositional parameter. Intuitively, the value of a parameter is the class of worlds where it is true. We next have the job of assigning similar values to all other sentences. We do this recursively. Conjunction and disjunction require little comment:

$$[\alpha \vee \beta] = [\alpha] \cup [\beta]$$
$$[\alpha \wedge \beta] = [\alpha] \cap [\beta]$$

The value of $[\alpha \rightarrow \beta]$ is the union of two components, N, the class of normal worlds where the sentence is true, and NN, the class of non-normal worlds where it is true. Assuming standard *S5* truth conditions, N is easily specified: $N = L$ if $[\alpha] \subseteq [\beta]$; otherwise $N = \phi$. NN requires a little more comment. This is the class non-normal worlds where the entailment is true. Which these are, is independent of any information we have used so far. Moreover, given that there is no (external) rhyme or reason as to what entailment does in a non-normal world, this has to be so. This is where f comes into the picture. f is a function of two arguments,which are both subsets of W, and its value is a subset of W-L. NN is exactly $f([\alpha], [\beta])$.

The final task is to define logical validity. α is logically valid iff it is true at all normal worlds of all interpretations. It is clear that we need to restrict the definition to normal worlds. Non-normal worlds are, after all, worlds where logical laws fail.

4. *Proof Theory*

What logical system is generated by these semantics? The answer is relatively easy to determine, since these techniques are already to be found in the literature (almost). In particular, Routley and Loparic [9] have used essentially these techniques to give semantics for the system P_+. This system is defined by the following axioms and rules. ($\vdash$ indicates a rule of inference.)

A1. $\alpha \rightarrow \alpha$
A2. $\alpha \wedge \beta \rightarrow \alpha\ (\beta)$
A3. $\alpha\ (\beta) \rightarrow \alpha \vee \beta$
A4. $\alpha \wedge (\beta \vee \gamma) \rightarrow (\alpha \wedge \beta) \vee (\alpha \wedge \gamma)$

R1. $\alpha, \alpha \rightarrow \beta \vdash \beta$
R2. $\alpha, \beta \vdash \alpha \wedge \beta$
R3. $\alpha \rightarrow \beta, \alpha \rightarrow \gamma \vdash \alpha \rightarrow \beta \wedge \gamma$
R4. $\alpha \rightarrow \gamma, \beta \rightarrow \gamma \vdash \alpha \vee \beta \rightarrow \gamma$
R5. $\alpha \rightarrow \beta, \beta \rightarrow \gamma \vdash \alpha \rightarrow \gamma$

The differences between the semantics here and those given there are twofold. The first is that in the semantics there, L is taken to be a singleton. This, in fact, makes no difference to the logic. (Any interpretation with many normal worlds is the same as an interpretation with one, where the behavior of $\rightarrow$ at all save the one is determined by f to be exactly the same.) The other difference is that, in Routley and Loparic, whether $\alpha \rightarrow \beta$ holds at a non-normal world depends on just the formula itself, and so on α and β. In the present construction, it depends on $[\alpha]$ and $[\beta]$. The construction therefore builds in the substitutivity of equivalents.

R6. $\alpha \leftrightarrow \beta \vdash (\alpha \rightarrow \gamma) \rightarrow (\beta \rightarrow \gamma)$
R7. $\alpha \leftrightarrow \beta \vdash (\gamma \rightarrow \alpha) \rightarrow (\gamma \rightarrow \beta)$

Building in substitutivity makes sense in the context: whether or not $\alpha \rightarrow \beta$ is true at a non-normal world should depend on the propositions expressed by α and β ($[\alpha]$ and $[\beta]$) and not just the sentences used to express them. Given the other axioms and rules, general substitutivity follows:

$$\alpha \leftrightarrow \beta \vdash C(\alpha) \leftrightarrow C(\beta)$$

where $C(.)$ denotes any context.

Let us call the system generated by A1-A4 and R1-R7, N_+. Its soundness with respect to the semantics is easily checked. Completeness is proved by an obvious modification of the argument given by Routley and Loparic. We do not, therefore, need to dwell upon it here.

The system N_+ is a subsystem of the basic positive relevant system B_+, as may be easily checked. This means that it is relevant too (i.e., if $\alpha \rightarrow \beta$ is logically valid, α and β share a propositional parameter). To obtain the system B_+ and stronger systems, Routley-Meyer semantics employ their ternary accessibility relation to determine the truth values of $\rightarrow$ statements at non-normal worlds. (See [6] and [7].) What, if anything, the ternary relation means, has been a matter of some controversy. I know of no interpretation that I find very satisfactory. Maybe one will be found. (Perhaps there is some feature of non-normality that it explicates.) But at the moment I see no considerations that make it worthwhile to add it to the semantical machinery used here.

I note (without proof) that N_+ is just the positive fragment of the logic of Priest [4], and hence that it has an equally natural algebraic semantics. None of the logics we shall obtain by adding negation is exactly the logic given there. However, it is not difficult to modify the algebraic techniques used there to provide algebraic semantics for all of the systems we will meet in the rest of this essay.

5. *Laws and Supervenience*

Before we move on and consider how negation should enter the picture, it is appropriate to say a few words about the semantics we have just been considering and their relationship to the informal motivating idea that non-normal worlds are worlds where logic fails. The notion of a world where logic fails is, in fact, ambiguous. A world where logic fails may be one where the laws of logic are different from the true laws; it may also be one where the logically impossible happens. The distinction itself has nothing particularly to do with logic. It makes perfectly good sense with respect to laws of nature also, provided only that one is a realist about these. Provided one is a realist, it makes perfectly good sense to suppose that there may be two worlds which have different laws of nature, but where the same things, *as a matter of fact*, happen. For example, it may be true in both worlds that nothing travels faster than the speed of light; but in the first world this is so because there is a law of nature to this effect, and in the second it is so because everything started out slowly and never made it up to a super-luminal speed. Thus, a world where the laws of physics are different (e.g., where acceleration through the speed of light is possible), need not be one where the physically impossible happens (where some things do so accelerate). Note that a world where the (logically or physically) impossible happens is a world where the laws (of logic or physics) are different. The converse, however, is not necessarily true: impossible things (logical or physical) need not happen in a world where the laws (of logic or physics) are different.

Now the non-normal worlds that have been employed in the semantics of the previous sections are worlds where the laws of logic are different, not worlds where the logically impossible happens. For example, though $\alpha \rightarrow (\alpha \vee \beta)$ may fail at a non-normal world, every world where α is true is a world where $\alpha \vee \beta$ is true; so there is no world where α is true but $\alpha \vee \beta$ is not. For a story of fiction, a world where the physically/logically impossible happens is, of course, more interesting than a world where the laws of physics/logic are merely different. However, we are not engaged in telling stories here, merely in charting "logical space". As such, it is unnecessary to consider worlds where the logically impossible happens. A consideration of worlds where logic is different will suffice.

There is a significant difference between physical and logical possibility here, however. If a natural law fails, there is, on the usual construction, a world accessible (from the world in question) where the physically impossible actually happens. For example, if the law is $\forall x(\alpha(x) \rightarrow \beta(x))$ there is an accessible world where for some a, $\alpha(a)$ is true but $\beta(a)$ is not. This does not happen with a logical law. The fact that truth values are assigned to conditionals in non-normal worlds, independently of what happens in other worlds, means that the laws of logic are, in a sense, intrinsic to these worlds, and do not supervene on the accessible possible-world structure. One could, of course, handle physical necessity in a similar way technically, but one need not, and standardly, does not. For logical necessity, however, this is absolutely essential if a non-trivial result is to be obtained. It would be quite possible to define worlds where the logically impossible actually happens. For example, we could give $\vee$ non-standard truth-conditions at non-normal worlds, and so ensure the existence of worlds where α is true but $\alpha \vee \beta$ is not. However, in that case, logic would break down even at normal worlds: every conditional of the form $\alpha \rightarrow \beta$, where α and β are distinct formulas, would be liable to fail at a normal world.

The fact that logical necessity (at a non-normal world) must be intrinsic to the world and not supervenient on the world-structure is, I think, significant; though what its significance is, I am unsure. The purpose of the present section has just been to point out this fact.

6. *Negation and Non-Normality*

So far we have looked only at positive formal logics. It is now time to consider negation. As we will see, this raises some important new issues. Let us suppose that negation, $\neg$, is added to the language. What should its semantic conditions be? A simple suggestion is that negation be evaluated as follows:

$$[\neg\alpha] = \overline{[\alpha]}$$

(where overlining is simple complementation). Such a definition will ruin the relevance of the logic since, as is easily checked, $\alpha \rightarrow (\beta \vee \neg\beta)$ is now valid. Whether or not this is undesirable, *per se*, is a question we need not go into here. Routley has argued (e.g., [8], p 232) that relevance should be the result of something more fundamental. And as we will see, he is correct about this. Of more importance in the present context is the following. If α is any logical truth of N_+ we can construct interpretations with non-normal worlds where α is false. (Simply assign every propositional parameter and maximal subformula of the form $\beta \rightarrow \gamma$ in α the value ϕ.) This is no longer the case given the above conditions for negation. The formula $\beta \vee \neg\beta$, for example, is now true at all worlds. This is important since the whole *point* of non-normal worlds was precisely to represent situations where logical truths may fail.

We could get around this problem by taking f to assign $\neg\alpha$ an arbitrary non-normal extension. However, this is unsatisfactory for at least two reasons. The first is that it will destroy any connection between negation and entailment. Thus, for example, α and $\neg\neg\alpha$ will have arbitrarily different truth-values at non-normal worlds. Hence $\alpha \leftrightarrow \neg\neg\alpha$ will fail (in both directions). The second reason is that it is part of the motivation of non-normality that we can assign arbitrary truth-values to *logical truths* (in accordance with the "laws of logic" that hold at a given non-normal world); but it is quite *ad hoc* and unmotivated in the context to suppose that we can assign arbitrary values to other things. And in general, negations, unlike entailments, are not logical truths.

A way around some of these problems is to employ the Routley * operation. We suppose that an interpretation comes with an additional component, a unary function from W to W, *. Negation is now given a value defined as follows:

$$[\neg\alpha] = \{w;\ w^* \notin [\alpha]\}$$

This construction is sufficient to validate De Morgan principles and the rule of contraposition:

A5. $\neg(\alpha \vee \beta) \leftrightarrow (\neg\alpha \wedge \neg\beta)$
A6. $\neg(\alpha \wedge \beta) \leftrightarrow (\neg\alpha \vee \neg\beta)$

R8. $\alpha \rightarrow \beta \vdash \neg\beta \rightarrow \neg\alpha$

as may easily be checked. Other principles concerning negation may be validated by adding further conditions on *. For example, if we insist that $w^{**} = w$ then the principle of double negation:

A7. $\alpha \leftrightarrow \neg\neg\alpha$

is validated. This is a route that Routley and Loparic [9] use to extend their semantics to negation. Using this construction (with no conditions on *) they obtain a system they call H, which is P_+ augmented by the contraposition and De Morgan principles just cited. If we employ the same construction here, we obtain the system H augmented by the substitutivity of equivalents, as one would expect—and a simple modification of their completeness proof demonstrates.

The major problem with this approach to negation is that it is not at all clear *why* negation should be treated in this way; why the * operator should poke its nose into the truth conditions of negation. * is, in fact, a device for ensuring that there are non-normal worlds that are inconsistent and incomplete, and so where certain "laws of logic" may fail. (To make this clear, we could, for example, require that if w is normal $w^* = w$, and hence that the truth conditions for negation at normal worlds are the familiar classical ones.) Moreover, it does this whilst retaining a certain amount of control over what happens at non-normal worlds (unlike the suggestion that $\neg\alpha$ be allowed to behave arbitrarily at non-normal worlds). However, it would seem to be *only* a technical device, and one, moreover, that both undershoots and overshoots the mark. It undershoots the mark since extra conditions on * have to be added to obtain clearly desirable negation principles, such as double negation. It overshoots the mark since for normal worlds it requires (negation-) consistency and completeness to stand or fall (actually, stand) together. This point will become clearer once we have considered a third approach to negation, to which we now turn.

7. *Four-Valued Semantics*

The third approach to negation that we will consider rejects the assumption, built in to the semantics so far, that the semantic values of truth and falsity are exclusive and exhaustive. There are many reasons to reject this assumption, which we need not go into here. All we need to note at present is that this is no mere technical device. In the context of the present discussion, this approach is best handled by replacing [.] by a pair of functions, $[.]^+$, $[.]^-$. Intuitively, the first delivers the class of worlds where a formula is true; the second the class of worlds where it is false. The semantic values concerning extensional connectives are now given in the obvious way (implementing the standard Dunn four-valued semantics):

$$[\alpha \vee \beta]^+ = [\alpha]^+ \cup [\beta]^+$$
$$[\alpha \vee \beta]^- = [\alpha]^- \cap [\beta]^-$$

$$[\alpha \wedge \beta]^+ = [\alpha]^+ \cap [\beta]^+$$
$$[\alpha \wedge \beta]^- = [\alpha]^- \cup [\beta]^-$$

For the positive extension of $\rightarrow$ formulas, we have:

$$[\alpha \rightarrow \beta]^+ = N^+ \cup f([\alpha]^+, [\beta]^+)$$

where $N^+ = L$ if $[\alpha]^+ \subseteq [\beta]^+$ and $[\alpha]^- \supseteq [\beta]^-$; and ϕ otherwise. The negative extension is less obvious. In fact, in one sense, it is not very crucial at all. For example, any definition will validate A1-7 and R1-8. Still, it is natural to suppose that:

$$[\alpha \rightarrow \beta]^- = N^- \cup f([\alpha]^-, [\beta]^-)$$

The second term in this *definiens* represents the class of non-normal worlds where $\alpha \rightarrow \beta$ is false. As with truth, this is quite arbitrary (the non-normal world is at liberty to say what laws of logic *don't* hold, as well as saying what laws *do* hold) and so simply specified by f. N^- is the set of normal worlds where $\alpha \rightarrow \beta$ is false. What this should be is less clear. A simple suggestion (taken from [5], section 6.3) is to set $N^- = L$ if $[\alpha]^+ \cap [\beta]^- \neq \phi$ and ϕ otherwise.

Validity is defined, as usual, as truth in all normal worlds of all interpretations. Let us call this system N. As I have observed, and as is easy to check, these semantics validate all the axioms and rules of N_+ together with all the negation principles we have met so far (A1-7, R1-8). This axiom system is not complete with respect to the semantics, however. For example, the semantics validate the rule $\alpha \wedge \neg\beta \vdash \neg(\alpha \rightarrow \beta)$, which is not an admissible rule in the system, as may easily be demonstrated. A complete axiom system is as yet an open problem.

This construction delivers an account of negation free from the objections I raised against the other accounts. First, it is a principled account. Next, it delivers principles concerning negation and implication in a natural way. Most crucially in the present context, every logical truth still fails in some non-normal world. In fact, we can construct a world where everything is untrue: let w be a non-normal world that is not in the extension or anti-extension of any propositional parameter or entailment statement. Then nothing is true at w, as a simple induction demonstrates.

It may be suggested that there are some logical principles that are not validated on the above account. A crucial pair are the following:

$\alpha \vee \neg\alpha$
$\alpha \wedge \neg\alpha \vdash \beta$

Whether or not these *are* principles of logic is another matter, and not one that I will address here. All we need note at present is that we can accommodate either or both of these principles. To accommodate the first, we need to ensure that all "gaps" are closed at normal worlds. We do this by, first, insisting on this fact for every propositional parameter, p:

$$[p]^+ \cup [p]^- \supseteq L$$

And, secondly, by modifying the falsity conditions for $\alpha \rightarrow \beta$ at normal worlds, thus: N^- is L if $[\alpha]^+ \cap [\beta] \neq \phi$ or $N^+ = \phi$, and ϕ otherwise. A simple inductive argument now establishes that every formula is either true or false at a normal world.

Dually, for the second, we require that for every propositional parameter, p:

$$[p]^+ \cap [p]^- \cap L = \phi$$

and modify the falsity conditions of $\alpha \rightarrow \beta$ at normal worlds. N^- is now L if $[\alpha]^+ \cap [\beta] \neq \phi$ and $N^+ = \phi$; and ϕ otherwise. Again, a simple inductive argument now establishes that no formula is both true or false at a normal world.

These constructions ensure that normal worlds are complete and consistent, respectively. (Note how these issues can be handled separately on this construction, in a way that they cannot be, using the * operator.) In the first, the Law of Excluded Middle, and, more generally, all classical tautologies, are valid logical principle (for whatever reason). There are, however, non-normal worlds where they fail (as there should be), namely incomplete ones. In the second, *Ex Contradictione Quodlibet* is a valid rule of inference. It is not, however, truth-preserving at all worlds.

8. *Conclusion*

The discussion has clearly left some loose ends. However, what it has shown is that there is a sensible philosophical story to be told in connection with non-normal worlds. Moreover, given that we want entailment to be truth-preservation in *all* situations, a possible-worlds account of validity almost mandates their use. Moreover, N and its two extensions just considered are both relevant. To see this, suppose that α and β do not share a propositional parameter. Construct an interpretation where there is a non-normal world in which every parameter or maximal entailment subformula of α is both true and false; and in which every parameter or maximal entailment subformula of β is neither true nor false. It is easy to check that α and

β inherit these properties from their subformulas, and hence that $\alpha \rightarrow \beta$ is not logically valid. Relevance *does*, therefore, fall out of more fundamental considerations: the fact that there are worlds where the laws of logic are different.

University of Queensland

References

[1] M.Creswell, 'The Interpretation of some Lewis Systems of Modal Logic', *Australasian Journal of Philosophy* 45 (1967), 198-206
[2] S.Kripke, 'Semantic Analysis of Modal Logic II; non-normal modal propositional calculi', in J.W.Addison *et al* (eds), *The Theory of Models*, N.Holland Publishing Co, 1965
[3] C.Mortensen, 'Anything is Possible', *Erkenntnis* 30 (1989), 319-37
[4] G.Priest, 'Sense, Entailment and *Modus Ponens*', *Journal of Philosophical Logic* 9 (1980), 415-35
[5] G.Priest, *In Contradiction*, Nijhoff, 1987
[6] G.Priest and R.Sylvan, 'Simplified Semantics for Basic Relevant Logics', *Journal of Philosophical Logic* 21 (1992), 217-32
[7] G.Restall, 'Simplified Semantics for Relevant Logics (and Some of their Rivals)', *Journal of Philosophical Logic*, 22 (1993), 481-511
[8] R.Routley *et al*, *Relevant Logics and their Rivals*, Ridgeview, 1982
[9] R.Routley and A.Loparic, 'Semantical Analyses of Arruda-daCosta *P* Systems and Adjacent Non-Replacement Systems', *Studia Logica* 37 (1978), 301-20