

## DAVID LEWIS: *ON THE PLURALITY OF WORLDS*

Phillip Bricker

The notion of a possible world is familiar from Leibniz's philosophy, especially the idea—parodied by Voltaire in *Candide*—that the world we inhabit, the *actual* world, is the best of all possible worlds. But it was primarily in the latter half of the twentieth century that possible worlds became a mainstay of philosophical theorizing. In areas as diverse as philosophy of language, philosophy of science, epistemology, logic, ethics, and, of course, metaphysics itself, philosophers helped themselves to possible worlds in order to provide analyses of key concepts from their respective domains. David Lewis contributed analyses in all of these fields, most famously, perhaps, his possible worlds analysis of counterfactual conditionals (Lewis 1973). But these analyses invoking possible worlds cry out for a foundation: how is all this talk about possible worlds to be construed? Do possible worlds exist? If so, what is their nature?

David Lewis boldly responded: this talk of possible worlds is the literal truth. Lewis propounded a thesis of *modal realism*: the world we inhabit—the entire *cosmos*—of which we are a part—but one of a vast plurality of worlds, or *cosmoi*, all causally and spatiotemporally isolated from one another. Whatever *might* have happened in our world *does* happen in one or more of these merely possible worlds: there are worlds in which donkeys talk and pigs fly, donkeys and pigs no less “real” or “concrete” than actual donkeys and pigs. Moreover, whatever you *might* have done but didn't *is* done in another possible world by a *counterpart* of you, someone just like you up until shortly before the

time in question, but whose life diverges from yours thereafter. According to modal realism, the actual and the merely possible do not differ in their ontological status. They differ only in their relation to us: merely possible worlds are spatiotemporally and causally inaccessible; we can't get there from here.

When David Lewis first endorsed modal realism in the late '60's and early '70's, it elicited "incredulous stares" from other philosophers, even from other practitioners of possible worlds analyses. But by the early '80's, a spate of papers had been published in which those incredulous stares were backed by argument, and in which seemingly "more sensible" approaches to possible worlds were presented, approaches, for example, taking possible worlds to be "abstract objects" of some sort. *On the Plurality of Worlds* is Lewis's response: an extended elaboration and defense of modal realism. The greatness of this work lies not so much in its power to persuade—Lewis himself did not think the case for modal realism was, or could be, decisive—but in the masterful presentation of positions and arguments in the metaphysics of modality, and in the many problems in outlying areas of metaphysics that are clarified along the way. It is systematic philosophy at its finest.

\* \* \* \* \*

Why, according to Lewis, should one believe in a plurality of worlds? In an earlier work, Lewis based his argument on a Quinean criterion of ontological commitment applied to ordinary language (Lewis 1973: 84). We say, for example: "there are many ways things could have been besides the way they actually are." Taken at face value, this commits us

to entities called, “ways things could have been,” which Lewis identifies with possible worlds. But it was soon pointed out that the phrase ‘ways things could have been’ seems to refer, if at all, to abstract entities—perhaps uninstantiated properties—not to Lewis’s concrete worlds (Stalnaker 1976). In *On the Plurality of Worlds*, Lewis abandons any attempt to defend possible worlds by way of ordinary language, and turns instead to systematic philosophy. The chief concern of systematic philosophy is *total theory*, the whole of what we take to be true. Possible worlds, if accepted, provide the means to reduce the diversity of notions that must be taken as primitive, thereby improving the unity and economy of our total theory. Moreover, possible worlds, Lewis claims, provide a “paradise for philosophers” analogous to the way that *sets* have been said to provide a paradise for mathematicians (because, given the realm of sets, one has the wherewithal to provide true and adequate interpretations for all mathematical theories). So, when asked “why believe in a plurality of worlds?” Lewis responds: “because the hypothesis is serviceable, and that is a reason to think that it is true.” (p. 3)

Lewis does not claim, of course, that usefulness, by itself, is a decisive reason: there may be hidden costs to accepting possible worlds; there may be alternatives to possible worlds that provide the same benefits without the costs. Lewis’s defense of modal realism, therefore, involves an extensive cost-benefit analysis. His conclusion is that, on balance, modal realism defeats its rivals: rival theories that can provide the same benefits all have more serious costs. A controversial underlying assumption of Lewis’s argument, that a theory that better satisfies the pragmatic virtues such as simplicity and unity is *more likely* to be true, is noted by Lewis, but never called into question.

\* \* \* \* \*

*On the Plurality of Worlds* consists of four lengthy chapters each divided into multiple sections. Lewis devotes four sections of the first chapter, “A Philosopher’s Paradise,” to an extensive survey of the uses to which possible worlds have been put. This provides him with an opportunity to present, and sometimes clarify, his view on such diverse topics as supervenience, counterfactuals, the analysis of belief, semantics for natural language, and theories of properties and relations.<sup>1</sup> But the survey begins with the most famous application of possible worlds: the analysis of the alethic modal notions, *necessity* and *possibility*. That will be my focus here.

Consider the modal statement: necessarily, all swans are birds. This statement can be analyzed in terms of possible worlds as: at every possible world, all swans are birds.<sup>2</sup> The necessity operator becomes a *universal* quantifier over possible worlds. Moreover, quantifiers in the embedded proposition are restricted to the domain of the world of evaluation: *all swans are birds* is true at a world just in case all swans *inhabiting the world* are birds. (Since possible worlds, for the modal realist, are like places, truth at a world is analogous to truth in some place: *all swans are black* is true in Australia just in case all swans *inhabiting Australia* are black.) Now consider the modal statement: possibly, there are blue swans. That statement is analyzed as: at some possible world, there are blue swans. The possibility operator becomes an *existential* quantifier over possible worlds. The embedded proposition, *there are blue swans*, holds at a world just in case some swan *inhabiting the world* is blue. These analyses of necessity and possibility have genuine explanatory power: they elucidate the logical relations between

the modal notions. For example, if a proposition isn't possible, it necessarily isn't so; and if a proposition isn't necessary, it possibly isn't so. The quantificational analysis allows these modal inferences to be explained in terms of familiar logical inferences involving 'every', 'some', and 'not'.

Thus far, I have considered only the analysis of modality *de dicto*: the modal operators, *necessarily* and *possibly*, were applied to entire propositions. What about the analysis of modality *de re*, the application of modal properties to things? Consider, for example, the modal property, *being necessarily human*, which is formed by applying the modal operator, *necessarily*, to the property, *being human*. One might think that to say that George W. Bush is necessarily human is just to say: Bush is human at every world he inhabits, at every world containing him as one of its parts. But on Lewis's conception of possible worlds as non-overlapping concrete universes, that won't do: since Bush inhabits the actual world, he fails to inhabit any other possible world. The proposed analysis, then, would wrongly make all of Bush's actual properties necessary. Lewis's solution is to analyze modality *de re* in terms of what properties one's *counterparts* have at other possible worlds. He writes: "Your counterparts resemble you closely in content and context in important respects. They resemble you more closely than do other things at their worlds. But they are not really you. ... [They are who] you *would have been*, had the world been otherwise." (Lewis 1968) Then, *Bush is necessarily human* is analyzed by quantifying both over possible worlds and counterparts: at every possible world, every counterpart of Bush is human. Similarly, *Bush might have been a plumber* can be analyzed by existentially quantifying over possible worlds and counterparts: at some possible world, some counterpart of Bush is a plumber. On Lewis's account, modality *de*

*dicto* is the central notion, depending only on what possible worlds there are. Modality *de re* is derivative, and more fluid: it depends also on a counterpart relation which, being a relation of similarity, is open to subjective and contextual factors.

That modal realism allows one to analyze modality *de dicto* and *de re* is for Lewis one of its chief selling points. Over and over, Lewis objects to alternative accounts of possible worlds on the grounds that they must accept primitive modality in one form or another. But isn't the modal realist also committed to primitive modality by taking the notion of *possible* world (or *possible* individual) as basic? No, for Lewis the 'possible' is redundant: there are no *impossible* worlds (or individuals). Thus, Lewis has no need of primitive modality to divide the worlds (or individuals) into two classes: possible and impossible. (Henceforth I will often drop the 'possible', and speak simply of "worlds.")

What is wrong with primitive modality? Two things. First of all, Lewis thinks that an important factor in the evaluation of metaphysical theories is *economy*, both ontological and ideological. To accept primitive modality is to take on a serious ideological commitment and thereby offend against economy. Lewis concedes, however, that this reason isn't decisive: in this case, as in many others, there is a trade-off between primitive ideology and extravagant ontology, and philosophers may disagree as to where the greater cost lies. Second, primitive notions, even though unanalyzed, should nonetheless be understood. A theory that invokes primitives that are mysterious fails this test. But modality *is* mysterious. Modal properties do not fit easily into an empiricist worldview: one can observe that Bush is human, but not that he is necessarily human. Modal properties do not seem to stand alongside fundamental qualitative properties as part of the furniture of the world. Thus, modality cries out for explanation in non-modal

terms. Theories that take modality as primitive, then, will sacrifice much or all of the explanatory power of modal realism.

\* \* \* \* \*

In the last four sections of chapter 1, Lewis presents some of the tenets of modal realism in more detail under the headings “Isolation,” “Concreteness,” “Plenitude,” and “Actuality.” I will say something about each of these in turn.

According to modal realism, worlds (in general) are large composite objects. In the section “Isolation,” Lewis provides demarcation criteria for worlds in terms of the relations between their parts. He asks the question: what are the necessary and sufficient conditions for two individuals to be *worldmates*, to be part of one and the same world? His answer is this: individuals are worldmates if and only if they are spatiotemporally related, that is, if and only if every part of one stands in some distance-relation to every part of the other. A world is *unified*, then, by the spatiotemporal relations among its parts.

One direction of the analysis of the worldmate relation (sufficiency) is uncontroversial. Whatever stands at some spatial or temporal distance to us is part of our world; contrapositively, non-actual individuals stand at no spatial or temporal distance to us, or to anything actual. In general: every world is *spatiotemporally isolated* from every other world. (The worlds are also for Lewis *causally* isolated from one another, as follows from Lewis’s counterfactual analysis of causation.) According to the other direction of the analysis (necessity), worlds are unified *only* by spatiotemporal relations;

every part of a world is spatiotemporally related to every other part of that world. This direction is more problematic, for at least two reasons. First, couldn't there be worlds that are unified by relations that are not spatiotemporal? Indeed, it is controversial, even with respect to the actual world, whether entities in the quantum domain stand in anything like spatiotemporal relations to one another; the classic account of spacetime may simply break down.<sup>3</sup> Second, couldn't a single world be composed of disconnected spacetimes, so-called "island universes"? Indeed, couldn't there be a part of actuality spatiotemporally and causally isolated from the part we inhabit? Lewis must answer "no." When Lewis's analysis of the worldmate relation is combined with the standard analysis of possibility as truth at some world, island universes turn out to be impossible: at no world are there two disconnected spacetimes. This is potentially a problem for Lewis, because other fundamental metaphysical principles that Lewis accepts seem to entail that island universes are possible after all.<sup>4</sup>

In any case, the analysis of the worldmate relation in terms of spatiotemporal relations allows Lewis to then provide an analysis of the notion of world: a *world* is any maximal spatiotemporally interrelated individual— an individual all of whose parts are spatiotemporally related to one another, and not to anything else. If one assumes with Lewis that being spatiotemporally related is an equivalence relation (reflexive, symmetric, and transitive), it follows that each individual (that belongs to a world) belongs to exactly one world: the sum (or aggregate) of all those individuals that are spatiotemporally related to it. Note that the notion of world has been analyzed in non-modal terms— spatiotemporal relations, mereology, and logic— thus vindicating the modal realist's claim to eschew primitive modality.

\* \* \* \* \*

It is natural to characterize modal realism as Lewis himself sometimes does as the acceptance of a plurality of *concrete* worlds. This captures the idea that the merely possible worlds do not differ in ontological kind from the actual world, the concrete universe of which we are a part. But Lewis is hesitant to say outright that worlds are concrete because the distinction between concrete and abstract, as used by contemporary philosophers, is fraught with unclarity and ambiguity. In the section “Concreteness,” Lewis distinguishes four different ways of drawing the abstract/concrete distinction, and then queries how each applies to his notion of world. It turns out that, on all four ways (with some minor qualifications), worlds do indeed come out as “concrete” for Lewis.

(1) *The Way of Example*. Worlds (typically) have parts that are paradigmatically concrete, such as donkeys, and protons, and stars. (2) *The Way of Conflation*. Worlds are particulars, not universals; they are individuals, not sets. (3) *The Negative Way*. Worlds (typically) have parts that stand in spatiotemporal and causal relations to one another. (4) *The Way of Abstraction*. Worlds are fully determinate in all qualitative respects; they are not abstractions from anything else. But even if “worlds are concrete” comes out true on all ways of drawing the distinction, to say simply “worlds are concrete” is to say something very ambiguous. Perhaps, Lewis suggests, it would be better to drop the abstract/concrete terminology altogether, and to list directly, as above, the fundamental features of worlds.

\* \* \* \* \*

If possible worlds are to serve in an analysis of modality, there will have to be enough of them: for any way a world could possibly be, there will have to be a world that is that way. Otherwise, there will be “gaps in logical space,” the space whose “points” are all and only the worlds; there will be possibilities that lack worlds to represent them. In the section “Plenitude,” Lewis asks what general principles would be sufficient to guarantee that there exists an appropriate abundance of worlds. His discussion focuses on a *principle of recombination*, roughly: anything can coexist, or fail to coexist, with anything else. The principle naturally divides into two halves. According to the first half, any two (or more) things, possibly from different worlds, can be patched together in a single world in any arrangement permitted by shape and size. To illustrate: if there could be a unicorn, and there could be a dragon, then there could be a unicorn and a dragon side by side. How will this be interpreted in terms of worlds? Since worlds do not overlap, a unicorn from one world and a dragon from another cannot *themselves* exist side by side. The principle is to be interpreted in terms of *intrinsic duplicates*: at some world, a duplicate of the unicorn and a duplicate of the dragon exist side by side.

According to the second half of the principle of recombination, whenever two distinct things coexist at a world, there is another world at which one of them exists without the other. This half of the principle embodies the Humean denial of necessary connections between distinct existents. (‘Distinct’, in this context, means non-overlapping, rather than non-identical.) To illustrate: since a talking head exists contiguous to a living human body, there could exist an unattached talking head, separate

from any living body. More precisely: there is a world at which a duplicate of the talking head exists but at which no duplicate of the rest of the living body exists.

The principle of recombination allows one to infer, given the existence of *some* possible worlds, the existence of whatever *other* possible worlds can be obtained by “cutting and splicing.” The principle is clear in theory, but somewhat murky in application. For example, the principle presumably is behind our belief that talking donkeys and flying pigs are possible, but it’s hard to see how applying the principle to macroscopic objects will give this result: a flying pig isn’t just a pig with wings stuck on. The relevant recombination presumably takes place at the genetic, or even the atomic, level. But then our confidence that the principle yields flying pigs is hostage to our confidence that what it is to be a pig can be analyzed in terms of DNA sequences, or fundamental particles; and that seems wrong, since the question of analysis doesn’t appear to play a role in the modal belief.

In any case, it is clear that the principle of recombination, when applied to the actual world and its parts, is sufficient to guarantee the existence of a vast plurality of worlds. Might the principle by itself provide for an appropriate plenitude of worlds, sufficient to ensure that there are no “gaps in logical space”? No, a great many worlds will still be left out. Two sorts of additional principles of plenitude will be needed to guarantee their existence. First, if one starts with a world of three-dimensional objects and applies the principle of recombination, any world that results is still (at most) three-dimensional. But it seems possible that there be a world with four or more spatial dimensions. An additional principle will be needed, then, to guarantee that a plenitude of spatial (and spatiotemporal) structures is represented among the worlds.<sup>5</sup> Second, if one

applies the principle of recombination to *actual* objects instantiating *actual* properties, one never arrives at an object instantiating *alien* fundamental properties, fundamental properties nowhere instantiated at the actual world. But it seems possible for there to be more or different fundamental properties than there actually are. An additional principle will be needed, it seems, to guarantee a sufficient plenitude of fundamental properties and relations. Although Lewis gestures at the end of the section on plenitude towards the need for additional principles of these two sorts, he does not attempt to provide formulations.

\* \* \* \* \*

In the final section of chapter 1, “Actuality,” Lewis asks how the notion of actuality should be understood by a modal realist. Is actuality a fundamental, absolute property that I (and my worldmates) have but that my counterparts in other worlds lack? If so, it would have to be a rather special sort of property: it could not be an (intrinsic) *qualitative* property because actual individuals have non-actual qualitative duplicates. (Moreover, if actuality were a qualitative property always shared by duplicates, then the principle of recombination would require that there be worlds at which actual and non-actual things coexist, which is absurd.) But, even if we put to one side the mysterious nature of such an absolute property of actuality, a more serious problem looms: how, if actuality is absolute, could I know that I am actual? I have counterparts in other worlds that are epistemically situated exactly as I am; whatever evidence I have for believing that *I* am actual, they have exactly similar evidence for believing that *they* are actual. But

if no evidence distinguishes my predicament from theirs, then I don't really *know* that I am not in their predicament: for all I know, I am a merely possible person falsely believing myself to be actual. Thus, Lewis concludes, modal realism together with absolute actuality leads to skepticism about whether I am actual. Such skepticism, however, is absurd. A modal realist, then, should reject absolute actuality.

Lewis proposes instead that actuality is an indexical notion: when I say of something that it is actual, I am saying that it is a part of *this* world, the world that *I* (the speaker) inhabit. In other words, given Lewis's analysis of world, when I say of something that it is actual, I am saying simply that it is spatiotemporally related to me. On the indexical account, I know that I am actual as a trivial matter of meaning: I know, trivially, that I am part of the world I am part of. Knowing I am actual, then, is analogous to knowing I am here, which also is trivial, analytic knowledge: I know, trivially, that I am located where I (the speaker) is located. In neither case do I need to examine myself to discover that I have some special property—*being actual, being here*. Nor is my counterpart off in some other world deceived when he thinks to himself that he is actual. For although in my mouth, 'actual' applies to me and not to him, in his mouth, 'actual' applies to him and not to me (if he speaks English, and so means by 'actual' what I do). My counterpart is not deceived when he thinks to himself, "I am actual," any more than someone off in another country (or planet) is deceived when he thinks to himself, "I am here."

On the indexical theory of actuality, actuality becomes a relative matter: no world is absolutely actual; every world is actual relative to itself (and its inhabitants), and non-actual relative to any other world (and its inhabitants). Someone might object: taking

actuality to be relative in this way fails to take actuality with metaphysical seriousness. “Right,” Lewis would reply: a deflationary account of actuality goes hand in hand with the modal realist assertion that the merely possible worlds do not differ from the actual world in ontological status. And only a deflationary account can explain how we know that we are actual.

\* \* \* \* \*

In the eight sections of chapter 2, “Paradox in Paradise?”, Lewis considers eight objections to modal realism and provides a reply to each. I will discuss four of these objections, the four that are most familiar, and most fundamental.

The first objection, in the section “Everything is Actual,” is that what Lewis calls “possible worlds” are not properly called possible worlds at all: Lewisian worlds, if they exist, would be parts of actuality, not alternatives to it. Modal realism, then, correctly interpreted, posits a massively bloated actuality, not a realm of *possibilia* existing separate and distinct from the realm of the actual. But if this is the correct interpretation of what the modal realist believes, then the analysis of modality in terms of quantification over these so-called “possible worlds” cannot be correct: modal statements have to do with alternatives to actuality, not parts of it.

Lewis concedes that “if the other worlds would be just parts of actuality, modal realism is kaput.” (p. 112) But, on his indexical analysis of ‘actuality’, the other worlds, being spatiotemporally isolated from our world, are not properly called “actual.” So, the objection, if it is good, must be that the indexical analysis is incorrect. There are two

ways that the objection might be pressed. One might hold that ‘actual’, like ‘entity’, is a so-called “blanket term”: it is analytic that ‘actual’ applies to whatever exists, to anything in the realm of being. Lewis argues, however, that even if it is part of our common sense view that whatever exists is actual, it is not plausibly taken to be analytic: it is coherent to posit non-actual, merely possible objects. But there is a second, more powerful way to press the objection that the modal realist believes in a plurality of actual worlds, a way that Lewis doesn’t consider. Plausibly, it is analytic that ‘actual’ is a *categorical* term: anything ontologically of the same basic kind as something actual is itself actual. In other words: a merely possible object and an actual object, even if qualitative duplicates of one another, belong to distinct ontological kinds. If this is accepted, then Lewis’s claim that there are non-actual worlds ontologically on a par with the actual world would indeed be incoherent. The only way to believe in a plurality of non-actual possible worlds, then, would be to combine it with absolute actuality, to hold that there is an absolute distinction between the actual and the merely possible□□ in which case the daunting problem of skepticism about actuality would have to be faced anew.<sup>6</sup>

\* \* \* \* \*

A second, powerful objection to modal realism is epistemic. Modal realism holds that we have substantial modal knowledge, both specific, such as that talking donkeys are possible, and general, such as is embodied in the principle of recombination. On analysis, that knowledge turns out to be knowledge of the goings-on at other possible worlds. But, the objection goes, we can’t have such knowledge because the other worlds

are causally isolated from us, and knowledge of any subject matter requires that there be some sort of causal connection.

Lewis's response is twofold. First, he rejects the premise that causal acquaintance is necessary for knowledge of a subject matter. Here he invokes mathematics as a precedent: we have knowledge of mathematical entities—numbers, sets, etc.—even though such entities are “abstract,” and stand in no causal relations to anything. Lewis's response appears to presuppose Platonism, the view that a realist interpretation of mathematics is correct; no doubt those who object to modal realism on epistemological grounds would object no less to Platonism. But Lewis's goal is not to refute the objector; achieving a standoff will do. He is content to argue that modal realism is no worse off epistemologically than realism about mathematical entities.

However achieving even a standoff requires a subsidiary argument to the effect that the mathematical and modal cases really are analogous. Is that so? Mathematical entities are abstract; Lewis's worlds are concrete. Isn't that a relevant difference with respect to how these entities can be known? “No,” says Lewis. The distinction that matters for epistemology is that between the contingent and the necessary: knowledge of contingent truth requires causal contact with what is known; knowledge of necessary truth does not. Mathematics and modality may differ with respect to the “concreteness” of their subject matter; but with respect to what matters for epistemology, they are the same.

\* \* \* \* \*

A third objection is that modal realism leads to moral indifference. If modal realism were true, one could have no moral reason to choose an act that leads to good over an act that leads to evil because, whatever one chooses to do, the same total of good and evil will occur throughout all the worlds. Thus, Robert Adams asks: “What is wrong with actualizing evils since they will occur in other possible worlds anyway?” (Adams 1974) If one chooses the act with the good outcome, one has a counterpart no less real who chooses the act with the evil outcome; if one chooses the act with the evil outcome, one has a counterpart who chooses the act that leads to good. This objection, unlike the two previously considered, does not claim that modal realism is incoherent; the modal realist has the option to simply embrace the demise of morality. But Lewis is conservatively inclined: if modal realism would require that we revise in fundamental ways our conception of ourselves as moral agents, that would, Lewis agrees, provide strong reason to reject it.

Lewis concedes that modal realism makes trouble for at least one ethical theory: universalistic utilitarianism, the view that that act is morally best which maximizes the sum total of utility (happiness, welfare, etc.) for *everyone, everywhere*, with the ‘every’ unrestricted. But such an ethical theory is implausible on independent grounds: it conflicts with common sense attitudes towards morality in at least two ways. First, the good and evil that we care about is the good and evil that occurs to those who stand in some special relation to us: our family, our friends, perhaps our countryman, or our fellow Earthlings. Common sense morality is *agent centered*, not *agent neutral*. Morality, as commonly understood, does not prohibit us from restricting our moral concern to our worldmates, or some portion thereof. Second, even if one allowed that,

contrary to common sense morality, good and evil everywhere should count equally in our calculations, there would still be an adequate answer to Adams' question: "what is wrong with actualizing evils since they will occur in other possible worlds anyway?" To actualize evils, Lewis responds, is to be an evildoer, a causal source of evil. Thus, even if one's acts cannot change the total sum of good or evil throughout the worlds, there may still be a moral reason to choose one act over another. One ought to choose an act that makes one a causal source of good rather than evil; and this is in no way undermined by the existence of counterparts who choose instead to actualize evil. What *I* ought to do depends not only on the range of possible outcomes, but on my causal relation to the outcome that results in my world. In conclusion, then, if a modal realist accepts that morality is agent centered in one or both of these ways, modal realism will not threaten morality by leading to moral indifference.

\* \* \* \* \*

The final objection to modal realism is what Lewis calls "the incredulous stare." Simply put, it is that modal realism, with its talking donkeys and flying pigs no less "real" or "concrete" than actual donkeys and pigs, is too incredible to be believed. No matter how great the theoretical benefits of modal realism, no matter how successful it is in systematically unifying and simplifying our total theory, the cost of believing such an incredible theory will always be too great. Lewis accepts that modal realism disagrees severely with common sense, and he accepts that this is a serious cost to the theory. But it is not a prohibitive cost. Not unless there are alternative theories of possible worlds

that can achieve most or all of the benefits of modal realism without incurring serious costs of their own. This leads Lewis to the third chapter of the book: an examination of alternative theories of possible worlds.

\* \* \* \* \*

Lewis's third chapter, "Paradise on the Cheap?", is devoted to an elaboration and criticism of the various views he calls *ersatz modal realism*, views that provide abstract surrogates to play the role of Lewis's concrete possible worlds and individuals. According to ersatz modal realism ("ersatzism," for short), there is only one concrete world. But there are countless abstract entities—the ersatz worlds—that represent ways the one concrete world might have been. (Calling them "ersatz," of course, is to take the modal realist perspective; those who don't believe in Lewis's concrete worlds may think of the abstract entities as the real thing—perhaps properly called "possible worlds"—not as sham substitutes.) One of the ersatz worlds correctly represents the concrete world in complete detail: it is the *actualized* ersatz world. The other ersatz worlds all misrepresent the concrete world in some respect; they are all therefore *unactualized*. Similarly, ersatzism posits ersatz possible individuals that are actualized or unactualized depending on whether or not they accurately represent any concrete individual. On the ersatzist account—unlike modal realism—there is a distinction between being *actualized* and being *actual*. Typically, ersatzists are self-proclaimed *actualists*, and so they hold that the ersatz *possibilia*, actualized or not, are all actual. Perhaps they are "metaphysical actualists," holding that ersatz *possibilia* are actual because they are

abstract entities, and abstract entities are actual by nature; or perhaps they are “analytic actualists,” holding that ersatz *possibilia* are actual because it is analytic that *everything* is actual. Either way, the ersatzist seems to have an advantage over the modal realist in agreeing with common sense that whatever exists is actual.

More importantly, the ersatzist has an advantage over the modal realist in agreeing with common sense about the extent of concrete reality: there are no more concrete donkeys, for example, than we ordinarily think there are. True, the ersatzist believes in countless infinities of abstract representations of donkeys; but common sense does not have a firm opinion as to the extent of abstract reality, and so the positing of abstract *possibilia* does not offend common sense beliefs. At any rate, so says the ersatzist.

Now, if the ersatzist can supply a sufficient plenitude of ersatz worlds and individuals, then she can take over the analyses proffered by the modal realist. For example, she can say that it is possible that a donkey talk just in case some ersatz world represents that a donkey talks. And where the modal realist constructs entities out of concrete *possibilia* to play various theoretical roles—for example, to serve as meanings or properties—the ersatz modal realist can mimic that construction using ersatz worlds and individuals. It seems, then, that the ersatzist can have the benefits of Lewis’s *possibilia* without bearing the costs. But, Lewis argues, appearances here are deceiving.

Lewis believes there are severe costs to ersatz modal realism, but different versions have different costs. So he divides the various versions into three sorts depending upon how the ersatz worlds *represent* (or *misrepresent*) the one concrete world. According to *linguistic ersatzism*, ersatz worlds are like stories or theories, they are constructed from the words or sentences of some language (called the “worldmaking language”), and they

represent by virtue of the stipulated meanings of these words and sentences. According to *pictorial ersatzism*, ersatz worlds are like pictures or scale models, and they represent by isomorphism, by being as structurally and qualitatively similar to a concrete world as is compatible with their being abstract. According to *magical ersatzism*, the ersatz worlds represent in a primitive and inexplicable way; they represent what they do simply because it is their nature to do so. Lewis's exposition and criticism of these three views is lengthy and involved. In what follows, I provide a brief account of his arguments against linguistic and magical ersatzism. (I will not discuss pictorial ersatzism. Lewis's chief argument against it is that, when properly and fully developed, it collapses into a version of modal realism, and so cannot really provide the benefits sought by the ersatzist programme.)

\* \* \* \* \*

Linguistic ersatzism is based on the following natural idea: although concrete, non-actual worlds are hard to believe in, there is no problem believing in *sentences* purporting to describe such non-actual worlds. Perhaps, instead of concrete worlds with flying pigs as parts, we can make do with ersatz worlds constructed from sentences such as the sentence 'pigs fly'. There will be no mystery as to how such ersatz worlds represent that pigs fly: the representational properties of the ersatz worlds derive directly from the meanings of the sentences they contain.

Here is one way of carrying out the idea.<sup>7</sup> Collect together into a set all the sentences that would be true if some Lewisian world were actualized; this gives an

abstract surrogate for that world. Since any Lewisian world possibly exists, the set of sentences that would be true at the world is *consistent*, implies no contradiction. And since any Lewisian world is fully determinate, the set of sentences that would be true at the world is *maximal* consistent, containing for any sentence, either that sentence or its negation. Conversely, any maximal consistent set of sentences is an appropriate surrogate for some Lewisian concrete world, the world at which all its sentences are true. So, the linguistic ersatzist holds that the ersatz worlds are just the maximal consistent sets of sentences of an appropriate worldmaking language.

What language should be used to construct the ersatz worlds? A natural language, such as English, won't quite do because its words and sentences are often vague, ambiguous, and context-dependent. So, let's suppose that these unlovely features have been purged from the worldmaking language: every sentence of the language is determinately true or false at any possible world. A more serious problem is that English, so purged, may be descriptively impoverished even with respect to its power to describe individuals and properties at the actual world. A simple solution is to enrich the language by letting actual things be *names* of themselves, and properties be *predicates* that express themselves. What Lewis calls a *Lagadonian language*. (Of course, a Lagadonian language is rather inconvenient for its users; one has to display an object in order to talk about it. But what matters for the purposes of ersatzism is just that the sentences of the worldmaking language have clearly defined meanings in virtue of which the ersatz worlds represent.) Finally, we do not want the worldmaking language to be *logically* impoverished; so let's suppose that there are no logical limitations on the language's expressive powers, even if that means adding infinitary logical connectives to the

language. Now, the ersatz worlds will be the maximal consistent sentences of this descriptively and logically enriched worldmaking language.

Lewis has two main objections. The first is that the linguistic ersatzist needs primitive modality; for the notion of consistency used in the construction is a modal notion, not reducible to any non-modal notion of (formal) logical consistency, whether syntactically or model-theoretically defined. (For example, although ‘some bachelors are married’ is consistent in formal logic, it is not consistent in the sense relevant to the construction of ersatz worlds: there shouldn’t be any worlds at which some bachelors are married.) Lewis is aware, however, that many ersatzists will gladly help themselves to this much primitive modality if that is the full cost of a ticket to paradise.

The second objection is not so easily discounted by the ersatzist. It is that, even after the worldmaking language has been enriched in the ways discussed above, it will still lack the descriptive resources to provide enough worlds to match the worlds of the modal realist. The problem arises when one considers the possibility of alien individuals instantiating alien fundamental properties. It seems hard to deny that the world could have satisfied different physical laws involving different fundamental properties; otherwise many physical theories that turned out false would be wrongly classified as metaphysically impossible. Now, alien fundamental properties present no problem for the modal realist because the alien worlds at which they are instantiated do not need to be reduced to anything else: the modal realist simply posits that, since it is possible that there be alien fundamental properties, there are worlds at which such properties are instantiated. But the linguistic ersatzist has to construct these alien worlds out of his worldmaking language, a language that lacks any predicates for alien fundamental

properties. The Lagadonian strategy is of no avail here, because the ersatzist doesn't believe there exist any alien properties to serve as predicates expressing themselves: there is only the one concrete world, with its actual properties. And to simply stipulate that the worldmaking language contains predicates expressing alien properties, without there being any account of how a predicate manages to express one such property rather than another, is to move away from linguistic ersatzism, and allow that representation is primitive and irreducible – the view called magical ersatzism, to be discussed shortly.

The problem isn't that the linguistic ersatzist cannot construct an ersatz world at which it is true that alien fundamental properties are instantiated. For, if the worldmaking language has the resources to quantify over properties, the sentence, 'there exists a fundamental property not identical to ... [here list all actually instantiated fundamental properties]', will be a consistent sentence of the worldmaking language, and so will belong to some maximal consistent set of sentences, some ersatz world. But there will not be enough such ersatz worlds to match the alien worlds of the modal realist. That's because the modal realist will hold (on the basis of the principle of recombination) that for each description of an alien world that the ersatzist can supply, there are many concrete alien worlds differing from one another either by containing different alien properties, or by permuting the alien properties they contain. Linguistic ersatzism, then, conflates distinct alien possibilities; it provides only one ersatz world to substitute for many Lewisian alien worlds. And this will have a detrimental effect on the truth conditions for modal statements. For example, the ersatzist cannot provide worlds to validate the (intuitively) correct modal inference: it is possible in many ways that  $p$ ;

therefore, there are many possibilities in which  $p$ . In short: the linguistic ersatz worlds cannot provide all the theoretical benefits of the modal realist's concrete worlds.

\* \* \* \* \*

Linguistic ersatzism runs into trouble with alien possibilities because it attempts to construct its ersatz worlds out of entities confined to the one concrete world. Magical ersatzism avoids this problem by positing, rather than constructing, its ersatz worlds: they are primitive entities with primitive powers of representing the ways the concrete world might have been. Magical ersatzism, then, is a form of realism about possible worlds, but one in which the worlds are abstract—perhaps properties, or states of affairs—rather than concrete.<sup>8</sup> Since the magical ersatz worlds represent in a primitive way, their internal structure is irrelevant. So we might as well suppose that they have no internal structure; they are mereological simples. One of these simples—the actualized ersatz world—bears the relation, *represents (in complete detail)* to the concrete world; the other ersatz worlds misrepresent the concrete world in some way, and so do not bear this relation to the concrete world. The representation relation is fundamental and primitive, not reducible to anything else.

Lewis's chief objection to magical ersatzism is that such primitive representation involves an odious form of primitive modality; it requires that there be necessary connections that violate the principle of recombination. Recall that, according to that principle, distinct existents can coexist in any arrangement permitted by shape and size. Lewis primarily had in mind *spatiotemporal* arrangements, since, for Lewis,

spatiotemporal relations are the clearest example of fundamental external relations. But the principle should apply no less to whatever other fundamental external relations there may be. For the case at hand, we are concerned with how the concrete world and the ersatz worlds are “arranged” vis-à-vis the relation of representation. Consider the ersatz world that correctly represents the concrete world, and some other ersatz world that does not. Lewis asks: why couldn’t it go the other way, with the second ersatz world standing in the representation relation to the concrete world, rather than the first? Shouldn’t it be possible for the concrete world and the ersatz worlds to be differently “arranged”? If one holds, as the ersatzist must, that there is one way they are “arranged,” and that that arrangement is absolutely necessary, then one will be saddled with necessary connections between distinct existents—the ersatz worlds and the one concrete world. These connections are “magical,” in that it is beyond our ability to understand how or why they should occur. Moreover, this sort of primitive modality, according to Lewis, is somehow worse than the sort of primitive modality needed by the linguistic ersatzist: the distinction between those linguistic representations that are possible and those that are not.

Lewis’s argument against the magical ersatz worlds is sweeping in its scope. Primitive intensional entities, such as propositions, properties, and relations, will likewise be swept away, because all such entities stand in fundamental external relations—truth, instantiation—to concrete entities, relations that violate Lewis’s generalized principle of recombination. But perhaps the argument is too sweeping to be credible. For it seems that sets, too, with their relation of membership to concrete entities, will violate Lewis’s constraints. And Lewis does not suggest doing without sets in his ontology, lest

mathematics lack a foundation. Whether Lewis's argument can be restricted in some way so as apply to magical ersatz worlds and primitive intensional entities, but not to sets, remains an open question.<sup>9</sup>

\* \* \* \* \*

The fourth and final chapter, "Counterparts or Double Lives?", is devoted to the infamous "problem of transworld identity," and related matters.<sup>10</sup> The chapter includes a thorough defense and elaboration of Lewis's counterpart theoretic solution: the idea, introduced above, that modality *de re* is to be analyzed in terms of a counterpart relation based on qualitative similarity.

Is there a problem of transworld identity? In one sense, the answer is "no": modal realists and ersatzists alike agree that one and the same concrete individual can truly be said to *exist at* more than one possible world, where an individual exists at a world just if the world represents *de re* that that individual exists. But the modal realist and the ersatzist will give different accounts of representation *de re*. For the modal realist, there are two ways for a (concrete) individual to exist at a world: one way is to be a part of the world; another way is to have a counterpart as a part of the world. Thus, a (concrete) individual can exist at more than one world without being a part of more than one world, without allowing that worlds overlap. Ersatzists, too, will need to give an account of representation *de re*. Although different ersatzists will give different accounts, on no ersatzist account will the concrete individual exist at an abstract ersatz world by being a

part of it. So the ersatzist, no less than the modal realist, rejects transworld identity in the literal sense of being a part of more than one world.

An ersatzist could choose to be a counterpart theorist, taking abstract ersatz *possibilia* to be counterparts of actual, concrete individuals. But most prominent ersatzists have argued that counterpart theory provides unacceptable truth conditions for *de re* modal statements. For example, Saul Kripke famously complained that, according to counterpart theory: "... if we say '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." (Kripke 1980: 45.)

Kripke's objection naturally falls into two parts. The first part is that, on the analysis of modality *de re* provided by counterpart theory, the modal property, *might have won the election*, is attributed to Humphrey's *counterpart* rather than to Humphrey himself. But surely, the objection continues, when we say that "Humphrey might have won," we mean to say something about *Humphrey*. This part of the objection, however, is easily answered. According to counterpart theory, Humphrey himself has the modal property, *might have won the election*, in virtue of his counterpart having the (non-modal) property, *won the election*. Moreover, that Humphrey has a winning counterpart is a matter of the qualitative character of Humphrey and his surroundings; so on the counterpart theoretic analysis, the modal statement is indeed a claim about Humphrey.

The second part of Kripke's objection is more troublesome. We have a strong intuition, not only that the modal statement, "Humphrey might have won the election," is

about Humphrey, but that it is *only* about Humphrey (and his surroundings). On counterpart theory, however, the modal statement is also about a merely possible person in some merely possible world; and that, Kripke might say, is simply not what we take the modal statement to mean. The first thing to say in response is that the charge of unintuitiveness would apply equally to the ersatzist's use of abstract ersatz worlds to provide truth conditions for modal statements; for our intuitive understanding of modal statements such as "Humphrey might have won the election" does not *seem* to invoke abstract worlds any more than counterparts of Humphrey. The objection, then, if it is good, would seem to cut equally against modal realism and ersatzism, and favor an antirealist view that rejected worlds, real or ersatz. But is the objection good? Should our pre-theoretic intuitions as to what our statements are and are not *about* carry much, or even any, weight? I think not. A philosophical analysis of our ordinary modal statements must assign the right truth values and validate the right inferences; but requiring more would fatally hamper philosophical attempts to attain theoretical systematization.

\* \* \* \* \*

The four remaining sections of chapter 4 address four important questions about how representation *de re* works. I have space here only to state the questions and indicate the gist of Lewis's responses. The first two questions pertain just to modal realism. The first is: if the modal realist were to allow overlapping worlds, were to allow a concrete individual to be part of more than one concrete world, could counterpart theory be

avoided? Could a modal realist say that the properties an individual has at any world are just the properties it has *simpliciter*, rather than the properties had by a counterpart inhabiting that world? (If so, representation *de re* would work by transworld identity in the literal sense.) “No,” replies Lewis. An individual’s intrinsic properties are sometimes accidental, in which case it has different intrinsic properties at different worlds. And that is impossible if representation *de re* works by literal transworld identity.

The second question is: could the concrete individuals we ordinarily refer to—people, and puddles, and protons—be *transworld sums*, partly in one world, partly in another? Then, the modal realist could say that an individual has a property at a world just if the part of it that is wholly contained in the world has the property *simpliciter*. Lewis rejects such transworld sums, not because he thinks they don’t exist—he puts no restriction on mereological composition—but because he thinks on semantic grounds that our ordinary names and descriptions do not refer to them. And, in any case, taking ordinary objects to be transworld sums would do nothing to satisfy the intuitions that seem to support transworld identity over counterpart theory.

The final two questions are for modal realists and ersatzists alike. The third question is whether representation *de re* is determined entirely by the qualitative nature of worlds, or whether instead there could be two worlds qualitatively alike that nonetheless differed as to what they represented *de re* of some individual? The *haecceitist* holds that representation *de re* is not qualitatively determined, and seems to have strong intuitions on her side. But Lewis argues that these haecceitist intuitions are better accommodated

in another way: by allowing that distinct possibilities may sometimes be realized within a single possible world.

Finally, Lewis asks: is representation *de re* a constant matter, fixed once and for all? Or does it vary with context, and sometimes have no determinate answer at all? Lewis argues for the latter approach according to which questions of essence and accident do not have the absolute metaphysical significance often attributed to them, but instead often shift with the wind.

\* \* \* \* \*

Since the publication of *On the Plurality of Worlds* in 1986, scores of articles have been published in philosophical journals responding to Lewis's arguments. I think it safe to say that modal realism is viewed today as more defensible, and ersatz modal realism as more problematic, than previously had been the case.<sup>11</sup> Still, only a small minority of philosophers are willing to give modal realism (or one of its close variations) unqualified support. Indeed, most philosophers, if driven away from ersatzism by Lewis's arguments, find themselves pushed not towards modal realism, but towards some anti-realist approach, an approach that rejects both concrete and abstract *possibilia*.<sup>12</sup> Perhaps Lewis's "paradise for philosophers" is simply not to be had, a will-o'-the-wisp. Perhaps belief in a plurality of concrete worlds is just too farfetched, supporting arguments notwithstanding. Perhaps. I invite the reader to engage with Lewis's compelling book in order to judge for herself.<sup>13</sup>

## ENDNOTES

<sup>1</sup> For a discussion of Lewis's views on these and other topics with modal underpinnings, see Nolan (2005).

<sup>2</sup> Here and below, I follow Lewis in speaking of what is the case *at* a possible world, rather than *in* a possible world. This usage grew out of "indexical semantics," where possible worlds and times are treated analogously at the formal level.

<sup>3</sup> Perhaps Lewis's introduction of analogical spatiotemporal relations (pp. 75-6) goes partway to answering this objection; but it doesn't seem to go far enough. For discussion, see Bricker (1996).

<sup>4</sup> See Bricker (2001: 35-7).

<sup>5</sup> See Bricker (1991) for a comparison of various formulations of such a principle.

<sup>6</sup> In Bricker (forthcoming), I argue that a believer in (concrete) possible worlds can combine indexicality of the *concept* of actuality with absoluteness of the *property* of actuality, and thereby evade the skeptical problem.

<sup>7</sup> For a rather different approach not addressed by Lewis, see Sider (2002).

<sup>8</sup> See especially Stalnaker (1976) and Plantinga (1974). Stalnaker calls his view "moderate realism" to distinguish it from Lewis's view, which he calls "extreme realism."

<sup>9</sup> Van Inwagen (1986) notes that Lewis's argument against magical ersatzism applies, with minor adjustments, to Lewis's own acceptance of sets. Lewis (1991: 35-8) contains a brief response.

<sup>10</sup> See Chisholm (1967) for an early, classic statement of the problem.

<sup>11</sup> For example, Divers (2002: xii), after an extensive review of the relevant literature, concludes: “In sum, I have come to think that the objections against [modal] realism, even taken collectively, are not convincing ... I here take [modal] realism to be more credible than [ersatzism] and I think that [modal] realism may be credible *tout court*.”

<sup>12</sup> Anti-realist approaches to modality were given short shrift in Lewis’s book. The two most prominent approaches are *fictionalism* (Armstrong (1989), Rosen (1990)), and *modalism* (Fine (1977), Forbes (1986)).

<sup>13</sup> Thanks to Jake Bridge and the editor, John Shand, for helpful comments.

## BIBLIOGRAPHY

*Suggested Reading.* For further background in the metaphysics of modality, I recommend Melia (2003) and Divers (2002). Melia (2003) is an introductory survey of approaches to modality; Divers (2002) provides a more comprehensive treatment that focuses on the debate between modal realism and ersatzism. The anthology Loux (1979) contains many of the classic articles to which Lewis is responding in his book. The text Giraldo (2000) provides an introduction to possible-worlds semantics for modal logic.

Adams, Robert. (1974) "Theories of Actuality," *Nous*, vol. 8, pp. 211-31. Reprinted in Loux (1979: 190-209).

Armstrong, David. (1989) *A Combinatorial Theory of Possibility*, Cambridge: Cambridge University Press.

Bricker, Phillip. (1991) "Plenitude of Possible Structures," *Journal of Philosophy*, vol. 88, pp. 107-19.

Bricker, Phillip. (1996) "Isolation and Unification: The Realist Analysis of Possible Worlds," *Philosophical Studies*, vol. 84, pp. 225-38.

Bricker, Phillip. (2001) "Island Universes and the Analysis of Modality." In *Reality and Humean Supervenience: Essays on the Philosophy of David Lewis*, Gerhard Preyer and Frank Siebelt (eds.), pp. 27-55. Lanham, MD: Rowman and Littlefield.

Bricker, Phillip. (forthcoming) "Absolute Actuality and the Plurality of Worlds."

Chisholm, Roderick. (1967) "Identity Through Possible Worlds: Some Questions," *Nous*, vol. 1, pp. 1-8. Reprinted in Loux (1979: 80-7).

- Divers, John. (2002) *Possible Worlds*, London: Routledge.
- Fine, Kit. (1977). "Postscript: Prior on the Construction of Possible Worlds and Instants." In *Worlds, Times, and Selves*, Arthur Prior and Kit Fine, pp. 116-61. London: Duckworth.
- Girle, Rod. (2000) *Modal Logics and Philosophy*, Chesham: Acumen.
- Kripke, Saul. (1980) *Naming and Necessity*, Cambridge, MA.: Harvard University Press.
- Lewis, David. (1968) "Counterpart Theory and Quantified Modal Logic," *Journal of Philosophy*, vol. 65, pp. 113-26. Reprinted in Loux (1979: 110-28).
- Lewis, David. (1973) *Counterfactuals*, Cambridge, MA: Harvard University Press.
- Lewis, David. (1991) *Parts of Classes*, Oxford: Blackwell.
- Loux, Michael (ed.) (1979) *The Possible and the Actual: Readings in the Metaphysics of Modality*, Ithaca, NY: Cornell University Press.
- Melia, Joseph. (2003) *Modality*, Chesham: Acumen.
- Nolan, Daniel. (2005) *David Lewis*, Chesham: Acumen.
- Plantinga, Alvin. (1974) *The Nature of Necessity*, Oxford: Clarendon.
- Rosen, Gideon. (1990) "Modal Fictionalism," *Mind*, vol. 99, pp. 327-54.
- Sider, Theodore. (2002) "The Ersatz Pluriverse," *Journal of Philosophy*, vol. 99, pp. 279-315.
- Stalnaker, Robert. (1976) "Possible Worlds," *Nous*, vol. 10, pp. 65-75. Reprinted in Loux (1979: 225-34).
- Van Inwagen, Peter. (1986) "Two Concepts of Possible Worlds," *Midwest Studies in Philosophy*, vol. 11, pp. 185-213.

