Modal Fictionalism[1]

GIDEON ROSEN

1. *The problem*

Like most people, Ed does not believe in blue swans. However, he does believe that such creatures *might* have existed. That is, he believes such things are possible.

Now according to the standard possible worlds account of modality,

(1) There might have been blue swans iff there is a world W such that, at W, there are blue swans.

So given what Ed already believes, if he embraces this analysis his beliefs will deductively entail that

(2) There is a (non-actual) possible world at which there are blue swans.

Thus for someone like Ed—someone with ordinary modal opinions—to accept the possible worlds analysis is to take on a *commitment* to believing in other possible worlds. Or at least, this is what our trivial deduction seems to show.

Unfortunately, again like many people, Ed just cannot bring himself to endorse this commitment. He cannot believe that other possible worlds exist. Thus it appears that Ed has no choice: on pain of inconsistency, he must reject the possible worlds approach to modality.

Now this may not seem like much of a tragedy. But we have left out one detail. What lends the situation its peculiar poignancy is that Ed is a philosopher; and as a philosopher he knows that in discussions of modal subtleties—discussions which can hardly be avoided nowadays—the language of possible worlds has become a nearly indispensible tool. For it permits the articulation of modal views with a clarity and vividness that cannot be achieved by other means—so much so that even philosophers who officially renounce the idiom often find themselves talking about possible worlds anyway when it becomes important to make a modal claim precise or a modal argument rigorous. There is a great risk of doublethink in such circumstances: asserting the existence of worlds at

one moment while denying it at another. Ed suffers a debilitating anxiety at the prospect of ontological hypocrisy. Something must be done. But what?

If we suppose that the employment of the possible worlds framework inevitably involves asserting, for instance, that there is a world with blue swans, then it would appear that Ed has only two choices short of abstinence: he can submit to whatever cognitive psychotherapy he might need to overcome his aversion to possible worlds; or he can look for a way to interpret his apparent quantification over worlds in such claims as a misleading *façon de parler*, lacking the usual implications of existential quantification.

If Ed chooses the psychotherapeutic route, he will find no shortage of willing therapists, and an abundance of useful self-help literature. These are the writings of those who openly advocate an ontology of things called 'possible worlds'; and here we should distinguish two fundamentally different approaches.

The first is what Lewis has called *ersatz modal realism.*[2] This is really a family of views, unified by the idea that talk of 'possible worlds' is not to be taken literally. There are indeed a vast number of 'worlds'; but they are not really *worlds* on a par with the spatially extended universe we inhabit. Rather they are *abstract representations* of various ways our world might have been. Thus one typical ersatzist proposal identifies worlds with sets of sentences which represent possible states of the World by describing them.[3] Other forms of ersatzism identify worlds with other, possibly less familiar abstract objects.[4] In each case, however, the response to the charge of ontological impropriety is much the same: people like Ed who are put off by the possible worlds approach have been misled by the colourful terminology into thinking that worlds must be bizarre things, more fit for science fiction than serious philosophy. But this is a mistake. So-called 'possible worlds' are really nothing but *actually existing abstract objects*, the likes of which most philosophers already believe in. To call them 'non-actual' is not to say that they inhabit some shadowy purgatory between being and non-being, but only that they fail to represent things as they are. From an ontological point of view, at any rate, nothing could be more straightforward.

The other approach is Lewis's own *modal realism*. For reasons I will not pursue, Lewis doubts that the ersatz modal realist can produce an account of worlds adequate to every task for which worlds might be wanted. (*Pl.*, ch. 3). He maintains instead that the full benefits of the framework can only be had if we suppose that the non-actual 'worlds' are literally

---

[2] D. Lewis, *On the Plurality of Worlds*, Oxford, Basil Blackwell, 1986; hereafter *Pl.*

[3] E.g., the state descriptions of R. Carnap, *Meaning and Necessity*, Chicago, University of Chicago Press, 1947, p. 9.

[4] Cf. A. Plantinga, *The Nature of Necessity*, Oxford, Oxford University Press, 1974; R. Stalnaker, *Inquiry*, Cambridge, MA, MIT Press, 1984; P. van Inwagen, 'Two Concepts of Possible Worlds', *Midwest Studies in Philosophy*, 1986, pp. 185–213.

worlds—distinct universes as real and robust as our own, yet wholly isolated from one another and from us.

Obviously the modal realist's response to sceptics like Ed cannot be as quick as the ersatzer's, since the realist actually believes what the sceptic finds incredible; namely, that besides our universe there are countless others, populated with blue swans, talking donkeys, and every other beast, conceivable or not. The realist must allay such worries while he answers head on every weighty metaphysical argument which purports to show that our world is the only world there is. And this is only the beginning. Critics have claimed to see in modal realism the seeds of every pernicious intellectual trend from inductive scepticism to moral nihilism. Lewis's recent book is devoted largely to the heroic task of answering these charges, as well as to positive argument for his position. The aim, ultimately, is to persuade people like Ed that despite our initial shock and horror, we all have good reason to revise our naive ontological view by taking on a commitment to robust possible worlds.

I will not stop to weigh the relative merits of realism and ersatzism here. In fact I am going to assume that this dispute has been settled in the realist's favour. That is, I will assume that Lewis is right to say that *if* there are entities fit to play the role of possible worlds in a general account of modality, they must be the robust, largely concrete objects the realist believes in. I will also suppose that Lewis has successfully blunted the more *philosophical* objections to modal realism, like the charge that it renders modal knowledge impossible or that it leads to paradoxes of various sorts.

The only objection I will not suppose answered is perhaps the most famous: the incredulous stare.[5] As Lewis says, 'modal realism *does* disagree, to an extreme extent, with firm common sense opinion about what there is' (*Pl.*, p. 133). To embrace it is to undertake a massive revision in the world-view we bring with us to philosophy. And it may be that in spite of everything Lewis says, in spite of the many ways in which he shows his hypothesis to be fruitful in systematic metaphysics, one still finds oneself unable to assent to it. This is the significance of the incredulous stare; and Lewis is right to say that he cannot refute it, because it is not an argument. It is rather an expression of the degree to which, given our starting point, modal realism, with its commitment to countless non-actual talking donkeys and the like, must strike us as utterly incredible.

When I said that Ed was unwilling to take on a commitment to possible worlds, this was the character I had in mind—a theorist who rejects ersatzism, perhaps for Lewis's reasons, but at the same time finds the realist's metaphysical picture impossible to accept. Such a theorist

---

[5] Cf. D. Lewis, *Counterfactuals*, Oxford, Basil Blackwell, 1973, p. 86; also *Pl.*, ch. 2, sec. 8.

has good reason to look into Ed's remaining option—to interpret his apparent quantification over possible worlds as an innocent *façon de parler*, involving no commitment to worlds of any sort.

I will call any interpretation with this feature a *deflationist* interpretation of the possible worlds framework.[6] According to deflationism, you can have all the benefits of talking about possible worlds without the ontological costs. You can legitimately say in one breath (perhaps in the course of explaining what you mean by some modal claim) 'there is a world where blue swans exist' and in the next breath; 'but really, I don't believe in possible worlds'. The trick is to explain why this is not a plain contradiction.

A complete survey of the deflationist's options would, I suppose, have to recognize as a potential starting point each of the myriad dodges philosophers have devised to explain away their apparent quantification over entities they profess not to believe in. Thus the deflationist might try to maintain that his quantifier over worlds is not the familiar existential quantifier but some more exotic bird. Perhaps it is a substitutional quantifier; or possibly Routley's Meinongian 'particular' quantifier which ranges indiscriminately over what there is and what there is not.[7] Alternatively, the deflationist might try to argue that the surface form of his assertion is more radically misleading, perhaps by claiming that his utterances of 'there are blue swan worlds' when properly understood contain no quantifier over worlds of any kind, much as claims about *the average family* are shown upon analysis to contain no term purporting to refer to such a creature.

I want to discuss the prospects for deflationism. But I will not survey the landscape here. Instead my procedure will be to describe at the outset a new deflationist strategy which strikes me as especially promising. The aim is to see what sort of balm it might offer for someone in Ed's sad condition.

2. Fictionalism

The deflationist strategy I want to discuss might fittingly be called 'fictionalism' about possible worlds.[8] The central idea is that sentences about, for example, blue swan worlds—sentences which look like straightforward existential generalizations—should be understood by analogy with (3):

---

[6] Cf. G. Forbes, *The Metaphysics of Modality*, Oxford, Oxford University Press, 1985, where such approaches are called 'anti-realist'.

[7] R. Routley, *Exploring Meinong's Jungle and Beyond*, Canberra: Philosophy Department, Research School of Social Sciences, Australian National University, Monograph #3, 1980, p. 176.

[8] As the name suggests, the view has affinities for certain so-called 'fictionalist' positions in the philosophy of mathematics. Cf. C. Chihara, *Ontology and the Vicious Circle Principle*, Ithaca, Cornell, 1973, ch. 2; H. Hodes, 'Logicism and the Ontological Commitments of Arithmetic', *Journal of Philosophy*, 1984, pp. 123–49.

(3) There is a brilliant detective at 221b Baker Street.

Taken as a straightforward existential claim, this is false, and anyone who asserts it with the intention that it be so understood speaks falsely. Yet we know that there are conversational contexts in which utterances of (3) may be perfectly *correct*—even true; namely, conversations about what happens in Sherlock Holmes stories. In these contexts, utterances of (3) are not intended straightforwardly. Rather, when the participants are mutually aware that the topic is a certain body of fiction, and not, for example, the history of British criminology, utterances of (3) will be read as elliptical renderings of (4):

(4) In the Holmes stories, there is a brilliant detective at 221b Baker Street.

Now this claim is perfectly true; and so, therefore, are utterances of (3) where the context can be counted upon to supply the silent prefix 'In the Holmes stories . . .'.[9]

For us, however, the important thing about (4) and its ellipsis (3) is not that they are true, but rather that someone may unflinchingly believe or assert what they say without committing himself to believing that a detective lives in Baker Street. The Holmes stories are *fictions*. And one characteristic feature of fictions is that we may legitimately have opinions about their contents without believing that what they say is true, or that the objects they purport to describe exist.

Call any sentential operator of the form 'In the fiction F, . . .' or 'According to such and such a story . . .' a *story prefix*. The moral of the last paragraph, then, is that as we ordinarily understand these expressions, quantification within the scope of a story prefix is not existentially committing. You can believe 'According to the fiction F, $\exists x P x$' without believing '$\exists x P x$'; for as a rule, the former does not entail the latter.

We may generalize further by noting that the story mentioned in a story prefix need not be a literary fiction, nor for that matter, any sort of fiction in the usual sense. Russell thought that *according to Leibniz's monadology*, the table is really a colony of souls. But we do not conclude on this basis that Russell was himself committed to the animist metaphysic. In general, the fiction mentioned in a story prefix can be *any representation whatsoever*: a story, a scientific theory, or a metaphysical speculation. The basic point is unaffected: so long as you are not independently committed to regarding

---

[9] This approach to fictional truth is articulated in D. Lewis, 'Truth in Fiction', *American Philosophical Quarterly*, 1978, pp. 37–46; cf. N. Wolterstorff, *Works and Worlds of Art*, Oxford, Oxford University Press, 1980. For another, rather different approach which may also lend itself to a fictionalist construal of possible worlds, Cf., K. Walton, 'Fearing Fictions', *Journal of Philosophy*, 1978, pp. 5–27.

this representation as true, when you assent to 'In *F*, *P*' you incur no obligation to assent to '*P*' by itself.

The fictionalist about possible worlds hopes to take advantage of this fact by interpreting his own apparent quantification over worlds as quantification within the scope of a story prefix. As with (3), the prefix will sometimes be *silent*; so the fictionalist will often sound just like the modal realist. Yet the fictionalist's claims about possible worlds will always be elliptical for claims about the content of a story; and the ellipsis can always be expanded. In this way, the fictionalist hopes to earn honest title to the language of possible worlds—that is, title to talk as if there were such things—while retaining a sensible one-World ontology.

The first and most obvious problem in developing this strategy is to specify the story the fictionalist plans to exploit. The chief constraint, of course, is that it be one according to which the usual claims about possible worlds—for example, the claim that there are blue swan worlds—are true. There may be more than one story or theory of possible worlds which meets this constraint. But one is particularly salient, especially in the present dialectical context. I have in mind Lewis's own modal realism, a theory explicitly designed to meet this constraint. This suggests a starting point: the fictionalist will explain that his own talk about possible worlds is to be understood, not as talk about what exists in fact, but rather as talk about what exists according to the realist's hypothesis of an immense plurality of robust universes.[10]

In other words, what the realist regards as true metaphysics, the fictionalist regards as a (probably) false story, to be mentioned *but not asserted* in his account of modality. For example, although the fictionalist does not for a moment believe that there are worlds with blue swans, he does believe (as anyone should) that

(5) According to the hypothesis of a plurality of worlds, there is a world *W* such that at *W* there are blue swans.

The fictionalist maintains that when he utters, as he might, 'there is a blue swan world' what he really means to assert is (5). This uncontroversial metafictional thesis is the fictionalist's paraphrase of the equally uncontroversial modal claim that there might have been blue swans.

More generally, let *P* be an arbitrary modal proposition. The modal realist will have ready a non-modal paraphrase of *P* in the language of possible worlds; call it *P**. The realist's assertions about possible worlds are guided by explicit adherence to the schema *P iff P**. The fictionalist's parasitic proposal is therefore to assert every instance of the schema: *P iff according to the hypothesis of the plurality of worlds, P**. Like modal realism, the theory would seem to provide truth conditions for modal claims in a

---

[10] For a theory of possible worlds that might serve as the basis for an alternative fictionalism, cf. D. M. Armstrong, 'The Nature of Possibility', *Canadian Journal of Philosophy*, 1986.

systematic way. Indeed, if the prefix is allowed to go silent, the two theories will be verbally indiscernible. The difference is that the fictionalist's account does not presuppose an ontology of possible worlds.

3. *The fictionalist's fiction*

Before we can begin to evaluate this proposal, we need to be a bit more explicit about this fiction I have called 'the hypothesis of the plurality of worlds'. We begin with a set of informal postulates.

(6*a*)   Reality consists in a plurality of *universes* or 'worlds'.

(6*b*)   One of these is what we ordinarily call *the* universe: the largest connected spatiotemporal system of which we are parts.

(6*c*)   The others are things of roughly the same kind: systems of objects, many of them concrete, connected by a network of external relations like the spatiotemporal distances that connect objects in our universe. (*Pl.*, pp. 2, 74–6)

(6*d*)   Each universe is isolated from the others; that is, particulars in distinct universes are not spatiotemporally related. (It follows that universes do not overlap; no particular inhabits two universes.) (*Pl.*, p. 78)

(6*e*)   The totality of universes is closed under a principle of recombination. Roughly: for any collection of objects from any number of universes, there is a single universe containing any number of duplicates of each, provided there is a spacetime large enough to hold them. (*Pl.*, pp. 87–90)[11]

(6*f*)   There are no arbitrary limits on the plenitude of universes. (*Pl.*, p. 103)[12]

(6*g*)   Our universe is not special. That is, there is nothing remarkable about it from the point of view of the system of universes.[13]

These postulates—whose significance will become clearer as we go

[11] The final proviso is required to forestall the paradox presented in P. Forrest and D. M. Armstrong, 'An Argument Against David Lewis's Theory of Possible Worlds', *Australasian Journal of Philosophy*, 1984, pp. 164–8. As Lewis understands this condition it has the effect that there is an (unknown) upper bound to the dimensionality of the largest spacetimes represented among the universes, and so an upper bound to the number of non-overlapping objects that inhabit the most populous worlds. This will be important below, sec. 7.

[12] As with recombination, adequate formulation of this principle is tricky. It is meant to guarantee, e.g., that the upper bounds mentioned in the previous note are *natural* ones. Thus it may be that the largest worlds possess a countable infinity of spatial dimensions. But it is ruled out that they have, say, $10^{17} + 1$ dimensions. To suppose so would be to imagine that the plenitude of worlds was in this sense arbitrarily restricted.

[13] Lewis does not distinguish this requirement from non-arbitrariness. But they srike me as importantly different. The present principle is concerned with the status of our universe with respect to the others. Thus (6*g*), but not (6*f*) entails that there must be properties or universals that are not instantiated in our universe, but are instantiated in others. If our universe contained an instance of every possible natural property, it would be a remarkably rich world, in violation of (6*g*).

on—are meant to capture in a preliminary way the ontological core of Lewis's modal realism. They describe the objects a modal realist *qua* modal realist must believe in.

In one important sense, however, the postulates fail to capture the full force of modal realism. To see why, it is enough to note that (6) is formulated without modal vocabulary. The postulates are all theses about what there *is*, not what there *must* or *might be*.[14] And this has the odd consequence that the theory they determine is (in some sense) compatible with the view that everything is actual. That is, we can imagine an extravagant actualist embracing (6) while adding that only actual things exist. There is no contradiction here. It is simply the eccentric view that actuality is much bigger than we ordinarily think—a vast sea of 'island universes'—a cosmology that might appeal to an actualist metaphysician driven by the converse of Quine's taste for clear skies and desert landscapes.[15]

I take it, however, that modal realism is incompatible with actualism by definition. To be a modal realist is to believe that some things are real but not actual. In order to capture this aspect, then, more needs to be said. For example, we might supplement the postulates with some remark such as (7):

(7) The universes are *possible worlds*; our universe and its parts are actual individuals; the others are *merely possible*.

(6) and (7) together clearly entail that some things are real but non-actual. The theory they determine deserves the name 'modal realism' (MR), and anyone who believes it is a modal realist.

I labour this point mainly to point out that the fictionalist's fiction is not exactly MR. With the addition of (7), MR contains overt (albeit defined) modal vocabulary in its formulation. For reasons of clarity, however, the fictionalist does well to keep his fiction non-modal. Still, he wants to retain the realist's idea that our universe is one of many. He therefore takes as his starting point the original postulates (6); the supplementary modal bridge laws in (7) will play no role in his account.

---

[14] That such a specification is possible is of course crucial for the modal realist, who hopes to exploit his theory of possible worlds theory in a fully reductive analysis of the modal notions.

[15] For a striking anticipation, cf. Poe's *Eureka*:

I myself feel impelled to the *fancy*—without of course calling it more—that there does exist a limitless succession of Universes, more or less similar to that of which we have cognizance—to that of which alone we shall ever have cognizance . . . . *If* such clusters exist, however—*and they do*—it is abundantly clear that, having had no part in our origin, they have no portion in our laws. They neither attract us nor we them. Their material—-their spirit is not ours—is not that which obtains in any part of our universe. They could not impress our senses or our souls. And between them and us . . . there are no influences in common. Each exists apart and independently, in the bosom of its proper and particular god!

In *The Complete Works of Edgar Allen Poe*, New York, AMS Press, 1965, 16: 256. Cf. P. Unger, 'Minimizing Arbitrariness: Toward a Metaphysics of Infinitely Many Isolated Concrete Worlds', *Midwest Studies*, 1984.

But these original postulates are not quite enough for the fictionalist's purposes. They tell us a great deal about the realist's general conception of the totality of universes. Yet they say rather little about what sorts of objects the various universes contain. Thus nothing in (6) guarantees that some of the universes contain people or clouds or stars, much less blue swans and tailless kangaroos. As we will shortly see, however, it is crucial that the fictionalist's fiction provide a rich and detailed picture of what goes on in the various universes. Towards this end, the fictionalist proposes that the original postulates be supplemented with an *encyclopaedia*: a list of the non-modal truths about the intrinsic character of this universe.[16] The encyclopaedia specifies, for example, that our universe contains objects which are rather like ordinary kangaroos but for the lack of a tail (at least in the form of undetached kangaroo-parts). So, by recombination (6e), other universes contain free-standing tailless kangaroos, duplicates of these residents of our world. This new theory—(6) together with the complete catalogue of non-modal intrinsic truths about our world—is the fictionalist's fiction. Call it *PW*.

4. *The fictionalist theory of possibility*

We can now state the fictionalist's proposal in a more precise way. Let P be an arbitrary modal claim, and P^* the modal realist's non-modal paraphrase of P in the language of possible worlds. (If the realist's claim to provide a reductive analysis is sound, P^* must always exist). The modal realist asserts the biconditional P *iff* P^*, which, as we have seen, leads directly to a commitment to possible worlds when conjoined with common-sense modal opinion. The fictionalist's ploy is to borrow P^*, the fruit of the realist's labour, without asserting it, and to assert instead instances of the schema:

P *iff according to PW*, P^*.

As Russell might have said, the method has (at least) the advantages of theft over honest toil. What remains to be seen is whether it has any others.

To illustrate, consider Ed's conviction that

(8) There might have been blue swans.

For the modal realist, this is equivalent to an existential generalization about possible worlds:

(8r)   There is a universe containing blue swans.

---

[16] For our purposes, S is non-modal intrinsic truth about our universe if it contains no modal vocabulary and entails neither the existence nor the non-existence of things outside our universe. It suffices for the latter that all quantifiers and names in S be restricted to the inhabitants of this universe.

The fictionalist therefore offers the following paraphrase:

(8*f*)   According to *PW*, there is a universe containing blue swans.

The example brings out the central contrast between fictionalism and modal realism. If, like Ed, you are at all sceptical about the real existence of blue swans, you should not believe (8*r*). But no one should doubt that (8*f*) is true. It should be utterly uncontroversial that *according to the realist's conception of the plurality of universes*, there is a universe with a blue swan in it. And this is as it should be, because it is equally uncontroversial that there might have been blue swans. And for the fictionalist these come to the same thing.

To get some sense of the intended range of the fictionalist analysis, consider a counterfactual.

(9) If swans were blue, ducks would be pink.

Glossing over certain subtleties, we may suppose that the realist construes this as a restricted universal generalization:

(9*r*)   In each universe that differs from ours as little as the blueness of swans permits, ducks are pink.

So the fictionalist offers the obvious parasitic analysis:

(9*f*)   According to *PW*, every universe that differs as little from ours as the blueness of swans permits is a universe where ducks are pink.

Here the sceptic about other universes can plausibly endorse the realist's paraphrase. If ours is the only universe, (9*r*) is vacuously true. In contrast, (9*f*), like the counterfactual itself, is obviously false. On the realist's view there are universes very much like ours with blue swans and ordinary ducks; and these are more like our world in the relevant respects than is any universe featuring both non-standard ducks and swans.

These examples are elementary, of course. They contain none of the iterated modalities or other syntactic complexities which the cognoscenti will recognize as potential sources of difficulty. Still, they bring out the central contrast: by and large, the fictionalist will paraphrase uncontroversially true (false) modal statements with similarly uncontroversial statements about the content of the theory *PW*. And other things being equal, this is an advantage, since the realist paraphrase of even the most trivial modal thesis is generally a highly contentious metaphysical speculation, sharply at odds with what most of us (including Ed) ordinarily believe.

I will come back to the relative merits of the two positions in a moment. But first I would like to stress one important respect in which the views are on a par. The examples suggest (without of course proving it) that in central cases, *the realist and the fictionalist will concur in their theoretically*

informed modal judgements. That is, whenever the realist believes that his paraphrase of a modal claim is true, the fictionalist should have the same view of *his* paraphrase. This is a straightforward consequence of the fictionalist's shameless parasitism. When the realist believes his account P^* of the modal claim P, this can only be because he judges P^* true on the basis of a theory he accepts. But this is precisely what the fictionalist is interested in: the truth value of the realist's paraphrase according to the realist's theory. The realist and the fictionalist therefore generate theoretically informed modal judgements by considering precisely the same range of facts. And this guarantees that the two will generally agree about the modal truths. (But see section 7 below.)

Now one plausible desideratum for a philosophical account of modal discourse is that it should *ratify* a substantial body of prior modal opinion. The theorist should do his best, in other words, to paraphrase modal claims we all believe with propositions that he is entitled to believe in light of his theory. The upshot of the last paragraph is therefore the following conditional correctness claim: if the modal realist is in a position to regard his analysis as adequate in this minimal sense, the fictionalist is as well.[17]

Of course, such fidelity to common-sense modal opinion is not the only constraint on an adequate account of modality. More generally, we should hope that our approach will be largely compatible with the rest of what we believe about the world and our place in it. For example, I take it as given that we possess a great deal of modal knowledge. Hence it is a grave fault in any philosophical account of the modal facts if it makes it mysterious that we can know them, given what we firmly believe about the limits of our faculties. We also bring a relatively firm ontological view to the analysis of modality; and any theory that requires us to abandon important parts of it is faulty for a similar reason. How then does fictionalism fare with respect to these further constraints on our project?

5. *Ontology*

As I have said, perhaps the most powerful objection to Lewis's view is that, seen in light of things we firmly believe, modal realism is simply incredible. Normally we do not think that there are countless universes besides our own, much less that they are populated by an infinity of blue swans, talking donkeys, and the rest. Yet modal realism tells us that these things exist. To embrace it is therefore to undertake a substantial revision in the world-view we bring with us to this metaphysical problem.

Now strangeness can never be a decisive objection. But I believe it must count for something. We have little use in the end for a philosophy we cannot seriously believe. Hence when we are exhorted to change our view

---

[17] Of course I do not claim to have shown that the realist's analysis is adequate in this sense. This is the burden of much of Lewis's book.

over to one which by our present lights seems fantastically unlikely, we are entitled to ask for correspondingly strong reasons. Lewis's main strategy is to motivate modal realism by showing how *useful* it is in systematic philosophy to suppose that there are many worlds. One may doubt that modal realism really is as useful as Lewis says; but even granting this, one may still wonder whether such considerations of utility could ever be enough to motivate so radical a revision in our conception of what there is.[18]

The first and most obvious advantage of fictionalism, then, is that it appears to have no revisionary ontological consequences whatsoever. At the very least, you can believe everything the fictionalist says about the truth conditions for modal statements without having to believe in possibilia. For as we have seen, the fictionalist's claims are claims about the content of a story. And because they do not presuppose the truth of that story, you can believe them without inheriting a commitment to the items the story purports to describe.

This is not to say that fictionalism is altogether silent on ontological matters. It may well be that in talking about stories, theories, and other representations as he does, the fictionalist takes on a commitment to these entities. And since it is conventional to regard these representations as *abstract* entities, fictionalism may not appeal to certain nominalists.[19] Now this conventional view is hardly sacrosanct. It may be possible, for all I know, to give a nominalistically acceptable treatment of the fictionalist's ontology. But even if this is not possible, it must be admitted that, given our starting point, a commitment to abstract objects is not nearly so weighty as the realist's commitment to possible worlds. Nominalism was no part of the sensible, naive ontology that Ed, for one, was so concerned to preserve. Most of us already believe in stories, theories, and the rest, even if we lack a neat theory of them. In any case, the point is not that fictionalism says nothing about what exists, but

---

[18] Indeed, one may wonder why such arguments from utility should be even prima facie compelling for Lewis given his other views. For Lewis, our aim in metaphysics as in physics is to believe the truth. And because he is a realist (in every sense) about truth, Lewis grants that there is no necessary connection between a theory's being useful, elegant, etc., and it's being true. Of course we may *hope* that the truth may be captured by a beautiful, fruitful theory, and our theory construction may be guided by this hope; but we have no independent reason to believe it must be so. Now the arguments for modal realism may establish that it is useful, elegant, and so on. So I can see how they might give someone reason to hope that it is true. As Lewis says, if modal realism is right we inhabit a paradise for philosophers, and we are all supposed to hope to dwell in paradise. The trouble is that reason to hope is not in general reason to believe. To suppose otherwise is to countenance wishful thinking as a sound policy for fixing belief in metaphysics.

[19] Chihara (op. cit.) maintains that fictionalism in the mathematical case is nominalistically acceptable because prefixes of the form 'In F, . . .' may be thought of as simple symbols, not open to quantification in the 'F' place. I have my doubts, if only because the fictionalist's explanatory remarks seem inevitably to involve quantification over representations. Of course even if this is right it should not bother a liberal nominalist like Chihara: anyone who admits an ontology of linguistic entities that outstrips the totality of concrete inscriptions can easily construe the fictionalist's fictions as entities of this type.

rather that, unlike modal realism, it requires no large scale change in view; and other things being equal, this must seem a distinct virtue of the fictionalist approach.

6. *Epistemology*

If fictionalism aspires (as modal realism surely does) to provide an account of the nature of modal truth, then the fictionalist must show his account to be compatible with the manifest fact of our modal knowledge. He must convince us that the modal facts as he conceives them are the sort of fact that creatures like us can know. In this section I will only take up part of this problem. I would like to ask whether fictionalism is compatible with the view that our usual methods for forming modal beliefs are generally a good guide to the modal truth. This is a central question for any account. A theory that cannot ground our confidence in the reliability of our best methods would seem to lead rather quickly to modal scepticism, the view that we have no modal knowledge; a claim which, like most strong sceptical theses, is very hard to believe.

Modal realism, of course, is often said to violate this condition. The worry is that if modal statements are understood as claims about a domain of physical objects from which we are causally isolated, it may seem hard to see why we should credit our world-bound faculties with the capacity to deliver the modal truth. I will not review Lewis's response to this challenge: I find it (sometimes) compelling (cf. *Pl.*, pp. 108–15). Instead I would like to focus on certain important remarks he makes along the way. Responding to the demand for a naturalistic account of the sources of modal belief, Lewis writes:

In the mathematical case, the answer is that we come by our opinions largely by reasoning from general principles that we already accept. . . I suppose the answer in the modal case is similar. I think our everyday modal opinions are, in large measure, consequences of a principle of recombination. . . One could imagine reasoning rigorously from a precise formulation of it, but in fact our reasoning is more likely to take the form of imaginative experiments. We try to think how duplicates of things already accepted as possible . . . might be rearranged to fit the description of an alleged possibility. Having imagined various arrangements—not in complete detail, of course—we consider how they might be aptly described. (*Pl.*, pp. 113–14)

This passage sketches a plausible account of the role of imagination in modal reasoning. Imagination does not put us in 'contact' with a special realm of being. Rather, when we imagine we consider the consequences of certain principles which guide our thought—in the modal case, principles which seem well captured by the postulates of *PW*. We construct representations of situations in accordance with our empirical knowledge, the principles of recombination, non-arbitrariness, and so

on.[20] And if one of the situations we imagine by this route is aptly described by the statement *P*, we conclude that 'Possibly *P*' is true.

Now if something along these lines is right, any account of modality must somehow explain why we should regard this imaginative method as a reliable way of finding out the modal truth. Discovering the consequences of principles that guide the imagination is, after all, a broadly psychological enquiry. Why should that be a good way of discovering what might have been?

The realist's answer is deceptively simple. (i) the modal truths are truths about a domain of universes; (ii) the principles which guide our imagination are true of that domain; so (iii) by and large, when we imagine in accordance with these principles the states of affairs we imagine are realized somewhere among the universes.

Straightforward, perhaps; but this line of response ought to seem profoundly puzzling. Letting the first premiss pass, (ii) remains a striking conjecture. We may ask: is it just a coincidence that the principles which guide our imaginations truly describe a domain of objects with which human beings had absolutely no contact when those principles were being shaped, presumably by a perfectly natural evolutionary process? After all, there might have been creatures whose imaginative principles were quite out of step with the distribution of worlds in modal space. How is it that we are so lucky as to have been given the *right* imaginative dispositions? Surely the realist owes us an explanation.

I will not presume to answer for the realist, who will no doubt begin by pointing to other cognitive domains (e.g., inductive reasoning) where the reliability of our faculties must be allowed to be an historically contingent matter. The only point I want to stress is that no analogous problem arises for the fictionalist. If the realist is right in suggesting that we are guided in the imaginative construction of possibilities by principles like the postulates of *PW*, then when we engage in imaginative experiments, *the least we discover is what is true according to PW*. But for the fictionalist, that is enough. The modal facts just are facts of this kind. Thus for the fictionalist there is no special mystery as to why we should trust our imaginations as a guide to the modal truth. And once again, this seems to me an important point in fictionalism's favour.[21]

---

[20] Lewis's remarks may suggest, by analogy with the mathematical case, that we accept or believe the principles which tacitly guide the imagination. But this seems wrong. The principles in question are principles about possible worlds. So if we believe them we are committed to modal realism from the cradle, simply because we reason about what might have been. It seems clear, however, that with one or two exceptions, modal realists are made, not born.

[21] This is not to say that there is not a deep epistemological question as to how we discover through imaginative reasoning the consequences of these imaginative principles. The point is only that the realist and the fictionalist both agree that such discovery is possible. So any problem this view presents is a problem for both of them.

7. *The incompleteness problem*

I hope fictionalism is beginning to look like a potentially attractive solution to Ed's dilemma: a way to exploit the possible worlds framework for various purposes without believing in worlds besides our own. Unfortunately, as the reader will no doubt have guessed, the picture is not so unreservedly rosy. In the sequel I shall discuss three problems for fictionalism. None is decisive against the view, but they give a fair picture of the complexities it may involve. For someone in Ed's position, they may be just the thing to tip the scales against the strategy we have been discussing.

Earlier I made the modest claim that fictionalism and modal realism agree with one another (and with everyday opinion) about a large body of modal doctrine—the modal statements that strike us either as obviously true or obviously false. A more general equivalence claim fails, however, and this may mean trouble for the fictionalist. It turns out that in some rather marginal cases, the fictionalist and the realist must disagree about what the modal truths are; and for a theorist whose intuitions clearly favour the latter, this may count against the fictionalist proposal.

To see the difficulty, we need as an example a modal claim which, from the realist's point of view, has a determinate truth value of which we are ignorant. Moreover, we require that this ignorance not be due to ignorance about the empirical facts of our universe. It must be a robust modal ignorance—an ignorance that would survive an arbitrary extension of our scientific and historical knowledge. Now it may not be obvious that such statements exist. But on Lewis's view, certain claims about how 'large' the universe might have been fit into this category. For example, let κ be a cardinal number larger than the number of space-time regions in our universe, and consider the claim that

(10)   There might have been κ non-overlapping physical objects.

For the realist, of course, this amounts to a claim about the size of the largest universes:

(10*r*)   There is a universe containing κ non-overlapping physical objects.

But this claim differs importantly from the others we have considered. As mentioned earlier, Lewis holds that there must be an upper bound κ^* to the number of non-overlapping objects that inhabit the most populous words (see n. 11). Yet we do not know what this upper bound is; and for all we know, even if we were fully informed about the empirical character of our universe we might still have no insight into this global aspect of the totality of universes. Thus for the realist, the modal claim (10) has a definite truth value; but we cannot say what it is, because we do not know—and may never know—whether $\kappa < \kappa^*$.

Now consider the fictionalist's attitude towards (10). For him, (10) is equivalent to

(10*f*)   According to *PW*, there is a universe containing κ non-overlapping physical objects.

What is the truth value of this sentence? Well, this much seems clear: the upshot of the last paragraph is that even if we had a full specification of *PW*—a list of the postulates, together with a complete account of the non-modal intrinsic nature of our world—we might still be in no position to affirm that according to *PW*, there are universes containing κ objects. *The story is simply silent on this point.* But (10*f*) implies that it gives a definite positive answer. So (10*f*) is not true.

The fictionalist therefore has two options: (10*f*) is false or it is truth-valueless. In either case, because (10*f*) is not true, the fictionalist must say that the modal claim (10) with which he takes it to be equivalent is not true. And this is already a departure from realism. The fictionalist who takes *PW* as his governing fiction can rule out in advance the possibility that there might have been κ objects; the realist cannot. So if for some reason you have a positive intuition that this ought to remain an open question, you may find this form of fictionalism counter-intuitive. But we can go even further. For it seems that whichever course the fictionalist takes he must say something implausible.

The more natural option is surely to call (10*f*) false. After all, (10*f*) says that the realist's theory settles whether there are worlds of a certain size; but the theory does not settle that question. So (10*f*) attributes to *PW* a property it does not have. To see why this natural verdict leads to trouble for the fictionalist, consider (11):

(11)   It is not the case that there might have been κ non-overlapping physical objects.

On the face of it, this is the negation of our original modal claim (10). And it is an important feature of our understanding of modal discourse that here, as everywhere, a statement and its negation are contradictories. That is, they must have different truth values if they have truth values at all. The fictionalist has to respect this appearance, I think. Hence if he declares (10) *false*, he must declare (11)*true*.

Unfortunately, this assignment of truth values is ruled out by symmetry considerations. The fictionalist's analysis of (11),

(11*f*)   According to *PW*, no universe contains κ physical objects.

has exactly the same form as (10*f*): the story *PW* is silent about the truth value of the embedded sentence. But since this was our reason for calling (10*f*) false, symmetry requires that we declare (11*f*)—and hence (11)—false as well. The intuitive claim that (10*f*) is false therefore seems to

force the fictionalist to concede that (10) and (11)—a modal claim and its negation—have the same truth value, in violation of what I take to be a fairly firm commitment of ordinary usage to the effect that such claims are always contradictory.[22]

The fictionalist can avoid this result by declaring that in general when the paraphrase $P*$ of a modal claim P is not determinately settled as true or false by the theory PW, the modal claim P is to lack a truth value. (10) and (11) would then both be truth valueless, which is compatible with their being contradictories.

I am going to suppose that the fictionalist makes this stipulation. Unfortunately, besides being less natural than the first, it has one or two worrying consequences. The first is that with this stipulation, the ordinary logical connectives when applied to modal statements are no longer truth-functional. (10) and (11) are truth valueless. But their disjunction is a logical truth, as is its paraphrase into the language of PW. And I assume that all such logical truths are true according to PW. So the fictionalist must allow that in the modal case we may have a true disjunction with neither disjunct true. A similar situation arises in other areas, of course, most notably in certain philosophical treatments of vagueness and the truth-predicate which employ a supervaluational approach to the seman-tics of truth value gaps.[23] Still it may seem troubling to see the phenomenon repeated here, in a domain where our ordinary ways of thinking give us no reason to expect a complicated propositional logic, and where the only reason for proposing one is, to say the least, generated by concerns rather distant from the linguistic practice in question.

A second problem is more troubling. Recall that in the intuitive explanation of the fictionalist's strategy I relied on an understanding of how story prefixes generally work. We understand what sort of expression the fictionalist's operator 'According to PW' is supposed to be only because we are told that it is a member of the familiar class of expressions that includes, for example, 'In the Holmes stories . . .'. Unfortunately, it seems fairly clear that the familiar story prefixes do not give rise to truth value gaps like those we have just proposed on the fictionalist's behalf. The Holmes stories are silent about Moriarty's blood type just as PW is silent

---

[22] It may be objected that we have a precedent for making just this claim in a domain quite close to our subject. According to many accounts of truth in fiction, the sentences 'Moriarty had blood type $A+$' and 'It is not the case that Moriarty had blood type $A+$' are not genuine contradictories; for as they are standardly intended they are elliptical for prefixed sentences where the prefix 'In the Holmes stories . . .' takes wide scope. 'In F, not-P' is not the negation of 'In F, P'. Hence both may be false. The fictionalist, however, cannot appeal to this precedent to explain how the modal theses (10) and (11) can both be false. For it is no part of his view that modal statements (in contrast with statements about possible worlds) are ever supplied with a tacit prefix.

[23] Cf. B. van Fraassen, 'Singular Terms, Truth-Value Gaps, and Free Logic', *Journal of Philosophy*, 1966, pp. 481-95; D. Lewis, 'General Semantics', reprinted in his *Philosophical Papers*, v.1, New York, Oxford, 1983; S. Kripke, 'Outline of a Theory of Truth', *Journal of Philosophy*, 1975, pp. 690-715.

about the size of the largest worlds. Yet if someone were to assert 'In the Holmes stories, Moriarty has blood type A +' we would naturally suppose that he had said something *false*, not something lacking in truth value. Again: quantum mechanics is silent on the panpsychism question. But when avatars of the new age say that according to quantum mechanics, the universe is conscious, they speak *falsely*. Quantum mechanics says no such thing. The solution to the problem of modal contradictories has the consequence that the fictionalist's operator departs from the paradigm by giving rise to gaps rather than falsehoods in analogous circumstances. And this should leave us wondering how well we really understand it. If it differs in this way from other story prefixes, how else does it differ and why? Unless he can answer, the fictionalist's view is infected by unclarity at its core.

8. *Primitive modality*

So far we have treated the fictionalist's operator as an intuitively understood but undefined primitive expression. Yet these last remarks suggest that it is in fact a potentially puzzling creature; and one may legitimately suspect that a serious effort to explain its semantic behaviour will reveal hidden complexity in the fictionalist's view.

Further reason to worry emerges when we consider what we might say by way of explication to someone who did not already understand the expression 'According to PW, P'. We might begin by offering any of the following glosses: *If PW were true then P would be true; If we suppose PW, P follows; It would be impossible for PW to be true without P being true as well.* These are not perfect paraphrases. None the less, each seems to give a fair preliminary indication of what we mean when we use the fictionalist's prefix. And the trouble is that in every case the key phrase is an overtly modal locution. This suggests that the fictionalist's device should itself be classed as a modal operator.

Two kinds of objection spring to mind at this suggestion. The weaker claims that, at the very least, the fictionalist cannot boast the comprehensiveness of the modal realist; for the realist claims to be able to analyse *all* modal locutions in non-modal terms—a full reduction of the modal to the non-modal—whereas the fictionalist's analyses invariably contain an unreduced modal component. The stronger maintains that the fictionalist's theory is not just less comprehensive than we might have hoped: it is altogether unhelpful. For the fictionalist's modal primitive is an especially obscure one—far more obscure, in fact, than the modal notions he would explain by means of it.

In response, the fictionalist should first allow that *if* the prefix is to be called a modal operator, then he has not furnished the materials for an eliminative reduction. What he has done (if his account succeeds in other

respects) is to suggest a way of reducing a wide variety of modal notions to this one. Now this may not seem as impressive as a thoroughgoing eliminative reduction. But it remains a non-trivial analytic advance. No similar claim can be made, for example, by those who would take the logician's standard modal operators as primitive. For as several writers have pointed out, many recognizably modal judgements cannot be expressed in a language whose modal vocabulary is restricted to boxes and diamonds.[24]

Second, the fictionalist should point out that if his prefix is a modal operator, then it is by no means obvious that the realist has given a fully reductive treatment of modality; for it remains unclear how the realist proposes to eliminate occurrences of this very locution. True, the realist does have a powerful strategy for eliminating a large class of story prefixes as part of his account of truth in fiction. This theory identifies the content of a story S with a set of possible worlds, C_S, determined by several factors, including the explicit text of the fiction, the intentions of the story teller, the circumstances of its telling, and so on. A sentence P is said to be true *according to S* iff P is true at every world in C_S. To take a simple example: the content of 'A Study in Scarlet' might be given by the set of all worlds least different from the actual world in which the text of that story is inscribed as a record of known fact. What is true in the story is what is true in all these worlds.[25]

This style of analysis admits of variations. The important point, however, is that even if one of them is ultimately successful as an account of truth in simple narrative fiction, the account does not generalize to metaphysical 'stories' like the realist's hypothesis. The fictions for which the possible worlds analysis is suited are in an intuitive sense stories about how things might *actually* have transpired. That is why they are naturally represented as sets of possible worlds. *PW*, however, is not a story in this mould. It is a representation of all of modal reality, not just our small corner of it. And this is an important difference.

To see this more clearly, consider a recent, justifiably ignored novel called *This Lonely World*. The author is a committed modal realist, and his book is a dystopia modelled loosely on certain anti-totalitarian fables of the thirties: in this case, a nightmare vision of the actualist's conception of reality. The reader is invited to imagine or pretend that there is only one world—ours—and then to contemplate the grim implications. In the story, whatever happens happens necessarily. Nothing is possible but what is actual. Thus, as the author makes quite clear, regret is misplaced;

---

[24] Cf. *Pl.*, p. 13; A. Hazen, 'Expressive Completeness in Modal Languages', *Journal of Philosophical Logic*, 1976, pp. 25–46. In the system of Forbes (op. cit.), the language of modal logic is extended with a countable infinity of modal operators to render it comparable in expressive power to the language of modal realism.

[25] Lewis, 'Truth in Fiction'; Wolterstorff, op. cit.

deliberation is pointless; and most tragic of all, the most elegant and fruitful of metaphysical theories is false.

Never mind whether these are the right implications. This is fiction after all. Our question is whether the modal realist armed with his reductive account of truth in fiction can represent the content of this story (which is, after all, not so different in its ontological aspect from what many of us actually think). That is, how does the realist propose to analyse statements of the form 'In *This Lonely World, P*' so as to eliminate occurrence of the story prefix?

If the analysis is to follow the paradigm, the content of the tale should be a set of worlds. But which set? Not the set containing the actual world alone: this represents a complete account of what is actually the case, and that is not what the book contains. Perhaps the set of lonely worlds— worlds according to which there are no other worlds? If there were such worlds it might make sense to say that everything that happens in them happens necessarily. The trouble is that the modal realist has no reason to believe in lonely worlds, if indeed he can make any sense of the idea.[26]

Clearly, the problem is that for the realist, a representation of the totality of worlds is either necessarily true or necessarily false. The content, if it is to be a set of worlds, must therefore be either the set of all worlds or the null set. And neither will do to capture the very determinate content of stories like *This Lonely World*, or for that matter, *PW* itself. Thus from the realist's own point of view, story-telling about the actual world is a very different activity from story-telling about all the worlds. The realist has an elegant reductive treatment for story prefixes of the first kind; but he lacks a parallel account for the second, a class that includes the prefix 'according to *PW*'.

Notice now that the fictionalist is in precisely the same position. Suppose, as we have, that he regards his operator as primitive. This done, he can parrot in his usual way the realist's account of truth in well-behaved fiction, for example, by saying that when *F* is an ordinary narrative fiction, '*In F, P*' is true iff *according to PW, the closest universes in which F is told as known fact are universes in which P*. This is borrowed, of course, but there is no circularity, since the fictionalist does not intend 'According to *PW*' to fall within the scope of the analysis.

If you are keeping score, the fictionalist can so far match the realist analysis for analysis, and both are without an account of the operator 'According to *PW* . . .'. The standoff might not last, of course. The realist

---

[26] I suppose a lonely world would have to be a world relative to which no others are possible. Presumably, however, any notion of relative possibility that might be introduced into a realist treatment would rely heavily on facts about similarity of worlds. The lonely worlds would then be worlds not relevantly similar to any others. Recombination guarantees, however, that for any world there are worlds very much like it in any respect that could possibly matter.

might some day be able to give an account of the fictionalist's prefix. And if that account crucially involves the apparatus of possible worlds, the fictionalist will not be able to borrow it without circularity. I stress, it is not at all obvious how this account might go; yet even if we suppose it given, it may still seem at this stage that the fictionalist has done remarkably well. He has offered a powerful reduction of a wide variety of modal notions to one—a streamlining of ideology—with no cost in ontology. His account is adequate to common sense modal opinion; it makes good sense of the sources modal knowledge . . .. Is not one modal primitive a fair price to pay for all that?

Some will no doubt be put off by the conception of philosophical method presupposed by the question. But let it stand. Even so, it may still be objected that our accounting is far too charitable. The difference between the two views is not just that the fictionalist must take as primitive a modal notion the realist defines, much as some treatments of propositional logic take negation and disjunction as primitive while others define both in terms of the Sheffer stroke. The trouble is that there is something deeply *unsatisfying* about the fictionalist's choice of primitive. First, as our discussion of the incompleteness problem shows, it is not nearly as clear as one might like a primitive to be. Insofar as the fictionalist tells us anything about how it works, he says that it works like the story prefixes we already understand. But the fact is that it differs importantly in two respects: in giving rise to truth value gaps and in resisting the standard possible worlds analysis. Second, and perhaps more importantly, this primitive just does not *feel* primitive. Truth relative to a story sounds like the sort of thing one ought to be able to explain, in stark contrast to more plausible primitives like negation. Lastly, it seems clear that a theory whose bedrock includes a notion as obscure as 'true according to *PW*' does little to address one traditional anxiety about the legitimacy of modal language. I have in mind the spectre of the Humeo-Quinean sceptic, who objects to the use of modal language in philosophy on the grounds that it is simply too poorly understood to serve a solid intellectual purpose. I have not stressed the project of answering this last, confessedly protean, sort of worry here. But pressure from this direction has persistently fuelled the search for a reductive account of modality in the empiricist tradition; and it would be a disappointment for many, I think, if the fictionalist were utterly powerless against it.

Considerations like these may lead the reduction-bent fictonalist to look for some more substantive account of his central notion—a procedure for analysing it away in terms of more basic non-modal notions. Unfortunately, I suspect the prospects are dim. As with all fictions, more is true according to *PW* than is given explicitly in the text—in this case, the postulates plus the encyclopaedia. And the problem is to specify the *consequence relation* which determines this larger class in a non-modal way.

I will simply report, however, that I do not know how to do this. We may count on firm intuitions about whether a sentence P is true according to the realist's hypothesis. But I know of no illuminating non-modal characterization of the relation between the hypothesis and P which these intuitions are meant to capture.

This may reflect nothing more than a lack of ingenuity on my part.[27] Still, we must face the fact that primitive modality may well be an ineliminable feature of the fictionalist's view. Whether one regards this as a serious flaw will depend very much on how one conceives the project we have been engaged in—that is, the project of producing an 'account' or 'analysis' of modal language that involves no commitment to possible worlds. If one seriously believes that, in the absence of a reductive account, the modal notions are somehow philosophically suspect, then fictionalism as I have presented it is not much of an advance. Alternatively, one might think (with Lewis) that although a reductive analysis of modality is a nice thing to have, it is best thought of as a virtue to be entered into the balance along with the others we have discussed. From this point of view, fictionalism's failure to provide for reduction may be outweighed by the utter plausibility of its ontology, by its compatibility with a broadly naturalistic epistemology, and so on. Finally, one may be simply uninterested in the reductive project. For many writers, a world-view purged of modality is no part of the interest of the possible worlds approach; and in this case the fictionalist's failure to ground a reduction may not matter at all.[28]

This is worth stressing. Various interests might be served by the sort of deflationism we have been discussing, and we have not kept them altogether separate. We began with the modest project of earning the right, as it were, to move back and forth between the modal idiom and the idiom of possible worlds without incurring egregious metaphysical commitments. This aim is separable from the realist's ambition to provide a reduction of the modal to the non-modal; and for all we have said, Ed's concerns would be perfectly well served by a deflationism that failed to

[27] For reason to think the problem intractable in principle, cf. *Pl.*, pp. 150 ff. The most obvious approach is to construe the relation as a logical one: i.e., to view the postulates and encyclopaedia (*PW*) as axioms and to define 'In *PW*, *P*' as *PW*⇒*P*'. Now given minimal abstract apparatus, various purely formal relations of logical consequence can of course be defined non-modally. But there is reason to think that ⇒ cannot be such a relation. First, the language of the theory is an unregimented fragment of English, for which no purely formal notion of consequence is available. But more importantly, even if we imagine the language regimented as, say, a first order language, a purely formal relation of consequence will not do justice to our intuitions, for the familiar reason that it will flout certain intuitively necessary but non-logical connections between terms in the language. Thus: necessarily, gold is a metal. And intuitively, according to *PW*, in all universes, gold is a metal. Yet the embedded sentence here is not one of the postulates, nor is it included in the encyclopaedia, since it involves quantification over other worlds. And what is highly implausible is that it should be a *purely formal* consequence of any regimentation of these.

[28] Cf. Stalnaker, op. cit.; and Kripke's remarks on possible worlds in *Naming and Necessity*, Cambridge MA, Harvard, 1980, esp. p. 19, n. 18.

provide such reduction. True, to the extent that the fictionalist's primitive operator is unclear, this is grounds for worry. Clarity is presumably a desideratum no matter what the project. But for Ed this may not seem at all conclusive. After all, the fact that we can engage in conversation about what's true according to the realist's hypothesis in order to test his theory against intuition is testimony to our understanding. We know how to use the prefix 'According to *PW*' tolerably well, and insofar as we do, we understand it. Whether this sort of understanding is enough to keep fictionalism in the running as a live option for someone like Ed is now a matter of somewhat delicate judgement.

9. *The argument from concern*[29]

I would like to close with a rather different challenge, one which invites further clarification of our project. The problem may be developed by analogy with Kripke's well-known objection to Lewis's counterpart theory. Kripke writes:

> The counterpart of something in another possible world is *never* identical with the thing itself. Thus if we say 'Humphrey might have won the election . . .' 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. Thus Lewis's view seems to me even more bizarre than the usual notions of transworld identification it replaces.[30]

Of course, there is a minor infelicity here. Lewis does not identify what might have happened to Humphrey with what *might* have happened to someone else. Rather, the truth conditions for claims about what might have happened to Humphrey refer to what *does* happen to others, namely, Humphrey's counterparts in other worlds. But this hardly affects the main point. Humphrey, we may suppose, cares a great deal that he might have defeated Nixon, but could not care less that someone rather like him—but still a complete stranger—did defeat someone rather like Nixon. One fact is a matter of immediate concern; the other, an occasion for indifference. Hence there is something paradoxical—or, as Kripke says, 'bizarre'—in the identification of the two.

Let us be clear about the nature of the objection. Humphrey apparently takes different attitudes towards the modal fact about himself and the non-modal fact about his counterpart; but this by itself is not a logical objection to the claim that the facts are identical. It is perfectly possible to have different attitudes towards a single thing under distinct descriptions. Lois

---

[29] Some comments from Simon Blackburn suggested what follows. Cf. his *Spreading the Word*, Oxford, Oxford University Press, 1984, pp. 212–16.

[30] *Naming and Necessity*, p. 45.

loves Superman; she is merely fond of Clark. This is possible because, even though Superman *is* Clark Kent, Lois has not figured this out. Similarly, it would seem possible for someone to care that P without caring that P^*, even when P and P^* pick out the same state of affairs, provided he does not believe that they do.

The objection is, I believe, more 'pragmatic' than logical. An 'account' of modality must answer to a range of phenomena. So far we have emphasized fidelity to our prior beliefs—fidelity to our beliefs about what might have been, to our views on the sources of modal knowledge, to our general conception of what there is, and so on. What Kripke has done is to point to a rather different but equally important phenomenon—our palpable *concern* for the modal facts—and to urge fidelity here as well. An adequate account will preserve the intelligibility of our caring about the modal truth as we do; of our responding emotionally, morally, and practically to our modal thoughts. The objection is that counterpart theory fails this constraint. We simply do not care about the fortunes of perfect strangers in the immediate way in which we care about our own modal properties. Hence it is implausible to identify the two sorts of fact.

In general, fidelity constraints may be seen as expressions of the methodological conservatism that guides our search for 'accounts' or 'analyses'. The requirement of fidelity to prior modal opinion, for example, may be seen as the requirement that anyone who takes the proposed analysis to heart should not thereby be forced to modify his corpus of modal belief in a radical way. Similarly, the constraint behind Kripke's objection may be viewed as the requirement that an adequate account must be acceptable without forcing a radical revision in our patterns of concern for the modal facts. To accept Lewis's counterpart theory, Kripke seems to be saying, would practically force the view that the modal facts do not much matter. Ordinarily, we have a vital moral and emotional concern for what might have happened to us. Yet to accept counterpart theory is to view these modal facts as no more vital than the day-to-day lives of strangers on another planet, that is, as matters of practical indifference. And this is a substantial revision.

Once the objection is put this way, there is an obvious—if not obviously correct—response. In taking counterpart theory to heart one might aim to revise, not one's interest in modality, but rather one's concern for certain distant strangers. Humphrey thinks he has no special reason to care about his counterparts. But when he accepts Lewis's analysis he will believe that the modal facts about him *just are* facts about his counterparts; and since he cares about the former, he will come to care about the latter.[31]

This is Lewis's preferred answer, I believe. And it would take us afield to evaluate it on its own terms. What needs stressing is that this is still a

---

[31] Or, perhaps, to insist that he has cared about the latter all along without knowing it.

revisionary stance. To respond in this way is to grant that to accept modal realism is to revise, not just one's ontology, but one's patterns of interest and concern. Now the requirement of fidelity to these practices is, like all such requirements, a defeasible one. A well-motivated analysis may actually give us grounds for changing our view of what is worth caring about. The question for modal realism—a question we need not answer— is whether we have been given adequate grounds for the revision the theory seems to force.

It should be easy to see how all this bears on the adequacy of fictionalism. Just as modal realism strains credulity by identifying facts of vital concern with apparently indifferent facts about distant simulacra, so fictionalism may be thought to strain credulity by identifying these same facts with facts about the content of an arcane story. Humphrey cares that he might have won. Perhaps he regrets some decision because he thinks that had he acted differently, he might have won. Now imagine that Humphrey comes to embrace the fictionalist's view of the content of his modal thoughts. He comes to believe that the fact that he might have won *just is* the fact that, according to the story *PW*, there is a world in which someone rather like him—his counterpart—does win. Could his pattern of concern conceivably survive this identification? Could he coherently feel *regret*—and this can be a powerful, crippling sentiment—at the thought that in some philosopher's story, someone resembling Humphrey won an election like the election Humphrey lost? Do not imagine that the story is taken as *evidence* for the distinct modal fact that Humphrey might have won. For the fictionalist, the fact about the story just is the modal fact; and a lucid fictionalist must have the same attitude towards both.

We begin, I suppose, thinking that what happens in fiction is not centrally important to us. Someone who does not care about what happens to Isabelle Archer may be a Philistine; but he need not be a moral or psychological monster. Indifference to the modal facts, however, is quite another thing. Consider the sentiment of regret, to take only a simple case. Regret is, in part, an emotional response to a modal thought. To regret having acted in a certain way, one must think that had one acted differently, things might have been better; and further, that one could have acted differently.[32] A sentiment is not regret unless such modal thoughts lie behind it. And if this is right, to be indifferent to the modal is to be incapable of regret.[33]

But regret is not a trivial thing. The character who feels no regret is a moral monster *par excellence*. That we think so is shown by the lengths we

---

[32] Note: this 'could' denotes, not metaphysical possibility, but rather the agent's freedom—still a modal notion, but one whose connection with the main topic of this paper is far from straightforward.

[33] Regret is a telling test case because here our concern for the modal is arguably non-derivative, in contrast, e.g., to our interest in counterfactuals like 'If I had slipped I would have broken my neck', which may well derive from an interest in what the counterfactual implies about the non-modal features of the actual world.

go to to inculcate a capacity for regret in children. Moreover, it seems to me that we do this, not simply for the cynical reason that someone with this capacity is easier to socialize, but because we also think that to feel regret when regret is called for is a sign of good character, worth having for its own sake.

If this is right—and I grant that it is only the barest sketch of an undoubtedly complex phenomenon—then a view which leads to indifference to the modal truth entails a revision in patterns of concern that is simply out of the question. Just as it is not a serious option for Ed to believe in a countless infinity of talking donkeys, it is not a serious option for any of us to abandon concern for the modal facts. If fictionalism really does require this, then taking fictionalism seriously is, if nothing else, a practical impossibility.

But does fictionalism require this? A response parallel to Lewis's is of course available. That is, the fictionalist may say that, although a revision is required, it is not the pernicious one just described. Rather, the lucid fictionalist may come to think that truth in the fiction *PW* is important to him in a way he had not previously imagined. The fact that things would have gone better if he had only acted differently *just is* the fact that according to *PW*, worlds most like the actual world in which his counterparts act differently are worlds where things go better. But this only means that he must care—perhaps desperately—about what happens to these characters in this fiction. True, he did not always care about such things. But now he does; and moreover, the transition was not that hard—not nearly as hard as the alternative, which is to become indifferent to the modal facts.

Now whether this transition is really possible for us is, I suppose, a psychological question; and I do not know the answer. But the objection may still be pressed. The fictionalist is now passionately concerned with the content of a certain fiction about alternative universes, namely, *PW*. Yet presumably he does not care about other similar fictions in the same way. Consider, for example, the story that says there are just two universes: ours and one populated only by blue swans. Just as modal truth is defined in terms of the fiction *PW*, we can define an analogous body of truth—the schmodal truth (?)—in terms of the new fiction. Thus we might say that *P* is *s-possible* iff, according to the two-word story, there is world at which *P*; and similarly for *s*-necessity, the *s*-counterfactual, and so on. Now schmodality is not modality. Yet for the lucid fictionalist, the modal facts and the schmodal facts are facts of the same kind: namely, facts about the content of a story about alternative universes. We may then ask the fictionalist: why do you care so much about the one and not at all about the other?

The question is designed to exploit what is plausibly a norm governing patterns of concern or interest. If you care especially for some members of

a class but not for others, you should be able to point to some feature which distinguishes the items you care about—a feature which *grounds* your difference in attitude. The worry is that the fictionalist cannot adhere to this norm in his concern for the modal and indifference to the schmodal. They are both facts about the content of certain false stories. And unless one of these stories has some *authority* which the other lacks, the fictionalist's concern may seem purely arbitrary—a *fetish* ungrounded in distinctive valuable features of its object.[34]

I know of no simple answer to this challenge. One response, I suppose, is to argue that the norm is too stringent—that by this criterion, all our concerns are fetishistic. It is chimerical to require that whenever we care about some things more than others we should be able to point to a feature that grounds this concern, since the question will always return: why do you care about things with that feature more than things which lack it? But if some of our legitimate concerns are in this sense ungrounded, it is no decisive objection that the lucid fictionalist's concern for the modal must fit this pattern. Obviously this line of response requires careful handling. I will not pursue it here except to say that if there is to be a practice of criticizing passions as arbitrary or vain, then it cannot depend crucially on this problematic model. So it remains possible that once we understand this critical practice better, the fictionalist's special concern for the fiction *PW* will still appear objectionably fetishistic.

Another response is to accept the norm and to look for some feature to distinguish the fiction *PW* from other fictions about possible worlds. This may seem hopeless. After all, what could possibly distinguish one bizarre science fiction scenario from another sufficiently to make one deserve the non-derivative concern we accord modality? Yet there may be an answer. Recall that in discussing the epistemology of modality I suggested in passing that the principles which guide the imagination when we construct possible states of affairs are in some sense well captured by the postulates of *PW*. I did not defend this claim—and it would be a subtle problem. But if there is something to it, then the fictionalist might try to argue that *PW* derives its 'authority' from being an explicit formulation of our own imaginative habits. So construed, fictionalism's affinities for 'conceptualist' theories of possibility is obvious. These theories aim to locate the source of modal distinctions in us, in our capacity to imagine or conceive alternatives to the actual state of things. Some of the problems of such approaches are well-known—and more would have to be said to show that fictionalism skirts them. For now I will only say that this strikes me as a potentially fruitful approach for the fictionalist bent on answering the challenge at hand.

One last response deserves mention.[35] Throughout I have supposed

---

[34] Here the objection parallels Blackburn's discussion of counterpart theory; Blackburn, op. cit.

[35] I owe the suggestion to Wm. Taschek.

that fictionalism, like modal realism, aims to be a *theory of possibility*, that is, an account of the truth conditions for modal statements, and hence of the facts that make modal statements true. But note that this assumption is not strictly necessary given the modest problem we began with. All Ed ever wanted was license to move back and forth between modal claims and claims about worlds. Such transitions are guided by biconditionals: for the realist, P iff P^*; for the fictionalist, P iff according to PW, P^*. But it is one thing to embrace these biconditionals—even to embrace them as a body of necessary truths—and another to regard them as providing analyses. The fictionalist who aims only to solve Ed's problem may therefore conceivably reject the stronger reading. He may claim to offer not a theory of possibility, but merely a theory linking the modal facts with facts about the story PW. The theory licenses transitions from one idiom to the other, without purporting to shed light on the nature of modal truth. This timid fictionalism of course raises as many questions as it answers. Still it must be granted that many of the objections we have mentioned, including the argument from concern, simply do not arise for this view. If the modal facts are distinct from facts about the fictionalist's fiction, there is nothing wrong with displaying divergent attitudes towards them.

10. *Conclusion*

Whether fictionalism is a viable deflationist alternative to modal realism depends very much on what one wants the theory for. Given the realist's interest in an account of the nature of modality, fictionalism has the advantage of being relatively modest in its ontology and sensible in its epistemology. It has the defect, however, of a primitive modal component (at least until a more ingenious exponent comes along) and perhaps of inviting what seems like an extraordinary revision in our patterns of concerns. The more modest one's aims, the more congenial fictionalism may seem. As a solution to Ed's dilemma, the timid fictionalism mentioned in the last section may in fact be wholly satisfactory. It is now up to theorists of modality who talk about possible worlds without believing in them to ask themselves whether, given their purposes, some form of fictionalism may not constitute an attractive option.

Philosophy Department   GIDEON ROSEN
University of Michigan
Ann Arbor, MI 48109
USA