Primitive Modality and Possible Worlds

JAVIER KALHAT

Abstract
This paper sets out a number of reasons for thinking that the framework of possible worlds, even when construed non-reductively, does not provide an adequate basis for an explanation of modality. I first consider a non-reductive version of Lewis' modal realism, and then move on to consider the ersatzist approach of Plantinga *et al*. My main complaint is that the framework of possible worlds gets the semantics and metaphysics of ordinary modal discourse wrong. That is, possible worlds do not give us an adequate answer to the semantic question of what ordinary modal claims *mean*, nor do they give us an adequate answer to the metaphysical question of what *makes* such claims *true*.

There is a growing consensus among philosophers that modality cannot be reductively explained. Modal notions are to be accounted for, not by trying to reduce them to non-modal ones, but rather by tracing their complex interconnections with one another. The most prominent *non*-reductive approach to modality is what David Lewis calls *ersatzism*.[1] Like Lewis, ersatzists attempt to explain modality in terms of possible worlds and their inhabitants. Unlike Lewis, they construe possible worlds as *abstract representations* of concrete worlds, rather than as concrete worlds themselves. And unlike Lewis, of course, they construe possible worlds as *primitively modal* in nature. Proponents of ersatzism include Adams, Kripke, Plantinga, and Stalnaker among others.[2]

[1] David Lewis, *On the Plurality of Worlds* (Oxford: Blackwell, 1986), Ch. 3. Ersatzism also goes by the name of 'ersatz modal realism', 'actualism about possible worlds', 'abstractionism', and 'moderate modal realism'.

[2] See: Alvin Plantinga, *The Nature of Necessity* (Oxford: Clarendon Press, 1974), Ch. 4, and 'Actualism and Possible Worlds', reprinted in Michael Loux (ed.), *The Possible and the Actual* (Ithaca: Cornell University Press, 1979), 253–73; Robert Stalnaker, 'Possible Worlds', reprinted in Loux, op. cit., 225–34, and *Inquiry* (Cambridge, Mass.: MIT Press, 1984), 43–58; Robert Adams, 'Theories of Actuality', reprinted in Loux, op. cit., 190–209; Saul Kripke, *Naming and Necessity* (Oxford: Blackwell, 1980), 15–20, 43–53.

doi:10.1017/S0031819108000855
©2008 The Royal Institute of Philosophy

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

In this paper, I will set out a number of reasons for thinking that the framework of possible worlds, even when construed non-reductively as the ersatzist does, does not provide an adequate basis for an explanation of modality. I will first consider a non-reductive version of Lewis' modal realism, and then move on to consider the ersatzist approach of Adams *et al*. My main complaint against using the framework of possible-worlds to account for modality is that doing so gets the semantics and metaphysics of ordinary modal discourse wrong. That is, possible worlds do not give us an adequate answer to the semantic question of what ordinary modal claims *mean*, nor do they give us an adequate answer to the metaphysical question of what *makes* such claims *true*. Both complaints will ultimately be underwritten by a version of the charge from irrelevance that Kripke and Plantinga originally levelled against Lewis' counterpart-theoretic account of *de re* modality.

1. Ordinary Modal Discourse vs. Possible-Worlds Discourse

Before setting out these challenges to the use of the framework of possible worlds, I need to say a little more about the notion of 'ordinary modal discourse' I am employing here. By ordinary modal discourse I understand a body of claims that are expressed by means of modal adverbs such as 'possibly', 'necessarily', and 'actually', as well as by modal verbs such as 'can', 'might', 'must', 'would', etc. These modal idioms are so-called because they qualify the way or *mode* in which a proposition is true or false, or the way or mode in which an object has a property or stands to another object. Examples of ordinary modal claims include: 'Possibly, there are unicorns', 'Necessarily, all bachelors are unmarried', 'Humphrey might have won the 1968 US presidential election', 'Socrates is necessarily human', 'The world could have been a better place', and so on. Ordinary modal discourse is *ordinary* in the sense that it employs the modal idioms that people use in everyday, non-specialised contexts to express modal claims. But, as some of the claims above illustrate, ordinary modal discourse includes claims that are typically made only in a philosophical context, and so claims that are in another sense anything but ordinary.

In contrast to ordinary modal discourse, there is the *possible-worlds discourse*. Possible-worlds discourse is a body of claims which involve quantification over possible worlds and possible individuals. Examples of possible-worlds claims include counterparts of our examples of ordinary modal claims above; thus, 'In some possible

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

world there are unicorns', 'There is a possible world in which Humphrey (or a counterpart of Humphrey) wins the 1968 US presidential election', and so on. Unlike ordinary modal discourse, possible-worlds discourse is largely the preserve of philosophers (and other specialists). Ordinary people do not typically express their modal commitments by quantifying over possibilia.[3]

In this paper, I will be concerned with the question of whether the framework of possible worlds provides the right semantics and metaphysics for *ordinary modal discourse*. Save for a remark towards the end of the paper, I will not address the question of whether possible-worlds discourse should be taken at face value, i.e., as quantifying over possibilia, or whether possible-worlds claims should be understood in terms of their renderings in the language of ordinary modal discourse. It is indeed a philosophically controversial question whether every claim involving quantification over possibilia is translatable into a claim that only employs ordinary modal idioms.[4]

2. Non-Reductive Modal Realism?

There are compelling reasons for thinking that Lewis' *modal realist* account of modality fails in its reductive aspirations.[5] The question remains nevertheless whether a *non*-reductive version of modal

[3] Possible-world discourse can mix quantification over possible worlds and possible individuals with the modal idioms drawn from ordinary modal discourse – e.g., 'There could have been more worlds than there are'. What is important for us, however, is only that possible-world discourse involves, while ordinary discourse does not involve, quantification over possibilia.

[4] Modalists claim that this can be done. See: Graeme Forbes, *The Metaphysics of Modality* (Oxford: Clarendon Press, 1985) and *Languages of Possibility* (Oxford: Blackwell, 1989) for a defence; and Joseph Melia, 'Against Modalism', *Philosophical Studies*, **68** (1992), 35–56, and *Modality* (UK: Acumen, 2003), Ch. 4 for dissent.

[5] See: Javier Kalhat, 'Has Lewis Reduced Modality?', *European Journal of Philosophy* (forthcoming); William Lycan, 'The Trouble with Possible Worlds', reprinted in Loux (ed.), op. cit. note 2, 274–316, 'Review of *On the Plurality of Worlds*', *Journal of Philosophy*, **85** (1988), 42–7, 'Two–No, Three–Concepts of Possible Worlds', *Proceedings of the Aristotelian Society*, **91** (1991), 215–27, and 'Pot Bites Kettle: A Reply to Miller', *Australasian Journal of Philosophy*, **69** (1991), 212–13; Scott Shalkowski, 'The Ontological Ground of the Alethic Modality', *Philosophical Review*, **103** (1994), 669–88; Colin McGinn, *Logical Properties* (Oxford: Clarendon Press, 2000), Ch. 4.

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

realism might constitute an adequate account of modality. Like reductive modal realism, non-reductive modal realism asserts that there is a plurality of *concrete* worlds and their inhabitants. Unlike reductive modal realism, non-reductive modal realism takes some of those worlds to have the primitive attribute of *being possible* and others to have the primitive attribute of *being impossible* (*mutatis mutandis* for their inhabitants). This means, of course, that the non-reductive modal realist cannot construe necessity and possibility simply as Lewis does, namely as quantifiers ranging over a domain of worlds and counterparts.[6] Nevertheless, non-reductive modal realism can offer analyses of modal notions in terms of the same biconditionals that Lewis gives. These biconditionals provide *de dicto* modal claims with truth-conditions that specify what is true in some or all possible worlds, and they provide *de re* modal claims with truth-conditions that specify what is true of some or all counterparts of a given individual. Thus:

(P_D) Possibly P iff for some possible world w, P is true at w

(N_D) Necessarily P iff for every possible world w, P is true at w

(P_R) x is possibly F iff for some possible world w, y is a counterpart of x in w, and y is F

(N_R) x is necessarily F iff for every possible world w, if y is a counterpart of x in w, then y is F

Is non-reductive modal realism a plausible account of primitive modality? I do not think so. While non-reductive modal realism avoids some of the problems confronting Lewis' reductive modal realism – chiefly those concerning the issue of reduction – non-reductive modal realism still confronts other problems that plague Lewis. I will set out three of those problems, mostly with reference to Lewis' own reductive brand of modal realism, though it will be clear that they apply equally well to non-reductive modal realism.

The first problem concerns the distinctive claim of modal realism, namely, that there is an infinite plurality of *concrete* worlds. No amount of *philosophy*, it seems to me, should convince us that there are such entities. Lewis' theory, fascinating and rigorous as it is, illustrates the pitfalls of following the argument wherever it leads, regardless of how implausible one's position becomes as a result of doing so.

[6] Lewis, op. cit. note 1, 5ff.

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

Lewis exploits a style of philosophical argument introduced by Quine called a *utility argument*.[7] Utility arguments postulate entities of a given kind (sets, possible worlds, etc.) on the grounds that doings so affords us adequate and systematic explanations of a range of philosophically central concepts – the implication being that the adequacy of the explanations is *evidence*, albeit indirect, for the existence of the entities. Perhaps this is the best that can be done in philosophy; but let us not be oblivious to the dangers of practicing philosophy in this way. As used by Lewis, one such argument is supposed to give us good reason to believe in an *infinite plurality of worlds*, each of which is just as real and concrete as the actual world we live in. But no amount of philosophical thinking (even if brilliant as Lewis' is) should convince us of *that*.

Lewis distinguishes two kinds of economy: qualitative and quantitative.[8] A theory is *qualitatively* economical if it keeps down the number of *kinds* of entities it postulates; so a theory that postulates bodies alone is more economical in this sense than a theory that postulates bodies and souls. A theory is *quantitatively* economical, on the other hand, if it keeps down the number of instances of the kinds it postulates; so a theory that postulates n electrons is more economical in this sense than a theory that postulates n_{+1} electrons. Lewis recognises 'no presumption whatever' in favour of quantitative economy; only in favour of qualitative economy.[9] And he claims that modal realism *is* qualitatively economical, for we already believe in the actual world, and he is only asking us to believe in more things of that kind. But this is disingenuous. For while Lewis is asking us to believe in the existence of more *worlds*, some of those worlds are populated by dragons, witches, unicorns, three-headed monsters, gods, indeed souls, and *every other* consistent creature of our imagination – and at least some of these *are* additional kinds of entities to the ones we are already prepared to countenance. Modal realism is thus not only quantitatively, but also qualitatively uneconomical.

Metaphysics is a genuine discipline, one that can contribute in our efforts to uncover the true nature of reality. But metaphysics cannot be in competition with science; in particular, the metaphysician cannot tell the scientist (or the explorer, for that matter) what there is. At least, he cannot do so when it comes to entities which are supposedly *no different in kind* from the entities that the scientist studies – 'empirical entities', we might call them. The metaphysician might

[7] Cf. Lewis, op. cit. note 1, 3–5
[8] David Lewis, *Counterfactuals* (Oxford: Blackwell, 1973), 87.
[9] Ibid.

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

uncover, for example, that there are substances, or that objects have essences. But neither substances nor essences are empirical entities: scientists are not in the business of discovering substances[10], or of characterising the essences of objects (neither 'substance' nor 'essence' is a scientific term.). Accordingly, when the metaphysician claims that there are substances or essences, he is not adding items to the *scientist's* inventory of what there is – substances and essences are not additional entities to, say, electrons, the Ebola virus, the planet Mars or the Milky Way. Lewis, on the other hand, *is* adding to the scientists' inventory of what there is, for his worlds are entities in approximately the sense in which Mars or the Milky Way are entities.[11] Yet this is implausible. Metaphysics is an *a priori*, conceptual discipline, and as such it is simply methodologically barred from pronouncing on what empirical entities there are.[12] (The fact that Lewis' worlds are spatiotemporally and causally isolated from us does not help; if anything, it adds to the puzzle of how we could ever be in a position to know of their existence – cf. the 'Benacerraf problem'.[13]) The first difficulty for non-reductive modal realism, then, is that just like its reductive cousin, it implausibly demands commitment to an infinite plurality of concrete worlds.[14]

A second reason why non-reductive modal realism cannot provide an adequate account of modality is that – again, like its reductive cousin – it arguably gets the order of explanation *the wrong way around*. Lewis' analyses of modal notions in terms of possible worlds and counterparts provide an answer to the metaphysical question concerning the source or ontological ground of modal discourse. The analyses specify truth-conditions for ordinary modal claims, and those truth-conditions in turn specify portions of reality – namely,

[10] In the metaphysical rather than in the chemical sense, of course.

[11] Cf. Lewis, op. cit. note 1, 2.

[12] In Quinean fashion, of course, the metaphysician might be able to tell the scientist what he – the scientist – *takes* there to be.

[13] See also Lewis' discussion of this matter, op. cit. note 1, 108–15. Tellingly, Quine defended the existence of sets on the grounds that they are central to mathematics, and mathematics is central to *science* (to physics, in particular) rather than to philosophy. For more discussion on the methodological issues touched on here, see e.g., John Divers, *Possible Worlds* (London: Routledge, 2002), Ch. 9, and Daniel Nolan, *David Lewis* (UK: Acumen), 203–13.

[14] I have ignored here the fact that *non*-reductive modal realism discerns *impossible* as well as possible concrete worlds among that plurality. This is, of course, an additional problem for non-reductive modal realism (though not for Lewis).

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

possible worlds and their inhabitants – *in virtue of which* ordinary modal claims are true. Stalnaker nicely puts the point as follows:

> The possibilist [i.e., Lewis] has a clear and simple answer to the central metaphysical question about modality: what is the source of modal truth? The merely possible worlds that make propositions necessary or possible have the same status as the actual world whose existence and character explains the truth of ordinary contingent propositions. It is true that some tigers have teeth because of the existence, in the universe in which we find ourselves, of tigers with teeth. According to the possibilist, it is possible that some tigers wear trousers for a similar reason: because of the existence in a part of reality (though perhaps not the part we find ourselves in) of tigers that wear trousers. Modal truths are made true by the same kind of correspondence with reality that makes empirical claims true; the difference is that contingent truths must be made true by local circumstances, while claims about what is necessary or possible concern reality as a whole.[15]

According to Lewis, then, ordinary modal truths concerning the actual world and its inhabitants are true in virtue of facts concerning other possible worlds and their inhabitants. Put differently, ordinary modal facts about the actual world and its inhabitants are, for Lewis, *constituted* by facts about possible worlds and their inhabitants. The same is true for the non-reductive modal realist; the only difference is that he takes the relevant facts about possible worlds and their inhabitants to be themselves primitively modal. The fact that (e.g.) Socrates is necessarily human is constituted, then, by the (modal) fact that every counterpart of Socrates is human. But this is to put the cart before the horse.[16] Socrates is *not* necessarily human *in virtue of* the fact that every counterpart of him is human; the fact that every counterpart of him is human is rather a *consequence* of the fact that he is necessarily human. If Socrates is necessarily human, then Socrates cannot exist without being human. A counterpart of Socrates in another world is an individual that Socrates *would have been* had that world been the actual world.[17] Thus, every counterpart

[15] Robert Stalnaker, 'Modalities and Possible Worlds', in Jaegwon Kim and Ernest Sosa (eds.), *A Companion to Metaphysics* (Oxford: Blackwell, 1995), 336–7.

[16] David Oderberg, *Real Essentialism* (London: Routledge, 2007), 2–4.

[17] David Lewis, 'Counterpart Theory and Quantified Modal Logic', reprinted in his *Philosophical Papers*, Volume 1 (Oxford: Oxford University Press, 1983), 28.

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

of Socrates is human because unless they were, *none* of them *would have been* Socrates in the counterfactual situation in which their world is actual. To reverse the order of explanation here is to make Socrates' essential humanity dependent on the circumstances in which those counterparts find themselves in – in particular, their finding themselves to be human in the world in which they exist. And this implausibly makes the fact that every counterpart of Socrates is human nothing short of a *metaphysical accident*.

The explanatory inadequacy of reductive and non-reductive modal realism can also be brought out by appealing to Fine's recent and provocative suggestion that Lewis was in fact as sceptical about modality as Quine was, for neither could 'understand modality except as a form of regularity'.[18] Like Quine, Lewis is an empiricist, and empiricists are generally suspicious of modal notions. They regard the world as what is the case, and allow no room for what *could* be the case, which falls short of what is the case, and for what *must* be the case, which goes beyond what is the case. Fine points out that

> [f]or empiricists, in so far as they have been able to make sense of modality, have tended to see it as a form of regularity; for something to hold of necessity is for it always to hold, and for something to hold possibly is for it sometimes to hold. But if there is not enough going on in the actual world to sustain the possibilities that we take there to be, then one strategy for the empiricist is to extend the arena upon which the possibilities are realized to include what goes on in each possible world.[19]

This is, of course, the strategy Lewis pursues. Fine's insight serves to highlight just how inadequate Lewis' account of modality really is. *For modality has got nothing to do with regularity*. To say that a proposition is necessarily true just in case it is true everywhere in logical space is essentially no more plausible than to say that it is true just in case it is true everywhere within the actual world. For it is not as if enough non-modal truths could *add up* to a modal one. Necessity is not the same as *universality*. Suppose that the correct interpretation of quantum mechanics involves violations of the law of the excluded middle. Does that show that (e.g.) it is no longer necessarily true that it is either raining or not raining? Not at all. It only shows that the *scope of application* of the law is less general than previously thought. In those domains where the law does apply, however, it

[18] Kit Fine, 'Introduction', *Modality and Tense* (Oxford: Clarendon Press, 2005), 2.
[19] Fine, op. cit. note 18, 1–2.

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

applies necessarily. For to say that the law applies necessarily is to say that it applies and could not have failed to apply, and the inapplicability of the law in one domain entails neither its inapplicability nor the possibility of its inapplicability in another domain (anymore than the inapplicability of Einstenian physics at the quantum level entails its inapplicability at the astronomical level). Parallel considerations show that possibility is not the same as 'being true somewhere' (in logical space) either. The modal status of modal truths has nothing to do with how pervasive they are – temporally, spatially or across possible worlds – but rather with the manner or *mode* in which they are true – whenever and wherever they are true.

Admittedly, *non*-reductive modal realism is slightly better off than it reductive cousin over the issue of regularity. The non-reductive modal realist takes worlds and their inhabitants to be primitively modal. Accordingly, the fact that possibly *p* is not constituted by the fact that *p* is true in some world *simpliciter*, but rather by the fact that *p* is true in a world that is (primitively) *possible*. Possibility, then, is not merely a matter of being true somewhere (in logical space). Still, non-reductive modal realism is guilty of confusing *necessity* with universality. For necessary facts *are* taken to be constituted by facts about what is true in *every* (primitively) possible world.

A final reason for thinking that non-reductive modal realism cannot provide an adequate account of modality has been implicit in what has been said so far: facts about possible worlds and their inhabitants are strictly speaking *irrelevant* to ordinary modal discourse. Possible worlds and their inhabitants are *semantically* irrelevant – i.e., even if there are such entities, they are not what we are talking about when we make ordinary modal claims. And possible worlds and their inhabitants are *metaphysically* irrelevant – i.e., even if there are such entities, ordinary modal claims are not made true by facts about them.

Kripke, of course, famously argued against counterpart theory on the grounds of irrelevance. He wrote:

> Thus if we say that 'Humphrey might have won the election (if only he had done such-and-such), we are not talking about something that might have happened to *Humphrey* but to someone else, a "counterpart"'. Probably, however, Humphrey could not care less whether someone *else*, no matter how much resembling him, would have been victorious in another possible world.[20]

[20] *Naming and Necessity*, op. cit. note 2, 45 n13; see also Plantinga, *The Nature of Necessity*, op. cit. note 2, 116ff.

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

Kripke's formulation of the objection from irrelevance unfortunately confuses the object language, which contains modal expressions like 'might have', with the meta-language in which the counterpart-theoretic truth-conditions for claims in the object-language are given, which does not contain any modal expressions.[21] The correct analysis of 'Humphrey might have won' is thus not 'A counterpart of Humphrey might have won' but rather 'A counterpart of Humphrey *does* win'. I also add that the issue of concern must be kept separate from the issue of irrelevance, which is logically prior. Humphrey does not care about the existence of a winning counterpart, because of the irrelevance of such a counterpart to the truth that he might have won – *not the other way around*. Concern is a matter of psychology; irrelevance is a matter of semantics and metaphysics. All this said, Kripke's essential point is correct. On Lewis' analysis, when we say that Humphrey might have won we are in fact talking about someone *other* than Humphrey.

Hazen and Divers resist this claim.[22] According to them, we *are* talking about Humphrey, for we are talking about a counterpart of *Humphrey*. Reference to Humphrey (and to winning) is thus essential to the correct assignment of counterpart-theoretic truth-conditions for 'Humphrey might have won'. But it is disingenuous to take this sort of reference to Humphrey to show that, on the counterpart-theoretic analysis, we are still talking about Humphrey. We plainly are not. We are making a claim about *another* individual, one who resembles Humphrey. Humphrey is, of course, relevant (indeed essential) to the selection of the individual in question. But that does not make the resulting claim one about Humphrey – anymore than the claim that Humphrey's favourite animal is threatened with extinction is a claim about Humphrey, on the grounds that reference to Humphrey is relevant (indeed essential) to the selection of the animal in question.

As Divers points out, the core of the objection from irrelevance is that counterpart theory wrongly construes possibilities (and necessities) as *relational* in nature.

> Even when an intuitively non-relational feature is attributed as a possibility, at a world, to an individual – say their being in pain – some other individual, and a relation get in on the (f)act: there is a counterpart who is in pain. It is the very idea that all de re

[21] Allen Hazen, 'Counterpart Theoretic Semantics for Modal Logic', *Journal of Philosophy*, **76** (1979), 321.

[22] Hazen, op. cit., note 21, 321–2, and Divers, *Possible Worlds*, op. cit., note 13, 129.

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

> possibility is misconstrued as having gratuitous relational structure – that an n-adic property is misconstrued as an (n + 1)-adic relation – that is ultimately the most plausible target for the irrelevance objection.[23]

When I say that Humphrey might have won I take myself to be stating a *non*-relational fact about Humphrey. That is, I take myself to be stating a fact involving Humphrey and *no other* person – or more carefully, a fact involving no other person who *wins*.[24] But on Lewis' analysis, that fact wrongly comes out as a *relational* one, namely, as the fact that Humphrey stands in a certain similarity relation to someone who exists in another possible world and wins there.

Are we mistaken in thinking that our intuitions concerning which facts we state when we make (true) modal claims enjoy any special authority? According to Hazen, counterpart theory must respect our intuitions concerning the truth-values of ordinary modal claims. But, he goes on to say, our ordinary modal intuitions have no jurisdiction over the correctness or otherwise of the counterpart-theoretic assignment of *truth-conditions* to those ordinary modal claims. For those assignments are *theoretical* claims and our ordinary intuitions have no jurisdiction over them.[25] But this is wrong. As Plantinga observes, there is no such 'facile bifurcation' between linguistic intuition and theory.[26] In assigning truth-conditions to ordinary modal claims, counterpart theory (supposedly) determines the actual *content* of those claims. But our intuitions concerning the correctness or otherwise of those content-ascriptions *are* relevant here. For radical semantic externalism aside, we surely have some grip on what our utterances mean. While we can all agree that there is much about *de re* modality that we do not know, we certainly know this much: when we say that Humphrey might have won we are *not* talking about the electoral success of an individual who, while closely resembling Humphrey, is nevertheless numerically distinct from Humphrey, and indeed lives in a world that is both causally and spatiotemporally sealed off from him. Can any amount of theory convince us that *that* is what, say, Humphrey's wife in fact

[23] *Possible Worlds*, op. cit. note 13, 131.

[24] Of course, the fact that Humphrey might have won entails the fact that *Nixon* might have lost, but this does not make either fact relational in nature.

[25] Op. cit. note 21, 323–4.

[26] Plantinga, 'Two Concepts of Modality: Modal Realism and Modal Reductionism', reprinted in his *Essays in the Metaphysics of Modality* (Oxford: Oxford University Press, 2003), 221; cf. Lewis, op. cit. note 1, 241.

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

said to her husband on the morning after electoral defeat when she uttered the words 'Humphrey, dear, you might have won the election'?

Divers complains that those who put forward the objection from irrelevance often do not say 'what application of [counterpart theory] the irrelevance objection is supposed to undermine. Is it that [counterpart theory] is supposed to be revealed as a failed attempt at conceptual analysis, a failed account of truthmaking or ontological identification ....?'[27] In my defence of it, the charge of irrelevance is intended to be both semantic and metaphysical. Counterpart theory assigns the wrong truth-conditions to ordinary *de re* modal claims. Insofar as truth-conditions determine content, counterpart theory gives us the wrong semantic analysis of ordinary *de re* modal claims. And insofar as those truth-conditions specify the features or parts of reality in virtue of which true *de re* modal claims are true, counterpart theory gives us the wrong account of the metaphysics of ordinary *de re* modality.

3. Ersatzism and Primitive Modality

In the preceding section, I argued that a non-reductive version of modal realism cannot provide an adequate account of modality, for such a version would still confront major difficulties that beset Lewis' own reductive brand of modal realism. It might be wondered why so much time was spent arguing these points, given that non-reductive modal realism is not a serious contender among theories of modality. The answer is that some of the same difficulties that beset reductive and non-reductive modal realism also beset what is probably the most popular non-reductive approach to modality: *ersatzism*. Ersatzists include Kripke, who construes possible worlds as 'counterfactual situations', Plantinga, from whom possible worlds are maximal possible states of affairs, Stalnaker, who takes them to be *sui generis* properties, namely, *ways things might have been*, and Adams, for whom they are sets of maximal consistent propositions.[28] In what follows, I will mostly ignore the specific differences that hold among these (and other) versions of ersatzism, since the difficulties I shall dwell upon hold independently of them.

Ersatzists do not make any claims to offer a reductive account of modal notions, and indeed quite openly construe possible worlds as

[27] *Possible Worlds*, op. cit. note 13, 132.
[28] See note 2 above for references.

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

primitively modal in nature – i.e., they construe possible worlds as possible as opposed to impossible, and treat the possible/impossible distinction as basic.[29] According to the ersatzist, possible worlds *actually exist*, and they are *abstract* objects of one sort or another. Each of these objects *represents* the actual world as being some way or other; we can therefore speak of what is the case *according to them*. Ersatz possible worlds are indeed *representatives* of what they represent, that is, they take the place of, do duty for, that which they represent (hence the word 'Ersatz').[30]

Proponents of ersatzism provide the following analyses of modal notions:

(P_D) Possibly P iff for some possible world w, P is true in w

(N_D) Necessarily P iff for every possible world w, P is true in w

(P_R) x is possibly F iff for some possible world w, x is F in w

(N_R) x is necessarily F iff for every possible world w, if x exists in w, then x is F in w

These analyses differ from those proposed by the modal realist, first, in the way the ersatzist understands the notion of a possible world, and secondly, in that the ersatzist allows individuals to exist in more than one world. But this said, the ersatzist also takes modal facts concerning the actual world and its inhabitants to be constituted by (modal) facts concerning other possible worlds and their inhabitants. Ersatzism thus invites the same charge of confusing the *order of explanation* that was earlier levelled against modal realism.

[29] Forrest's conception of possible worlds as structural universals might be an exception; see Peter Forrest, 'Ways Worlds Could Be', *Australasian Journal of Philosophy*, **64** (1986), 15–24. For discussion, see Forbes, *Languages of Possibility*, op. cit. note 4, 80–2.

[30] To clear up a potential elementary confusion: for the ersatzist all possible worlds actually exist, but only one of them is *actual*. For a world to be actual (non-actual) is for it to represent (fail to represent) things the way they are. Some ersatzists, such as Plantinga, take the actual world to be the same kind of entity as the merely possible ones, viz., an abstract entity. (Plantinga thus distinguishes between the actual world and physical reality – i.e., 'us and our surroundings'). Other ersatzists take the actual world and merely possible worlds to be different kinds of entity – Stalnaker, for example, takes a merely possible world to be a property, but takes the actual world to be the entity that instantiates one such property.

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

According to the ersatzist, Socrates is necessarily human in virtue of the fact that he is human in every possible world in which he exists. But that gets things the wrong way around. Socrates is human in every possible world because he cannot exist without being human, and so, in any possible world in which he exists, he is human. To maintain otherwise is to make Socrates' essential humanity dependent on the circumstances in which he finds himself in – in particular, his finding himself to be human in every world in which he exists. And this implausibly makes the fact that every counterpart of Socrates is human a *metaphysical accident*. To repeat, then, the fact that Socrates is human in every world in which he exists *follows* from the fact that he is necessarily human; it is not *constitutive* of it.[31]

As a corollary, ersatzism is also guilty of confusing necessity with universality, and hence with a kind of regularity. For, as the analyses above make clear, the modal status of a necessary truth amounts to its holding true in *every* possible world. But, again, to say that a necessary truth is one that holds true in every possible world is in principle no more plausible than to say that a necessary truth is one that holds true everywhere within the actual world. The modal status of a necessary truth concerns its *mode* of being true, and that is not a function of how pervasive or universal the truth is.

Ersatzism also faces a version of the charge of irrelevance earlier levelled against modal realism. Unlike the modal realist, of course, the ersatzist believes that individuals can exist in more than one world – they are 'trans-world individuals'. Thus, what is possible or necessary for an individual is analysed in terms of what is the

[31] The sense that the ersatzist confuses the order of explanation is particularly acute in Adams' version of ersatzism, according to which possible worlds are sets of maximal consistent propositions ('Theories of Actuality', op. cit. note 2). On Adams' view, 'For me to feel a pain in some possible world is just for a proposition, to the effect that I feel pain, to be a member of a certain kind of set of propositions (namely, of some world-story)' ('Theories of Actuality', 205). It follows from this that I am possibly in pain just in case the proposition that I am in pain is a member of a certain kind of set (a world-story). But I am not possibly in pain *in virtue of* the fact that the proposition 'I am in pain' is a member of some world-story, anymore than I am in pain *in virtue of* the fact that 'I am in pain' belongs to the set of actually true propositions. Rather: the proposition 'I am in pain' belongs to the set of actually true propositions in virtue of the fact that I am indeed in pain. And similarly, the proposition 'I am in pain' is a member of some world-story in virtue of the fact that I am indeed possibly in pain.

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

case for that *very same* individual in some or all of the possible worlds in which it exists. For example, Socrates is necessarily human in virtue of the fact that Socrates – *our* Socrates (the only Socrates!) – is human in every possible world in which he exists. Where, then, is the irrelevance in the ersatzist's analysis?

To see where, let us begin by noting that, for the ersatzist, a possible world is not really a *world*. A possible world is not a *place*, and individuals do not exist in a world in the literal sense of being *parts* of it.[32] For the ersatzist, it will be recalled, a possible world is an *abstract* object. It *represents* the actual world as being a certain way, a way which includes having certain individuals, such as Humphrey, for whom things go differently than in the actual world, e.g., Humphrey wins the 1968 election. But the question is why should facts about (such) representations tell us anything about the things represented? Why should the existence of a representation of a winning Humphrey tell us anything about what is genuinely possible for Humphrey? In short, then, what is the *relevance* of such representations, which are themselves abstract objects, for what is genuinely possible for Humphrey, himself a concrete object?

Here, of course, we must distinguish between what is true *of* a possible world and what is true *according to* it.[33] What is true *of* a possible world is that it is an abstract object, that is has the capacity to represent things, etc. The ersatzist will agree that such truths are irrelevant to questions about what is possible or necessary for the actual world and its inhabitants. But the ersatzist maintains nevertheless that what is true *according to* a possible world is relevant to such modal matters. For he thinks that, say, the possibility of Humphrey's winning the election is *constituted* by the fact that there is a possible world according to which Humphrey wins the election.

But even when we distinguish between what is true of a possible world and what is true according to it, the charge of irrelevance remains. For the bottom line is that, according to the ersatzist,

[32] An object x exists in a world w only in the following sense: if w had been actual, x *would have* existed *simpliciter* (Plantinga, *The Nature of Necessity*, op. cit. note 2, 46ff; cf. Stalnaker, 'Modalities and Possible Worlds', op. cit. note 15, 336).

[33] Oderberg appears not to distinguish between the two when he makes a similar complaint to mine: '. . . the question arises as to how what is true *of* that kind of thing [a possible world construed as an abstract object] can have any bearing on the modal properties of a concrete material object such as a man, a mouse or a mountain' (*Real Essentialism*, op. cit. note 16, 3; emphasis added).

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

possibilities and necessities are constituted by the existence of *abstract objects which represent* the realisation of those possibilities and necessities. But the existence of, say, a possible world which represents Humphrey as winning the election is no more constitutive of the possibility of his winning than is the existence of a consistent novel or film which represents him as winning. To insist otherwise is ultimately to confuse a representation with what it represents. Possible worlds can represent necessities and possibilities, but they are not identical with those necessities and possibilities themselves. The point is so basic that it should hardly need making – no one, after all, would confuse a portrait with its subject, etc. And yet philosophers enchanted by possible worlds are happy to conflate representations with the things they represent when it comes to modality.[34]

The charge against the ersatzist is not merely that he identifies facts of one sort with facts of a different sort. That, after all, is a very common phenomenon in philosophy. To give one obvious example, some materialists in the philosophy of mind identify mental facts with physical facts about brains. This identification is implausible in my view, but let us suppose it were correct. Then the right analogy with ersatzism would be this: identifying the fact that (e.g.) Humphrey might have won the election with the fact that there is an abstract object that represents Humphrey as winning (or, for that matter, with the fact that there is a counterpart of Humphrey who wins the election in another world) is like identifying the mental facts associated with one person (i.e., the facts about that person's mental life) with physical facts about a *representation* of his brain (or with physical facts about the brain of someone who is mentally very much like him). If mental facts are identical with physical facts at all, then the mental facts associated with a given person must surely be identical with physical facts about *that person's* brain. Similarly, if modal facts are identical with facts about possible worlds and their inhabitants at all, then the fact that Humphrey might have won must be identical with the fact that *Humphrey himself* wins in another possible world – not with the fact that there exists an abstract object which represents Humphrey as winning (nor with the fact that there is a counterpart of Humphrey who wins in another possible

[34] Indeed, not just when it comes to modality. Witness, for example, Lewis' and Stalnaker's identification of propositions with sets of possible worlds where they hold true. This is patently inadequate: propositions can be true or false; sets can be neither (cf. Plantinga, 'Two Concepts of Modality', op. cit. note 26, 207).

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

world).[35] The charge of irrelevance, then, is directed at the identification of modal facts with facts of the *wrong* sort, and not at the identification of modal facts with facts of a merely different sort.

Lewis, and various writers after him make the point that if Kripke's charge of irrelevance applies to the modal realist's analyses of *de re* modality, then it also applies to the ersatzist's analyses of *de re* modality. Writing about the related issue of concern and care, for example, Melia says that such considerations 'would also refute the view that possible worlds are abstract entities, since Humphrey is not presumably interested in whether or not there is an abstract entity according to which he wins'.[36] Melia, of course, intends this remark to count in favour of Lewis' modal realism, for if ersatzism is no better off when it comes to the charge of irrelevance, then that is not a stick with which it can beat modal realism. This is dialectically correct. But unlike Melia *et al.*, the conclusion I draw is that *both* modal realism and ersatzism should be rejected. Regardless of whether we construe possible worlds as actually existing or as merely possible, as abstract or concrete, as primitively modal or non-modal in nature, facts about what goes on in them are simply irrelevant to the content and truth of the ordinary modal claims we make about the actual word and its inhabitants. The possible-worlds framework, therefore, cannot serve as an adequate basis for an account of modality.[37]

[35] I do not mean to imply, of course, that it is indeed correct to identify the fact that Humphrey might have won with the fact that Humphrey himself wins in another possible world (anymore than it is correct to identify mental facts with physical facts). In order for Humphrey himself to win in another possible world, that world would have to be concrete, since Humphrey is concrete and he would be literally a part of it. But as I argued in section 2, there are no concrete possible worlds. And even if there were, Humphrey could not exist in two of them (as the present identification would require), for no concrete object can be in two different places at the same time.

[36] *Modality*, op. cit. note 4, 108; cf. Lewis, op. cit. note 1, 196 and Divers, *Possible Worlds*, op. cit. note 13, 129, 134–5.

[37] Rosen's fictionalism about worlds, which is itself a form of *anti*-realism rather than realism about possible worlds, is also susceptible to the objection from irrelevance (Rosen himself observes this; see his 'Modal Fictionalism', *Mind*, **99** (1990), 349–54). According to fictionalism, facts about what is possible and necessary are identical with facts about the content of the story *PW* – the hypothesis of a plurality of concrete worlds. Thus, the fact that Humphrey might have won is (supposedly) identical with the fact that according to *PW*, there is a world in which a counterpart of Humphrey wins. Rosen candidly brings out the implausibility of this identification by noting that it calls for a radical shift in our patterns of modal

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

4. Two Legitimate Uses of Possible Worlds

Despite what has been argued so far, I do not wish to deny that possible worlds *can* be useful. I conclude this paper by considering two such uses.[38] First, possible worlds can be *rhetorically* useful: they can provide a vivid and colourful way of expressing modal claims. This use of possible worlds is deeply entrenched in contemporary metaphysics; so much so, indeed, that even philosophers who would baulk at the suggestion that *there are* possible worlds are happy to use the idiom themselves.[39]

The rhetorical use of possible worlds is perhaps comparable to the rhetorical use of God-talk. Both uses are nicely illustrated by a pair of informal characterisations of the idea that the mental supervenes on the physical:

[1] Ask what Oskar's mental life would be in a counterfactual situation or possible world which duplicated all physical aspects of Oskar's actual state. Supervenience maintains agreement in all mental aspects also.

[2] Imagine the gods constructing the world from scratch. They pick the individuals and set all their physical properties. According to supervenience the Gods have thereby also set the individuals' mental properties.[40]

---

interest and concern. Thus, if fictionalism is correct, Humphrey must now care deeply about what happens to a certain *fictional character* (his winning counterpart) in the story *PW*. This is implausible. I also note that, as observed earlier, the issue of concern presupposes the issue of irrelevance. Humphrey cannot be expected to care deeply about the fate of a certain fictional character because facts about that character are simply *irrelevant* to what is true of Humphrey, namely, that *he* might have won. The fictional character, and the world he inhabits, represent Humphrey as winning, and are therefore at best *evidence* that Humphrey might indeed have won. (For a general overview of fictionalism, see Nolan, 'Modal Fictionalism', *Standford Encyclopedia of Philosophy*, E. Zalta (ed.), (http://plato.stanford.edu/entries/fictionalism-modal/).)

[38] I will not consider the issue of whether possible worlds (construed as abstract objects) can be useful in philosophical accounts of concepts *other* than modality, though I have already indicated that one such additional use (propositions) is problematic (see footnote 15 above).

[39] Whether this lack of 'ontological seriousness' is harmless is an important question, but one which I leave for another occasion.

[40] Paul Teller, 'Supervenience', in Kim and Sosa (eds.), op. cit. note 15, 485.

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

Possible worlds and God afford vivid ways of illustrating the supervenience claim, which is itself an ordinary modal claim – namely, the claim that there *can be no* difference in the mental without a difference in the physical. But possible worlds are no more essential to determining the *content* of supervenience (and other ordinary modal) claims than God is. If putting ordinary modal claims in the idiom of possible worlds is helpful, that is only because they bring out a *consequence* of those claims. Thus, if the mental supervenes on the physical, then it *follows* that no two worlds that are physically alike can differ in the mental facts that obtain in them – a claim that is true even if there are no possible worlds. Drawing out this particular consequence can be useful in the way in which, quite generally, the content of a claim can be clarified by considering its consequences.

A second, more substantial use of possible worlds is in the area of *modal logic*.[41] In the first half of the 20th century, a number of logicians, notably C.I Lewis, developed systems of modal logic, the aim of which was to formalize modal inference. These various systems were non-equivalent; they differed mostly by including or excluding controversial axioms about iterated modalities, such as 'If p, then necessarily possibly p', 'if necessarily p, then necessarily necessarily p', and so on. While it was possible to investigate the logical relations among the resulting systems, it was not clear which were correct, or what the differences between them amounted to exactly. These matters could not be settled simply by invoking modal intuitions, since there was no widespread consensus on them either. Furthermore, for some formulae there were no intuitions at all, e.g., 'if possibly necessarily p, then necessarily possibly possibly necessarily p'. As Loux points out, the problem with such a formula is not merely that we have no settled intuitions about it; it is that we have no idea what it *means*.[42] Indeed, systems of modal logic had a syntax but no *semantics*, i.e., they did not specify models for those systems, sets of objects in terms of which the logical formulae could be interpreted.

In the late 1950s and early 60s the situation changed.[43] A number of logicians, including Kripke, used the framework of possible worlds to

[41] For an accessible account of this use of possible worlds, see Michael Loux, 'Introduction: Modality and Metaphysics', in Loux (ed.), op. cit. note 2, 16–30.

[42] Loux, op. cit. note 41, 19.

[43] Although Carnap had already done some pioneering work in his *Meaning and Necessity*, first published in 1947 (Chicago: Chicago University Press).

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

provide systems of modal logic with a formal semantics.[44] Kripke's proposed semantics for a propositional modal logic consists of an ordered triple {*G*, *K*, *R*}, where *K* is a set of objects, *G* is one of those objects, and *R* is a relation defined over the members of *K*. Intuitively, said Kripke, we can think of *K* as the set of all possible worlds, *G* as the actual world, and *R* as a relation of 'accessibility' (or 'relative possibility') between possible worlds – a world w is accessible relative to a world w_1 just in case every true proposition in w is *possibly* true in w_1. For a system of quantified modal logic, we add a function assigning a set of objects to each member of *K* – intuitively, we think of that set as the set of objects existing at a given possible world. With this semantics, the differences among controversial axioms of iterated modality, and hence the differences among the various systems of modal logic they give rise to, can be understood as arising out of the placing of different *restrictions* on the accessibility relations holding among worlds. Thus, the question of whether, say, what is actual is necessarily possible can now be understood as the more manageable question of whether the accessibility relation is symmetric. The possible-world semantics also made it possible to define the key logical notion of validity and with it to give completeness proofs for different systems of modal logic.

This is all well and good. The trouble begins only when the framework of possible worlds is taken to provide not only a formal semantics for modal logic, but also an *applied* semantics for ordinary modal discourse.[45] Modal logic is an artificial language; we are at liberty to stipulate meaning for it in pretty much whatever way we like. Indeed, as Lewis observes, a Kripke-style semantics for (e.g.) propositional modal logic need not in fact make use of the framework of possible worlds; all it needs is some set of objects or other, and a binary relation defined over them.[46] On the other hand, ordinary modal discourse is evidently embedded in a natural language. As such, it *already* has a semantics. The job of the philosopher is to understand

[44] Saul Kripke, 'Semantical Considerations on Modal Logic', reprinted in Leonard Linsky (ed.), *Reference and Modality* (Oxford: Oxford University Press, 1971), 63–87.

[45] Here a remark by Kripke is apposite: 'The apparatus of possible words (sic) has (I hope) been very useful as far as the set-theoretic model-theory of quantified modal logic is concerned, but has encouraged philosophical pseudo-problems and misleading pictures' (*Naming and Necessity*, op. cit. note 2, 48 n15).

[46] Lewis, op. cit. note 1, 18ff.

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press

it – not to *stipulate* it. That semantics, I have urged, is not adequately given by the framework of possible worlds.[47]

Lewis, of course, (in)famously motivated the commitment to possible worlds partly by saying that 'possible world' is just another name for 'ways things might have been', and we *ordinarily* (and correctly) believe that there are countless ways things might have been (the so-called 'paraphrase argument for possible worlds'[48]). The ersatzist agrees.[49] However, it is one thing to say that the belief in possible worlds is implicit in what we ordinarily believe, and quite another to say that our ordinary modal beliefs (e.g., 'It could have been sunny today') are beliefs about possible worlds and their inhabitants.

Furthermore, it is disingenuous to claim that the belief in possible worlds is indeed implicit in what we ordinarily believe. In Lewis' mouth, 'possible world' is (also) a name for 'causally and spatiotemporally disconnected concrete world' – and we do not ordinarily believe in *that*. The ersatzist might claim an advantage here, since he does not construe possible worlds as concrete objects. But the advantage is illusory, for he construes possible worlds as abstract objects that represent the different ways things might have been – objects which are also causally and spatiotemporally isolated from us – and we do not ordinarily believe in *them* either. Both Lewis and the ersatzist turn our ordinary belief that there are countless ways things might have been into a belief in the *existence of objects* of one sort or another. In doing so, they turn an ordinary belief into one that is anything but ordinary. To then claim support for the non-ordinary belief by appealing to the ordinary one is disingenuous at best.

University of Zürich

[47] To deny that the possible-worlds semantics is the right semantics for our ordinary modal discourse is not to deny, of course, that ordinary modal claims can be about worlds. If I say that the world could have been a better place, for example, I am evidently making a claim about the actual world. But according to the possible-worlds semantics, I am in fact making a claim about a *different* world – one which exists, but is not actual – for what I am saying is that there is a possible world which is indeed a better place than the actual world. It is this sort of analysis that I am rejecting.

[48] See Lewis, op. cit. note 8, 84.

[49] Cf. van Inwagen, 'Two Concepts of Possible Worlds', reprinted in his *Ontology, Identity and Modality* (Cambridge: Cambridge University Press, 2001), 208; Loux, op. cit. note 41, 30ff. Stalnaker is an exception; see his 'Possible Worlds', op. cit. note 2, 226–227.

https://doi.org/10.1017/S0031819108000855 Published online by Cambridge University Press