ACTUALITY, POSSIBILITY, AND WORLDS

ALEXANDER R. PRUSS

Actuality, Possibility, and Worlds

Actuality, Possibility, and Worlds

Alexander Robert Pruss

For my father,
who taught me also the value of philosophy

Contents

Preface

I would like to thank Robert Brandom, James Conant, Troy Cross, Trent Dougherty, James Dreier, John Earman, Gregory Fitch, Richard Gale, Jeremy Heis, Jonathan Jacobs, Robert Koons, Steve Kuhn, John Leslie, Mark Lance, Kenneth Manders, David Manley, William McEnaney, John Norton, Tim Pawl, Alvin Plantinga, Nicholas Rescher, Joshua Rasmussen, Ernest Sosa, Donald Turner, Peter van Inwagen, and the anonymous referees of various papers for encouragement, interesting discussions, and/or comments on various parts of this project. I would like to thank the National Endowment for the Humanities, Georgetown University and Baylor University for support of this work. I particularly would like to thank Daniel Johnson for a careful reading and helpful comments; all remaining errors and infelicities are fully my own responsibility.

And above all, I would like to thank my wife and parents for all their loving encouragement and patience.

There are several paths through this book. All of them start with Part I for background and terminology. Readers already familiar with applications of possible worlds can skip Part II. Parts III and IV present and criticize the best current theories of modality and possible worlds. Readers only interested in this critical analysis can stop reading after Part IV, while readers who already take themselves to have reasons to reject the best current theories might simply start reading with a skim of Part V and then a detailed reading of Part VI, since Part VI develops the positive proposal of this book, while Part V develops a view that while inadequate in its full version will be taken up in a moderate form in response to some of the objections in Part VI.

PART I

INTRODUCTION

Section 1 *Generic definitions and basic modal realism*

Modal assertions involving possibility and necessity are not only a part of our ordinary languages, but also a part of our philosophical patrimony. There are many things we could not say if we confined ourselves to non-modal language. We could not mark the difference between a unicorn,[1] which *could* exist, and a square circle, which *could not*. Modality is a natural way of marking the difference between, on the one hand, the relation of Smith being a bachelor to Smith being unmarried, and, on the other hand, the relation of Smith being 50 feet tall to Smith not being a mammal. Someone *could not* fail to be unmarried if he is a bachelor, but he could be a mammal even if he were 50 feet tall — though *in fact* no mammal is that tall.

It is important for ethical purposes to say what *could* have been done but was left undone, and what would have happened *had it* been done. It is plausible that a human being can only be held responsible for an act if it was at least logically possible that he avoid it. When we say that moral worth supervenes on actions and non-moral circumstances, we are saying that it *could not* be the case that someone's moral worth was different though his actions and the non-moral circumstances were the same.

When we discuss the problem of evil, we sometimes wonder whether it is *possible* for God and evil to co-exist, a different problem from the *de facto* question of whether the evils of this world make the existence of God probable or not.

---

1 Following Lewis's (1986a) practice, by a "unicorn" I shall simply mean an animal that looks much like the "unicorns" of our stories; I shall also stipulate that for ten generations, the animal was born to an animal that looked like the "unicorn" of our stories, just in case somebody tells me that at some country fair she attended there was a "unicorn" made by affixing a horn to a horse. I do not mean any specific natural kind of such animals, and hence avoid Kripkean objections.

When we talk of natural objects, we often cannot specify the kind that the object falls into without talking of dispositional properties. Something might in fact live all its life just like a horse, but if it is true that *were* it poked in the underbelly, where in fact it never was poked, it *would* suddenly and naturally sprout wings and fly away, then it is not a horse.

Our expressive capabilities would be greatly impoverished without "can be," "might be," "must be," "is possible," "is necessary," "would be," and their ilk. We need these terms to talk of the reality around us. Yet, paradoxically, talk involving possibility often does not appear to be *about* anything real. The unicorn that is possible does not exist, I have not done otherwise than I have, and the actions and non-moral circumstances are only as they are.

A popular and vivid way to organize one's modal notions is to think of a "possible world," a way (with "way" understood so broadly as not to prejudice the ontological question of *what* possible worlds are) that a cosmos could have been. Different possible worlds are different ways that our world could have been.

The main alternative to thinking of modality in this global sense is thinking of it in a local sense, of thinking of alternative ways that *portions* of this world could have been. It does appear that such piecemeal modality is what ordinary language users are often interested in. When we say that Hitler might never have been born, we do not seem to mean that there is some possible world in which he doesn't exist — e.g. a world at which the universe has always had an unchanging constant energy density. We mean that that portion of this world which consist of the birth of Hitler might not have been, even though much of the rest of the world, especially at least the distant past prior to Hitler's birth, was almost the same, and the laws of nature were those that we have. However, what exactly is to be kept fixed in this "might never have been born" claim depends on the context. Thus, while apparently speaking only of portions of worlds, the context determines what *whole* worlds we are speaking of, namely what portions of the actual world are supposed to be imagined as remaining in that possible scenario in which Hitler had never been born. Hence, to disambiguate our ordinary piecemeal talk of possibility, we bring in whole possible worlds.

The need to talk of whole worlds is shown particularly clearly when we make counterfactual utterances. For we can ask questions like: "How might or would have the course of history gone had Hitler never been born?" And on a plausible account of how to think about such questions, we should think of *whole* worlds in which Hitler was not born, and to say what holds in such worlds. Given what our context fixes, namely most events prior to Hitler's birth and the laws of nature, we can easily say certain things about

what happens in those worlds at present (in our reference frame). For instance, the course of events in other galaxies up to the present would be the same as in the actual world whether the awful events of the twentieth century occurred on earth or not, if only because the information about these events, traveling at the speed of light, has not yet arrived there. But the arrangement of matter would be slightly to significantly different within 120 light-years of earth. On earth, the arrangement of matter would be significantly different, while on Pluto there would only be minor differences, for instance due to the tiny variations in gravitational force (even an infant Hitler exerts a gravitational force, and the tanks of the Third Reich exert a greater one, so that had Hitler not been born, Pluto would have been slightly differently located). Our comfort in saying in the same breath that events in other galaxies would have been the same, but events on earth here would have been different, does indicate that it is appropriate to analyze counterfactual situations holistically.

Moreover, what is possible in a portion of the world may well depend on global features of the world, such as laws of nature. It is impossible for there to be a world with exceptionless laws of nature like ours but where things do not fall when dropped under appropriate conditions; however, apart from such laws, it is certainly possible. And it is arguably impossible that there be unjustified evil, i.e. evil the permitting of which serves no moral purpose in any portion of the universe if there is an all-powerful, all-knowing and all-good being anywhere in space (if this is possible for such a being) or outside of space. Moreover, in worlds where there is such a deity, what evils can exist in a portion of the world may well depend on what happens elsewhere in the world, since the justification of some evil in one portion of the world can depend on events elsewhere. Our ordinary modal claims need to be contextually disambiguated, and when thus disambiguated are seen to involve whole possible worlds. Because of all this, possibility and necessity prima facie require reference to be made to whole possible worlds, and so one should try to make sense of possible worlds.

Given a basic notion of possible worlds, whatever their ontology, we need some correlative notions. By "the (or our) cosmos" I shall mean the aggregate of all actually existing things. If one is fine with arbitrary mereological sums, one can take that aggregate to be a mereological sum. By "the (or our) universe" I shall mean the aggregate of all actually existing spatio-temporal things. If one is worried about the existence of this mereological sum, one should be able to paraphrase talk of the cosmos, universe, and other aggregates, into plurally quantified talk.

Each world corresponds to or represents a way the cosmos could have

been. In what way this representation works is one of the central questions for our investigation. One of the worlds shall be distinguished as "the actual world," i.e. the world that represents the way *our* cosmos in fact, or actually, is. An individual "exists in" a world *w* if, were that world actual, that individual would exist, or, equivalently, if *w* represents the cosmos as containing that individual. A proposition is "true at" a world *w* if, were that world actual, that proposition would be true, or, equivalently, if *w* represents the cosmos as described by that proposition.

What the notions of "represents," "actual," "exists in," and "true at" really signify will depend on what our ontology of possible worlds is. There are many possible such ontologies. In particular, there is the crazy one, which nonetheless will be conceptually useful at times to keep in mind, that there necessarily is an infinite Platonic library somewhere that contains physical books, each of which gives a maximal consistent description of a cosmos in some fixed language. On this view, a world is one of these books. A world *represents* a given way of being a cosmos if the book that the world is correctly describes the way that that cosmos would be. A world is *actual* if everything written in it is true. A proposition is *true at* a world if it is expressed by some sentence in the book. An individual *exists in* a world if the world describes the individual as existing.

Other theories will have other renderings of the basic notions. For instance, David Lewis (1986a) thinks that each possible way that the universe could be is a way that some concretely existing universe really is. Moreover, cosmoi and universes are the same for him. Thus, worlds are concrete universes. A world represents some cosmos if it *is* that cosmos. The *actual world* is the world we inhabit. A proposition is *true at* a world if it truly describes a state of affairs obtaining in that world. An individual *exists in* a world if it inhabits that world.

A Propositional Ersatzist may take a world to be a maximal collection of compossible propositions. The *actual world* is that particular collection all of whose propositions are true. A world corresponds to a cosmos by having as its members propositions true of that cosmos. A proposition is *true at* a world if it is a member of it. An individual *exists in* a world if some proposition in that world says that the individual exists.

Leibniz, on the other hand, thinks that worlds are maximally consistent ideas in the mind of God. The actual world is the idea that God has chosen to actualize. An idea corresponds to a universe by being a mental representation of it. A proposition is *true at* a world if it is a part of, or maybe represented by, that world. An individual *exists in* a world if the idea represents him as existing.

We can now give a possible worlds semantics for possibility and necessity claims. It is (metaphysically) possible that p providing there is a world w at which p is true. It is (metaphysically) necessary that p providing p is true at every world. This yields the standard duality between possibility and necessity: p is possible if and only if $\sim p$ is not necessary. Possible worlds now let us consider "local" and "global" modalities in a uniform way. When I say "Hitler might not have existed" in an ordinary way, and not by way of stating a merely metaphysical possibility, I am saying that the proposition, that Hitler does not exist is true at some world that matches ours in various relevant respects. When I say "It is logically possible that unicorns exist," I may just be making the claim that the proposition that unicorns exist is true at some world, without putting any restriction on which worlds are relevant here.

Some further terms are useful. A proposition is *contingent* providing it is true at some but not all worlds, i.e. providing that neither the proposition nor its negation is a necessary truth. An individual x is a *necessary being* if it exists in all worlds. An individual is a *contingent being* if it exists at some but not all worlds. Occasionally, I shall use $\square$ and $\lozenge$ to indicate necessity and possibility, respectively.

But of course all of this raises two basic ontological questions. The first is the grounding problem. What ontological features of reality make the right modal assertion be *true*? The second is the worlds problem. Possible worlds are very useful, but what are they? This book is devoted to these two problems.

I will begin with a clarification of the kind of modality that we are after: metaphysical modality rather than strictly logical modality. For we are interested in the kind of modality that is objective and independent of the vicissitudes of our language, and as shall be argued, this kind of modality is metaphysical modality. Next, I will sketch several different approaches to the grounding and/or worlds problems. Doing this should help clarify what exactly is at issue in the two problems.

Since the main argument for possible worlds is from their theoretical utility, in Part II we will examine a number of uses that are made of possible worlds. Some applications fail and others succeed. Overall, we will see that there is some reason to accept possible worlds. Starting with Part III, we will examine in detail four proposals for dealing with the grounding problem. Of these proposals, all but the Spinozistic–Tractarian one will offer solutions to the worlds problem. First I examine the two most promising contemporary approaches, the Lewisian approach that claims all possible worlds to exist as concrete, physical universes (Part III) and the various "ersatzist" approaches

on which possible worlds are "ersatz" linguistic or Platonic entities rather than concrete universes (Part IV). The Lewisian account will be seen to fail, both because it does not actually provide the reductive analysis of modality it promises and because it gives rise to multiple paradoxes. Of the ersatzist approaches, some, such as the Platonic, will be seen as at least offering a useful account of possible worlds, but none will give a satisfactory answer to the grounding problem. The next approach (Part V) is inspired by Spinoza's account of knowledge and by Wittgenstein's *Tractatus*, and makes the radical claim that there are no impossible propositions. This view, though on the face of it most implausible, is actually more attractive than might otherwise seem. Although I will argue that this approach does not in the end succeed as a solution to the grounding problem, nonetheless when I finally defend the Aristotelian account of modality in Part VI, a version of the Spinozistic–Tractarian approach will be seen as complementary to it.

In the end, an Aristotelian approach will provide a causal account of modality inspired by Aristotle's analysis of change. This approach is capable of giving a satisfactory solution to the grounding problem, if we are either willing to surrender intuitions such as that the *whole* history of the world could have been different or accept that there is a First Cause outside of time. However, the Aristotelian approach does not yield possible worlds, unless supplemented either in a Platonic or Leibnizian theistic way. The final view that will be defended will be the supplemented Aristotelian–Leibnizian view.

Section 2 Metaphysical versus logical possibility?

The modality in connection with which the possible worlds are possible is what is often called "metaphysical" or "broadly logical" possibility, with the paradigmatic example being that if Kripke (1980) is right, then it is metaphysically impossible that water fail to be H_2O. It is commonly thought that there are in fact two *different* kinds of modality. Some propositions, such as <H_2O contains hydrogen atoms> (where "<*s*>" denotes the proposition *that s*) are *logically* necessary since it is logically necessary that anything that has two atoms of hydrogen and one atom of oxygen in each molecule (and that, after all, is the *definition* of "H_2O") contains hydrogen atoms. But <Water *is* H_2O> has a different kind of necessity, since it is not one that follows from the logic of the terms involved.

Consider the claims:

(1) H_2O contains hydrogen atoms
(2)   Water contains hydrogen atoms.

The defender of a distinction between logical and metaphysical possibilities claims that (1) and (2) have different modal status.

As an opening gambit, one can reply that they cannot have different modal status, because modal status belongs to propositions, not to sentences, and (1) and (2) express the same proposition, and hence have the same modal status by Leibniz's law. The defender of the distinction between necessities can either deny that (1) and (2) express the same proposition, or claim that they differ in modal status as *sentences*. The latter claim I have no need to dispute, since I can simply confine my account to that of the modal status of propositions.

But in fact claiming a difference in the modal status of the two sentences is dubious when there is no difference in the two propositions or in their modal status. What could one mean by claiming a difference in the modal status of the *sentences*? That sentence (2) could have expressed a false proposition? Yes, doubtless, but so could (1): after all, it might have been uttered in a language where H_2O means "two electrons and one photon." Or does one mean that (2) might have been true in *our* language? But it could not: the language spoken on Twin-Earth in which "water" means XYZ is not English — it is not *our* language. Of course, one could say, along with two-dimensionalists (e.g. Chalmers 2006), that there is a world where (2) in a language relevantly like ours (e.g. with the same internal structure and the same connections to immediate awareness) expresses a false proposition, but there is no world where (1) expresses a false proposition in a language relevantly like ours. But while that is indeed true, and does state a fact about variation between languages that may be interesting, it is not clear that it really does state a modal difference between the two sentences. For it seems plausible that *sentences*, whether considered as types or as tokens, are individuated in such a way that physically indistinguishable inscriptions or sounds in different languages are *not* the same sentence, unless perhaps the languages overlap in that sentence.[2] Granted, there is an interesting question as to the conditions under which a homograph or homophone of a sentence of English in a language relevantly like English would also be true,

---

2   There is no sharp distinction between dialects and languages. And dialects *do* overlap: while some sentences of British English are not sentences of American English (e.g. the British "He is a superb football player" is a different sentence from its American homograph), many are.

but answers to that question do not seem to tell us about the modal status of the English sentence.

One might also say that the difference between the sentences is that we can know *a priori* that (1) expresses a true proposition. If this is what is meant by claiming that the modal status is different, I concur, but note that the difference is epistemological, not ontological: it is a matter of epistemic modality, while we are interested in alethic modality. For Gödelian reasons, there will be truths of arithmetic we cannot know *a priori* if our resources for proving arithmetical truths are limited to those of a formal system, but it is plausible that these truths of arithmetic do not have an ontological and alethic modal status different from the others. Only their epistemic status relative to us is different.

Consider now the alternative of claiming that (1) and (2) express different propositions. If one claims that (1) and (2) express different propositions, by the same token one should claim that

(3)   Cicero is Cicero

and

(4)   Tully is Cicero

express different propositions. After all, names are fixed by ostension just like natural kind terms like "water" are, and so it is plausible that what one says about natural kind terms one should say about names as well. But if one claims that (3) and (4) express different propositions, one can no longer accept the plausible thesis that the only thing names contribute to a proposition is the individual which they name.

But suppose we grant that (1) and (2) express different propositions. Nonetheless, it is implausible that the claims in each pair differ in modal status. Recall that the modality we are interested in is supposed to be independent of the vicissitudes of language. But while most English speakers use "water" in the Kripkean way as referring to the kind of thing that has historically been ostended to and dubbed "water," it is very plausible that many English-speaking chemists use "water" as a two-syllable synonym for the three-syllable term "H_2O." When they think of "water," they think of its molecules, and in their use of the word, it is analytic that its referent contains hydrogen atoms. (Even if this empirical surmise about the practices of chemists is false, it could easily be true.) But then if we are asked about the modal status of the claim that water is H_2O, we would need to further

query which sub-community the given speaker is a member of. It does not appear, in light of this, that a modal status in respect of which (1) and (2) differ is a modal status independent of the vicissitudes of language.

Moreover, if (1) and (2) express different propositions, there would appear to be indeterminate cases. For instance, if a child whose parents are both chemists, but who also has friends whose parents are not scientists, uses the word "water," it does not appear determinate whether the meaning of the word is inherited from her parents or from her friends. If (1) and (2) express different propositions, this much indeterminacy needs to be lived with. But it is implausible that the modal status of (2) should be indeterminate in this case. Moreover, it is implausible that a historical investigation into which sub-community one inherited the word "water" from would be needed to ascertain the modal status of one's utterance of (2). Granted, there may be cases where a historical investigation into one's linguistic acquisition *is* needed to determine modal status. But the "water" case does not seem to be such. We move too easily between chemical and ordinary discourse.

Still, we can do justice to our intuition that there is some sort of a difference between the two *sentences* by adverting either to the epistemological difference or to the following distinction. Some terms in English are defined by ostension and some verbally, and in the case of some it may be indeterminate. "Bachelor" is defined verbally as an "unmarried man." "Water" is defined ostensively as *that* natural kind. For any sentence S, let V(S) be the sentence obtained from S by first replacing each unquoted word that is verbally defined by its definition, iterating as many times as possible, and then replacing every remaining item of non-logical vocabulary by an undefined logical constant, a different constant for each word defined by a different ostensive act. Then, we can say that S is *verbally necessary* if and only if V(S) is a tautology. Thus, (1) is verbally necessary. To see this, suppose for simplicity that "H_2O" is defined as "a chemical constituted by molecules containing two atoms of hydrogen and one of oxygen" where each non-logical term here is not itself verbally defined.

Then, V(1) is something like "A C constituted by Ms containing two As of H and one of O is a C constituted by Ms containing two As of H and one of O," where capital variable letters are logical constants. And this is a tautology. But "water" is not a verbally defined term, so V(2) is "W is a C constituted by Ms containing two As of H and one of O," which is plainly non-tautologous. So we *can* do justice to both the intuition that there is a difference in the logical status of (1) and (2) and the argument that the propositions they express have the same modal status.

But the concept of verbal necessity is not our main subject. It depends

too much on historical accidents, such as whether a secondary name was defined expressly a synonym for the first or was independently ostensively bestowed. These are important issues for the philosophy of language, but have little ontological significance in them for the structure of possible worlds or the modal status of propositions. I will talk of logical necessity, necessity *simpliciter* and metaphysical necessity as synonymous, for I do not think useful ontological distinctions can be made between them. None of these necessities are verbal. They are all "real necessity," to use Kant's term.

We now have two choices. We can think of "real necessity" as more like the notion of "metaphysical necessity" or as more like the notion of "formal strictly logical necessity."

A basic problem with taking real necessity to be strictly logical necessity defined in a formal way is that of the modal status of the axioms and the rules of inference. A formal system always presupposes a particular system of axioms and a particular collection of rules of inference, and then claims that a proposition is necessary provided it is a theorem, i.e. provable from the axioms by use of the rules of inference. On this view, the axioms are automatically necessary, and if the rules of inference include the $\supset$-introduction rule that whenever q can be inferred from p, then $p \supset q$ is a theorem, then a material conditional corresponding to each rule of inference is also automatically necessary. But what is it in virtue of which a proposition counts as an axiom and a rule counts as a rule of inference within the system? As long as we confined our axioms to verbal definitions, we had an answer to the first question. But it will not do to confine ourselves to verbal definitions, since as we have seen, what claims come out necessary would then depend on historical vicissitudes about what was in fact defined how, and hence Kripkean necessities would disappear, as would philosophical necessities, such as that there necessary are some Platonic entities or that there necessarily are no Platonic entities.

Moreover, on such a strictly formal view, the necessity of Gödelian unprovable mathematical truths would disappear. These would end up being necessary only in a weaker metaphysical sense, while not being "strictly logically necessary." That there should be such a non-epistemological distinction amongst mathematical truths does not appear plausible.

For Gödelian mathematical truths cannot be accounted for in terms of strict logical necessity understood in a formal way. Let me go into a little detail here so that we might see what assumptions are being made about the formal system. The easiest way to see the philosophical issue is to start with decidability. Any sentence can mechanically be assigned a unique Gödel number by enumerating all the characters of the language as $a_1,...,a_N$ and

then replacing the sentence $a_{i(1)}a_{i(2)}...a_{i(n)}$ with the number $p_1{}^{i(1)}p_2{}^{i(2)}...p_n{}^{i(n)}$, where p_k is the kth prime. A set of sentences is *decidable* provided there is a recursive procedure for generating the Gödel numbers of these sentences. Then it is a theorem that the collection of all true first-order arithmetical sentences is undecidable (Boolos and Jeffrey 1995, Theorem 15.3). *A fortiori*, so is the collection of all true arithmetical sentences in any higher-order logic, since the first-order sentences will also be sentences of the higher-order logic, and the decision procedure for the higher-order case would give us the answers for the first-order case.

Now, for any formal system of the sorts we have, we can find a recursive procedure such that if the procedure is given a string of characters as input, the procedure can decide whether the string constitutes a proof or not. This in fact seems a very reasonable definition of "formal system," and is the one I shall adopt: the axioms and rules of inference must be specified in such a way that we can recursively judge the validity of a putative proof. Assuming "can mechanically judge" reduces to "can judge via a recursive procedure," which it does if Church's Thesis is true, and which anyway surely does reduce for any formal system relevantly similar to the ones we actually work with, a "formal system" is just a system where we can mechanically judge of the validity of a proof. But then it follows that there is no formal system, in this sense, together with a sufficiently rich language with a finite alphabet, such that all the true sentences of arithmetic are provable and no false ones are. To see this, note that if a language has a finite alphabet, we can recursively enumerate all the strings of characters. Since we can, *ex hypothesi*, recursively judge whether a string of characters is a proof, by recursively enumerating all the strings of characters and applying the decision procedure for checking whether a string is a proof, we will thereby be able to generate a list of all the theorems, i.e. of all provable sentences, since the theorems are precisely the concluding clauses of the proofs. But if the system is such that of the arithmetical sentences all and only the true ones are provable, then we will thereby have recursively generated a list of all true arithmetical sentences, contrary to the undecidability theorem.

Therefore, within any formal system that has a language sufficiently rich for arithmetic, there will be "Gödelian" mathematical truths, i.e. mathematical truths that cannot be formally proved. This is true in a very general sense of "formal system," not limited to first-order systems or systems with a finite set of axioms. The plausible assumption was that the availability of a recursive procedure for telling whether a string is a valid proof is the basic characteristic of a formal system.

But perhaps one can just abandon the necessity of Gödelian mathematical

truths. Maybe some arithmetical truths can vary across possible worlds. (This is bizarre, but perhaps no more bizarre than the possibility that I might have been the number seven, which formalist accounts of necessity must anyway countenance.) It will not do just to say that it is hard to see what kind of a contingent truth a mathematical proposition could express. For, as the debate between the Platonists and the anti-Platonists shows, it is certainly also hard to see what kind of a *necessary* truth a mathematical proposition expresses, especially in the light of Gödelian truths where the truth is no longer identical with provability. One reason why we think it is mysterious to think of mathematical truths as contingent is that it is difficult to see what they are contingent *on*. But this objection only applies on views of possibility on which contingent propositions are contingent *on something*. The final account of possibility argued for in this book will be such, but many competitors are not.

It is perhaps Kripkean and philosophical cases, then, that are more problematic for the formal theory. Suppose, for instance, that (contrary to fact, I believe) the correct theory of mind is functionalism. Then, surely, it is *necessarily* the case that every mind is a functional system of a certain type. But this is not a definitional truth. We do not in fact *have* a definition of mind. One arrives at functionalism through a series of philosophical arguments, such as an inference to the best philosophical explanation of some facts, positive arguments based on imagining what would happen were we to replace parts of a mind with black boxes, and the like. None of these arguments can be plausibly thought of as purely logical derivations from definitional truths. And the same is true, surely, of the question of what theory of mind is correct, and of much of the rest of philosophy. For instance, if the number seven exists as a Platonic entity, it does so necessarily, and this can only be a definitional truth if there is a sound version of an ontological argument starting with only definitional truths and concluding with the existence of numbers.

The deep issue here is this. As we saw, the axioms are necessary and the rules of inference necessitate *automatically*. It will not do to include just purely definitional truths among the axioms. Consider the question whether we should, say, include the law of excluded middle (either as an axiom or as implicit in the system of rules of inference).There is presumably a fact of the matter about whether this law has necessity or not. But its necessity or lack thereof cannot be analyzed in terms of whether it is a theorem or not, since whether it is a theorem or not depends precisely on whether it is included explicitly or implicitly in the axioms or rules of inference. The formal account cannot tell us what it is about a proposition that fits it for

being an axiom. If the account says that in fact some propositions just *are* axioms and others just *are not*, then this is no longer a formal account. It is some kind of Platonic realism about axiomhood.

And in fact it is not a very plausible Platonic realism. For any given formal system can be re-axiomatized in many ways. We can take the parallel postulate as an axiom in Euclidean geometry and derive the Pythagorean theorem, or we can take the Pythagorean theorem as primitive and derive the parallel postulate. It is implausible that in the case of every set of equivalent axiomatizations there is a fact of the matter as to which axiomatization is "the right one." Given this, one might find congenial the claim that any proposition that *could* be used as an axiom of a formal system equivalent to the right formal system has some special Platonic property. But *any* theorem could also be used as an axiom. Thus, the view now comes to this: Some propositions simply have a special Platonic property. Since these propositions are the theorems, and the theorems are all the necessary propositions, the view seems not to substantially differ just from the view that necessity is a special property of propositions. This *is* a formidable Platonic view, one we will tackle in Part IV, but it is not a formal view. And, of course, what I said about axioms can be said about which rules can be chosen as rules of inference.

Thus a formal view of the truth of modal claims is unsatisfactory, as long as we mean it to be an analysis of objective modal truths. But if we mean it to be an analysis of how we speak, of which truths are definitional relative to our contingent language and which follow from definitions, then in that limited sphere it is a useful tool. In one logical system, it may let us say that there is a contingency about the claim "I am not the number seven" that is not to be found in the claim "This bachelor is unmarried." In another, with different axioms, "I am not the number seven" might be quite necessary. This is all very useful. But it is not an account of the kind of objective modality that this book is after. That is why this book about alethic modality is a book of metaphysics, not of logic.

Section 3 S5

The modal logic assumed through most of this book is S5, i.e. a logic satisfying the axioms:

(5) $\Box(p \supset q) \supset (\Box p \supset \Box q)$
(6) $\Box p \supset p$
(7) $\Diamond p \supset \Box \Diamond p,$

for all propositions p and q, and where $\Diamond$ is short for "it is possible that" while $\Box p$ is short for $\sim\Diamond\sim p$, together with the Rule of Necessitation that if a formula is an axiom or theorem, then that formula prefixed by $\Box$ is also an axiom or theorem.[3] This system is known as S5 and technically is characterized by an accessibility relation that is reflexive, symmetrical, and transitive. Moreover, the standard possible worlds interpretation of the system assumes all worlds are mutually accessible.

The most controversial axiom here is (7) which says that if something is possibly true, then it is impossible for it to fail to be possibly true. In modernity, the axiom goes back at least to Leibniz's discussion of Descartes's ontological argument. Descartes had defined God as a being that has all perfections, and one perfection is the property of necessary existence. Leibniz (Ariew and Garber 1989: 237) noted that Descartes's argument was missing a crucial premise, namely that it was *possible* for God to exist, and argued that once that premise was added, the argument became valid.

Now, there certainly are kinds of modality for which (7) fails. For instance, suppose we consider a forward branching temporal structure, and say that p is physically possible at some point z in the structure providing p is already true at z or at some future point that can be reached from z without violating any of the laws of nature. Then, p is physically necessary at z providing that p does not fail at z or at any future point that can be reached from z. Then, it is possible (here and now) that I will at some point in my life run a marathon. But it is certainly not necessary that this is possible, because there is a future I can reach where my legs are cut off before I run a marathon and at a point in the future of that accident there would no longer be any reachable points at which I could run a marathon.

One way to argue for (7) in the case of metaphysical possibility would be to start with two intuitions. The first is that things could not have been such that it would have been impossible for things to have been as they in fact are. However things might have gone, it still would have been true that they might have gone the way they in fact have gone. If things could have gone a certain way, then had they gone that way it would have been true that they could have gone the way they in fact went. This is the Brouwer axiom: $p \supset \Box\Diamond p$. It tells us that the accessibility relation is symmetric.

The second intuition is that we when we talk about metaphysical possibility, we are talking about "ultimate" possibilities. Now, if we have a possibility operator $\Diamond$ such that $\Diamond p$ can hold without $\Diamond\Diamond p$ holding, then this operator does not tell us about *ultimate* possibilities. If it could have been

---

3   Loux (1979: 16).

that it could have been that p holds, then there is a real sense in which p "could have held." If we then deny that $\Diamond p$, we are saying that $\Diamond$ does not tell us of the ultimate possibilities there are, but of possibilities relativized to some way that things have been. Indeed, in such a case there *is* a reasonable more ultimate metaphysical possibility operator, namely $\Diamond\Diamond$. Thus, if we are talking of ultimate possibilities, it is reasonable to require that $\Diamond\Diamond p$ should imply $\Diamond p$. This is the S4 axiom; it tells us that the accessibility relation is transitive. We can also make the equivalent point using $\Box$: the ultimate necessities are necessities that couldn't have been different. So, if something is necessary, it has to be necessary — i.e. $\Box p$ implies $\Box\Box p$, which is equivalent to S4.

Here is another argument for S4. Suppose that the rule of necessitation holds together with (5) and (6). Moreover, add the following axiom:

(8) $\Diamond \exists n F(n) \supset \exists n \Diamond F(n),$

where n ranges over the positive integers. If the quantification in (8) were to range over all objects, this would be the Barcan formula:

(9) $\Diamond \exists x F(x) \supset \exists x \Diamond F(x).$

The Barcan formula is surely unacceptable. For instance, let F be the property of being essentially a member of a species of one-horned equines. Since donkeys could evolve into a new equine species that has a single horn and since the members of that species would then be essentially members of it, it is possible that there exist an entity that has F, but we may assume that in fact there is no such entity. (If the assumption is false, just vary the number of horns until one comes to a case that works.) If (9) holds, then it follows from the possibility of there being an F that there actually exists an entity, call it Sam, that is possibly an F. But if there is nothing that is essentially a member of a species of one-horned equines, it is exceedingly implausible that there actually be an entity that could be such a member.

However, while the Barcan formula is implausible for contingently existing beings, it is very plausible for mathematical entities like positive integers. After all, replacing F with $\sim F$ and contraposing shows that (8) is equivalent to:

(10) $\forall n \Box F(n) \supset \Box \forall n F(n).$

But $\forall n F(n)$ is essentially an infinite conjunction: $F(1)\&F(2)\&\dots$. Thus, (10) simply says that if every conjunct in an infinite conjunction is necessary, so

is the whole conjunction. In the case of finite conjunctions, this follows from (5) and Necessitation,[4] but the infinite case is also surely highly plausible.

Now, if $\Diamond$ does not satisfy S4, but does satisfy (5), (6), (8), and Necessitation, and working in a logic that allows mathematical induction, we can construct a new modal operator, $\Diamond^*$, that not only satisfies (5), (6), (8), and Necessitation but also satisfies S4. To do that, let $\Diamond^n$ be $\Diamond...\Diamond$, where there are n diamonds. Then define $\Diamond^*p$ if and only if $\exists n\Diamond^n p$. The corresponding necessity operator is $\Box^*p$ which holds if and only if $\forall n\Box^n p$, where $\Box^n$ is the nth iteration of $\Box$. That $\Diamond^*$ satisfies (5), (6), (8), Necessitation and S4 is shown in the Appendix (assuming the availability of mathematical induction). If $\Diamond$ satisfies S4, then $\Diamond^*$ is equivalent to $\Diamond$. But if a candidate $\Diamond$ for a metaphysical possibility operator does not satisfy S4, though it does satisfy (5), (6), (8), and Necessitation, then $\Diamond^*$ will be a better candidate for a *correct* metaphysical possibility operator, since $\Diamond^*$ has more of the ultimacy that we expect of metaphysical possibility.

But of course the Brouwer and S4 axioms, together with (5), imply (7). To see this, apply the Brouwer axiom to $\Diamond p$ to conclude that $\Diamond p \supset \Box\Diamond\Diamond p$. S4 says $\Diamond\Diamond p \supset \Diamond p$, and applying Necessitation we get $\Box\Diamond\Diamond p \supset \Box\Diamond p$. Putting this together with $\Diamond p \supset \Box\Diamond\Diamond p$, yields (7).

Another intuition in favor of S5 is that broadly logical possibility cannot have been different, since it is a matter of what propositions follow from what propositions (a proposition is possible if and only if its negation does not follow from it), and what follows from what could not have been different. Therefore, if $\Diamond p$, then it could not have been the case that $\sim\Diamond p$, i.e. $\Diamond p \supset \Box\Diamond p$.

Or as a variant of this intuition, one might say that precisely those propositions are broadly logically possible which the fundamental laws of metaphysics allow. But the collection C_0 of all the fundamental laws of metaphysics could not be different from what it is — that is central to its being the collection of the *fundamental* laws of metaphysics — and the "could not" here is surely metaphysical. Moreover, what C_0 allows cannot have been different. If it were different, that would presumably be because a collection of laws might permit different things in different circumstances. Suppose that C is some collection of fundamental laws that permits different

---

4   The finite case says that if $\Box p_1\&...\&\Box p_n$, then $\Box(p_1\&...\&p_n)$. To prove this, observe that it suffices to prove it in the case where $n=2$, since the general case follows by iteration. The proof of the binary case is due to an email from Steve Kuhn. It is a theorem of propositional logic that $p_1\supset(p_2\supset(p_1\&p_2))$. By Necessitation and (5), we have $\Box p_1\supset\Box(p_2\supset(p_1\&p_2))$. Using (5) again and doing some propositional logic, we get $\Box p_1\supset\Box(p_2\supset\Box(p_1\&p_2))$. With a bit more propositional logic, we get $(\Box p_1\&\Box p_2)\supset\Box(p_1\&p_2)$.

things in different circumstances. But then there would need to be further metaphysical laws as to what the laws in C collectively permit under what circumstances, and barring a vicious regress of more and more basic laws, there would have to be fundamental laws specifying what the laws in C permit. And *these* laws couldn't be in C, since then the laws in C would not permit different things in different circumstances. Therefore, if C permits different things in different circumstances, then C does not contain *all* the fundamental laws, in the way that C_0 does. Thus, what C_0 permits could not be different, and hence modality could not have been different. And that is what axiom S5 says.

The S5 system of modal logic will be in the background for most of this book. It is worth noting that some of the most prominent views of possibility to be considered, namely the Lewisian one and many ersatzist ones leave little room for the denial of S5. The Aristotelian account argued for in this book will also have this property, at least if one makes a plausible assumption (see Section 2.5 of Part VI). One could thus also give the following argument for S5: the best metaphysical accounts of possibility that we have require or imply S5.

I shall feel free to use S5 in my arguments for and against various metaphysical views of possibility, on account of the prima facie plausibility of S5 holding in the case of a notion of ultimate metaphysical possibility. A good theory of possibility should either yield S5 or some close approximation.

Section 4 Eight views of possibility

4.1 *Parmenides, Leslie, and Rescher*

In his poem *On Nature*, Parmenides learns from the goddess that there are only two

> ways of enquiry that are to be thought of. The one, that [it] is and that there is no non-being [*ouk esti mê einai*], is the path of Persuasion (for she attends upon Truth); the other, that [it] is not and that it is needful that there be non-being [*esti mê einai*], that I declare to you is an altogether indiscernible track: for you could not know [*gnoiês*] what is not [*to ge mê eon*] — that cannot be done — nor indicate it.[5]

---

5   Fr. 291. Throughout, I am following, with modifications, the translation of Kirk et al. (1983).

> What is there to be said and thought must needs be: for there is being, but [there is] nothing [that] is not [*esti gar einai, mêden d' ouk estin*].[6]

The argument, insofar as it is more than just an assertion, is that non-beings plainly do not exist and if we speak and think, we are speaking of *something*.

We can put this argument in a more modern form by considering *truth-makers*. According to advocates of truthmakers, realism requires that propositions be *made true* by something real. The proposition that there are horses is made true by the horses of this world. The proposition that Socrates is sitting is made true by Socrates's sitting, or the sitting Socrates *qua* sitting. An item in the world that a proposition is made true by is called its *truthmaker*. I will take it that the existence of the truthmaker entails the proposition it makes true, but this entailment is not sufficient for being a truthmaker (otherwise every entity would be a truthmaker for a necessary truth, since a necessary truth is entailed by any proposition). What exactly the truthmakers of propositions are depends on one's ontological system. For instance, if one is committed to an Aristotelian worldview on which all there is are substances and their attributes, broadly construed, then the truthmaker of every true proposition will ultimately be a number of substances and their attributes. An event ontology, on the other hand, may have the truthmakers be mereological sums of primitive events. But whatever the truthmakers are in one's ontology, in the case of propositions giving concrete facts about concrete entities, the truthmakers are going to be made up of concrete things like tables, chairs, dogs, cats, sittings, and shoutings.

Moreover our language provides a way of referring to the truthmaker of a proposition in a way that is neutral between ontological systems. To every declarative sentence there corresponds a participial nominalization. To "Socrates is a philosopher, was a war hero and taught Plato" there corresponds "Socrates' being a philosopher, having been a war hero and having taught Plato." To "Brutus betrayed Caesar" there corresponds "Brutus' having betrayed Caesar." To "There are horses" there corresponds "There being horses." If a sentence expresses a proposition, then the referent of its participial nominalization is the truthmaker of that proposition, or a plurality of truthmakers (plausibly, each horse makes it true that there are horses; there being horses then might simply be the plurality of all horses). But what kind of an item "Brutus' having betrayed Caesar" denotes, whether it is ultimately a complex ultimately of substances and their attributes,

---

6   Fr. 293.

or of events, or a fact in a world that is all that is the case, is a difficult metaphysical question.

In any case, then, according to a modernized Parmenides a proposition is true if and only if it has a truthmaker that really exists. This gives us a sense we can attach to Parmenides's cryptic remarks. If we know or speak truly, there must be an object of our knowledge or speech, namely the truthmaker of the proposition we know or express. It is this object that we know or speak of. The assertion that we cannot know or speak of what is not, then, becomes the claim that if we are to be right, there must *be* something we are right *about*: something that makes our affirmations true. Where the truthmaker is not, neither is there anything true.

Of course the notion of a truthmaker is going to be pointless unless we have some substantial theory about *what kinds* of entities can play that role. I can always say that the truthmaker of *p* is just *its being the case that p*, and if I do this for every true proposition, every true proposition will have a truthmaker in an apparently trivial way.[7] However, saying this would not be so completely trivial. It is after all a substantial ontological claim that there are such things as *its being the case that p*. But this slight trivialization of the truthmaker theory does suggest that a criticism that some theory cannot provide a truthmaker for some proposition is shorthand for an argument that we are not satisfied with just this trivial truthmaker for the proposition, i.e. that we want a further analysis of what *its being the case that p* consists in for this *p*. Moreover, it is in general preferable in a philosophical theory of some proposition *p* that one be able to say more about the truthmaker of *p* than that it is identical with *its being the case that p*. Being able to say more about this truthmaker is itself a reason in favor that theory. Thus, even if we do not want to insist that *always* more can be said, we will *ceteris paribus* prefer a contentful theory that says more.

While Parmenides did not deal with modality *per se*, we may be able to find in his writings an argument against change that easily generalizes to an argument against modality.

> And how could something that is [*to eon*] be in the future? How could it come to be? For if it came into being, it is not: nor is it if it is ever going to be in the future.[8]

A claim about the future must be made true by a truthmaker that is in the

---

7   I am grateful to Robert Brandom for discussion of this issue.
8   Fr. 296, my translation.

future. But then there is nothing *now* by which the claim is to be made true. Hence, when we say something *will be*, we are perforce speaking of something that is not, and thus not speaking truly.

The argument construed in this way may be criticized by a B-theorist for conflating existence *simpliciter* with merely *present* existence, but it is much more interesting in the modal case. We can say that the proposition that there *will* be a sea battle tomorrow is made true by tomorrow's sea battle, which exists *simpliciter*. But there is a much deeper problem in the case of modal propositions. What makes true assertions of mere possibility? Suppose no sea battle in fact occurs tomorrow. What, then, makes true the proposition that there *can* be a sea battle tomorrow? If there will be a sea battle tomorrow, then maybe tomorrow's sea battle can make the proposition reporting its future occurrence true now. So, in parallel with this, is it that the merely possible sea battle makes true the proposition that there can be a sea battle? But this will not do, because the adjectival phrase "merely possible" is truth-canceling in the way "fake" is — fake money isn't money — while "tomorrow's" is not truth-canceling. A merely possible sea battle is not anything that exists. If it is not anything that exists, it cannot make anything true. But what else could the assertion that there can be a sea battle be about, one asks, other than the future sea battle?

Parmenides, not having a clear notion of modality, merely claims that his one reality is atemporally unchanging. But he could have used the same arguments to arrive at the further claim that this one reality *must* be as it is and can be no other, and doubtless if he were asked the modal question clearly, he would say this. And this is the Parmenidean puzzle of modality or what I have earlier called the problem of the ontological ground of possibility. It comes as a paradox and a problem. It seems that the proposition that there can be unicorns is, if anything, about unicorns — its truthmaker would have to be comprised, at least in part, of unicorns or their existing. Thus, its truthmaker does not exist, there being no unicorns and no existing of unicorns, and so the proposition is false. But it is paradoxical to admit that the only things that could be are the things that are.

If we were to solve this problem on Parmenidean terms, we would need to explain *what* the truthmakers of modal propositions are, and what it is about these truthmakers that makes them suitable to be such. It will be the purpose of this book to attempt something like an answer to the problem. The attempt will be made within the confines of a broadly Aristotelian ontology, where the basic entities are substances and their modifications (properties and relations), an ontology which thus will tend to be unfriendly to the idea that such entities as *the state of affairs of it being possible that*

unicorns exist could be fundamental. However, although the intuitions behind this kind of ontology drive much of the project, they are not presupposed by all of the individual arguments that will be given for the preferred answer to the Parmenidean problem and against the non-preferred answers.

That said, we can, and probably should, in general, reject the truthmaker theory as it stands, and opt for one of two more sophisticated versions. The first of these builds on Aristotle's famous definition that to speak truly is "to say of what is that it is, and of what is not that it is not [*legein … to on einai kai to mê on mê einai*]" (Aristotle 1984: 1011b27). There are, thus, two fundamental ways of being true. A positive proposition is true when it has a truthmaker. A negative proposition is true when it lacks a falsemaker, i.e. when nothing that would make the proposition's negation true exists. And, presumably, what grounds the truth of non-fundamental propositions is a combination of truthmakers and lacks of falsemakers.

Or we can follow David Lewis (2001) in adding a *third* kind of fundamental proposition, one that attributes a property or relation to one or more objects, which is made true by an object's having that property or the objects' standing in that relation.

The Aristotelian truthmaker–falsemaker theory generates much the same ontological grounding puzzle for modality as the Parmenidean theory does. It is still puzzling what existent thing makes it true that there could be unicorns even if we allow that the proposition that there could be unicorns could be made true, not just by the existence of something, but also by a combination of a thing's existence with another's non-existence. And even if we extend to the Lewisian theory, it is still puzzling what things' having properties or standing in relations helps ground modal claims.

In general, thus, what we are searching for in this book is the *grounds* of alethic modal claims. These grounds might be truthmakers. Or they might be a combination of truthmakers and lacks of falsemakers. Or they might be a combination of truthmakers, lacks of falsemakers, and attributions. And even if one does not think that *all* propositions have the relevant grounds (e.g. Merricks 2009), one would prefer a theory of modality that yields grounds to one that leaves modal truths ungrounded — bruteness should be a last resort. A call for grounds is an extension of the Parmenidean challenge.

But first let us consider two modern developments of the Parmenidean idea that only the actual is possible, those by John Leslie and Nicholas Rescher. Leslie proceeds through an "axiarchic principle," or principle of ethical requiredness (Leslie 1997). This principle corresponds to Plato's Form of the Good, and imposes on the world the necessity of satisfying certain conditions that make it be for the best. One argument for this principle

would be through the considerations that fall under the head of "the anthropic principle" (cf. Leslie 1990). The constants in the laws of nature (masses and charges of elementary particles and strengths of basic forces) appear to be calibrated in such a way as to make life possible. If they were somewhat different, and physics gives us no reason to think they could not be, life as we know it would not be possible. This provides evidence for the axiarchic theory, in that if the axiarchic theory is true, such fine-tuning is unsurprising, while if it is false, it is more surprising. However, obviously, this also provides evidence for other alternate theories, such as traditional theism, or multiverse theories.

Rescher (2000), on the other hand, has argued for a metaphysically necessary principle of optimality as a theory that explains why we find orderly laws of nature that can be mathematically formulated and understood by us. This principle ensures that, necessarily, things are for the best, understood in a Leibnizian sense as a balance between variety and lawlike unity. Of course, there are other theories that, if true, explain the same explanandum. Theism provides one such theory.[9] Another would be a more limited version of Rescher's theory, which merely claims that the *laws* of nature are necessarily for the best, while at least some of the contents of the world are contingent. This more limited version does not overturn modality, and rather than counting as a general view of possibility *simpliciter*, it is simply a substantive view about what possibilities there in fact are.

Rescher's view in its unlimited form appears to be subject to the following objection. First of all, if there is only one possible world, then saying that our world is the best of all possible worlds is not saying anything interesting. One could say with equal propriety that it is the worst of worlds. Consequently, the optimality principle cannot explain why the laws of nature are orderly, because if, *per impossibile*, the only possible world were one where they were disorderly, that world would also be the best.

Rescher's own reply (personal communication) is to distinguish a notion of logical possibility from a notion of metaphysical possibility. There is more than one logically possible world, and of these the best one is the one that

---

9   Actually, both Rescher and Leslie are willing to draw theological conclusions from their theories. It is for the best that there be a God, after all. Leslie (2001), though not Rescher, even goes so far as to posit an infinite number of deities, since he think that the more, the better. (This is of course problematic in light of the set theoretic fact that there is no largest infinity, as the same reasoning would say that it is always better to have a higher infinite cardinality of the set of deities than a lower infinite cardinality. In correspondence, Leslie invoked Cantor's murky notion of an absolute infinity as an answer.)

is metaphysically necessary. One might make non-trivial sense of the claim that the one and only possible world is optimal, for instance by considering worlds that are metaphysically impossible recombinations of things in this world, but nonetheless are modeled by mathematically coherent structures and hence capable of comparison to our world. But then the problem of evil rears its ugly head. The argument from evil against the existence of an omnipotent, omniscient, and omnibenevolent deity, difficult enough as it is, takes a particularly difficult form if it is claimed that this world is in fact not just worthy of being made by such a God, but is the *best* conceivable world. Even if we could answer the original argument from evil,[10] defending the claim that this world is the *best* one is a yet further task. Moreover, the evidence from the apparent non-optimality of this world weighs against the evidence from the lawlike orderliness of the world. And there *are* theories that are better supported by the conjunction of these two pieces of evidence than Rescher's full theory; e.g. the more limited theory that says that the laws of nature are necessary and optimal, but the events in the universe, including freely done human actions, are contingent.[11]

4.2 *Leucippus, Democritus, Meinong, Lewis, and Aristotle*

The Parmenidean extreme holds that merely possible worlds do not exist in any way, because our world is the only possible one. The other extreme view is that all possible worlds must exist. Leucippus's and Democritus's atomism could be an early representative of this view.

> Leucippus holds that the whole is without bound . . . part of it is full and part

---

10 The best published answer I know of is still the combination weak theodicy championed by Alston (1991; see the discussion in Gale and Pruss (2003: xxxi–xxxii).

11 There is one version of the optimality principle that is capable of answering the problem of evil, and this is Donald Turner's (forthcoming) view according to which all universes which have a favorable balance of good over bad are created by an all-good deity, the separate universes being unrelated spatio-temporally. The fact that our universe is non-optimal is not a problem for the theory as long as the balance of good over evil is favorable. And, overall, the world, which is the mereological sum of all these universes, is optimal. Rescher himself does not agree with this approach, since this large world consisting of many universes as a whole lacks the coherence and order of an optimal world. Some of the objections I shall levy against Lewis in Part III, below, can be retooled to work against Turner.

void . . . Unboundedly many [*apeirous*] worlds arise hence and are resolved again into these elements.[12]

If one takes the "unboundedly many" in the most extreme sense as involving all possibilities, then indeed we do get a view that all possible worlds exist.

This view would have been of interest merely to historians were it not for Alexius Meinong and David Lewis. Meinong sought to explain the intentionality of thought by invoking objects that correspond to all of our ideas, even ideas not exemplified in our world. Thus, there are some things that don't exist.

David Lewis does this, too, at least for *possible* objects, but further organizes the things that don't actually exist into worlds. More precisely, Lewis posits that every possible world exists, and that these worlds are ontologically on par with one another. What makes two entities be a part of the same world is that they are spatio-temporally related. Thus, the different worlds are not spatio-temporally related, presumably unlike the worlds of Leucippus and Democritus. But, because of the ontological parity thesis, as in Leucippus and Democritus, the worlds are actualized as concrete entities just like our world. The horses in the other worlds are horses in exactly the sense in which the horses we know are, except that they are not spatio-temporally related to us.

Material reality is for Lewis much richer than we used to think. There exist dog-headed "men," and chimeras and unicorns — but not in our world. Fortunately, most of our language is relativized to our world, the *actual* world, which for Lewis is set apart from other worlds only indexically: the actual world is nothing but the mereological sum of all things that are spatio-temporally related *to us*. When I say, speaking ordinarily, that there are no unicorns, I mean that no unicorns *are actual*, that the *actual* world does not contain unicorns, i.e. according to Lewis that no unicorn is spatio-temporally related to us. In ordinary non-modal language, quantifiers are restricted to the speaker's world.

But of course, and this is the point of the theory, we can also speak with unrestricted quantifiers. Thus, we can translate the assertion "Unicorns are possible" as "There is a world *w* such that unicorns exist in *w*." These sorts of quantifications give sense to modal language. Moreover, Lewis believes his theory of possible worlds makes it possible to give an account of various other philosophical notions. Thus, a proposition is a set of possible worlds

---

12   Hippolytus, in: Kirk et al. (1983: 417), translation modified.

— those worlds that it is true at — and a property is a set of individuals, with the set being allowed to extend beyond one world if desired.

This theory is elegant, solves many problems, including the extended Parmenidean problem of ontological grounding, and appears coherent. But why should we think it *true*? Why should we think that reality is so much richer in material objects than we had thought? Lewis's answer follows in the footsteps of Leibniz's answer to Lady Masham's worry about Leibniz's system. Lady Masham wrote:

> But it appears not yet to me that [your system] is more than a Hypothesis; for as God's ways are not limited by our conceptions; the unintelligibleness or inconceivablness by us of any way but one, dos not methinks, much induce a Beleefe of that, being the way which God has chosen to make use of.[13]

Leibniz replied *inter alia* with the following methodological observation:

> (1) it seems that it is quite a considerable thing when a hypothesis manifests itself (*paroisse*) as *possible*, when none of the others manifests itself so at all, and (2) . . . it is extremely *probable* that such a hypothesis is the true one. Likewise one has always known in astronomy and physics that the most intelligible hypotheses are eventually found to be true: such as, for example, that of the movement of the earth to save the appearances of the stars . . .[14]

The very fact that a theory gives a coherent account of difficult problems, where other theories have failed, is evidence for the theory's truth, in philosophy as in science. Thus, Lewis thinks, we should believe his theory because it is elegant, solves many problems and, Lewis thinks, appears coherent.

It should not come as a surprise that such a drastic revision of the account of what we think there is as Lewis provides carries with it counterintuitive consequences. For instance, as we shall learn in Part III, were we to believe Lewis, we would have to become inductive sceptics and revise some important moral notions. This price would be too high. For just as the fact that a theory gives a coherent account of one thing provides evidence for the theory's truth, likewise the fact that the theory gives rise to seemingly absurd consequences elsewhere, e.g. in epistemology or ethics, gives evidence against it. The theoretical benefits of Lewisian possible worlds theories will be critically considered in Section 4.7. If another theory can be found that

---

13 Gerhardt (1960–61), vol. III, p. 350.
14 Gerhardt (1960–61), vol. III, p. 353, my translation, emphases in original.

has all or most of the benefits that survive this critical examination, but lacks the demerits of contradicting induction or morality, then that other theory is to be preferred. It shall be argued in Part VI that there is such a theory.

One final theorist who uses the same strategy as Lewis for grounding modality in a larger totality of existent things should be mentioned. This is Aristotle. The inclusion of Aristotle here may seem surprising, but in fact we can find two different Aristotelian threads of thought about modality. One of these threads involves a modal logic based on time, and used in *De Interpretatione*. A proposition is *necessary* if and only if it holds at all times and possible if it holds at *some* time. Observe that in order to avoid the Parmenidean consequence that all true propositions are necessary, we need to take an analysis of indexical sentences which makes a sentence like "It is now noon" express *one and the same* proposition at different times, which one proposition is true at noon but false at other times.[15]

This Aristotelian theory is structurally quite similar to Lewis's, except with time-slices in the place of concrete universes. In both cases, modal claims are analyzed in terms of quantification over concreta. It is also true on Aristotle's theory that the difference between mere possibility and actuality is indexical. The actual is what is *now*, just as for Lewis the actual is what is *here*, i.e. in *our* world. There is, however, a difference. Aristotle does not see, as far as I can tell, the full ontological parity between other times and the present that Lewis sees between other worlds and ours.

And there is a single objection that can be made both against Aristotle's theory and against Lewis's. Both theories share a crucial feature with the fatalist accounts of Parmenides, Leslie, and Rescher. The *whole* of reality could not be different than it is. In the case of Lewis's theory, the whole of reality, i.e. the mereological sum of all universes, is fixed. There are no semantic resources in his theory for making it possible for this sum to be different. For were it possible for this sum to be different, it would *be* different at some world, whereas all worlds are parts of the *same* total reality.

---

15 This would not be the only way to analyze indexical claims if we were not committed to this account of modality. For instance, we might take "It is now noon" to express a different proposition at each different time during which it is uttered, and there is support in ordinary language for this. It may, after all, seem absurd to think the propositional content of our beliefs changes as I go from believing that tomorrow there will be a sea battle to believing that yesterday there was one: surely I do not "believe differently" and have not "changed my mind" in such a case. But given an Aristotelian account of modality, if "It is now noon" were taken to express a different proposition at each time that it was uttered, then it would be an eternal and hence necessary truth.

In the case of Aristotle's theory, contingency is only possible in propositions that change in truth-value. Hence, a proposition that reports in a timeless way the sum total of what happens over time would have to be necessary for Aristotle — though Aristotle is apparently unaware of such propositions.

But surely the sum total of reality could have been different in some respect. If Lewis is right, there are infinitely many universes. But it seems quite possible for there to have been only one.[16] This coherence Lewis must reject as merely apparent. Similarly, it is quite coherent to imagine the possibility that in fact there never was any change, a possibility Aristotle must, and does (in *Metaphysics* Λ.6), reject as merely apparent. Both Lewis and Aristotle thus go against common sense modal claims. Lewis, however, can argue that the theoretical benefits of his theory are worth it. But if an alternate theory were found which had the same benefits and fewer paradoxical conclusions, then on his principles Lewis would have to prefer it (cf. Lewis 1986a: 5).

4.3 *The linguistic view*

Possible worlds have much theoretical value. Thus, it would be nice to have them without paying the price of Lewis's extravagant ontology. One suggestion that has been made in many forms (see, e.g. Roper 1982; Jeffrey 1983) is that possible worlds can be taken to be maximal consistent sets of sentences. A proposition holds at a world if it is entailed by the propositions expressed by the sentences that the world consists of. The actual world is the world all the sentences in which are true. This does not mean that all the sentences *uttered* "in that world" are true. We need to distinguish between two senses of a sentence *s* being in a world: first, when *s* is a member of the set of sentences that constitute that world, i.e. when *s* is true at that world, and, second, in virtue of its being true at that world that that sentence is uttered.

The linguistic view commits one ontologically to the existence of sentences, which we already probably believe in, and sets, which have various

---

16 I am not denying S5 here by claiming that it might have been that there was only one possible *world* — though Lewis will see it this way. Rather, I am claiming that it surely might have been that there was only one *universe*, namely only one maximal spatio-temporally interconnected object. Lewis will take that as tantamount to a claim that there might have been only one world, but the intuitions behind my claim are intuitions about possibilities for physical reality — for what sorts of spatio-temporal objects might exist — and not for the space of possibilities.

theoretical benefits and which Lewis, too, needs in order to reap all the benefits of his possible worlds theory. Thus the price is low.

Of course one needs to be more precise about what one means by sentences. It will not do to limit ourselves to *actually uttered* sentences, since that would lead to the absurd conclusion that were there in fact no speakers, nothing would be possible. But it seems that if we speak of *possible* sentences, then our account of possibilia becomes circular, since we were supposed to be clarifying the ontological status of possible individuals. Fortunately, there is a simple solution to this dilemma. By "sentences" we mean *types* of sentences. Now, there is no great ontological extravagance in positing such types. Languages like those of humans can, to a good approximation, be reduced to sequences of discrete symbols, and types of sequences of symbols can easily be modeled set-theoretically. So this account in fact needs nothing more in the ontology beyond set theory. Since we may well want set theory for independent reasons, this is cheap.

But, Lewis has argued, you get what you pay for, and we shall see in Section 2 of Part IV that we do not get enough by this method. We can put the point briefly with a dilemma. If we use an *actual* language, we have the problem of alien properties: basic properties for which our language has no words, but which are instantiated at other possible worlds. But if we use a non-actual language, then we need to have some way of specifying what that language is, and that is impossible for us unless by specifying the language we create it, thereby contradicting the fact that it was supposed to be non-actual. Moreover, unlike Lewis's account, the linguistic account does nothing to *illuminate* the meaning of modal propositions, because it presupposes modality in the requirement that we talk of maximal *consistent* sets of sentences, whereas of course a set of sentences is consistent if and only if the conjunction of the propositions expressed by the sentences is possible.

4.4 *A Platonic primitive modality view*

One can escape the arbitrariness and limited expressiveness of linguistic representations through abstract propositions. For various theoretical purposes, it is useful to introduce entities known as *propositions*, which are what our sentences express. Two sentence tokens have the same "content" if and only if they express the same proposition. Moreover, it is the proposition which is the carrier of truth, because truth does not belong to a sentence, which is language-relative, but to the language-invariant proposition expressed by the sentence.

Now that we have introduced propositions by ostension as entities that sentences express, we can speak of the whole collection of propositions. Not all members of this collection are expressed by some sentence actually uttered. Nor even are all the members actually express*ible* by some sentence of an actual human language. Whatever can be expressed by any possible language is a proposition. Note, however, that we escape the circularity objection because we are not actually *defining* propositions in terms of possible languages. We are defining them by ostension in terms of our language. But then we realize, on theoretical grounds, that propositions have a life of their own going beyond our actual language, rather as the electrons we posit on theoretical grounds to explain some actual phenomena have a life of their own and possess dispositional properties not exhausted by the actual circumstances of this world. Admittedly, the kind of explanatory role the two serve is different: electrons play a causal role while Platonic entities such as propositions do not. But nonetheless the propositions do explain various facts about sentences and propositional attitudes.

The above should have given us a grasp of the notion of a proposition. Propositions, moreover, enter into logical relations, and we can talk of propositions being consistent or not. Now we can define a possible world: it is simply a maximal consistent collection of propositions, assuming there are such maximal collections (I shall argue in Section 3 of Part IV that this assumption can be justified). Or, alternately, we can define a possible world as a class of logically equivalent maximally strong propositions, where a proposition is "maximally strong" if it entails every proposition compatible with it. These approaches have been championed most notably by Robert M. Adams (1974) and Alvin Plantinga (1974).

Of course, both of these definitions presuppose modality whether in the notion of consistency or in that of entailment, and so we will not get a reductive Lewis-type analysis of modality. But possible worlds may still be a useful construct to have, even if they do not give such an analysis. In Section 3 of Part IV, I shall argue that approaches to modality along these lines objectionably fail to answer the Parmenidean objection to modality. Presumably, on this account, the ground of the proposition that it is possible that there are unicorns is the having of some property by the proposition that there are unicorns. But how does the having of some abstract property by some abstract proposition relate to the possibility of there being unicorns? I shall argue that this is an insoluble problem if one limits oneself to the resources of this Platonic theory.

Moreover, there is a mystery as to the ontological status of propositions. Are they substances? If so, what sorts of substances are they? Someone who

is not enamored of Platonism will shrink from propositions. But on the positive side, propositional and linguistic approaches both avoid the paradoxes that plague Lewis's theory, and the propositional approach is not tied ad hoc to a particular language.

4.5 *Aristotle again and branching*

While one of Aristotle's notions of modality was seen to be unsatisfactory, there is also another implicit in his work to choose from. Parmenides was worried that change involved something's coming to be out of nothing. For when A comes to exist, then earlier A did not exist. To answer this concern, Aristotle developed his tripartite account of change. There is a substance, a form and a privation. In the case of generation, the substance goes from having a privation of a form to having that form. Thus, a man may go from having a privation of beardedness to having a beard. But the beard does not come from nowhere. Rather, the man at the beginning of the process of change was *potentially* bearded, though *actually* clean-shaven. The privation that he had not a *mere* lack: it was a potentiality for beardedness.

On this account, there is something in the substance which can be identified as a potentiality for the alternate states of the substance. If we further accept the general Aristotelian thesis that potentiality is grounded in actuality, we have to say that there is something *actual* in the substance in virtue of which that substance can change. But this account not only helps to solve the Parmenidean puzzle about change, but it may also help with the extended Parmenidean puzzle about the grounds of modality. Even if I never grow a beard, it is true to say it is possible for me to grow a beard because there is in me and in the environment around me something in virtue of which the growing of a beard is possible, say, a power (of course further scientifically analyzable) in the hair-follicles on my chin to produce hairs together with the capability for refraining from shaving. The ground, on an Aristotelian account, of the proposition that it is possible for me to have a beard is to be found in such powers or capabilities. (I prefer the term "capability" to the term "power," since "capability" immediately implies a capability for a particular activity or effect. The term "power" seems more common in the literature. I shall use these terms interchangeably, except when talking of Shoemaker's "powers.")

As an account of modality in general, this appears insufficient. For one, at first sight it only applies to local *de re* modalities. This approach will not give us possible worlds in any obvious way. Moreover, the account is

not reductive, since it accounts for modality in terms of ability, and ability is a modal term. However, while ultimately not reductive, the account is illuminating. For in ordinary language, the notion of capability or ability is arguably more basic than that of metaphysical possibility (cf. Place 1997) and we obtain the general notion of possibility by reflecting on ability. We have personal knowledge of abilities. For instance, as Kant outlines in the second *Critique*, we recognize ourselves as morally responsible for an evil act and thus as having been capable of doing otherwise. There is also less mystery about capability than there is about modality in general since capabilities are actual properties of actually existing things, and so a capability account of modality is indeed helpful. And at the very least, if this approach worked, it would reduce modal talk in general to a particular subset of it.

Aristotle's account when generalized in a global way may lead to branching theories of modality (see, e.g. Mackie 1998). When a substance has more than one alternative before it, these alternatives can be thought of as presenting a world-branch, though unless we want to make Lewis's move of making all worlds concretely existent, we should not think of there being concretely existing worlds corresponding to all branches. If we look at all modality as induced by such branchings, so that we see a proposition as possible if it is true somewhere on the full tree and necessary if true everywhere, then we will in fact be building on Aristotle's account of the change in a single substance.

There is, however, still a problem: in fact, the same problem as was the decisive consideration against Aristotle's temporalized approach. It is, it seems, possible for all events in time to have been different — while there might be necessary entities such as mathematicals, there surely are no necessary temporal events. This intuition can be thought of as a constraint on theories of possibility, one that it seems prima facie difficult to accommodate in a branching theory, unless there is a branching outside of time.

4.6 *Leibniz*

Leibniz gave life to the notion of possible worlds. On his view, God necessarily exists, and possible worlds are maximal self-consistent ideas or concepts in God's mind. One could also talk of these worlds as maximal self-consistent thoughts entertained by the divine mind, and this would for all practical purposes be equivalent.

Leibniz in fact gave an argument for the existence of God from the existence of necessary truths, and hence from the existence of modal truths (assertions

of necessity and possibility are necessary truths by S5, and Leibniz appears to accept S5). Necessary truths, Leibniz argues, must be grounded in some reality, and the only reality Leibniz can see as capable of this is a necessarily existent mind. Of course, the argument leaves much for discussion. Why can't the necessary truths be grounded in the thoughts of different minds in different worlds? Why can't they be self-subsistent in a Platonic way?

Positing divine ideas as possible worlds gives one the benefits of the linguistic and propositional theories of possible worlds. Like the linguistic theory, this approach allows something that we all have some ordinary pre-theoretical understanding of, the ideas of a mind, to constitute the collection of possible worlds. Admittedly, mystery is introduced by the fact that these are the ideas of a *divine* mind and by the maximality involved in them. But at least we do not have the dark mystery of the Platonic propositions, whose ontological status is almost completely opaque — for what indeed are "abstract" things? And like the propositional theory, the possible worlds have a representational power that a merely human-like language will not.

Of course, the theory does carry the ontological commitments of theism. But it is not revisionary of our ethical and epistemological notions in the way that Lewis's theory will be seen to be. Moreover, there is independent evidence for these commitments, namely all the evidence for theism.[17]

Unfortunately, Leibniz's account fails to answer the need for truth grounds — the extended Parmenidean worry. Granted, the proposition that such and such a world is possible is an idea in the mind of a necessarily existent God. But what makes *this* idea true if the world in question is non-actual? And, in virtue of what is the idea that is identical with a possible world *self-consistent*?

In Part VI, I shall argue that the limitations of Leibniz's approach can be neatly supplied by the merits of the Aristotelian approach in the previous section. The resulting theory will be the most satisfactory account of the above accounts of the nature of possible worlds and the meaning of modal propositions.

4.7 Modal irrealism

The modal irrealist, on the other hand, denies the objective existence of modal facts. She might be a subjectivist. Modal facts are simply up to us. Or she might simply think that the modal notions make no sense at all.

---

17   See, e.g. Swinburne (2004) for a survey of the evidence.

However, modal claims are relevant to other issues. Here is a scattering of cases. First some ethical ones. It is very plausible that an action is only obligatory if it is *possible*. While probably everyone should admit that a wrong is done to someone by conceiving her in circumstances in which one is unable to care for her in the way in which she deserves, nonetheless the question of what exact wrong is done depends on whether this very person *could have been* conceived in other circumstances. Promises of *impossibilia* are not morally binding, and do not require one to request release from the promise; when one has promised something impossible, one should not apologize for non-fulfillment, but for making the promise. Counterfactuals appear to be important to evaluating consequentialist moral claims, and counterfactuals may be modal animals. And it is widely held that two situations *could not* be the same in all non-moral respects and different in moral respects. We seek objective answers to questions about what would happen if we did something, answers we need in order to decide how to act.

Contemporary physics is often couched in terms of stochastic laws that entail the physical possibility of non-actual states of affairs, since the laws assign non-zero physical probabilities to such states of affairs, and anything physically impossible has zero physical probability. But, surely, anything that is physically possible must be *metaphysically possible* as well.

If ethical or scientific concepts require or entail modal truths and are objective, and arguably they play a role that only objective concepts could, then the required modal truths will be objective as well. For instance, if modal truths were merely a matter of our (individual or societal) fiat, we could set the truth value of the counterfactual "Had Jones got up earlier, she still would have been late for work" according to our (individual or societal) preference: if we wanted to blame Jones for her lateness, we would say it's false, and if we wanted to excuse her, we would say it's true.

Another way to argue against the irrealist in general is to point out particular modal truths whose objectivity appears obvious. Napoleon could have won at Waterloo. Had Napoleon won at Waterloo, it still would have been the case that he could have lost. It is possible that Napoleon lost at Waterloo. It is necessary that if he did not lose, he did not lose. Had Napoleon won at Waterloo, he still would have been a short man. It is possible for any ordinary sugar cube to melt. It is possible that I will not finish this book. It is necessary that either someone will read the previous sentence or no one will read it. I cannot become a point in space or the number seven. Necessarily, every horse is a mammal. Necessarily, if there are no modal truths, then the irrealist's view is not a necessary truth. Possibly, there was a man named "Sherlock Holmes." And of course we can now make a familiar move. Any

premises the irrealist points out from which she argues against the existence of modal truths are unlikely to have as good warrant as these claims, and any premises the irrealist argues from against the objectivity of modal truths are unlikely to be as obvious as the objectivity of the above claims.

One can also conceive this book is an answer to one powerful irrealist argument, the argument that we cannot make ontological sense of modality, an argument parallel to the argument for moral irrealism from the queerness of moral facts. For we shall see that there is an ontology on which modality makes sense.

Another objection to modal facts is epistemological. We observe actual, occurrent facts. We do not observe modal truths. But as Nancy Cartwright (2001: 70) points out, such an objection is tied to an outmoded epistemology of sense-data. Once we have rejected the myth of a pure realm of sense-data, of shapes and color patches, we have no reason to deny that we might in fact observe modal truths. The experienced coach might say of someone practicing: "I can *see* that she can lift ten pounds more." No coach concludes this from propositions about apparent color patches. The inexperienced coach bases her judgment, not on observations of color patches but of bodies, arms, muscle flexings, and so on, while the experienced coach can just *see* how much more the athlete can do. And if we insist on a causal role for belief formation, we can say that just as the width of the trainee's arms causes the coach to acquire certain beliefs, so too the power of her arms causes the coach to acquire other beliefs. In both cases there is a causal sequence. The width of the trainee's arms is a partial cause of certain reflections of light. The power of the arms is a partial cause of certain flexings which in turn are a partial cause of certain reflections of light. There is a difference in the causal chains, but in both cases we in fact have long causal chains (the light needs to be refracted in the eyes, trigger electrical impulses in rods and cones, etc.), and there is no *a priori* reason to rule one of them out as perhaps being partially constitutive of perception. This only gives us a start towards an epistemology of *causal* modality, but it is enough to show that some modal truths can be known in an empirically respectable way. And, fortunately, as we shall learn by the end of the book, causal modality is what we are looking for when we are looking for metaphysical modality.

4.8 *Conventionalism*

A different way to evade some of the metaphysical questions is modal conventionalism. A naïve formulation is:

(11)   Convention has set the meaning of "is possible" and
       possibility is in fact the property that we attribute to
       propositions when we apply the predicate "is possible" to
       them.

This is not the view that we can analyze "*p* is possible" as "*p* has the
property that we predicate of propositions with the predicate 'is possible.'"
For then square circles could be possible, *pace* S5, since there are possible
worlds where the words "*p* is possible" mean that *p* is impossible. Rather
the second conjunct in (11) claims that possibility is the property that we
actually attribute with "is possible." In some other worlds, other properties
are attributed with these words, but those properties are not possibility.

Claim (11) is the conjunction of two claims: (a) "is possible" received
its meaning by convention, and I know no one who denies that;[18] and (b)
the meaning of the English words "is possible" is set by convention and we
attribute *possibility* with "is possible," which no realist with a mastery of
English can deny. In particular, the major parties in the modality debate,
maybe with the exception of modal irrealists, agree with (11). So, if (11) is
the conventionalist position, it is not an *alternative* to metaphysical theories.

Perhaps, however, we should take the conventionalist to supplement (11)
with:

(12)   There is no substantive characterization of possibility or
       necessity besides something like (11).

First, observe that a deflationary theory of the above sort is *excessively*
deflationary. Compare with the deflationary theory of truth. According to
Horwich (2005), truth is completely characterized by the proposition that
only propositions are true, together with the axiom schema:

---

18 I would not be surprised if there were someone who thought that the English language
    is designed by a supernatural being who non-conventionally assigned meanings to
    words, but I would be surprised if there were a philosopher who thought that, and
    I, in fact, know no one who thinks it. There are theistic philosophers who find it
    plausible that the *first* human words were non-conventionally assigned their meanings
    by God, and it is not completely beyond the bounds of empirical possibility that some
    word of English like "mama" that is used by infants has a non-conventional meaning.
    But the English word "possible" was not a part of the first human language (though
    I suppose a cognate might have been, though I wouldn't count on it) nor is it a word
    used by infants.

(13)   The proposition that *s* is true if and only if *s*

(with instances that lead to the liar paradox omitted). But observe a crucial feature of the deflationary theory of truth: it does present a genuine substantive axiom schema for truth. Granted, the instances of the schema are all uncontroversial, but uncontroversiality of axioms is a plus. But what would be directly analogous to (11) is not the above deflationary theory of truth, but the claim:

(14)   Convention has set the meaning of "is true" and truth is in fact the property that we attribute to propositions when we apply the predicate "is true" to them.

The suggested deflationary theory of possibility is radically less substantive than the deflationary theory of truth, since the attractiveness of, and argument for, Horwich's theory of truth is that so much work can be done by the axiom schema (13).

Second, we can ask for the *evidence* for (12). Perhaps *some* evidence would be provided by the failure of all extant substantive theories, if indeed they all fail. But I hope that the reader who completes this book will see that not all extant accounts fail — and in particular that the account of *this* book does not fail. However, even if all extant attempted theories do in fact fail, that is quite weak evidence for the claim that no theory *can* succeed. It would be premature to conclude from the failure of all extant attempts at proving or disproving the Goldbach Conjecture that the Goldbach Conjecture can be neither proved nor disproved. Yes, the failure of attempts incrementally confirms the insolubility thesis, but the degree of confirmation is small. Likewise, in scientific matters we ought not assign much weight to arguments of the form: *all proposed scientific accounts of phenomenon* Φ *fail, so phenomenon* Φ *has no account*. The failure of a significant number of past attempts by smart people gives significant evidence mainly to two claims: (a) the problem is difficult; and (b) new ideas are needed.

One might attempt to motivate the deflationary view without using a pessimistic induction. Instead, one might argue that the modal cannot be reduced to the non-modal. But even if a successful argument could be given for that thesis, it does not follow that there can be no substantive account of possibility and necessity. All that follows is that there can be no substantive *non-modal* account of possibility and necessity, but it is left open that possibility and necessity could be given an account in terms of some more fundamental modal concepts, or that a privileged set of possibilities and

necessities could be given which generate all others in some explanatorily helpful way. Here is a simple illustration of the first approach — the other main illustration will be the Aristotelian account defended in this book. Suppose that subjunctive conditionals involve a more fundamental and less mysterious modal concept than possibility and necessity. Then one might say that a proposition q is necessary if and only if for every proposition p it is the case that were p to hold, q would hold. How plausible one finds this account will depend on what logic of counterfactuals ones thinks is right, and whether subjunctive conditionals are less mysterious and/or more fundamental than possibility and necessity. I do not find this toy account plausible — for instance, I think subjunctive conditionals are too context-sensitive to do the job here — but it shows that the idea of explaining one set of modal notions in terms of another is not absurd.

Ross Cameron (2008), building on the work of Sider (2003) and Peacocke (1999), offers a somewhat different version of the deflationary view. Instead of saying that there is no substantive account of modality available, Cameron says that *possibility* might turn out not to be a *natural* property. Here, the naturalness of properties is to be understood in Lewis's (1983a) sense in which the property of being positively charged is more natural than the disjunctive property of being positively charged or having spin up, and the property of being green is more natural than the property of being green if it's before the year 2100 and blue otherwise. It could be that the concept of possibility is just a gruesome mess.

Here, for instance, is an account of modality that really is such a mess. Let us work in a sufficiently rich deductive system D. Then say that a proposition is possible provided it can't be disproved from D conjoined with a collection C of axioms, where there is no unity among the axioms in C. For instance, one could put in as one's axioms (a) all mathematical truths (whether provable or not); (b) the claim that the initial material constitution of any material object is essential to it; (c) the claim that any collection of objects has a fusion; and (d) all axioms generated by the schema "the property of *F*ness exists and an object is *F* if and only if it instantiates *F*ness," where *F* ranges over all predicates or all privileged predicates, and so on.

There are two ways to produce this sort of deflationary account. One approach is to explicitly specify the members of C, or provide a schema that specifies them, and be hopeful that there are no undiscovered families of axioms that need to be added. But there is little reason for this hope — metaphysical discoveries, like Kripke's discovery of the essentiality of origins (assuming he is right about that), surely continue to be made. A second approach is simply to assert that the correct account has some such messy

form, without actually giving what one thinks the account is.[19] Or, more weakly, one could claim that for all that we know, the correct account has such a messiness.

However, possibility and necessity appear to be fairly natural. For instance, they appear to enter into explanations in a way in which we do not expect quite unnatural properties to do so. There are no square circles because square circles are *impossible*. George is not obligated to square a circle because it is *impossible* to square a circle and one *can* only be obligated to do something *possible*.

Moreover, the best of the substantive accounts of possibility — say, Lewis's account — do make possibility and necessity be fairly natural as properties. To positively claim that possibility is quite unnatural as a property would require the failure of these accounts (and Cameron does endorse the claim that the extant accounts fail). But in this book we shall develop an account that does not fail in this way, despite Cameron's criticisms.

Furthermore, if possibility and necessity are gruesome messes, then it is not all that likely that there is such a thing as *the* properties of metaphysical possibility and necessity. For instance, Peacocke can be read as simply including a version of the essentiality of origins among the axioms generating his account of possibility. But there are philosophers who deny any doctrine of the essentiality of origins. If Peacocke's account makes possibility into a mess without some sort of unity, and essentiality of origins is simply built into the account, then probably the thing to say about philosophers who deny the essentiality of origins is just that they are attributing a different property with the predicate "is possible" than the property that Peacocke attributes. But then the deflationary will be unduly eirenic. There really does appear to be a substantial metaphysical disagreement about essentiality of origins.

---

19 Peacocke (1999) does something intermediate between the first two approaches. He gives a particular account but is open to its being supplemented. Cameron (2008) seems closer to the second approach.

PART II

APPLICATIONS AND PSEUDO-APPLICATIONS

Section 1 Modality

1.1 Box and diamond

The most obvious application of possible worlds is to furthering the under-standing of modal claims. This can be either as a useful logical device to make it easier to grasp complex modal assertions and maybe even to express assertions that cannot be otherwise expressed, or one may more ambitiously see possible worlds theories as giving an *analysis* of all modal claims or at least as being items that are closely connected with the grounds of modal claims. Whether one can take the more ambitious approach or not depends on whether one's construction of possible worlds presupposes the modal claims one wishes to analyze or not

The main account of possible worlds, which does not appear to presup-pose any sort of modality and hence that supports the more ambitious use of possible worlds, is that of David Lewis. Unfortunately, I shall argue in Part III that it leads to too many paradoxes for it to be at all acceptable. Just having counterintuitive consequences is not enough to refute a view, but the sheer number and weight of these in the case of Lewis's system is enough. Just as Lewis's case for his account is a cumulative one based on the multiplicity applications, my case against his account is a cumulative one based on the multiplicity of serious paradoxes.

The general way in which modal claims are expressed in terms of possible worlds is by quantifying over all worlds: for instance, $\Box p$ holds if and only if $\forall w(p$ is true at $w)$, while $\Diamond p$ holds if and only if $\exists w(p$ is true at $w)$. But as it turns out, the expressive power of possible worlds goes beyond box and diamond operators as Lewis (1986a, Section 1.2) claims and Melia (1992) proves (see also Section 2.7.3 of Part IV for a proof of a similar result). We shall see some plausible examples in Sections 1.3 and 1.4.

1.2 *The global nature of modal claims*

Moreover, as mentioned in Section 1 of Part I, the notion of a possible world is correlated with our intuition that the box and diamond modalities have a global component. To tell whether some proposition is possible, it is plausible that one has to think whether it could be made to fit into a story of a whole world, unless one has Humean intuitions that a world is made up out of parts such that any part is compatible with any other part (maybe with the qualifier: considerations of shape space and time permitting).

There is good reason to reject a view of possibility that does not have the resources for discussing global possibilities of some sort. Many ordinary language modal claims are of an apparently local nature and for disambiguation require globalization, as we saw in the case of "Hitler might not have been born" in Section 1. The ordinary language assertion "Hitler might not have been born" does not simply claim that the proposition <Hitler was not born> is logically possible. The ordinary language modal claim is not made true by an empty world, or a world with laws radically different from ours, or a world like ours except that there are no planets and there is no life. Rather, we are claiming that <Hitler was not born> is true *in a possible world very much like ours*. This requires one to talk of worlds *as a whole* rather than piecemeal of the possibility or necessity of an isolated proposition.

1.3 *Supervenience*

Another standard example of the usefulness of possible worlds is provided by the notion of supervenience: *A*-type states of affairs (modally) supervene on *B*-type states of affairs (the *locus classicus* being the claim that goodness supervenes on natural facts — see Hare [1964: 80ff]) if and only if any two worlds which are indistinguishable in respect of *B* are indistinguishable in respect of *A*. David Lewis (1986a, Section 1.2) has argued that in fact such claims cannot be expressed with ordinary box and diamond operators. If so, then possible worlds are indeed a useful tool.

Of course one could also do the same thing with quantifications over "aspects" and occurrent states of affairs:

(15) $\Box(\forall a \forall b((a$ is an *A*-type state of affairs and *a* obtains and *b* is the whole *B*-aspect of the actual world) $\supset \Box(a$ does not obtain $\supset b$ does not obtain))).

However, if we take (15) to be an analysis of the claim that A-type states of affairs supervene on B-type states of affairs, we have not gained much over a possible worlds analysis, since we have allowed for quantification over complete aspects of worlds.

And if one thinks that the modal account of supervenience is insufficient, because one thinks supervenience should be an asymmetrical relation (if the A-type states of affairs supervene on the B-type ones, the B-type ones can't supervene on the A-type ones), one can supplement the modal definition, for instance by adding that an appropriate explanatory relation holds.

1.4 *Transworld comparison*

One might wish to define the notion of, say, x's being an entity than which no greater is possible or a picture that which no picture can be uglier. For instance, take the Anselmian case. As Lewis (1970) has noted, the notion of maximal greatness is prima facie ambiguous. One could reasonably understand the claim "nothing greater than x is possible" as any one of the following:

(16) $\forall w \forall y((y$ exists in w and x exists in $w) \supset (y$ is not greater at w than x is at $w))$

(17) $\forall w \forall y((y$ exists in $w) \supset (x$ exists in w and y is not greater at w than x is at $w))$

(18) $\forall w \forall y((y$ exists in $w) \supset (y$ is not greater at w than x is at the actual world)).

Arguably, (18) is the best interpretation in an Anselmian context. Of course if one allows oneself quantification over greatnesses, then one can do without possible worlds even in (18), just as if one allows quantification over aspects, one can do without possible worlds in analyzing supervenience. Thus, (16)–(18) appear equivalent to:

(19) $\square(x$ exists $\supset \forall y(y$ is not greater than x is))

(20) $\square(\forall y(x$ exists and y is not greater than x is))

(21) $\forall g(g$ is the actual greatness of $x \supset \square(\forall y(y$ does not have greatness exceeding $g)))$.

Note, however, that introducing quantifications over greatnesses or aspects is moving to a second order logic. Given possible worlds on the ground level,

one could do all this in first order logic. Moreover, one may plausibly argue that (18) is not only easier to understand than (21) but is closer to what is *meant* by the assertion that nothing greater than x is conceivable. For, it seems more natural to say that we are comparing individuals in respect of greatness rather than comparing greatnesses.

Section 2 Counterfactuals and causality

2.1 *Lewis's account of counterfactuals and counterexamples*

Perhaps the biggest feather in the possible worlds theorist's cap would be the Lewisian analysis of counterfactuals, if this analysis were correct:

(22)   "A counterfactual 'If it were that A, then it would be that C' is (non-vacuously) true if and only if some (accessible) world where both A and C are true is more similar to our actual world, overall, than is any world where A is true but C false."[1]

When there is a world w where both A and C are true and that is more similar to the actual world than any world where A is true but C is false, I shall say that w "witnesses" to the counterfactual "If it were that A, then it would be that B."[2]

The challenge then is to come up with an account of similarity that assigns correct truth values to counterfactuals. Unfortunately, Lewis's own account of similarity fails at this task. The problem is that in ordinary cases (not ones involving time travel, say), we reject counterfactuals of the form "If event A were to happen at t_A, then event B would happen at t_B," when t_B precedes t_A and B did not actually happen. One reason we reject such counterfactuals is our intuition that whatever happens now cannot affect the past. The past is what it is, and it is to be kept fixed in our ordinary counterfactuals.

The future, on the other hand, would in many respects be different had the present been different. There is thus a temporal asymmetry in counterfactuals. But there is none in the Lewisian possible-worlds analysis. Nonetheless, Lewis (1979a) claims the asymmetry in counterfactuals can be derived from

---

1   Lewis (1979a: 465). I have argued in Pruss (2007) that the account should be complicated somewhat, but those complications do not need to concern us here.

2   The terminology is due to Steve Kuhn, personal communication.

his possible-worlds account of counterfactuals as long as we assume that we are dealing with a world whose contingent distribution of matter is like that of our world. Moreover, Lewis claims he can derive the asymmetry even if the laws of nature are two-sidedly deterministic, i.e. a full description of the world on any one time-slice fully determines all the states of the world before and after this time-slice. Further, Lewis thinks he can still do this if, further, these laws are time-reversal symmetric so that reversing the order of states, and making appropriate adjustments like replacing a velocity v with $-v$, preserves nomicity.[3] If correct, this would be quite a surprising result.

To get out of an objection of Fine's[4] to his account of counterfactuals, Lewis proposes a measure of similarity of worlds that has four factors ranked as follows:

(23) "It is of the first importance to avoid big, widespread, diverse violations of [physical] law."

(24) "It is of the second importance to maximize the spatio-temporal region throughout which perfect match of particular fact prevails."

(25) "It is of the third importance to avoid even small, localized, simple violations of law."

(26) "It is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly."[5]

One might of course have objections to these four factors and/or to their mutual ordering.[6] But even if they are implausible, it would be very impressive if Lewis could derive a time-asymmetry from them and from his definition of counterfactuals, since (23)–(26) are time-reversal symmetric, as is (22).

However, there is an intuitive argument that Lewis's measure of similarity

3 Contemporary physics suggests an invariance under simultaneous reversal of time, charge and parity.

4 Fine (1975).

5 Lewis (1979a: 472).

6 For instance, intuitively one might think that very close approximate similarity of particular fact over a much larger area (factor 4) outweighs exact match throughout a small area. The world which is exactly like ours for all times in our future but whose past was *radically* different from the past of the actual world is surely further from our world than is a world which is the same as ours for all time except that the background radiation is everywhere and everywhen about $10^{-1,000,000,000}$ percent higher.

of worlds is likely to give the wrong answers. The past is shorter than the future. Astrophysics suggests that either our universe will keep on expanding forever or it will eventually collapse. If it will keep on expanding forever, then the past, which as far as we can tell is finite, is infinitely shorter than the future. If it will eventually stop expanding, and collapse back, then the past is presumably still shorter than the future as we haven't yet reached the point where the expansion turns to collapse. So, condition (24) favors worlds that exactly match the actual world over the whole of the future over worlds that exactly match the actual world over the whole of the past. Thus, condition (24) is apt to introduce an asymmetry that goes the wrong way.

Moreover, there is a definitive counterexample to Lewis. Elga (2001) considers the scenario where "[a]t 8:00, Gretta cracked open an egg onto a hot frying pan," and considers the plausibly true counterfactual: "If Gretta hadn't cracked the egg onto the pan, then at 8:05 there wouldn't have been a cooked egg on the pan." Now, imagine running the whole egg-cracking and frying process in reverse, backwards from 8:05 down to 8:00, starting with time-reverse of state of the cooked and cracked egg at 8:05, then the egg uncooking, the shells coming together, and finally the egg flying up into Gretta's hand. A process whereby broken shells come together to form an egg is evidently very sensitive to initial conditions, since it is a process in which entropy *decreases*, and such processes are highly improbable. A slight variation in the "initial" conditions will ensure such a process will not occur. Thus, if we slightly vary the "initial" conditions for the time-reverse of the breaking-and-cooking from being those conditions that the time-reverse of the state at 8:05 in fact satisfies, we will not arrive at a solid shell. Thus we can insert a small miracle into the time-reversed process to ensure that it does not arrive at a solid shell: this small miracle makes the world different, but not very different — small miracles only introduce a difference in respect of the *third* of Lewis's criteria of similarity. Then consider the time reversed world, w_1^*, where this modified process occurs. It differs only by a small miracle from the time reverse, w_0^*, of our world, w_0. Thus, the time reverse of w_1^*, which we can denote by w_1^{**}, differs only by a small miracle from our actual word, w_0. But now w_1^{**} is a world with laws very much like ours except for a small deviation around 8:05, and exactly matching our world after 8:05, but where there never was a solid egg and hence Gretta had never cracked it.

Moreover, w_1^{**} is extremely close to our world, and indeed closer to our world than a world like w_1, where Gretta does not crack the egg and where there is no cooked egg in the pan. For in w_1 there may have to be a medium-sized miracle inserted before 8:00 to ensure that Gretta does not

crack the egg (imagine that Greta really wants to crack the egg), while in w_1^{**} the only deviation from the laws of nature is a *small* shift in the position of a shell. Moreover, w_1 exactly matches our world in the past of these events while w_1^{**} matches it in the future, and as we noted, the future is bigger than the past. Consequently, if the case is correctly set up, a world like w_1 will fail to be a witness to "Were Gretta not to have cracked the egg, there would not have been a cooked egg." And it is hard to see what other witness there could be. Hence, the counterfactual is false on Lewis's analysis, though surely it is in fact true.

Lewis could try to restrict his account to cases where the antecedent of the counterfactual reports a *positive* state of affairs, but another example could then be given (Pruss 2003). Besides, Lewis's account of causation requires counterfactuals whose antecedents are negative states of affairs.

None of this contradicts the plausible claims (a) that there is an asymmetry in our counterfactuals, even though Lewis's analysis has failed to give a proper objective grounding to it; (b) that this asymmetry may well be responsible for the future counting as open and the past as closed; and (c) that there may be an asymmetry in causal overdetermination. But the asymmetry in counterfactuals remains unexplained at this point.

2.2 *A fix and the coat thief problem*

We can make Lewis's account yield the right values for the problematic counterfactuals above by simply building the arrow of time explicitly into the definition of the counterfactual. To do that, we simply demand that any worlds invoked in the analysis should closely match the actual world in the past (cf. Davis 1979), or the past light cone, perhaps, of the event described in the antecedent except perhaps for the very recent past.[7] Though if we do this, we can no longer use counterfactuals to explain the arrow of time as Lewis (1979a) attempts to do.

Here is one way to make this somewhat precise. For some antecedents *A* of counterfactuals, we can talk of "the time of *A*" in a non-actual world *w*, and denote it by "time(*A*, *w*)." When *A* reports simply the occurrence of an event *E* that occurs at *w*, time(*A*, *w*) is the time at which *E* begins at *w*.

---

7   If we are willing to confine our account to worlds that are indeterministic, on the grounds that counterfactuals in deterministic worlds are inescapably problematic, this exception can be omitted. If we keep it, we can either make it a restriction, or weight the value of match in a time-dependent way, with earlier events weighted much less than later ones.

And when A reports the non-occurrence of an actual event E, then time(A, w) is stipulated to be the time of E at the actual world. In cases where A does not report the occurrence or non-occurrence of an event, and where there is no good way to extend the above definitions, we can leave the old Lewisian account unchanged. Suppose from now on that we can make sense of time (A, w).

Next, for any world w, define $M(A, w)$ as the largest number u such that w and the actual world exactly match at all times prior to time(A, w)$-u$, if there is any such u, and otherwise set $M(A, w) = 0$. We then say that when measuring the closeness of worlds w_1 and w_2 for purposes of evaluating a counterfactual with antecedent A, that w_1 is automatically closer to the actual world than w_2 if $M(A, w_1) > M(A, w_2)$. In other words, we privilege worlds that initially exactly match the actual world up to as close as possible before the time of A. We then restrict (24) to comparing the regions of exact match to the actual world of worlds w_1 and w_2 only *after* the later of time(A, w_1) and time(A, w_2). More precisely, our modified statement of Lewis's criteria, in the case where it makes sense to talk of the time of A, is:

(27)   If one of $M(A, w_1)$ and $M(A, w_2)$ is greater, then the corresponding world is closer to the actual world.

(28)   If the antecedent of the preceding condition is not satisfied, and one of w_1 and w_2 does a better job "avoid[ing] big, widespread, diverse violations of [physical] law," then that world is closer to the actual world.

(29)   If the antecedents of the preceding conditions are all unsatisfied, and one of w_1 and w_2 has a greater extent of the spatio-temporal region of "perfect match of particular fact" after the later of time(A, w_1) and time(A, w_2), then that world is closer.

(30)   If the antecedents of the preceding conditions are all unsatisfied, and one of w_1 and w_2 does a better job avoiding "small, localized, simple violations of law," that world is closer.

(31)   If the antecedents of the preceding conditions are all unsatisfied, and one of w_1 and w_2 does a better job with respect to "approximate similarity of particular fact," it *might* count as closer, depending on context.

Because the perverse world w_1^{**} in our exposition of Elga's counterexample to Lewis does *not* match the actual world anywhere in the past of 8:00, and

hence $M(\text{antecedent}, w_1^{**}) = 0$, the account survives Elga's counterexample. It also appears to survive some other known counterexamples (namely, those of Pruss 2003), and the worry that the actual past is shorter than the actual future is neutralized.

The present proposal may seem unduly complex. Would it not be easier to eschew all mention of the time of the antecedent, and the complications in modifying (24), and simply say that if w_1 exactly matches the actual world over a longer initial sequence of times than w_2 does, then w_1 is closer than w_2, and then use Lewis's original conditions? But then we would run into the coat thief problem which is inspired by a case of John Pollock (see Bennett 1984; Edgington 1995). Suppose two people enter a room, one before another, before noon. My coat was in the room all along, and at noon it is still there. Then if we either privilege length of initial match or unrestricted extent of spatio-temporal match, it will be automatically true that were one of these two people to have stolen the coat, it would have been the later one. But if the later one is someone known for her moral rectitude while the earlier is a notorious coat thief, then it is clearly mistaken to give an account of counterfactuals that automatically makes the later one the counterfactual thief.

It is likely that even with the above modifications to Lewis's account, a counterexample can be found. However, even if so, something may survive of the Lewis–Stalnaker approach to counterfactuals: to check whether it is true that were A to hold, B would hold, we need to examine the *relevant* A-worlds (i.e. the worlds where A holds) with the B-worlds, whether the relevance is defined in terms of similarity or not, and check if B holds at the relevant A worlds. Indeed, this may be pretty close to how we intuitively think about counterfactuals. The notion of relevance, however, will be time-reversal asymmetric: perverse worlds like w_1^{**} will rarely count as relevant.

2.3 *Causation and the order of time*

2.3.1 The basic argument

To fix Lewis's account of counterfactuals, we needed to make use of the order of time. But it is very plausible that the order of time derives from the predominant direction of causation. And if the order of time derives from the predominant direction of causation, then causation cannot be analyzed in terms of counterfactuals. I shall argue for the major premise shortly, because the conclusion that causation is not susceptible to a Lewisian analysis is important for forestalling a potential objection to the Aristotelian

account that I shall eventually give of modality. My account shall presuppose the concept of causation. If Lewis's analysis of causation were to work, that account would be circular. But if the order of time depends on the order of causation, then there is no hope for Lewis's analysis.

The most promising extant accounts of the direction of time are that arrow of time is (a) primitive; or (b) illusory; or (c) derives from the direction of increase of entropy; or (d) derives from the directionality of counterfactuals; or (e) derives from the direction of causation; or (f) derives from the openness or unreality of the future. Option (d) is either hopeless or collapses into (e) unless we have a plausible account of counterfactuals independent of both the direction of time and that of causation, and it is hard to imagine a plausible such view in light of the examples of Elga (2001) and Pruss (2003).[8] Option (b) will not help with a counterfactual analysis of causation. Option (e) is what we are arguing for. That leaves options (a), (c), and (f) to argue against. But before doing that, I shall sketch a quick plausibility argument in favor of option (e).

The Special and General Theories of Relativity have changed how we think about time. On the standard reading of Relativity, simultaneity only makes sense relative to a reference frame. But now there comes a question of how we should understand the *earlier-than* and *later-than* relations. One could understand them as relative to a reference frame. But it has also proved natural to introduce the notion of the absolute future and absolute past of a space-time point x. The absolute future (past) is the future (past) half of the light-cone with vertex at x (not including x itself). And it is very natural to think that a space-time point y is absolutely earlier (later) than a x provided that y lies in the past (future) light-cone centered on x. Even if neither the Special nor the General Theory of Relativity turns out to be true, it is plausible that *if* they were true, that would be the way to think about temporal relations.

And this plausibility judgment calls out for an explanation. After all, why should *light-cones* be central to the understanding of temporal relations? Why not pyramids, or spheres, or cones of some other sort, or complements of cones? Basically, the light-cone centered on x is the collection of points that can be reached from x by traveling at or below the speed of light or from which x can be reached by traveling at or below the speed of light. But why should *that* collection matter? Certainly, I can define the set of points in space-time that can be reached from my present location by traveling at or

---

8   Bennett (1984) attempts such an account, but it is shown in Pruss (2003) that it falls
        prey to a variant of Elga's example.

below Obama's highest running speed and from which I can be reached by traveling at or below Obama's highest running speed, but I am not tempted to use that set as part of the definition of past and future. What makes the speed of light special? I submit that it is *causation*: it is not physically possible for causality to propagate at any higher speed. And that is why we chose light-cones rather than some other regions of space-time to base the concepts of *earlier-than* and *later-than* on.

Thus, the light-cone centered on x is the set of points y in space-time such that (i) an event at y physically could be causally affected by an event at x, or (ii) an event at y physically could causally affect an event at x, or (iii) $y=x$. And once we understand a light-cone this way, we also have a template for extending the concept to other kinds of physics. For instance, Newtonian physics has no speed limit, and thus the set of points satisfying at least one of (i), (ii) and (iii) will yield all of space-time in a Newtonian framework.

The above physically significant causally-based definition of a light-cone is already disjunctive and it is now very natural simply to take (i) to characterize the future of x (perhaps with an added qualifier that $y \neq x$, if simultaneous causation is possible) and to take (ii) to characterize the past of x (with the same added qualifier if necessary). One might also wish to replace "physically could" by "normally physically could," to allow for possible worlds where time-travel happens in exceptional circumstances.

Now let us quickly sketch why the present proposal is preferable to (a), (c) and (f).

2.3.2 Taking the direction of time as primitive
Primitiveness should in general be a last resort. However, the causalist about the direction of time takes causality as primitive, while the primitivist about the arrow of time can explain causation by using an improved version of Lewis's story that privileges past similarity in evaluating counterfactuals. So each takes one notion as primitive and explains the other in terms of it. The primitivist does, however, have to explain why it is that we take light-cones, whose significance appears to be causal, to be central to a relativistic understanding of temporal relations, and that is a count against her view, and so primitivism about causation seems preferable to primitivism about the direction of time.

2.3.3 Deriving the direction of time from increase of entropy
The entropy in our world is increasing with time. One might take this to be the defining feature of the direction of time: the direction of time is that direction which is the predominant direction of the increase in entropy (Kutach 2002).

But now suppose that the universe began with the Big Bang 12 billion years ago, and consider the following hypothesis, which, in a universe with massive quantum indeterminacy, will be physically possible but exceedingly unlikely. In the year 5 billion AD, the entropy starts to decrease, and it continues to decrease for 30 billion years, at which point the world ends in a singularity. If this were to happen, then the predominant direction of entropy would be other than we think it is — it would go in one direction for the first 17 billion years, and in another for the next 30 billion, and 30 is bigger than 17, so the way it would go for the 30 billion would predominate. And so the Boer War would be before the Wars of the Roses. But it is implausible that whether the Boer War counts as before or after the Wars of the Roses should depend on what happens outside the interval of years from 5 billion BC to 5 billion AD.

One might try to modify the account to say that the direction of time is the direction of *local* entropy increase, where the locality is both spatial and temporal. This would neatly allow for bubbles where the arrow of time points differently from how it points elsewhere, and would save the intuition that whether the Boer War is earlier or later than the Wars of the Roses depends only on what happens over a relatively small interval of times and on earth.

The localized account, however, is implausible for a different reason. Suppose that scientists managed to carefully place a million deterministic particles in a large container, and carefully arranged them so that they would have high initial entropy, but the entropy would then decrease. This kind of an arrangement is physically possible, though unlikely (just take an arrangement of a million deterministic particles whose system entropy had increased and then reverse their velocities, and switch charges and parities). But it is implausible that if this happened, then after ten minutes of the experiment the particles would be *younger*.

Maybe then we should consider entropy increase that is localized temporally but not spatially. But if we graphed the increase of entropy of our universe with infinite precision, the graph would surely be serrated and fractal-like on a microscopic scale — there would be tiny decreases of entropy over tiny intervals of time balanced by larger overall increases. And hence there would be short periods of time (remember, we are localizing temporally) over which the arrow of time points in the opposite direction from its predominant direction. Time's arrow on the microscopic scale would look something like this:

→→→→←→→→←←→→→→→→→←

But if so, then as time goes on, we don't monotonically become older. We only on the average get older. Now, granted, if age is a state of mind and health, that is true anyway. But, intuitively, age is not just a state of mind and health, and it is monotonic. We should, thus, reject an entropy account.

2.3.4 The openness or unreality of the future
Growing Block theorists think that reality consists of the past and present, but contains no future events or objects. Rather, future events and objects continue to be added to the four- (or higher-) dimensional block that constitutes physical reality. If Growing Block is correct, we can try to define the direction of time in terms of this fundamental asymmetry.

One problem for Growing Block is in reconciling it with Relativity. Moreover, there is a problem with inductive reasoning. On Growing Block, there are no future ravens. But why, then, should we think we can safely inductively reason from existent past ravens being black to future ravens being black? If the future ravens are unreal, isn't that like reasoning from real detectives being often completely stumped to fictional detectives being often completely stumped?[9]

Open Futurists, on the other hand, hold that while either it is true that yesterday there was a sea battle or it is true that yesterday there was no sea battle, it is not the case that either it is true that tomorrow there will be a sea battle or tomorrow there will be no sea battle, because whether there is a sea battle tomorrow or not depends on contingent events — especially people's free choices. The cost of Open Futurism is that it forces one to deny one of the following rules of logic (the second being a law of a plausible tense logic):

(32)   Bivalence: either p is true or $\sim p$ is true.

(33)   It will be the case at t that $\sim p$ if and only if (the actual timeline includes t and it is not the case that it will be the case at t that p).

I find these laws of logic rather more plausible than the conjunctions of the premises in the arguments for Open Futurism. Moreover, Open Futurists have a serious difficulty in explaining probabilistic claims about the future. Given the present world political situation, and stipulating that an encounter with pirates is not a battle, we should in fact say that probably tomorrow there will be no sea battle. But if it is probable that tomorrow there will be no sea battle, then it is probable that it is *true* that tomorrow there will be

---

9   For a full development of this argument, see Pruss (forthcoming).

no sea battle. But according to Open Futurists, given that we are pretty sure that people could choose to fight a sea battle tomorrow, we should also be pretty sure that it is *not* true that there will be no sea battle tomorrow. And that is absurd.[10]

Section 3 Propositions

3.1 *Unstructured propositions*

On Lewis's first approach, a proposition is a set of possible worlds (Lewis 1986a; Section 1.4). We say a proposition *p* is true at *w* if *w* is a member of *p*. This won't quite work, because as we shall see later (Section 7.2 of Part III), on no reasonable account of possible worlds is there a set of all possible worlds. However, it may still be open to us to say that a proposition is a class or collection of possible worlds, and this may be just as good. I will use "collection" as a neutral term.

Providing we know what collections are and have an account of possible worlds, we thus have an account of propositions. Unfortunately, as Lewis certainly realizes, this account does not distinguish between propositions that are logically equivalent. But the standard criticism of Lewis here, made forcefully by philosophers like Alvin Plantinga, is that for many purposes such a distinction is necessary. For instance, if one thinks that propositions are both the bearers of truth and what sentences express, then one will be uncomfortable with saying that all the necessary truths are one and the same proposition, identical with the collection *W* of all worlds. In particular, all mathematical truths are the same. Matters are even worse if we accept essentialist claims that genus-species relations are necessary: the necessary *a posteriori* proposition that horses are mammals will turn out to be the same as the proposition that spiders are invertebrates, and both are identical with the *a priori* necessary truth that Fermat's Last Theorem is true.

If one thinks that knowledge is of propositions, one then wonders what we have learned when we learned that Fermat's Last Theorem was true that went beyond the proposition we already knew that 1=1. One might want to say that we just learned something about our language, namely that when English speakers use the words

---

10 Pruss (2010). Rhoda (2010) in trying to get out of this objection ends up saying that *p* can be very probable while the proposition that *p* is *true* is impossible, but that does not appear plausible at least in this kind of a case.

(34)   "There are no positive integers a, b, c and n such that $n>2$ and $a^n+b^n=c^n$"

they express a necessarily true proposition, i.e. on Lewis's view the proposition W. But this answer won't do on either of the two reasonable interpretations of the word "English." Either "English" is a rigid designator here of the language which I am now using or it is a definite description in sociological terms of a language spoken by a group of people who are qualitatively described, where a designator is "rigid" provided it has the same referent in every possible world in which it has a referent ("Socrates" is rigid, but "Plato's teacher" is not since in some worlds it is Parmenides who is Plato's teacher). In the first case, the proposition that when English speakers use the words in (34) they express a necessarily true proposition is itself a necessary truth, and hence on the above Lewisian account of propositions, to learn <English speakers express a necessary truth with (34)> seems to be the same as to learn <1=1>.

On the other hand, consider the case where "English" is a definite sociological description. Then there is a possible world w_1 where "English" does uniquely pick out a language but where (34) means that $1234 \times 4321 = 5332114$. Evidently then what we have learned in discovering Fermat's Last Theorem is different from what the English speakers in w_1 have learned upon discovering that the words in (34) express a necessary truth. But if what we learned was that (34) expresses a necessary truth in English, then we have indeed learned nothing other than those people have. And this is absurd.

Nor are propositions that are logically equivalent the same proposition, as this Lewisian account would make it out. We can see this when we observe that explanation is a relation between propositions. Now let w be a forwards- and backwards-deterministic world. Let L be a proposition reporting the laws of nature of w. Let S_t be a proposition reporting the complete physical state of w at a time t. Then, two-way determinism ensures that S_t and L jointly entail S_u regardless of what t and u are, so that the conjunctions $(S_t$ and $L)$ and $(S_u$ and $L)$ are always logically equivalent. Now, the conjunction of S_0 and L evidently explains S_1 in a deductive nomological way. Hence, if propositions that are logically equivalent are to be identified, likewise $(S_1$ and $L)$ explains S_1, since $(S_0$ and $L)$ and $(S_1$ and $L)$ are logically equivalent. But this is absurd: the present state of a deterministic world when conjoined with the laws does not explain the present state of that world. We can also apply similar reasoning to conclude the absurdity that $(S_1$ and $L)$ explains S_0, since the logically equivalent proposition $(S_{-1}$ and $L)$ does.

3.2 *Structured propositions*

Realizing the need for an account of propositions that allows for differences
between logically equivalent propositions, Lewis calls those propositions
that he identified with collections of possible worlds "unstructured proposi-
tions," and suggests that we also define "structured propositions" as set
theoretic constructions out of the unstructured ones.

> For instance we could associate the modifier "not" with the unstructured
> relation [collection of all pairs of individuals in all possible worlds that fall
> under the relation] N that holds between any unstructured proposition and its
> negation, that being the set of all worlds where the original proposition does
> not hold. Then a negative structured proposition could take the form ⟨N, P⟩,
> where P is a (structured or unstructured) proposition. (Lewis 1986a: 57)

This process can be continued with other connectives. If A is the relation
that holds of a triple $\langle p, q, r \rangle$ of unstructured propositions if and only if r is
the conjunction of p and q, then in the context of the argument at the end
of the previous section we can say that $\langle A, \langle S_0, L \rangle \rangle$ explains S_1 but $\langle A, \langle S_1,
L \rangle \rangle$ does not, even though the unstructured propositions corresponding to
$\langle A, \langle S_0, L \rangle \rangle$ and $\langle A, \langle S_1, L \rangle \rangle$ are identical. Explanation is a relation where
structure matters. Similarly, when one learns that Fermat's Last Theorem
is true, one learns that a certain complicated structured proposition is true,
and that proposition is different (and more complicated) than the structured
proposition one learns when one learns that 1=1.

However, there are many set-theoretic constructions that will work
equally well or equally badly for these purposes. Why should we call $\langle N, P \rangle$
"the proposition which is the negation of P" instead of bestowing that title
on the pair $\langle P, N \rangle$? Moreover, even if we choose an order for the ordered pair,
there are multiple set-theoretic constructions for ordered pairs. For instance,
we can use { N, {N, P} } to set-theoretically represent the ordered pair $\langle N, P \rangle$
or we can use { ∅, N }, { {∅}, P } or even { N, {N}, P } . It is up to us — all
do the job. So which constructions should we choose?[11]

But can we not just choose whichever one we want? Concerning a
somewhat related issue with properties, Lewis writes:

> It's not as if we have fixed once and for all, in some perfectly definitive and
> unequivocal way, on the things we call "the properties" . . . (1986: 55)

---

11   cf. Jubien (2001).

But propositions are theoretical entities introduced as that which we mean by our language. They are supposed to provide criteria for sameness of content within and between languages: two sentences (in the same or different languages) have the same content if and only if they express the same propositions. It is essential that propositions be language independent — otherwise, for most intents and purposes we could just define propositions as sentences of some fixed language, say Latin. The multiplicity of set theoretic constructions that "do the job" mirrors the multiplicity of languages, and hence Lewisian set-theoretically structured propositions are no great improvement over the situation we have when we just stick to languages, except for the advantages of formalization and the availability — for many semantic purposes unnecessary — of alien properties that our languages have no terms for.

There are two possible responses here. The first is that we should choose a particular set theoretic approach, and call the resulting constructions "the propositions." These propositions will do the job that propositions are supposed to do. However, if this is done, then it becomes mysterious how it is that propositions are supposed to be what we *mean* by language. Suppose I affirm the negation of an unstructured proposition. Let us grant that I affirmed ⟨N, P⟩. But how could we possibly find out that when I affirmed this, my meaning in fact used the one privileged set theoretic construction for ordered pairs rather than another? Indeed, it is not even plausible to suppose that it *did* use one construction rather than another. What a queer fact it would be about our language that when we affirm negations our meanings are one kind of set theoretic construction of ordered pairs rather than another!

It is tempting to say that this is a misunderstanding of the role that the set theoretic constructions are supposed to play. Take a particular axiomatic rendition of the general theory of relativity. This theory *models* our space-time as some kind of a Riemannian manifold. But there are, of course, many set theoretic ways of expressing Riemannian manifolds, just as there are many set theoretic ways of expressing real numbers (one can express them as pairs of lower and upper Dedekind cuts, or as lower Dedekind cuts, or as upper Dedekind cuts, or as equivalence classes of Cauchy sequences, and so on). Which one of these constructions is "the right one"? Well, surely, the question matters little more than the question of the color of the ink in which the physics textbook is written — the same physical reality is modeled. And likewise, it might be suggested, it does not matter which set-theoretic construction is used to model propositions. But if Lewis were only giving a *model* of propositions, then by analogy there should then *be* propositions out

there of which the set-theoretic constructions are models, just as (assuming that the correct physics includes particles) there really are electrons that are not merely the mathematical constructions in quantum physics textbooks. But if this were so, then Lewis's account of propositions would no longer provide an ontological reduction of propositions to set-theoretic constructions out of possible worlds. One could just as well consider the objectively existent propositions themselves, whatever they may turn out to be, instead of the items in his model. The model might be theoretically helpful, but it would not remove the ontological puzzlement that Lewis himself admits to feeling (Lewis 1986a; Section 3.4) about what propositions are.

The other possible approach is to go in the opposite direction. A proposition now is not just *some* set-theoretic construction, but *all* of them in some sense. Suppose that C_1 and C_2 are systems of set-theoretic methods for modeling propositions, and p_1 and p_2 are variables that range over the set-theoretically constructed entities within these respective systems that model propositions. Then, there is an equivalence relation that holds between the pairs $\langle C_1, p_1 \rangle$ and $\langle C_2, p_2 \rangle$ if and only if p_1 and p_2 play the same role in their respective systems (e.g. if C_1 is the system where the negation of P is represented by $p_1 = \langle N, P \rangle$ and C_2 the system where it is represented by $p_2 = \langle P, N \rangle$).[12] Perhaps, the suggestion goes, *the* propositions are equivalence classes of pairs under this relation. And perhaps there is some underlying intuitive and natural notion of a "relation" and "equivalence class of pairs" that does not require a specific set-theoretic construction (given that these are the only two notions we are dealing with, this is more plausible than the suggestion that there are always such notions for all the set-theoretic constructions that would be involved within a single system C for constructing structured propositions) — since otherwise we haven't gotten rid of the arbitrariness in the definition.

But the difficulty here is in the equivalence relation. There does not appear to be a way to explicate it. We could take it as primitive. But if we can help ourselves to "the relation" that holds between two system–entity pairs when the entities "play the same role," then why can't we just help ourselves to "the relation" that holds between two language-sentence pairs when the sentences have the same content, and dispense with propositions altogether?

Some of the arbitrariness in the construction of structured propositions is particularly problematic in light of the fact that different constructions give different truth values to semantic claims. On the account of conjunctive

---

12 Compare the (also problematic) concept of "relational structure" used by Ladyman et al. (2007).

propositions that was introduced above, the sentences "The sky is blue and grass is green" and "The grass is green and the sky is blue" do not have the same content, because the former expresses the proposition $\langle A$, <The sky is blue>, <Grass is green>$\rangle$ and the second expresses the proposition $\langle A$, <Grass is green>, <The sky is blue>$\rangle$. But one might also give an alternate construction for the structured conjunction of p and q: one might define this conjunction as $\langle A,\{ p,q \}\rangle$. On this construction, the two sentences will end up expressing numerically the same proposition. The choice between $\langle A,p,q\rangle$ and $\langle A,\{ p,q \}\rangle$ cannot be arbitrary because it has substantive philosophical consequences.

My own intuition is that the conjunction of two sentences, at least ones without pronouns, expresses the same proposition regardless of the order of conjoining;[13] I also think the parallel point is clear in the case of properties (surely there is no difference between an electron's spinning and being charged and an electron's being charged and spinning). But whether this is so should not be legislated by constructional fiat. If, as realists, we think there actually is a fact of the matter as to whether the conjunction of two propositions is the same regardless of the order of the conjuncts, then this fact of the matter will be grounded in some reality independent of our construction of "structured propositions." But if so, then Lewis's theory of structured propositions, if it be correct, actually presupposes as primitive a fact about whether conjunction depends on the order of conjuncts.

One might not, however, want to be a realist about sameness of content or identity of propositions, holding instead that there is no fact of the matter that determines whether two sentences have the same content. For various theoretical purposes one might choose various accounts of "propositions" and "content." This may even be the parallel of the point that Lewis is making above for properties. But it is significant to note that this is a point that Lewis in the large structure of *On the Plurality of Worlds* cannot afford to make. Lewis is offering an argument for why we should believe that there are possible worlds by arguing that they have great philosophical utility, and offers his accounts of propositions and properties as examples. But in order for this argument to provide justification for *realism* about possible worlds it is surely necessary that the possible worlds should figure in an *explanation* of the philosophical phenomena in question. That possible worlds are useful for producing an account of propositions is only going to

---

13 To maintain this, I have to say that "Bob and Jane got married and Jane got pregnant" is not merely the conjunction of "Bob and Jane got married" and of "Jane got pregnant." The "and" between the two is an "and then."

give a reason to be a realist about worlds if one is a realist about propositions or at least about the sort of talk that propositions are used to clarify. If the propositions are arbitrary constructions useful contextually, the worlds can be such, too. Admittedly, as we shall see in Section 2 of Part IV, no one linguistic construction of worlds does everything we want, but perhaps we can use different ones for different contexts, just as it is alleged we might have different constructions of propositions for different purposes.

There is another difficulty with the arbitrariness involved in the construction of structured propositions. The purpose of propositions is to be language-independent analogues of assertoric sentences, one per equivalence class of sentences with the same content. But Lewisian structured propositions are basically linguistic entities, albeit ones constructed in a language whose symbols are various sets, classes or collections, instead of marks on pages or vibrations in the air. Indeed, the fact that they are linguistic entities is seen from the fact that there are *different* languages that all provide constructions of propositions, e.g. a language that specifies that the grammatically correct form is $\langle p, q, A \rangle$ instead of $\langle A, p, q \rangle$.

But if moving from sentences to propositions is just moving from sentences in natural languages to ones in a more rarified language whose symbols are abstracta, then the move fails to provide the language-independence that the notion of propositions was supposed to yield. And, indeed, the move threatens a classical third-man regress. If we primarily introduce propositions to be that whose expression sentences with the same content have in common, and if propositions are *themselves* sentences, albeit ones in a rarified language whose terms are Platonic abstracta, then either:

(35)   there should be a yet third class of sentences to explain the sameness of content between the sentences of natural language and the propositions which on this account are sentences,

or

(36)   we have no principled reason to introduce propositions, at least for this purpose.

For if (35) is false then sometimes sameness of content between sentences can be explained without reference to any entities other than the sentences themselves (at least in the case where one of the sentences is a proposition). But if we can do it sometimes, we do not seem to have an argument against the possibility of doing it in general, and (36) follows.

Option (35) leads to a vicious regress while (36) removes a good part of the epistemic ground for supposing there to be propositions — and hence removes much of the warrant that the availability of Lewis's account of propositions had added to the real existence of possible worlds. Of course, one might have *other* reasons for positing propositions. But if propositions are really sentences, albeit in a rarified language, then one doubts whether there will *ever* be a principled argument for why one needs something other than the "more ordinary" sentences, or at least sentence types.

Finally, observe that Lewis's account of structured propositions presupposes that there are *simple* or *atomic* propositions that cannot be analyzed as connected combinations of other propositions and that out of these simple propositions all other propositions are composed — call this thesis *general atomism*. For the Lewisian structured propositions are built out of unstructured propositions, and the unstructured propositions cannot be analyzed into other propositions. If general atomism is false, some proposition p is analyzable into a combination of simpler propositions, some of which in turn are analyzable into a combination of yet simpler propositions, and so on without termination. To model such a p in his system, Lewis would have to cut off the infinite complexity of p at some point, making some level of analysis consist entirely of unstructured propositions, thereby destroying the finer structure of p and hence failing to model it. If we think that the existence of propositions of such infinite propositional complexity is a genuine possibility, then we will have reason to be dubious of Lewis's theory of propositions prejudging against this. (As an *ad hominem* point, recall that Lewis (1986b) himself worries about the epistemic possibility of infinite complexity and uses it as an argument against Armstrong's view of structural universals.)

If Lewis's account were to give one an intelligible ontology for propositions in terms of possible worlds, this would be a big asset for those possible-worlds theories that do not construct possible worlds out of propositions. But, alas, Lewis's account has proved unsatisfactory, whether in its unstructured or structured incarnation. However, the unstructured incarnation may have some limited theoretical uses. For some intents and purposes, the differences between logically equivalent propositions can probably be ignored. For these purposes, Lewis's account may be a useful tool. But since the "unstructured propositions" are not really propositions, the anti-realist about possible worlds can accept unstructured propositions just as much, considering them to be mere useful fictions.

We can observe that this attack on Lewis's account of propositions takes much of the wind out of Lewis's sails when he assails those accounts that

construct possible worlds out of propositions for their failure to give an account of what propositions are. For Lewis cannot give a satisfactory account of them either.

Section 4 Properties

Lewis offers an account of properties that is similar to his account of propositions. Thus, a property may be seen in an unstructured way as the collection of the individuals in all possible worlds that exemplify it, and we can use set-theoretic constructions to produce "structured properties." Lewisian propositions can then be seen to be properties of worlds.

The difficulties with the unstructured and unstructured properties are much the same as in the propositional case. For instance, the unstructured account is unable to distinguish co-extensive but distinct properties, whereas we have the following serious problem with structured properties. Granted, one can in a quasi-linguistic way distinguish a number of co-extensive properties. For instance, if $\wedge$ is the ternary relation between unstructured properties which holds if and only if the third relatum is the intersection of the first two, then we can distinguish the conjunctive property of being a horse *and* a mammal from the necessarily co-extensive property of being a horse by using $\langle \wedge, \langle H, M \rangle \rangle$ for the first property and H for the second, where H and M are the properties of being a horse and a mammal respectively. The main objection against structured propositions was the arbitrariness of the set-theoretic construction. One might level a similar objection against structured properties now. If one thinks of the properties of an object as those entities in virtue of possessing which various predicates are predicated of the object, then one might wonder why a horse should possess $\langle \wedge, \langle H, M \rangle \rangle$ as opposed to, say, $\langle \langle H, M \rangle, \wedge \rangle$. Properties are those aspects of reality that are pointed out by an attribution of a predicate to a subject. If we are to be realists about them, these should thus be uniquely defined, even more so than propositions should.

On the other hand, one might hold to a plausible sparse account of properties on which not every predicate corresponds to a property. Thus, there might not be any disjunctive properties though there are disjunctive predicates. On such a theory, there might not be any need for properties to be structured and it might be that in the end there are no distinct but necessarily co-extensive properties. But if one takes such a sparse view, then the problem is that Lewis's theory supplies us with *too many* properties. For instance, on Lewis's view there is a property had by the actual world's Empire State

Building and the actual world's Napoleon but by nothing else: namely the unstructured property { Empire State Building, Napoleon }. But the person who dislikes disjunctive properties will hate this property.

Section 5 Overall assessment

Possible worlds let one formulate in a uniform way various modal notions that seem to intrinsically involve consideration of and comparison between more than one world: e.g. supervenience, transworld comparison of individuals, and counterfactuals. Moreover, ordinary modal claims like "I might have been a physicist" or "Hitler might never have been born" are naturally disambiguated against a background of possible worlds, with the context determining which worlds we are quantifying over. Since there is good reason to think that all these modal claims make sense, and since a very natural and intuitive way to make sense of them is possible worlds, this gives us good reason to believe there are possible worlds. Granted, not all uses of possible worlds stand up to critical scrutiny — in particular, the analyses of propositions and properties in terms of possible worlds do not appear to be helpful — and the Lewis–Stalnaker account of counterfactuals in terms of possible worlds needs some additional refinement. But there is enough use that it counts in favor of a philosophical view if it allows us to have possible worlds.

But there is more than one theory of possible worlds. If we should find that only one extant theory is a serious option that withstands all criticism, then the fact that there is good reason to believe there are possible worlds would provide us with good reason to believe that this theory is true. I shall eventually argue that of the theories under consideration, only one has the hope of being satisfactory: an Aristotelian modification of Leibniz's theistic theory. Until a better theory should be found, this gives us some reason to think the theory true (and, importantly, to think that God exists). But first we must consider the alternatives.

PART III

THE LEWISIAN ONTOLOGY OF
EXTREME MODAL REALISM

Section 1 The Lewisian account of possible worlds

For purposes of discussion of Lewis, I will call a maximal mereological sum of objects that are spatio-temporally related to one another a "universe." I shall assume that being spatio-temporally-related-to is a transitive and symmetric relation. By definition, if there are two distinct universes, there are no spatio-temporal relations between them. David Lewis's extreme modal realism (EMR) then claims that possible worlds are existing physical universes ontologically on par with the universe we in fact inhabit and every possible way for a universe to be is a way that some universe is. In a Lewisian theory, thus, "possible world" and "universe" are interchangeable. Not so, of course, on other theories: e.g. if possible worlds are maximal sets of compossible propositions, then a possible world is not a universe, since there are no spatio-temporal relations between propositions or sets of propositions.

A proposition, then, is true *at* a world providing it truly describes a state of affairs obtaining in that world. Quantifiers in many ordinary language propositions are restricted to the world at which we are evaluating their truth value. The tokening of a proposition expresses a truth if and only if the proposition is true at the world at which it is tokened. It is true at w that there exists a unicorn if and only if in the concretely existing universe that w is, there really is a unicorn. My ordinary utterance of "There exists a unicorn" expresses a truth if and only if in *our* universe, i.e. in the maximal aggregate of things that are spatio-temporally related, and that includes me, there is a unicorn. It is irrelevant whether there are such creatures elsewhere, and so it is indeed false that there exists a unicorn. However, context may indicate a wider scope for quantifiers, allowing them to range over objects in multiple worlds. "There are worlds" is an obvious example. Likewise, in

an appropriate context, a Lewisian will willingly say "There exist infinitely many unicorns."

According to Lewis, ontologically all worlds are on par. Our world and universe is just one among infinitely many. We correctly say it is actual, but this only says it is *home* — being actual is no more an absolute non-relational property of our world than something's being *home* (as opposed to being *a* home) is such a property of a house. Like something's being home, a world's being actual is but an indexical claim. The actuality of the actual world consists in nothing but its being the maximal spatio-temporally connected aggregate that contains *us*. Moreover,

> at any world *w*, the name "the actual world" denotes or names *w*; the predicate "is actual" designates or is true of *w* and whatever exists in *w*; the operator "actually" is true of propositions true at *w*, and so on for cognate terms of other categories.[1]

In the standard Lewisian model each individual exists in only one possible world. For if some individual *x* is in worlds w_1 and w_2, then all items in w_1 are spatio-temporally related to *x* and likewise all items in w_2. But since being *spatio-temporally-related-to* is symmetric and transitive, it follows that all items in w_1 and w_2 are spatio-temporally related to one another, and since the worlds are supposed to be maximal, w_1 and w_2 are identical.

But if every possible way for a world to be is a way that some world is, then surely there is a world at which Margaret Thatcher had dyed her hair green in 2002. Since it is false at the actual world that she has done so, one might think that it follows that Thatcher exists in some other world, a world where she had green hair in 2002. However, on Lewis's account, it is not Thatcher herself who exists at that other world, but someone very much like she, a *counterpart* to her, albeit with green hair in 2002. Lewis can then enrich the semantics of ". . . is true at . . ." by saying that a proposition about some particular individual *x* is true at some other world (where *x* does not exist by the above argument) if and only if a *counterpart* to *x* exists at that world and what the proposition had asserted of *x* is true of the counterpart. On this semantic move, one can say that indeed there is a world at which Thatcher has green hair in 2002. That is a world at which Thatcher's *counterpart* has green hair in the year that a counterpart to 2002.

To say that one individual is a counterpart of another is to affirm that

---

1   Lewis (1970: 184–5).

there is a certain contextually determined similarity between them. If we consider *being-a-counterpart-of* as a relation that, for any given individual *x* in a world, picks out the unique individual, if there is one, in another world who most closely resembles *x*, assuming this resemblance is "close enough," then *being-a-counterpart-of* will be a non-symmetric non-transitive relation for Lewis. The non-transitivity appears to follow from the possibility of a gradually varying sequence going between very different individuals (see Chisholm 1967). The non-symmetry can be seen as follows. Suppose that I have an almost identical twin in the actual world. The only significant difference between him and me is that he has a big birthmark on his arm which I do not have. Now, there is a world, *w*, where there is a perfect copy of this twin of mine, but where there is nobody else even remotely like me or my twin. The copy of my twin is then a counterpart of mine and makes it true to say that I have a big birthmark on my arm at *w*. But I am not the counterpart of that copy of my twin — my twin is, for my twin resembles him more than I do, since it is he who has the birthmark.

Actually, for the interpretation of many claims Lewis (1979b) thinks we will need to speak not just of truth at a world but truth at a world and at an individual in that world. Supposing that I have a twin, like in my example above, this is how Lewis would make sense of the fact that it is possible that both I have a birthmark just like my twin and there is another person just like me but who does not have that birthmark. If we imagined a universe with someone just like me and with the birthmark and someone just like me but without the birthmark, my counterpart would seem to be the fellow *without* the birthmark since he is more similar to me. Hence, that universe is *not* one at which I have a birthmark. Here, we need to interpret the statement that at *w* it is true that I have a birthmark and there is someone just like me but without the birthmark by assigning it a truth value at the ordered pair ⟨*w*, the fellow in *w* just like me with the birthmark⟩ and not just at *w*. This consideration shows that a Lewisian needs to allow that a given person can have more than one counterpart at a world, which he is anyway committed to, since he thinks that the counterpart relation should be contextually understood (I do not know if he can handle the above case by mere contextuality).

Since Lewis wants to reap the benefits for modality of the available of possible worlds, he will define possibility as truth at some world and necessity as truth at all worlds. To take care of the sorts of issues mentioned in the previous paragraph, one might have to sometimes modify this by saying that possibility is truth at some world-individual pair (or even world-individual$_1$-...-individual$_n$ (n+1)-tuple), where the individual (or *n*-tuple of individuals)

is a counterpart to the individual (or individuals) that the proposition is ostensibly about.

Section 2 Identity vs. counterpart theory

If we are going to believe that every way that some world could be is a way that some world really is, which we may call "*basic* EMR," we still have a choice whether to accept identity theory or counterpart theory for the individuals, *pace* Lewis's packaging of basic EMR with identity theory. If we accept identity theory, then the same individual ends up existing in multiple worlds and what I said in the previous section about counterparts is not true. Lewis refers to this as an "overlap of worlds," in that different worlds overlap in having the same individuals in them. If we accept counterpart theory, then each individual exists in exactly one world but may have counterparts in some others.

2.1 Arguments for counterpart theory

Consider an argument for counterpart theory, perhaps implicit in Leibniz's reasoning behind his view that each individual existed in only one world: (a) there evidently are such things as essential properties; and (b) there is no intelligible way of drawing a distinction between essential and other properties (cf. Leibniz's "Discourse on Metaphysics"). As a result of (a) and (b), it is concluded that all properties are in fact essential. But "existing in world *w*" is a property, and hence an essential property, and thus no individual that has it can lack it, and so if an individual exists in world *w* it exists in no other world.

   This argument is not the usual reaction to someone's claiming (b). The natural reaction is not to claim that all properties are essential, but that all properties are inessential. However, I shall not take the route of this objection to the argument, because there are very good reasons to believe in essential properties: in no world am I omnipotent or a number. And the fact that we do not always *know* how to draw the distinction does not militate against the existence of a distinction. Moreover, if basic EMR is true and if identity theory is true, there is a very simple and highly intelligible way of drawing the distinction: a property of an individual is essential if and only if the individual has this property in all worlds. To deny the intelligibility of this is to beg the question against the identity variant of EMR and not to give an argument against it.

Lewis has given a more formidable argument. We normally distinguish between relational properties, such as *being a father*, and non-relational or innate properties, such as *being square*. But if the identity variant of EMR is true, then something (e.g. a sponge) may be square in one world and round in another, and so being square is a property we must relativize to a world. Otherwise, we violate the law of non-contradiction. But, generalizing, it follows that *every* non-essential property is relative to a world.

Lewis himself has rejected the theory of numerical identity over time, preferring a theory of temporal stages that make up an individual considered as a space-time worm. However, it is certainly open for an identity theorist to insist on the older Aristotelian understanding of the law of non-contradiction according to which things are barred from having contradictory properties *at the same time*, but are allowed to have them at different times. By the same token, perhaps, the same being can have contradictory properties in different worlds.

But what does it *mean* to say that something is square relative to one world and circular relative to another? What is it *really* like? The mystery about what properties like shape are if they turn out to be relational does not even disappear if we say that space itself is relational. For in a given extended object, there will presumably be internal spatial relations as well as external ones. The internal ones will differ depending on the shape of the object. Thus, a circular disc has the *internal* property of containing a point such that all peripheral points are internally equidistant from that point.[2] But since the disc could have been square-shaped, it follows, assuming basic EMR and identity theory, that these internal distances in the disc are not merely binary relational properties between points in the disc, but are ternary relations between the two points and the world (or maybe other items in the world). This seems to be an unpleasant complication.

But a complication is not a contradiction. It makes a theory more expensive, but does not knock it down. Maybe we are wrong about internal distances being binary relations. Maybe we are wrong about electric charges being innate properties. Perhaps everything is more holistically relational.

But now here is a more serious problem. What are the innate properties relational *to*? First to show the problem more clearly, suppose our identity theorist accepts a non-relational view of space-time. She will then presumably accept identity theory for space-time as well, since the same arguments

---

2   The internal distance between two points in an object or region is the shortest distance (or the infimum of the distances, if there is no shortest distance) along a path wholly within the object or region and joining the two points.

as she advances for identity theory in other cases apply,[3] so that a given point will exist in the space-time of several different worlds, and different worlds might have the same space-time. But if she does this, then she cannot give what might be the most natural solution to the problem, namely that shapes are relational properties with the other relatum being space-time. For if something exists in more than one world, it might be that both worlds have numerically the same space-time. So when we say that something is square relative to one space-time and circular relative to the other, we are affirming something contradictory — since there is only one space-time in the two universes, and nothing can be square and circular relative to the same space-time, at least not in the same way.

If we think space-time is relational, the same problem reappears. For unless we think that "the world" is something over and beyond the mereological sum of its parts, something that could itself stand in relations, we will want to say, e.g. that this sponge is square in relation to the other individuals in the world, while in another world it is circular in relation to the individuals of that world. But it is conceivable that the two worlds have the same individuals, the only difference between them being the shape of this sponge. Then, we have said that the sponge is square in relation to the same individuals in relation to which it is circular. And this is no help.

The remaining solution for the advocate of the identity variant of EMR is to insist that "the world," which basic EMR insists is nothing else than "our universe," is something that things can stand in relation to. Thus, the sponge is something that is square in relation to one world and circular in relation to another. But now another problem appears. If the world is itself a concrete individual, as for Lewis it is, then the same kind of counterfactual reasoning that would make one think that *this* very sponge which is actually circular could be square apply to the *world* at large. *This* very world, which in fact is populated by people, might have been populated by mere hydrogen gas. *This* very world or universe, which in fact is populated by our circular sponge, might have contained this same sponge which, though, was square. But then our sponge is circular and square in relation to the very same world. This objection works even if a world is a mereological sum of its individuals. For two worlds could have the same individuals but differ in the shapes of some of these individuals. The identity theorist who accepts basic EMR thus needs to exempt worlds from her identity theory. But this is

---

3   The main argument is, of course, the insistence that counterfactual properties apply
     to the same individual. But a point in space-time also has counterfactual properties
     (e.g. there being an electric charge-density ρ_1 at it, whereas in fact the electric charge-
     density at it is ρ_0).

not satisfactory, given that basic EMR equates worlds with universes, which are concrete individuals, and identity theory should surely hold either for all or for no individuals.

Admittedly, we could identify the other relatum in the relative account of shapes, not as a world but as a world-type. But once one has world-types in the picture, the cost–benefit argument in favor of EMR is weakened, for it seems that a good deal of the work done by worlds could be done by world-types instead.

Another EMR-based argument against identity theory is that based on Lewis's analysis of actuality. If for a world to be actual is for it to be the world in which I exist, and if I exist in more than one world, then more than one world is actual, which is absurd. Moreover, if in another world I ride a unicorn, then it seems that the unicorn is spatio-temporally related to me, and hence the unicorn is actual, which is absurd.

On the other hand, *without EMR* we have no good argument against identity theory, this time construed as the assertion that when we are saying that something could have had that property we are not saying this in virtue of any individual other than the one we are talking about.

2.2 Arguments for identity theory

2.2.1 General arguments

As has often been argued, someone very much like me becoming, say, a spelunker in another world cannot be a truthmaker of the proposition that it is possible that I become a cave explorer. How indeed, it is asked, is it at all even relevant to the proposition? This way of putting the issue is not sharp enough since that someone in *this* world who is very much like me becomes a spelunker is good *evidence* for the proposition that I could become one, too. But it is not *because* someone like me becomes a spelunker in this world that it is possible for me to become a spelunker. Why, then, should things be different when that fellow who is like me in fact dwells in another world? His example may inspire my imagination, thinking about him might lead me to regret me not having taken his exciting road in life, but his example plays an essentially different role than the role of *making* it true that I could have become a spelunker.

Suppose we have a system that emits a light when a button is pressed. We should not say that the fact that light was emitted at t_0, t_1, t_3, and t_4 when the button was pressed, together with the situation at these times being closely analogous to t_2 in terms of set-up, is what *makes it true* to say that at t_2 the

system also had the dispositional property that it would respond with light were the button pressed. Certainly, the facts about the activation of a like dispositional property t_0, t_1, t_3 and t_4 are evidence, perhaps even conclusive evidence, that the property was had at t_2. But one must not confuse the evidence with the truthmaker. In this case, such a confusion would lead one to buy into a Humean reduction of dispositional properties (and laws, too) to occurrent states. Lewis appears to be guilty of a similar reduction, albeit the occurrent states in his account are in other worlds. Nonetheless, our argument should not be limited to pointing out that there is such a reduction going on.[4]

2.2.2 Attributions of ability

A person is only free to do something if she *can* do it. A Lewisian analysis of the modal claim that someone can do something will involve statements about what some of the person's counterparts in other worlds in fact do. For instance, if one's notion of ability is of the sort incompatibilists introduce, then a necessary condition for my now being able to do something is, in Lewisian terms, that some counterpart of me who shares a copy of my past up to now in fact does it.

However, here the objection discussed in Section 2.2.1 comes in: That I *can* do something is surely a statement reporting a fact specifically about me, not about another person such as a counterpart of me. After all, my being able or unable to do something has normative import for *me*: if I am unable to do something, then I was not free to do it, and hence I cannot be held responsible for not having done it. The objection, however, as given appears to merely beg the question. After all, the statement that my counterpart does something is, on Lewis's theory, a statement about *my* modal properties, and hence in a straightforward sense of "about" a statement about *me*.

If the objection is not to beg the question, a more precise sense of "about"

---

4   Plantinga (1987) and Lycan (1994, 2002) have both criticized Lewis by saying that what happens in other Lewisian worlds has nothing to do with modality. This is no different from saying that what happens in other similar circumstances has nothing to do with causality or nomic modality, which is in danger of begging the question against the Humean. Moreover, one must be careful here. At least as read by Lycan (1994: 84), Plantinga claims that discovering facts about other concrete universes would not tell one anything about modality. Actually, *epistemically* it would provide *some* modal information. That someone very much like me becomes a biologist in some world is very good evidence for the claim that *I* could have become a biologist. And what better evidence can there be for the claim that there could be a unicorn than that there *is* a unicorn? What discovering facts about other concrete universes would not do, I will argue, is provide *truthmakers* for modal claims.

must be brought in. The notion of a truthmaker helps. We can say that a true proposition is *about* some existent thing, if that thing is one of the items *involved* in the proposition's truthmaker,[5] where the notion of being "involved" in the kind of thing that is the truthmaker of a proposition is illustrated by propositions such as:

(37)   The butler, the master, the stabbing and the knife all are *involved* in the butler's having stabbed his master to death with a knife.

(38)   Socrates is *involved* in the truthmaker of the proposition that Socrates existed.

(39)   Clinton is *involved* in the truthmaker of the proposition that there was a 42nd President of the United States.

(40)   If[6] pain is (reductive identity) nothing but a firing of one's C421, then my C421 is *involved* in my being in pain.[7]

One might take involvement in a truthmaker as simply being a part of the truthmaker, but perhaps that isn't the only option.

The objection now is that surely I must be involved in any truthmaker

---

5   What a false proposition p is about presumably needs to be defined counterfactually, e.g. by saying that a false p is about x if and only if, were p true, p would be about x. Or one might want to relativize what entities a proposition is about to worlds. Thus, a proposition p is about x in w if p is true at w and it is true at w that p is about x. If a proposition is not actually true, then on this approach we cannot say what the proposition is about *simpliciter*, but can say what it is about in any given world in which it is true. This kind of relativization is necessary to give an account of disjunctive propositions which might be about different things in different worlds: perhaps <There are horses or there are unicorns> is about horses and not about unicorns in our world.

6   *Per impossibile*, I believe.

7   These examples are not sufficient to determine what we are to say about disjunctive cases such as "A horse exists or a dog exists," or even "A horse exists," but we do not need to determine the answer to this for the purposes of the following discussion. One solution would be to allow that a given proposition can have more than one truthmaker. It is, after all, reasonable to say that any horse is a truthmaker of the proposition that *a horse exists or a dog exists*, just as any dog is. If so, then we have two possible readings of the claim that p is about x if and only if x is involved in p's truthmaker. We could say that p is about x if and only if x is involved in every truthmaker of p. Or we could say that p is about x if and only if x is involved in some truthmaker of p. The stronger reading is preferable if we do not want to say that the proposition that *a horse exists or a dog exists* is about Genghis Khan's third horse, and the spirit of the stricter definition fits better with the sense of "about" in the claim that the proposition *that I can jump* is about me.

of the proposition that I am now able to, say, jump. But the truthmakers of that proposition on Lewis's account, it seems, are the jumpings of relevant counterparts of mine, and this involves merely the actions of certain *counterparts* of mine and not me.

This objection, however, fails because it commits a *de dicto/de re* modal fallacy. Suppose x, y and z are rigid designators of those counterparts of mine whose actions are involved in making it true that I can jump, where we stipulate that a rigid designator has literally one and the same referred in every possible world in which it has a referent. Given counterpart theory, a rigid designator of a concrete entity has a referent in at most one world.

But then it is not only "x"s, "y"s, and "z"s jumping that makes it true that I can jump, but also their being relevant counterparts of *me*. It is true that for Lewis that I can jump is made true by the jumping of certain counterparts of mine, but here "counterparts of mine" must be read *de dicto*: one cannot substitute the rigid designators "x, y, and z" for "certain counterparts of mine." But I am in fact involved in the truthmaker of the proposition <x, y, and z are relevant counterparts of me>, since the truthmaker of that proposition is whatever makes for the relevant similarity between me and x, y, and z, and that surely involves me. Thus, I am also involved in the truthmaker of the proposition that I can jump, even on Lewis's view.

But despite this analysis, one may have a feeling that x, y, and z are interlopers in the truthmaker of the proposition that I can jump. Why should their actions be *at all* relevant to the truth of this proposition? The worry now is not that I am uninvolved in the truthmaker, but that some people who have no business being involved are. One may think that in some sense the proposition that I can jump should be *only* about me, my intrinsic properties and my immediate surroundings (including my relations to my immediate surroundings). For, instance, why should the properties of other people causally isolated from me have any bearing on all the normative claims that follow from my being able or unable to jump?

There is, however, a way in which Lewis can accommodate some of the background of the intuition that I, my intrinsic properties and my surroundings are the only things that are involved into the truthmaker of the proposition that I can jump. Lewis can say that I, my properties and my surroundings are the only *actual* things that enter into this truthmaker. But this would only be a plausible way to account for the intuition if one thought, as Lewis denies, that to be actual and to exist are the same thing. Given Lewis's distinction between the actual and the existent, the mystery as to why some non-actual but nonetheless existent persons should be involved in the truthmaker of the proposition that I can jump remains.

Moreover, the answer that the other persons involved in the truthmaker are non-actual will not do on Lewisian principles. For we can imagine a case where many of those persons whose actions make it true that I can jump are, in fact, actual. Imagine I am in an almost deterministic world in which at the beginning of every century one new human being has come into existence, lived for a hundred years, and then disappeared. Then a new human being, qualitatively indiscernible from the previous, appeared, and the cycle went on. This would of physical necessity go on forever in the future, except that there is one and only one bit of freedom and contingency that each human being is allowed. At the beginning of the 27th year of his life he can jump or refrain from jumping. If he jumps, everything will go on as before, a new human being appearing at the end of this one's life, and so on. But if he refrains from jumping, the laws of nature and initial conditions are thus constituted that after he dies, no more humans will live, and there will no longer be any contingency allowed by the laws of nature. Suppose now I am one of these human beings, and I refrain from jumping. Up to my time, the events of every century repeated, century after century. After me, everything is different and the laws of nature allow no more contingency.

Consider then the true claim that I could have jumped at the beginning of my 27th year. In an appropriate context where one is not stating a merely metaphysical possibility but a physical possibility, this is made true by the jumping of my counterparts in worlds with the same causal structure. Now, first assume that there are no qualitatively identical (i.e. exactly alike) but distinct worlds — the only duplicate a world has is that world itself. Then in fact, there are only two relevantly similar different worlds with the same causal structure as the actual world. One of these worlds, w_1, is a world of endless recurrence where in fact every human being jumps at the beginning of his 27th year. The other world, w_2, is one where an infinite number of human beings jump, but then one doesn't, and then everything is different. The proposition that I could jump at the beginning of my 27th year is made true by my counterparts in either of these two worlds jumping (i.e. by the jumping of all the people in w_1, and of all the people in w_2 prior to the first who did not jump). But now observe that w_2 is in fact qualitatively identical with the actual world, and hence w_2 *is* the actual world since we have assumed that qualitative identity implies identity for worlds. Hence, among those counterparts whose jumping makes it true that I could have jumped, are counterparts living in my world, i.e. *actual* persons. The same conclusion holds even if qualitative identity does not imply identity of worlds, because nonetheless the actual world will be among those worlds that are relevantly similar to the actual world, and my worldmates in the past will still be

good counterparts for me. Hence in any case, some of the people involved in some of the truthmakers of the claim that I could have jumped are *actual* people who existed prior to me in the actual world — and *these* are surely interlopers. These people's existence and their jumping may *illustrate* my being able to have jumped, but surely my ability does not even in part *consist* in their having jumped.[8]

The best strategy for Lewis here would be the *tu quoque*. On competing views of possible worlds, entities other than I, my properties and my surroundings are also involved in the truthmaker of the proposition that I can jump. For instance, views which construct possible worlds out of abstract propositions will have abstract propositions involved in the truthmaker. For instance, the truthmaker of the proposition that I can jump might be <I jump>'s being compossible with circumstances similar to those in which I in fact find myself, which clearly involves <I jump>, an entity not part of me, my properties or my surroundings, and in no way causally relevant to anything concrete. Lewis can ask: Why should the intrinsic properties of this abstract entity be involved in *my* concretely being able to jump? Thus, perhaps, whether we be Lewisians or not, we simply must accept that my being able to jump involves things other than me, my properties and my surroundings.

Lewis's opponents here can argue that the entities *they* involve in my being able to jump are *abstract*, and hence one should not complain of their presence here. But why not? Similar intuitions to those that say that the truthmaker of the proposition that I can jump should only involve me, my properties and my surroundings will also say that the truthmaker should be something entirely concrete — which on Lewis's view it is.

However, *tu quoque* answers always have the weakness that they become useless when a theory comes into view that does not share the bad features of the old theories. In Part VI, I shall argue for a theory on which the truthmaker of the proposition that I can jump involves only me, my qualities and

---

8   Note that it would not be easy for Lewis to disallow this-worldly counterparts if there are no numerically distinct but qualitatively exactly similar worlds. For instance, imagine a world w with two individuals, Castor and Pollux, whose first ten years of life are exactly alike (Lewis's commitment to the possibility of arbitrary arrangements allow this), but then Castor coughs in his sleep while Pollux doesn't, and yet their lives after that are all exactly alike again. It should be true at w that everything could have gone as it did, except that Pollux coughed just as Castor did in w and that Castor failed to cough just as Pollux did in w. But any world that grounds this possibility will be exactly like w, and hence will have to be w itself if no numerically distinct worlds are qualitatively alike. To get this possibility, then, we will have to take Pollux as the counterpart of Castor and *vice versa*.

my surroundings, and is in a relevant sense a *concrete* part of the world. At present, however, we can observe that there is a theory in view that *almost* does this: Lewisian multiple worlds with an identity theory of transworld identity. For, then, although other-worldly properties of me are involved in my being able to jump on this theory, the only *substances* involved are ones that exist in this world: I and those substances that exist in my surroundings. For, that I can jump, on this theory, just says that *I* jump — in some world with relevantly similar surroundings. Hence it is more understandable why *I* could be held responsible for not jumping if I do not.

Further ethical considerations will be brought to bear on this issue in Section 8.

2.3 *Conclusions about identity and counterpart versions of basic EMR*

Neither the identity, nor the counterpart versions, of basic EMR are completely satisfactory, and this provides a dilemma argument against basic EMR: either the identity or counterpart version is true. But in each case problems ensue.

Section 3 Indiscernible worlds?

Some of my arguments against EMR will presuppose the thesis that there are no indiscernible worlds, i.e. that no two worlds are duplicates of each other, while the argument in Section 8.3 can be made to work in the case where there are only finitely many copies of each world. On the other hand, the argument in Section 9 will work if each world has an *equal* number of copies, finite or infinite. Thus, to evaluate some of the arguments against EMR, we should first examine whether there are indiscernible worlds or not.

Lewis (1986a: 84) himself is not committed to either the view that there are indiscernible worlds or to the view that there are not.

First of all, however, there is a simple argument based on the apparent absurdity of the following claim:

(41)   It is possible for everything to be as it is while the actual world is not actual.

This claim is true if there are multiple copies of this world (note that it *is*

possible that the actual world not have been actual, since that would just mean that some other world than *this* one had been actual; compare the fact that the tallest man in the world might not have been tallest). If one thinks (41) to be absurd, then one will not accept multiple copies of worlds.

One could argue against indiscernible worlds based on the Principle of Identity of Indiscernibles (PII), which states that no two distinct objects are indiscernible, i.e. share all the properties describable in purely general terms. But one can also argue against indiscernible worlds on weaker assumptions about individuation, using an argument of Donald Turner (2003). Turner says that indiscernible objects can be individuated only if they are in a common space and/or time and thus capable of spatial and/or temporal separation. If one takes this view, then distinct Lewisian worlds could not be indiscernible, because there are no spatio-temporal relations between Lewisian worlds.

One difficulty with this argument is that current physics appears to provide a counterexample in the form of bosons (French 1988) such as photons. Two or more bosons can have the same wave-function. Consequently, it seems possible that they be indiscernible in all respects, including in spatio-temporal relation. The defender of Turner's argument against indiscernible worlds can proceed in two ways. The first is to adopt a controversial interpretation of quantum mechanics according to which while normally one describes the universe as containing n photons with k degrees of freedom each, it would be more correct to speak of a single "n-fold Photon," existing in up to n places at once, and having nk degrees of freedom.[9] The second is to weaken the principle that indiscernibles can be distinguished only by actual spatio-temporal relations to a much more reasonable principle that says that individuals can be distinguished only by actual *or potential* spatio-temporal relations or actual *or potential* differences in properties. Multiple photons that in fact share a wave-function can be distinguished by virtue of being *potentially* spatio-temporally separated — namely, by virtue of it being possible that they are thus separated. But even this weak PII does not allow one to distinguish indiscernible worlds. For on Lewis's account there are no alternate possibilities for *pairs* of worlds, but only for individual worlds. So it is impossible to speak of a joint possibility for the two worlds where they are distinguished, e.g. by having different qualitative properties.

Another argument against indiscernible worlds might be from Ockham's

---

9   Note that although my arguments in Section 7 will presuppose the possibility of different infinite numbers of photons existing in the same place, they can be formulated in a way that is consistent with this interpretation, by talking of a single "n-fold Photon," with n infinite.

razor. Why posit multiple worlds where a single will do? But there is a reply: a single world *will not* do in view of the absurdities resulting from a no-duplicate-worlds version of EMR as we shall see in Section 8.3.

However, there is a strong plausibility argument against indiscernible Lewisian worlds based on arbitrariness considerations. If there are multiple copies of some world, one asks: How many? If 227, why not 327? One answer might be that the number is proportional to the objective probability of that world being actual. This might even neatly solve the problem of induction that will be raised in Section 9: the worlds where induction holds maybe *are* more numerous because there are a lot more copies of them. The problem with this probabilistic answer is that probabilities of various worlds being actual surely can be irrational numbers (certainly objective probabilities of events can be irrational, if one thinks that there are events of the sort that non-deterministic interpretations of quantum mechanics talk about), while this approach would make them only rational numbers, i.e. ratios of numbers of world copies. So this will probably not do.

Multiple finite numbers of copies of a world thus introduce a very unfortunate element of arbitrariness into EMR. And, plausibly, if some world has exactly 167 copies, could there not have been 168? However, this possibility is one that EMR has no conceptual resources to analyze. For suppose there are exactly 167 copies of this world (counting this world as a trivial copy of itself), and we wish to say that there could have been 168 copies of that very world. How could we express this possibility? First of all, EMR as it stands has no conceptual resources for possibilities for multiple worlds at a time. But maybe we can extend EMR. We can say some proposition about multiple worlds is possible if there is some set of multiple worlds that is aptly described by that proposition. Thus, the claim that there could have been 168 copies of this very world is true if and only if there *is* a set of 168 worlds which makes it true. But these 168 worlds would all have to be copies of this world, thereby contradicting the assumption that there are precisely 167 such worlds. So, no, even an extension of EMR cannot handle this conceptual possibility. But unless there is a non-arbitrary account of the numbers assigned to each world, this possibility is very plausibly there.

Nor will it do to posit *infinite* sets of copies of each world. For the cardinalities of these sets will have the same arbitrariness that the finite numbers would. Why should there be $\aleph_{167}$ instead of $\aleph_{168}$ copies of this world? Is there not then a conceptual possibility of this cardinal number being different, a possibility EMR cannot countenance? This arbitrariness would make the theory of multiple copies of Lewisian worlds very implausible, even if the argument based on the PII is unsound.

There is a solution to the arbitrariness problem, and this is to posit that there is a Cantorian Absolute Infinity of copies of each world. Cantor believed that there was an Absolute Infinity, in which all the paradise of sets was contained.[10] However, the notion of Absolute Infinity is so murky that to have to escape to it is unacceptable. And of course if one escapes to it then one presumably can no longer even say that the collection of all worlds is a proper class. And this is surely absurd if the worlds are concrete things.

Another solution to the arbitrariness problem would be to say that the number of indiscernible copies of each world depends on the number of individuals in it. Suppose we have a situation where in the actual world there exist two individuals, A and B, perhaps identical twins, and there is a non-actual possible world w_1 which has an individual C who plays such a role R in w_1 that it was both possible for A to have played R and for B to have played R. Here, I stipulatively take the "role" of an individual so expansively as to fix all the qualitative aspects of the world inhabited by the individual. The role is a conjunction of all of an entity's intrinsic and extrinsic properties specifiable in purely general terms without mentioning identities of particular individuals, i.e. it is a Leibnizian complete individual concept. Then, given identity theory, there is good reason to think that there is at least one additional indiscernible copy of w_1, since there must be one world where A plays that role and another where B plays that role. This gives us a natural way of accounting for the number of copies of a world. The number of copies of a given type of world is equal to the number of ways that n-tuples of individuals, where n is the number of individuals of that world type, could be assigned to the roles of the individuals of that world type. So the identity version of EMR may have a non-arbitrary answer to the question of how many copies of a world there are.

And perhaps the counterpart theorist could suppose indiscernible worlds for a similar reason. Suppose that that copy of w_1 is actual in which the given role is played by A. Let that copy be denoted by w_A. Then it is true at w_A that *someone else* could have played the role that A in fact plays. It is difficult to make sense of this on standard counterpart-theoretic Lewisian semantics with worlds having only single copies. To see this, suppose for simplicity

---

10 John Leslie (personal correspondence, 1999) used a similar solution to solve a problem for his theory according to which, in order that the world be the best possible, there are infinitely many indiscernible omniscient deities. For then one wants to know: What infinity of deities is there? So as to preclude the possibility of there being more, the answer is Absolute Infinity. The objection I will give in the text will also apply to Leslie's case.

that w_A is actual. Then a natural way to try to make true the possibility that someone else could have played the role that A plays is to find a possible world where someone who is not a counterpart of A plays that role. But given the expansiveness of roles, anybody who plays the role of A will be indiscernible from A, and hence will surely be a counterpart in every context: it seems one cannot get more similar than being indiscernible.

However, actually, there is a sense in which one can be more similar than an indiscernible copy. One can be *numerically identical*. Thus, a counterpartist defender of EMR could hold that when we say that someone else could have played the role of A, we are setting a uniquely strict contextual standard for counterparts. This standard is so strict that only A qualifies. But then it *is* true that someone other than A could have played the role of A if there is an indiscernible copy of w_A. For at that copy world, the individual who is the indiscernible copy of A will not be a counterpart of A by the extremely strict standard adopted for the counterpart relation, and hence relative to that standard, that world will make it be true that someone other than A could have played the role of A. This is not completely satisfactory, because when we say that someone else could have played the role of A we probably do not say it in this super-strict context, but let us grant this option to Lewis.

Now, we might go on to say that the number of copies of a world represents the number of n-tuples of possible individuals that could play the same role that the actual world's individuals in fact play. This makes the number of copies of a world actually tell us something about the space of possibility, and hence not be arbitrary.

Unfortunately, this attractive story does not really work, because this account still appears to underdetermine the number of individuals that could fill a given role, leaving some arbitrariness. To see this, first, given an actual individual A, define $i(A)$ as the number of possible individuals that can fill A's actual role. Suppose that the world contains n individuals, $A_1,...,A_n$. The number, call it N, of indiscernible copies of the actual world then is the number of n-tuples of individuals that can jointly fill the respective roles of $A_1,...,A_n$. What is the relation between the numbers $i(A_1),...,i(A_n)$ and N then? Obviously, N cannot exceed the product of the numbers $i(A_1),...,i(A_n)$, since each distinct n-tuple of possible individuals is a different combination of n possible individuals for the respective roles. Actually, typically, N will be less than the product, because not all combinations will be viable if $n>1$. For instance, it may not be possible that A_1 fill the role of A_1 while A_2 does as well, because a part of the role of A_1 might be to be the tallest person, and two people cannot each be the tallest person.

Thus,

(42) $N \leq i(A_1)i(A_2)\ldots i(A_n)$

There is also an inequality going the other way. Clearly, $i(A_k) \leq N$ for every k. Therefore:

(43) $i(A_1)i(A_2)\ldots i(A_n) \leq N^n$

Now remember that the number of copies of a world was chosen to make sense of the intuition that there is a fact of the matter about how many individuals could fulfill a given role, and the number of copies can be interpreted as the number of n-tuples of individuals that can fulfill the roles of the world's individuals. However, while one can use the number of copies of a world to ground an answer to the question of how many n-tuples of possible individuals can respectively jointly fulfill the roles of the world's n-tuple of individuals, this number does not appear sufficient to ground an answer to the question of how many possible *individuals* can fulfill the role of, say, A_7: the only constraints we seem to have on, say, $i(A_7)$ are (42) and (43) But if there is a fact of the matter as to how many n-tuples can fulfill the roles of the world's n-tuple of individuals, by the same token there should be a fact of the matter as to how many individuals can fulfill the role of A_7, i.e. a fact of the matter about what the number $i(A_7)$ is.

Thus, while the identity variant of Lewis's account gives a solution to the arbitrariness problem for thinking that there can be more than one copy of a world, the counterpart variant does not give a satisfactory solution. Thus, if we are looking at the counterpart version of Lewis's theory, by Ockham's razor and out of a desire to avoid arbitrariness we will be justified in assuming that there is only one copy of each world. Since the counterpart variant is anyway the orthodox version of EMR, we can safely assume in many of our arguments against Lewis that there are no indiscernible worlds. And an attentive reader will notice that at least some of the arguments can be made to work in the indiscernible worlds case, too.[11]

---

11 For examples, the inductive argument in Section 9 will work as long as the induction-friendly worlds do not have more indiscernible copies than induction-unfriendly ones, and it would be quite surprising metaphysically if they did.

Section 4 Lewis's arguments for his ontology

Lewis presents a cost–benefit argument for the truth of his theory. On the benefit side, it solves the problem of what modality is and, he thinks, has other useful applications. Let us consider the benefits.

Recall the extended Parmenidean challenge. It seems that when we talk about things that are non-actual, we are neither talking of anything that is nor saying that something is not, and hence our sentences are either false or meaningless. Moreover, even if it is granted that there is some way of rendering talk of unactualized possibilities meaningful by providing it with objects, nonetheless it is not clear *what* sorts of things modal language affirms of these objects. EMR handles both concerns. What is a possible but non-actual horse? It is nothing but a horse, albeit in a different world. We know what horses are, and the non-actual horse satisfies the description "horse" in exactly the same way the familiar horses in this world do. The non-actual horse exists, though not here it. Neither is there any problem in talking of a unicorn or an unphilosophical analogue to Socrates. All these things exist elsewhere. There is no difficulty about their ontological status, because they have the same ontological status as the things we are familiar with. They only differ relationally from more familiar things in that they fail to be spatio-temporally related to us.

Moreover, EMR tells us what we are actually asserting of a situation when, e.g. we say it is possible but not actual. Rather than predicating some mysterious shadowy existence of a shadowy object, we are simply affirming that the object exists and lacks the relational property of being spatio-temporally related to us. EMR allows us to reap in the fullest sense the benefits of a possible worlds rendition of modal language (cf. Section 1). On the official story, there is no primitive modality left on the ontological level: everything modal is analyzed in terms of the relations between concrete and unproblematic entities. The mystery of modality is removed.[12]

It would be a mistake to object that what worlds exist depends on what worlds count as possible, since the only way we can know about what worlds exist is in terms of what worlds are possible. Epistemological

---

12 Skyrms (1976) criticizes this kind of cost–benefit argument for EMR. Possible worlds are physical entities on Lewis's account. Thus only what *physical* theories require should count as evidence in favor of them (1976: 326). However, physical theories have counterfactual implications, and thus presuppose modality, and if modality requires EMR, then physical theories require EMR. That something is water entails certain claims about how it would behave in counterfactual situations, and thus the theory that there is water has modal force.

priority does not imply ontological priority, as Aristotle had already noted. According to EMR, while it may well be that modal claims are "more knowable to us," the plurality of worlds is "more knowable by nature," to put it in Aristotelian terms. Thus, a significant advantage of EMR is that it does explicate modality, as long as we are satisfied with counterpart theory or can make identity theory work.

We have seen in Section 2.3 of Part II that the use of possible worlds to explicate propositions and properties is not helpful. However, we have also seen that possible worlds can be used to clarify other modal concepts, e.g. supervenience and counterfactuals, which are difficult to clarify without invoking possible worlds. Lewisian possible worlds are eminently suited for this clarificational task. Admittedly, it does not seem that the order of time cannot be analyzed completely in terms of possible worlds — Lewis's counterfactual asymmetry approach has failed, we now know — and likely neither can causality. But there are enough applications left that there is good prima facie reason to believe Lewis's theory, unless there are some strong objections. But, in fact, there are decisive objections.

Section 5 Objections to Lewis's indexical account of actuality

5.1 Common sense

Normally we make no distinction between actuality and existence. Of course, there is a sense of the word "actual" in English which carries connotations of presentness, and this differs from existence, but I shall scrupulously avoid using "actual" in that sense. We say interchangeably that no unicorn exists and that no unicorn is actual. Lewis thus draws a distinction that common sense does not make. But of course, as Aristotle already observed, the *hoi polloi* are not very good at making distinctions, and so we philosophers must often introduce distinctions where there were none. This in no way counts against Lewis.

However, what does count against Lewis is that the distinction he makes is one that the language of educated English speakers is already capable of marking, and it is they who can make a distinction where Lewis sees none. For, we can already speak in English of something "being spatio-temporally related to us." And this is a concept that is different from our concept of actuality. Even if we think that as a matter of fact everything actual is spatio-temporally related to us (and many of us don't: after all, many English speakers are theists of the sort who believe God is beyond space and time,

and many are open to the possibility of us living in a multiverse of spatio-temporally disconnected island universes), we do not analyze actuality into such spatio-temporal relatedness. Lewis thus *drops* a distinction which our language has: the distinction between being spatio-temporally related to us and being actual, and the distinction he is apparently introducing is not new. Thus, he cannot defend himself against the accusation of violating common sense by saying that he is introducing a distinction.

Lewis says that a "spokesman for common sense" would affirm three theses:

(44)   "Everything is actual."
(45)   "Actuality consists of everything that is spatio-temporally related to us, and nothing more (give or take some 'abstract entities'). It is not vastly bigger, or less unified, than we are accustomed to think."
(46)   "Possibilities are not parts of actuality, they are alternatives to it."[13]

Lewis then goes on to say that his critics take (44) to be analytic and (45) to be "up for grabs." However, Lewis writes:

[T]he two theses, indeed all three, are on an equal footing. Together they fix the meaning of 'actual', but they go far beyond just fixing meanings. I don't see any evidence that the analyticity is concentrated more in some of them than in others.[14]

Therefore, Lewis does not think he goes against common sense any more by violating (44) than he would by violating (45). Presumably the point here is that if the alternative to Lewis's view of actuality is one on which the whole "paradise" of EMR with all possible worlds is seen as actual, then (45) would be violated just as unacceptably. But even if this were so, the serious alternatives to Lewis's account are not in violation of (45). The serious alternatives keep all three theses basically in place (though in regard to (45), we need to remember that most people believe in non-spatial entities like deities, souls, or spirits, and many even believe in a God who is not spatio-temporally related to anything), but replace the "paradise" of EMR with some other account of possible worlds.

---

13   Lewis (1986a: 99).
14   Lewis (1986a: 99–100).

Moreover, it is worth noting that the evidence of common sense for (45), which Lewis so much wishes to respect, is no greater than the evidence of common sense for:

(47)   *Existence* consists of everything that is spatio-temporally related to us, and nothing more (give or take some 'abstract entities'). It is not vastly bigger, or less unified, than we are accustomed to think.

And *this* Lewis rejects. Since common sense affirms *both* (45) and (47) with equal force, indeed seeing them as equivalent, and since Lewis must go against (47), his acceptance of (45) is not much comfort. It would be best if we could keep all four theses, at least to some extent.

Another commonsensical objection to Lewis's account is that it has the absurd consequence that there are people with whom we do not disagree, though they believe propositions we disbelieve. For observe that if Smith in some other world believes that there actually are unicorns and I believe that there actually are no unicorns, then the proposition Smith believes is the negation of the proposition I believe. (This is particularly clear on Lewis's account of unstructured propositions, since the proposition I believe is the collection of all worlds that do not contain unicorns, and the proposition Smith believes is the complement of that collection of worlds.) But Smith and I do not disagree. And if his world contains unicorns, then I will say that his doxastic system is just as doxastically successful as mine in respect of the existence of unicorns. But it is absurd that two people should believe logically incompatible propositions and both be doxastically successful in respect of these propositions.

5.2 Is "the actual world" indexical?

Richard Gale (1991, chapter 5) has objected to Lewis's view that actuality is indexical essentially on the grounds that this gives the wrong truth values to sentences. The easiest way to see the problem is just to observe that just as the tallest woman might not have been the tallest woman, so likewise:

(48)   The actual world might not have been the actual world.

But indexicals rigidly designate their referents — to assert literally "I might not have been I" is to assert an impossibility. On the other hand, "the actual

world" often is a definite description that designates non-rigidly, as in (48). And so Lewis must judge (48) false, while it is true.

Nonetheless, *sometimes* "the actual world" is used rigidly. Thus, we might say:

> (49)   "There could have been someone taller than everybody in the actual world."

Here, if we took "the actual world" as a definite description, we would get a claim that entails the absurdity that there could be someone taller than herself.

The same point applies to cognates like "actually" or "actual." "There could be someone distinct from every actual person" says basically that if S is the set of actual people, there could be someone who isn't (wouldn't be?) a member of S. On the other hand "If Sally had finished her masterpiece, there would have been an actual example of a science-fiction novel that rises above the great Russian novels" does not entail that there are any actual examples of science-fiction novels.

The fact that often "the actual world" and cognates are used non-rigidly, and naturally so, should be an embarrassment to the view that these are indexicals. That we use indexicals non-rigidly in humorous statements "Tomorrow never comes" (because once it comes, it's today) underscores my point: there is no humor in (48). Perhaps, however, Lewis could take refuge in the observation that we *do* sometimes use indexicals and demonstratives non-rigidly in serious speech. After all

> (50)   "You always say that you'll drop in on me tomorrow"

seems to uses "tomorrow" non-rigidly, and

> (51)   "You're never there when I need you"

seems to use "there" non-rigidly. However, (50) is better transcribed as:

> (52)   "You always say that you'll drop in on me 'tomorrow'"

while the "there" in (51) is anaphoric with an implicit antecedent ("the place where I need you") rather than demonstrative.

But if "the actual world" is not indexical, then it seems it has to be either some other kind of rigid designator — which will run into the same problems

as the indexical reading — or a definite description. For instance, if worlds
are maximal consistent sets of propositions, then we might say that "the
actual world" is to be analyzed as "the world all of whose members are
true." But if "the actual world" and its cognates are definite description,
then cases like (49) will seem to be an embarrassment. So perhaps neither
side in the debate can avoid embarrassment.

However, the definite description side can avoid embarrassment, and in
two ways. We do in fact rigidify definite descriptions. Thus, one might say:

(53)   I wish that the greatest butcher of the 20th century,
        whoever he is, had died in infancy

without wishing something that would entail that the greatest butchery of the
20th century have been performed by an infant. In such cases, the article "the"
is really the rigidified definite description particle "dthat" of Kaplan (1991).

The second approach is to note with Donnellan (1966) that definite
descriptions can be used referentially in an assertion, without the assertion
claiming that the description is correct. "The man who is talking on his
iPhone is a superb physicist" can be true even if the physicist has a cheap
knock-off instead. It is very plausible that referential uses are always rigid.
"I wish that the man who is talking on his iPhone had left his mobile at
home" seems to be fine English and a sensible wish whether or not the man
is talking on an iPhone or on some other phone. It is, thus, quite possible
that in at least some cases — and (49) could be an example — we might have
"the actual world" be rigidified by being used referentially.

All in all, the definite description view of actuality is better able to make
sense of our language than an indexical view. The definite description view
is incompatible with the Lewisian account of modality, however. For if only
one world satisfies the definite description "the actual world," then it would
be false according to EMR that a world with unicorns could have been the
actual world. For according to EMR, a world with unicorns could have been
the actual world if and only if there is a world such that (a) it has unicorns;
and (b) it is the actual world. But there is no such world, since the only actual
world has no unicorns.

5.3 *Substitution*

However, perhaps there is another sense in which actuality can be taken to be
broadly indexical, or at least not a definite description. Rather than trying to

paraphrase "the actual world" by a single phrase, we might simply explain how this term behaves semantically, context by context.

To do this, first paraphrase all modal statements into quantifications over possible worlds; e.g. "It is possible that p" is paraphrased into "$\exists w(p$ is true at $w)$." It is important here not to leave any implicitly modal operators unparaphrased (e.g. we must take care to translate "p entails q" into "$\forall w((p \supset q)$ at $w)$"). We now proceed recursively, limiting ourselves for simplicity to sentences not involving any propositional-attitude or modal verbs, or any cognates of "the actual world" other than "the actual world" itself. Suppose s is a sentence where all modalities got paraphrased into quantifications over possible worlds as above. Then, those occurrences of "the actual world" that occur outside the scope of the left hand sides of any "... is true at ..." operators are replaced with "*this* world." In the case of the remaining occurrences, we replace each "the actual world" by "w" where "... is true at w" is the *innermost* "... is true at ..." operator in the scope of whose left hand side the given occurrence of "the actual world" is found, when "the actual world" is used non-rigidly, and we replace "the actual world" with "*this* world" when it is used rigidly. Thus,

(54)   "((That (that the actual world contains unicorns is true at w) and horses exist in the actual world) is true at v) and z is the actual world"

is to be paraphrased into:

(55)   "((That (that w contains unicorns is true at w) and horses exist in v) is true at v) and z is *this* world."

If this paraphrase works, then it may give some sense in which "the actual world" is indexical, namely that "the actual world" is a variable symbol that picks up different objects depending on context. And it is easy to convince oneself that this approach preserves truth values.

The paraphrase has the unfortunate property, however, that if O is a modal sentential operator, s a sentence, Tr the translation of a sentence r according to the above scheme, then $T(O(Ts))$ and $T(Os)$ can have different truth values. To see this, let s be the sentence "The actual world contains horses." Then, Ts is "*This* world contains horses." If O is $\Box$, then in fact $O(Ts)$ is true, because the rigidity of "this" ensures that it is true at every world that *this* world contains horses. But then likewise $T(O(Ts))$ is true because T preserves truth value. But, Os is "Necessarily the actual world

contains horses," which appears false, and hence $T(Os)$ is likewise false. Thus, the translation fails to behave well under embedding.

So perhaps one should not, in charity, characterize Lewis's view as saying that "actuality" is to be translated as an indexical expression. Rather, I characterize Lewis's view as making an ontological claim that we can abbreviate with a phrase like "actuality is nothing but an indexical property." By this we mean that any sentence involving "the actual world" and cognates, while perhaps not paraphrasable into sentences involving just indexical operators, can nonetheless be given *truth conditions* that do not involve any occurrences of "the actual world" or its cognates, but instead either substitute demonstratives or variable letters coming from ". . . is true at w" operators. This situation is not in any obvious way absurd. Michelle Beer (1988) has argued that in the theory of time, the B-theorist's best bet is to claim that A-propositions can be given truth conditions expressed in terms of B-propositions, even if there is no possibility of a paraphrase. If one further thinks that the B-propositions are ontologically prior, then this might count as an ontological reduction. We can take Lewis to be making a similar claim about "the actual world" and cognates, and the s to Ts operation shows how the truth conditions are to be evaluated.

More weakly, one could at least say that Lewis is claiming that one could give complete instructions as to when, where and in what world assertions involving "the actual world" are correct, with the instructions not involving the term "actual" or any non-relativized cognate such as "actually true."[15] This is parallel to the way that I can use B-language to give instructions to completely specify when the expression of an A-proposition is appropriate.

Section 6 The possibility of spatio-temporally unrelated co-actual entities

A consequence of Lewis's view that worlds are universes, i.e. maximal spatio-temporally related aggregates, is that it is a necessary truth that all entities are spatio-temporally related. For indeed, at every world, each pair of entities is thus related. Lewis writes:

> A first, and simplest, objection is that a world might possibly consist of two or more completely disconnected space-times . . . and one world with two disconnected space-times is a counterexample against my proposal. Against

---

15   I am grateful to Robert Brandom for this suggestion.

this objection, I must simply deny the premise. I would rather not; I admit some inclination to agree with it. But it seems to me that it is no central part of our modal thinking, and not a consequence of any interesting general principle about what is possible. So it is negotiable. (1986a: 71)

However, it is not clear that this is negotiable if one is to hold that EMR is intelligible. For let us start off from an ordinary pre-EMR understanding of modality. Either the notion of two completely disconnected space-times is intelligible or not. If it is intelligible, then there is a presumption in favor of its possibility. But if it is unintelligible, then EMR itself will be unintelligible. After all, EMR crucially depends on the idea that distinct worlds are completely disconnected space-times.

This argument against EMR appears to commit the fallacy of equivocation. Lewis distinguishes between actuality and existence, and can say that although it is intelligible that two completely disconnected space-times *exist*, it is unintelligible that they be both *actual*. However, ordinary modal thought makes no such distinction, and consequently it finds two existent completely disconnected space-times intelligible if and only if it finds two actual completely disconnected space-times intelligible.

Of the two horns in the dilemma argument, Lewis cannot accept the horn that according to ordinary modal thought two completely disconnected space-times are unintelligible, for then someone who subscribes to ordinary modal thought would simply fail to understand EMR and it would be irrational for him to become an extreme modal realist. Thus, Lewis must insist that although ordinary modal thinking renders the notion of two completely disconnected space-times intelligible, nonetheless this is one of those cases where intelligibility holds, but possibility does not. Violating the presumption that the intelligible is possible is an additional price Lewis can pay, but it is not a decisive consideration.

What is more serious, however, is Lewis's inability to fit into his theory entities that are not spatio-temporal at all. Lewis (1986a) himself discusses the case of spirits and comments that perhaps he does not need "to defend the possibility of spirit tales — after all, people have been known to accept impossible theories, as witness naive set theory" (1986a: 73). But surely there is a presumption of possibility when dealing with statements in a well-established language game where the statements have not been shown, and do not appear to be in any imminent danger of being shown, to be incoherent. Lewis, however, has a better answer than just denial of the possibility of spirits. After all, the spirits will be temporally related, which is a special case of spatio-temporal relation. Temporal relatedness is not as

tight a relation as exists between material objects, but it is good enough for binding the spirits into worldmates.

Presumably the sorts of temporal relations that Lewis is thinking of between ghosts are things like *later than*, *earlier than*, and *simultaneous with*. But why does it make sense to talk of non-spatial entities as standing in such relations? Is it any less mysterious that two ghosts are temporally related than to say that they are worldmates in some primitive sense? We can make sense of what it means for ordinary spatio-temporal objects to be spatio-temporally related. Objects occupying a single spatio-temporal continuum can be brought together in contact with one another, made to interact with one another. Apart from this possibility, it is not at all clear that it is intelligible to suppose that they *are* spatio-temporally related.

Of course, likewise, non-spatial objects like ghosts can be conceived of as interacting, too. One ghost might in some way communicate with another, for instance. And if the pattern of causal interaction is tightly enough knit together, this might suffice for the intelligibility of the claim that they inhabit one time. I claim that in fact it is the capacity for causal interconnection rather than occupation of a single temporal order that is more plausibly considered as rendering two ghosts worldmates. Imagine, perhaps *per impossibile*, two non-spatial ghosts between whom there is no possible causal interconnection.[16] Is it any less mysterious to say that they occupy one temporal sequence than to say they occupy one world? If Lewis does not think the idea of a primitive worldmate relation is acceptable (I suppose it would be "magical," in the way that he thinks the representation relation posited by his ersatzist opponents is supposed to be magical — see Section 3.4 of Lewis 1986a), neither should we accept a primitive co-temporality relation. Besides, if, as I have argued in Section 2.3 of Part II, the arrow of time is grounded in the direction of causation, it is plausible to suppose that all of temporality is causal in nature, given that the directedness of time is essential to the concept of temporality (this seems to be one of the hard to dispute points in McTaggart 1908; it is because of this that McTaggart rejects the C-series as an account of *time*).

Unfortunately, the suggestion that the capacity for causal interconnection is what makes two ghosts occupy the same temporal continuum is one that Lewis cannot adopt. For, first, it is highly plausible that *unless* there

---

16 Ultimately, I do believe that this is a *per impossibile* consideration. For, it will be essential to my full account of modality in Part VI that there be a God who necessarily is the creator of all things that exist. But then all things that exist are causally related to God, and hence causally interconnected. Thus, this argument is ultimately *ad hominem*.

necessarily is a single creator to whom all things are necessarily causally interrelated, something that Lewis does not admit, there *could* be two ghosts that are not capable of causal interrelation. But this is a possibility Lewis cannot admit if the *capacity* for causal interconnection is what makes two ghosts live in the same world. Moreover, if this suggestion were taken, then possible worlds would no longer be the logically primitive basis of an analysis of modal language. For, the boundaries between possible worlds would depend on causal *capacities*, and capacities are modal.

It does not, thus, appear possible for Lewis satisfactorily to accommodate non-spatial concrete entities in his system. But such entities are prima facie possible, and Lewis has not given any argument against their possibility.

Abstracta, however, are a different matter for Lewis: these exist in all possible worlds and are worldmates to everyone by stipulation.

Section 7 Cardinality and the "set" of all possible worlds

7.1 Introduction

Partly in response to a counterexample by Forrest and Armstrong (1984), Lewis has modified the principle of recombination on which their argument was based. The principle of recombination is a crucial ingredient of Lewis's view in that it gives a criterion for what worlds are possible, based on the idea that anything could coexist with anything,[17] i.e. any number of duplicates of things could be patched together into a world. The original principle was that

> [n]ot only two possible individuals, but any number should admit of combination by means of coexisting duplicates. Indeed, the number might be infinite. (Lewis 1986a: 89)

I shall call this the "unrestricted principle of recombination." To meet Forrest and Armstrong's counterexample, Lewis then added the restrictive

---

17 An idea that is in general questionable, but which in this section I shall accept for the purposes of argument. Moreover, even though I myself would reject the *general* claim that anything could coexist with anything (e.g. the God of traditional monotheism *cannot* coexist with another copy of himself), nonetheless I do think that some restricted principle sufficient for my argument may be correct. The only thing I ask from this principle is that for any cardinality *n* there be a possible world containing exactly *n* photons.

proviso: "size and shape permitting" (Lewis 1986a: 89), thereby opening up the possibility that there is a logical limit on the cardinality of space-time which would, according to Lewis, impose an upper limit on the number of things that can be found in it, thereby blocking Forrest and Armstrong's gigantic world.

I shall use a variant of an argument that Nolan (1996) has formulated but defends Lewis against, and argue that it works even with Lewis's proviso. Although the argument will be formulated in terms of Lewis's EMR, it still does lead to a substantial conclusion on *any* reasonable theory of possibilia: the collection of possible worlds cannot be a set, if actual infinites are possible.

7.2 There is no set of all possible worlds

I will need a slightly stronger unrestricted principle of recombination than the one literally given by Lewis. The stronger principle is that for any thing and any number n (possibly infinite), there is a possible world that contains exactly n duplicates of that thing. As Nolan (1996: 245) notes, this is stronger than saying that there is a world that contains n duplicates of that thing, since, e.g. the latter principle opens up the possibility that maybe there is no world that contains exactly seven photons but there is one that contains exactly eight, as the eight photon world also contains seven photons. However, the stronger form of the principle of recombination is surely what Lewis intends. Otherwise, the principle fails to do what it was designed to do: be a powerful tool for constructing possible worlds according to the idea that anything can coexist with anything. Our intuition says that a world with *exactly* seven photons is possible, and we will not be satisfied with a world with eight. Henceforth, the "unrestricted principle of recombination" shall be this stronger one.

For a *reductio*, suppose that the collection of all possible worlds forms a set W. Let n be any cardinality greater than the cardinality of this set (for instance, by Cantor's diagonal argument, we can let n be the cardinality of the power set of W). Assuming the Axiom of Choice, there then exists an infinite number n^*, which has the property that there are exactly n different infinite cardinalities between $\aleph_0$ (the cardinality of the set of integers), inclusive, and n^*, exclusive. The number n^* is denoted by $\aleph_n$, the nth element in the aleph sequence.[18] In other words, there exist exactly n distinct cardinal

---

18   We can show this in the notation of Kuratowski and Mostowski (1976): the number n

numbers m satisfying the inequality $\aleph_0 \leq m < n^*$. The unrestricted principle of recombination then says that for every such number m, there is a possible world, call it w_m, which contains exactly m photons. Now, if m and m' are different, then w_m and $w_{m'}$ are distinct since then they contain different numbers of photons. Hence, there are at least n distinct worlds of the form w_m, as there was one such world for every value of m satisfying $\aleph_0 \leq m < n^*$, and there were n values of m satisfying that inequality. Hence, there are at least n distinct possible worlds, which contradicts the assumption that n was *greater* than the cardinality of W. Therefore, we have a *reductio ad absurdum* of the assumption that the collection of all possible worlds forms a set.

Hence, the unrestricted principle of recombination entails, *pace* Nolan, that the collection of all possible worlds is not a set. Note that in the derivation of the aleph-sequence, this argument has implicitly used the Axiom of Choice (AC), which states that for any set S of disjoint non-empty sets, there exists a set which contains exactly one member from every set which is a member of S.[19] There is intuitive plausibility to this axiom. It functions in mathematics as a kind of principle of plenitude for sets. Even though it has been shown that the AC cannot be proven from the other axioms of set theory, mathematicians tend to accept the AC because it is needed for important theorems like the Hahn–Banach theorem on the extension of linear functionals on infinite dimensional vector spaces. Of course the AC does lead to some paradoxes like the Banach–Tarski theorem that a solid sphere can be disassembled into a finite number of pieces which can be reassembled to form two spheres of the same size. However, arguably, such paradoxes are not much stranger than the "usual" paradoxes of infinity like Hilbert's hotel. If we were forced to give up the AC in order to hold on to Lewis's EMR, the price would be too high.[20]

---

is the cardinality of some set, say, S. By the Axiom of Choice (1976: 254, Theorem 1), S is well-orderable, and so we can find an initial ordinal ω_n (often also denoted $\aleph_n$) with index equal to the cardinality n of W (1976: 273, Theorem 5). By definition of the "index," it follows that there are precisely n infinite initial ordinals less than ω_n (1976: 273, Definition 1). This means that there are precisely n cardinalities between $\aleph_0$ (inclusive) and ω_n (exclusive). Then let $n^* = \omega_n$. (*Note:* Throughout their p. 273, Kuratowski and Mostowski talk of "initial ordinals" where they mean "infinite initial ordinals.")

19 See previous note.
20 Note that the use to which the AC is put in this paper is the construction of very large cardinals like $\aleph_n$ for an infinite n. Even though one might think that mathematicians do not need large cardinal numbers constructed in such ways, and hence do not need the AC at least in this connection, they in fact do. With regard to the worry about the uselessness of these kinds of large cardinal numbers, J. D. Monk writes: "it turns

Furthermore, it is worth noting that one can also prove that the collection of possible worlds is not a set without assuming the AC, providing one is willing to grant three new assumptions:

(56)   There is a predicate S such that Sw holds if and only if w is a possible world such that the collection of all photons in w is a set;

(57)   This predicate S can be used in defining sets in the sense that whenever A is a set, there is a set $\{ x \in A : Sx \}$ of all elements of A satisfying S; and

(58)   Sets are necessary beings and essentially have the properties of being sets and possessing their members; i.e. if S is a set in a world w and it is true at w that x is in S, then likewise it is true at any other world w_1 that S exists, is a set, and x is in S.

The first two premises are quite plausible. It is the third premise here that is the most controversial and limits the applicability of this argument. Some possible worlds theorists like Plantinga (1982) will deny (58) on the grounds that if x is an individual that fails to exist in w_1, then one cannot assert at w_1 a proposition like the one that x is in S. Lewis, however, would surely accept (58). Given (56)–(58), assume for a *reductio* that W is indeed a set, and define $W_1 = \{ w \in W : Sw \}$ by (56) and (57). If w is a world in W_1, let P_w be the set whose members are all and only the photons in w. By (58), P_w is also a set in the actual world and has the same members there (though of course only some of the members exist *actually*). In the actual world, let U be the union of all these sets P_w for w in W_1. This union will itself be a set in the actual world, by the Axiom of Union. Let n be the cardinality of U and let n^* be any strictly greater cardinality. Then by our principle of recombination we can find a world w which has exactly n^* photons. Clearly, w will be a world in W_1, since the collection of photons in w will have cardinality n^* and hence be a set, namely P_w. Thus, all the photons in w will be members of U and so U contains P_w. Since n was the cardinality of U and n^* that of P_w, it follows that $n \geq n^*$. But we had assumed that $n^* > n$, and so we have both $n^* > n$ and $n \geq n^*$, a contradiction. Therefore, it cannot be the case that W is a set if (56)–(58) hold.

---

out that many important mathematical questions, for example, in abstract measure theory or in the theory of Abelian groups, essentially involve very large cardinals" (1969: 134).

7.3 *Lewis's proviso*

My arguments in the previous section require the construction of a world containing a very large infinite number of photons. But maybe worlds aren't big enough to have so many photons? After all, Lewis restricts his principle of recombination to allow only combinations which "size and shape" permit. He thinks that if there is a limit on the cardinality of space-time, an option he leaves open, then this induces a limit on the number of distinct things in a world:

> Only a limited number of distinct things can coexist in a space-time continuum. [This number] cannot exceed the infinite cardinal number of the points in a continuum. So if we have more than continuum many possible individuals to be copies, or if we want more than continuum many copies of any single individual, then a continuum will be too small to hold all the coexisting things that our principle [of recombination] seems to require. (Lewis 1986a: 89)

But why should size and shape bar more than continuum many photons existing in one space-time continuum? Of course, since there would be fewer points in the continuum than there would be photons, some (indeed, infinitely many) photons would have to be in the same place in space-time, but why not?

Lewis's speed in arguing indicates that he must be tacitly making use of an intuitive assumption to the effect that no two distinct things can exist at the same point in space-time. However, this assumption is contrary to our current physics which claims that those elementary particles which are bosons "can be in the same state" (Dirac 1987: 210). Photons are bosons, and this is the reason why in the arguments above I made use of photons rather than the electrons of some other authors, electrons being fermions and thus barred from sharing states (ibid.). Photons have "size and shape" and other properties such that more than one is permitted in the same place, and hence Lewis's proviso "size and shape permitting" fails to rule out the possibility of there being more than continuum many, indeed any arbitrary cardinality, photons in a space-time continuum.

Lewis has two ways out if he wants to maintain a limit on the number of photons, a limit he must maintain if he is to continue claiming that the collection of possible worlds is a set. First, Lewis could say that current physics is wrong about what is physically possible. He could insist that there are *meta*physical reasons why multiple things cannot share a place in space-time. But unless the metaphysical arguments had *very* plausible

premises, one would just take the evidence for current physical theory to be a disconfirmation of the conjunction of the premises of these arguments.[21] Alternately Lewis could insist that although there is no impossibility about bosons sharing the same place, nonetheless there is an impossibility of an *infinite* number (or a sufficiently large infinite number) of them doing that, which would be necessary if there were to be more than continuum many of them in a space-time continuum.[22] However, unless this was grounded in some general argument against actual infinites, this would be ad hoc and contrary to the spirit of the principle of recombination. And Lewis has no objection to actual infinities.

Lewis's present proviso fails to rule out the large cardinal numbers of photons in one world that my arguments against him deploy. Any strengthening of the proviso that would rule it out would either be ad hoc or contrary to current physics or opposed to Lewis's friendliness to actual infinities.

One physical worry about this argument is that when an infinite cardinality of photons is found at the same location in space-time, this will create an infinite energy density, and perhaps an infinite energy density is not possible. If this worry is successful, then the argument will have to be run either in a world with laws different from that of our world's laws, laws that allow for an infinite energy density (e.g. perhaps in a world where energies are calculated using hyperreal arithmetic), or with a hypothetical possible particle that lacks energy and that there can be multiple copies of in one place. It seems possible that a material entity could lack energy, and we know it is possible for more than one material entity to be in the same place, because photons are such. There seems little reason to deny that the two properties could co-exist: that there could be an energy-free material entity that there can be many copies of in one place.

7.4 What should Lewis do about the cardinality problem?

If Lewis is committed to the possibility of an actual infinite, he should just abandon EMR, since there is much intrinsic plausibility to the idea that if possible worlds are concretely existing material objects, as they are on

---

21 Note, too, that one doesn't need the *truth* of the current physical theory in order to disconfirm the claim that two particles cannot occupy the same place. All one needs is the *logical possibility* of its truth.

22 If each point of a continuum held at most a finite number of particles, then the total number of particles would still, as can be easily proved, be at most the cardinality of the continuum.

Lewis's account, then their collection is a set. If he is not to abandon EMR, and wants to allow actual infinities, then he must firmly hold that the collection of possible worlds is not a set, but, say, a proper class.

The worries involved here are somewhat less pressing if one does not take possible worlds to be *concretely* existent as they are in EMR. For although it is very plausible that a well-defined collection of substantial existing material objects such as Lewis's possible worlds would form a set, it is somewhat less clear that the collection of all merely possible worlds considered as pure possibilia is a set — not every collection of abstracta is a set, after all. The cardinality argument as an objection thus weighs more heavily against EMR than against other theories of possibilia. Nonetheless, the argument does demonstrate that any reasonable theory of possibility needs to deny that there is a *set* of all possible worlds, if actual infinites are possible.

Section 8 Ethical issues

8.1 *The ethical objection*

Robert M. Adams (1974: 215–16) has objected to EMR on ethical grounds:

> We may be moved by the joys and sorrows of a character known to be ficti-
> tious; but we do not really believe it is bad that evils occur in a nonactual
> possible world, or good that joys occur in a nonactual possible world, though
> of course it would be bad and good, respectively, for them to be actual. I think
> that our very strong disapproval of the deliberate actualizing of evils similarly
> reflects a belief in the absolutely, and not just relatively, special status of the
> actual as such. Indeed, if we ask, "What is wrong with actualizing evils, since
> they will occur in some other possible world anyway if they don't occur in this
> one?," I doubt that the indexical theory [of actuality] can provide an answer
> which will be completely satisfying ethically.

A similar objection was raised by D. C. Williams in regard to why one should strive to get rid of evils, since the evils one has pushed out of this world just reappear in another.[23]

Lewis, could, of course adopt a revisionary view of ethics that accepts Adams's and Williams's allegations. However, instead he claims that far

---

23 Lewis (1986a: 123n6) attributes this to a lecture of Williams at the University of
    Notre Dame in 1974.

from demolishing ordinary morality: "If my modal realism has any bearing at all on matters of value and morality, it pushes me towards common sense, not away" (Lewis 1986a: 128). I shall argue that while Lewis's arguments do succeed in part in countering Adams's and Williams's charges, and thus allow one to save a *part* of common sense morality, nonetheless there are other important parts of morality that EMR would undermine. Given that those other parts are correct, Lewis's EMR is false.

8.2 *The indexicality of morality*

Consider the following case, which I give on behalf of Adams and Williams[24]:

> (59)   It is possible for me to murder Mr. Smith. Therefore, if I do not murder Mr. Smith, then a counterpart of mine in another possible world will murder the counterpart of Mr. Smith. Having just heard of Lewis's extreme modal realism, I murder Mr. Smith, reasoning that in another possible world — just as *real* as ours — a counterpart of Mr. Smith will die if I do not murder Mr. Smith, and hence by murdering Mr. Smith I do not cause any *overall* decrease in the sum total of good and evil since the life lost in the actual world is saved in another possible world.

Lewis (1986a: 123) agrees that in such a case "[t]here would indeed be the same sum total" of good and evil. He holds that the evils in the other worlds are every bit as *real* as those in ours, though they lack the indexical property of being *actual*, i.e. of being in the same world as me. However he says:

> If you actualise evils, you will be an evil-doer, a causal source of evil. That is something which, if you are virtuous, you do not want to be. Otherworldly evils are neither here nor there. They aren't your evils. Your virtuous desire to do good and not evil has nothing to do with the sum total of good and evil throughout reality. It has to do with what befalls you and your worldmates, and in particular it has to do with the way in which what befalls yourself and others depends causally on what you do.
>
> For those of us who think of morality in terms of virtue and honour, desert and respect and esteem, loyalties and affections and solidarity, the

---

24   See Adams (1974) and Lewis (1986a: 123n6).

other-worldly evils should not seem even momentarily relevant to morality. Of course our moral aims are egocentric. And likewise all the more for those who think of morality in terms of rules, rights, and duties; or as obedience to the will of God. (Lewis 1986a: 127)

There seem to be two distinct replies here. The first insists on moral aims being "egocentric" and on morality being indexed to the moral agent. Moral obligations are indexical and addressed to the individual: "*Thou* shalt not kill." That is why the indexical "you" appears over and over in the first paragraph quoted from Lewis. The second reply insists on the importance of virtues. By murdering Mr. Smith, I make myself into a murderer. I neglect the obligation of respect for the life of Mr. Smith. In doing so, I deprive myself of virtue. And this is bad.

However, the two replies are actually not distinct. For the importance of virtues is also an *indexical* importance for Lewis. My primary task is to promote *my* virtue: "our moral aims are egocentric." Lewis has to claim this, since if I do not murder Mr. Smith and thereby do not engender in myself the vice of murderousness, then my counterpart in another world will murder the counterpart of Mr. Smith, and the vice of murderousness will be engendered in that counterpart so that I could look at my deed of murdering Mr. Smith as a selfless sacrifice of my own virtues for the sake of allowing my counterpart to retain them.[25] Considered impersonally, then, murdering Mr. Smith seems to be a very good thing to do, since by doing so I am saving another human being — my counterpart — from decay of moral character, at great personal cost, and I am not increasing the total amount of evil in any way since, although now Mr. Smith dies, still his counterpart would have died had I not murdered Mr. Smith. To avoid this unsavory conclusion, Lewis needs to insist that it is not *virtue* that should be my primary moral aim, but *my* virtue, just as on deontological theories it is *my own* moral duties that it is impermissible for me to violate.

Lewis's view not only fits well with standard deontological views, but as Turner (2003) notes, a theory like Lewis's EMR is in a somewhat better position than the standard deontological views, because it avoids the deontologist's conclusion that sometimes one ought to act in a way that results in much lower overall utility. For the overall utility, counted over all worlds, is fixed by EMR.

The challenge for Adams and Williams would be to come up with a clear-cut moral objection to David Lewis's view without relying on

---

25 I am grateful to Richard Gale for this observation.

consequentialistic intuitions that Lewis expressly rejects.[26] Such an objection would have to take account of the force of Lewis's observations of the essential indexicality of morality, the apportionment of moral duties to each individual *I*. I shall attempt at least a partial fulfillment of this challenge in the next section.

But before doing this, observe that if instead of counterpart theory we accept identity theory, the Lewisian reply is weakened. For then by murdering Mr. Smith in this world I ensure that *I* do not murder him in another. Lewis's indexicality of morality reply is much less compelling in this case. It is still true that "[i]f you actualise evils, you will be an evil-doer, a causal source of evil," but it is also true that if you do *not* actualize evils in this world, *you* will still be an evil-doer, a causal source of evil in another existent world in which *you* exist. Therefore, even if the moral objection in this section of the paper fails against Lewis's considered view, it damages a "transworld-identity" variant of his theory.

Of course, in reply a transworld-identity Lewisian might claim that morality is not only indexical on the "I," but also on "now" and "here in this world." However, even if this might have some plausibility in the deontological case of avoiding murder (e.g. it might be that one has an obligation to refrain from murdering *now* even if one should somehow know that one will commit it later if and only if one refrains from it now), it has less plausibility for other virtues. Must one strive hard to improve some virtue if one knows that this will result in one not having that virtue at some point in the future or in another world?[27]

8.3 *Ethical problems for counterpartist EMR*

8.3.1 A transworld counterfactual

The objection in Section 8.2 worked for an identity variant of EMR, but failed for counterpart theory. I will now formulate a complex of logically possible ethical counterexamples to counterpart theory. All of these will be built out of the following simple counterexample to Lewis's claim that there are no transworld counterfactuals (cf. Lewis 1986a: 78).

Imagine a possible world where the laws of nature and initial conditions allow for one and only one free *and* indeterministic choice, everything else

---

26 Lewis (1986a: 127) writes: "If modal realism makes a problem for anyone, it is for utilitarians," but neither he nor I sees a disadvantage in this.

27 This section owes much to a number of discussions with Richard M. Gale.

being deterministic and there being no other choices (deterministic or not) nomically possible. At time *t*, I get to decide whether to freely perform action *A* or whether to freely refrain from *A*. There are two variant worlds involved here. Let w_p be the variant world in which my counterpart does perform *A*, and let w_R be the variant world in which my counterpart refrains from *A*. I will say that one world "matches" another if it shares the same initial conditions and laws of nature.

Then, the following subjunctive conditional is true at w_p on standard Lewisian semantics:

(60) Were I to refrain from *A*, my counterpart in the other
      world that matches the actual world ("other matching
      world" for short) would perform *A*.[28]

For, to evaluate (60) at w_p, we look at the closest world to w_p at which the antecedent holds, which is nothing else than w_R. But at w_R it is true that "my counterpart in the other matching world would perform *A*" because it is true at w_R that "the other matching world is w_p," and my counterpart *does* perform *A* there.[29]

8.3.2 The first case: saving a life

Now that the basic situation has been set up, we can formulate two specific ethical cases in this setting. The first case is where *A* is the trivial action of e-mailing a petition to a dictator saying "Free Smith!" which I somehow know would result in Smith, an innocent stranger sentenced to death by a dictator on another continent, being freed. Should my note not be sent, Smith will die; for some reason, I am the only person in a position to send such a note.

Being a Lewisian, I deliberate over the decision as follows. Were I to send the note, a dictator's innocent prisoner would be spared. Were I *not* to send the note, however, my counterpart in the other matching world would send the note according to (60), and hence it would be true that in the other matching world the counterpart of the innocent prisoner would be spared.

---

28 I am assuming throughout the ensuing discussion that the worlds in question have only one counterpart of me.

29 Observe that Lewis's arguments against interworld causality do not apply to this case, because those arguments (Lewis 1986a: 78) assume that one is talking of causing a proposition of the form "*p* is true at *w*" to be true where "*w*" is a *rigid* reference to a world. It is obvious that there is no interworld causality of this kind, since if a proposition *p* is true at *w* where "*w*" is rigid designator, then <*p* is true at *w*> is a necessary truth.

What should I do? Whatever I do, somebody dies. Admittedly, in one case, a worldmate of mine dies, and in the other case someone in another world dies. However, this worldmate is a stranger to me. I know no more about him than I do about his counterpart. Admittedly, he is spatio-temporally related to me (perhaps, though, very distantly — we can change the case so he be in another galaxy and in the far future), while the counterpart is not, but spatio-temporal relatedness does not appear to be a particularly morally salient relation, in the way that friendship or parenthood is.

Lewis's answer, presumably, will be that were I to send the letter, *I* would be causally responsible for saving Smith's life. This, however, will not do if we analyze causation in a Lewisian manner in terms of counterfactuals. For it is just as much true that were I *not* to send the letter, I would cause it to be the case that the counterpart of the prisoner in the other matching world lives. For suppose that I refrain from sending the letter. Then, the counterfactual "Were I to have performed *A*, my counterpart in the other matching world would have refrained from *A*" is true by the reasoning behind (60). But then it is true that "Were I to have performed *A*, the counterpart of the prisoner in the other matching world would have lived." And this counterfactual relation is all there is to causation on Lewis's account. Hence, on Lewis's account, if I do send the letter, I cause the prisoner in the actual world to live, and if I do not send the letter I cause the counterpart of the prisoner in the other matching world to live.

Since it seems to make no difference, I might as well take the easier course, that of not sending the e-mail. But in doing so, I go against ordinary morality, which says that in the situation as set up, if a trivial action of mine can save a person in the actual world — the only concretely existing world ordinary morality knows of — then the action is prima facie obligatory. (Ordinary morality cares that I am working with a case where the choice I make is the only free choice ever made in the history of the world.)

One might think that there is a *de dicto/de re* modal fallacy lurking in the above account. Suppose that "Smith" is a rigid designator for the prisoner in my world. Then, while not sending the letter would make it be that the prisoner's counterpart in the other matching world lives, it would not make it be that Jones lives, where "Jones" is a rigid designator of Smith's counterpart in the other matching world. For, if Jones is a rigid designator, then on counterpart theory propositions such as <Jones will be executed in his world> and <Jones will die in his world> are *necessary* propositions, since they are true at all worlds. So, if I send the letter, I perhaps will make it be the case that Smith lives; if I do not send the letter, I will not make it the case that Jones lives. Thus, I should send the letter.

But this reply fails, because if "Smith" is a rigid designator in exactly the same sense that "Jones" is, then it will be false that if I send the letter, I will make it be that Smith lives. For that Smith lives or is executed in his world becomes once again a necessary proposition on counterpart theory. So given counterpart theory I cannot base my decision whether to send the letter on whether x lives or dies, if "x" is a rigid designator in the strict sense of designating the same individual in every possible world. The only way out is to allow a finite person to bring about necessary states of affairs, which seems implausible.

In fact, Lewis's method for evaluating counterfactuals and *de re* possibility claims makes names not into *rigid* designators, but into quasi-rigid designators. We may call "x" a *quasi-rigid* designator of y if "x" designates y in the actual world, and in every possible world where y has a unique counterpart, "x" designates that counterpart. "Smith" in "If I send the letter, I will thereby cause it to be the case that Smith lives" should be read as a quasi-rigid designator if counterpart theory is correct. I now claim that if we do this, then it is true that

(61)   If I do not send the letter, I will thereby cause it to be the case that Smith's counterpart in the other world matching Smith's world lives,

given Lewis's analysis of causation. For suppose I do not send the letter. The actual world then is w_R. The nearest world in which my counterpart fails to refrain from sending the letter is w_P. At that world <Smith's counterpart in the other world matching Smith's world lives> is false. Recall that "Smith" is quasi-rigid and at both w_R and w_P it is true that Smith is the (relevant) prisoner of the dictator. But then at w_P, it is true that Smith's counterpart in the other world matching his is the prisoner in w_R, who dies. Therefore, the counterfactual

(62)   Were I not to have refrained from sending the letter, Smith's counterpart in the other world matching Smith's world would not live

is true at w_R, and hence it is true that I cause it to be the case that Smith's counterpart in the other world matching Smith's world lives.

Why then should I prefer a scenario on which I cause it to be the case that Smith's counterpart in the actual world (i.e. Smith himself!) lives over a scenario on which I cause it to be the case that Smith's counterpart in the

other world matching Smith's world lives, if all worlds are ontologically on par? If you cut Smith's counterpart, will he not bleed just as much as Smith?

One way out for Lewis is to abandon his account of causation and insist that there is more to causation than the truth of the counterfactuals discussed above, and this "something more" makes it impossible that I cause anything in another world. I do not object to such an account of causation. However, this approach has another weakness: it rests everything on the importance of the fact that by sending the letter I *cause* Smith to live. But ordinary moral thought need not consider this very important. If the people in danger are strangers, I may not care whether *I* save a life or whether someone else saves a life, providing overall a life is saved. It would be nice to save a life, but why shouldn't I let someone else have the pleasure? While there is a moral imperative to take a trivial action if it is true that were that action taken, a life would overall be saved, and were it not taken, a life would not be saved, that moral imperative does not apply here.

And there is no moral imperative to ensure that *I* be the one who saves a life, assuming that the overall consequences are the same. Imagine two Samaritans is walking along and coming upon a wounded man by the side of the road. Both Samaritans would like to help the man, but only the help of one is needed. One of the two Samaritans is zealous to be the helper *himself*. The second Samaritan only cares that the wounded man's condition be ameliorated, and he is fully willing to help the wounded man, but being less zealous, he acquiesces in the first Samaritan's request to be himself the one who helps. Now, on an account on which causation is the main thing that counts, it seems that the first Samaritan is the one who has done a significantly better thing by insisting on *his* being the helper than the second has done in acquiescing. And this seems absurd.

Lewis's moral intuition would lead to an unseemly contest between the two Samaritans as to which one of the two should help the wounded person. Since it is supposed that oneself being the helper is the most important thing, one ought to put in *strenuous* effort to ensure that someone else does not take one's place. But while a short disagreement may be in order, a long discussion as to who should help, with both trying to volunteer, seems out of place even if time is not of the essence.

Of course, sometimes being oneself the cause also matters. It may matter that *I* be the one who saves my child. It may matter that I be the one to save someone whom I actually met with, whose face was an appeal to me (Lévinas 1991: 291). But in the case of a complete stranger whom I have never met, *my* being the savior does not seem to matter very much, and certainly not nearly as much as the fact that a life *is saved*. While with Lewis we should

reject consequentialism as a general account of morals, consequentialistic reasons have a proper place in our deliberation. I may save the life of a stranger solely because I may want fewer people to suffer. Nor is there anything morally blameworthy in this reasoning. On the contrary, one might think that there is something selfish in being motivated by wanting to be the one who saves the life. But Lewis's theory leads to the revisionary result that consequentialistic reasons are *never* relevant.

The indexicality of moral obligation is irrelevant here. For while I do have the prima facie obligation to perform a trivial action if it can save a stranger's life, it is a defeater for this obligation that if I were to take the action, then someone, who otherwise wouldn't die, would die. But were I to save Smith, then Smith's counterpart would die, given Lewis's theory. Likewise, it is a defeater for the prima facie obligation to perform a trivial action to save a complete stranger's (someone that I have met is not a complete stranger) life if there is no consequentialistic improvement *on the whole* from one's action.

Andrew Beedle (1996) has argued that Lewis's reply to Adams's ethical objection implies fatalism. According to Beedle, Lewis is committed to the view that we do not choose which world we inhabit but merely discover it, and since the character of the totality of worlds is fixed, fatalism necessarily ensues. Note that this argument is closely similar to the reply to my self-torture example considered at the end of Section 8.3.3.

My arguments above assumed that we *do* choose which world we inhabit by our actions. But I could instead present a dilemma for Lewis: either (a) we choose which world we inhabit; or (b) we do not. In case (a), my arguments work, and Lewis's view runs counter to ordinary deeply-engrained notions of morality. In case (b), my argument fails, but then Lewis's view runs counter to ordinary deeply-engrained notions of morality even more strongly, since fatalism is contrary to these notions.

8.3.3 The second case: self-torture
Suppose again that I am in one of the two worlds that are deterministic and I am unfree except for one choice I can make. But now this choice is whether to perform act *A*, which is to stick my finger in the light socket, thereby causing myself an excruciating bit of pain, but, I somehow know, not actually killing or causing permanent damage to me. I know nobody in this world would benefit from this. It would clearly be irrational, indeed crazy, to perform *A*. Certainly, the act would not be praiseworthy. However, I shall argue that on Lewis's account, this act could be seen as praiseworthy.

When, and only when, I perform *A*, my counterpart in the other matching world refrains from *A*. Thus, if I perform *A*, my counterpart will be spared

the pain of the electric shock. Moreover, on Lewis's analysis of causality, if I do perform A, then my performing A causes my counterpart in the other matching world to be spared the shock. For then, were I not to have performed A, my counterpart in the other matching world would have — this fact is just a variant on (60). But it is a heroic and praiseworthy thing to spare someone else pain at the cost of receiving serious pain oneself. If I and a stranger were held prisoner by a tyrant, and I were told that if and only if I do not stick my finger in the light socket, my fellow prisoner will be forced to do so, then it would be a heroic thing to do it.[30]

Hence, if Lewis is right, then the act is heroic and praiseworthy. But if ordinary morals are right, the act is neither heroic nor praiseworthy, but insane. Therefore, the implications of Lewis's view in this situation contradict ordinary moral notions. The fact that we are working in a world whose history allows for only one decision fork is irrelevant — our moral intuitions do not care about how many choices are available to how many people.

One might try the following reply on Lewis's behalf. Whatever I do, the numerically same people are shocked and unshocked. After all, the total space of possible worlds is fixed. Let "Smith" be a rigid designator of my counterpart (recall that I define myself as my own counterpart, and so I might *be* Smith) in w_p and let "Jones" be a rigid designator of my counterpart in w_g. Then, Smith gets shocked and Jones does not, whatever I do. So I should not bother to shock myself, given that nothing changes by it. However, if this reply is made, then by the very same standard, in the case considered in the previous section, there is no point in my going to the trouble of e-mailing the dictator, given that nothing changes by it. Hence, the reply considered here is one that Lewis cannot make if he is to maintain ordinary moral views in other situations.

8.3.4 Forking

One might also reply to the last two cases that our moral beliefs implicitly depend on something that at first sight seems irrelevant to them, the fact that our world-history forks in more than one place (not just at the choice between doing A and refraining from A), so that there are many comparison worlds for counterfactuals whose antecedents are our actions. There is nothing surprising about the fact that there is an implicit dependence of our moral beliefs on a merely contingent fact about our world. However, apart from EMR, the multiple forkedness of our world-history seems to have no

---

30 If, as I believe, there is a moral prohibition against suicide or self-maiming, the assumption that the damage to the body is not permanent or fatal is needed.

moral significance for the question of whether I should stick my finger in a light socket or send an e-mail that would free an innocent prisoner. The fact that this contingent fact about our world has the moral significance that EMR forces on it itself shows that EMR is revisionary of our moral beliefs.

Moreover, to escape from the paradoxes of Sections 8.3.2 and 8.3.3 one needs to assume a world-history that not only is multiply forked, but is *infinitely* forked. And surely our moral intuitions don't require infinite forking.

Suppose that we are working with a world that is only finitely many times forked. This means that the world-history is deterministic, except possibly for finitely many points in time when it forks indeterministically, and let us suppose that each fork is between finitely many alternatives. Suppose further that any possible free choice must (according to the way the initial conditions and laws of nature are set up) happen at one these forking points and the choice is between those alternatives that the fork goes into. I shall only discuss the self-torture paradox here, but what I say goes over *mutatis mutandis* to the other case.

Suppose, for example, that there are 460 worlds allowed by the laws of nature. One of these worlds is mine. I am now choosing whether to fry my finger in the light socket or not. Some of the 460 worlds are ones that are no longer available to my choice, having been closed by previous choices that I and others (and perhaps nondeterministic physical processes) have made. But there are, say, 312 worlds left open. Of these, my counterpart (who in the case of the actual world is myself) fries his finger in, say, 112, and avoids the pain in 200. By a "matching world" I shall now mean one of the worlds that are nomically open given all the choices and forkings in the past of the present world. There are thus 312 matching worlds. If I choose to stick my finger in the light socket, then it will be true that 111 people other than me in the matching worlds will suffer this terrible electric-shock pain — for, there are 112 matching worlds where my counterpart feels the pain, but on this choice, I am one of them. But if I don't stick my finger in the light socket, then 112 people other than me suffer this pain in these worlds. Assuming there are no other relevant effects, by sticking my finger in the light socket, I bring it about that one fewer person other than me in a matching world suffers the pain. This is a heroic deed. But of course it is paradoxical for it to be heroic, given that it is crazy.

8.3.5 The need for some consequentialistic considerations in ethics

As a general matter of fact, also, ethics does require some consequentialistic considerations. Suppose you meet a stranger on the street who happens to tell you that he is thinking of giving $1000 to the Helpers of the Hungry

(HH). The stranger is not good at calculation and hates thinking about numbers. But you know that HH would save five lives with the $1000. And, moreover, you know of a charity, the Efficient Helpers of the Hungry (EHH), which would indeed save ten lives with a $1000 donation. The two charities serve the hungry in different distant countries, and neither you nor the stranger has any ties to either country or either country's residents. You also notice that the stranger is someone who just gives to the last charity he's heard of. If you tell him: "Why don't you give the money to EHH instead?" he will do just that, without asking you any questions. If you try to explain anything about numbers of lives saved, he will not listen to you — he just wants to help, but does not care how many he helps.

Now, whether or not one has a duty here, certainly it would be *better* to make the suggestion that the stranger donate to EHH. The natural justification of this judgment is that there is nothing deontologically wrong with speaking out whereas the overall consequences of speaking out are very good — five more lives are saved. Even more clearly, in a reversed case where the stranger originally planned to give money to EHH, it would be wrong to suggest that he give it to HH instead.[31] Again, the grounds for this judgment are consequentialistic.

Lewis cannot make these consequentialistic judgments, since, if EMR holds, then whatever one does, there are infinitely many worlds where money is given to HH and infinitely many where it is given to EHH, and one's own actions make no difference to the overall result. Instead, Lewis is committed to analyzing the matter in terms of causality. Initially, one might try to say that what is good about telling the stranger to give to EHH in the original case is that one then causes ten lives to be saved. But this account fails, for although the saving of ten lives is a good thing, first of all one is at most responsible for five overall lives saved, but more crucially if one admits that in the original case the action is rightly morally described as "saving ten lives," which is good, then one will have to describe the action of suggesting that the stranger give money to HH in the reversed case as "saving five lives," which is *also* good, whereas it is evident that in the reversed case it is wrong to make the suggestion to the stranger. So the description "cause ten lives to be saved" is not the morally relevant one for the action of telling the stranger in the original case to give to EHH.

---

31 The assumption that neither you nor the stranger has any ties to the countries served by HH and EHH may be relevant here. Some might think that if your family is among those that HH would save, you might be justified in suggesting the stranger give to HH instead, assuming he does not actually care about the exact number of people saved.

One might describe the action of suggesting the donation to be given to EHH as "causing five lives to be saved." However, that is inaccurate, because it is only five lives *overall* that are saved, while as a result of one's action five lives are lost that would otherwise have been saved. If one says "five lives were saved," without adding the word "overall," then one is begging for the unanswerable question: *Which* five lives were saved? For after all, there were *ten* saved, though only five overall. If the five that HH would save were a part of the ten that EHH would save, then there would be no problem with specifying which five were saved. But the supposition was that HH and EHH serve different countries.

Now Lewis cannot say that by directing the donation to EHH one has "caused five more lives *overall* to be saved." For the "overall" brings in a consequentialistic consideration. The value of the action described in this way is that in the grand scheme of things more lives are saved. But Lewis cannot make this judgment, unless he makes it in a parochial way by restricting it to our world, which fails to do justice to our moral intuition that the lives of all *existent* strangers in need who are in objectively similar circumstances are of equal significance. More significance does not depend on spatio-temporal relatedness.

Lewis could say that by directing the donation to EHH one has "caused it to be the case that ten lives are saved and five are lost." That is literally true on his counterfactual account of causation: were one not to have told the stranger about EHH, it would not have been the case that ten lives were saved and five were lost. However, this causal formulation again fails to do justice to the moral issues here. After all, a deontologist will say that there are actions that fall under the description "cause it to be the case that ten lives are saved and five are lost" that are in fact immoral: for instance, killing five innocent people to use their organs to save ten lives.

So for the description to carry the moral weight it does, it has to be modified as follows: one has "in a permissible manner caused it to be the case that ten lives are saved and five are lost." Arguably, any action described in this way is a good action, and the description is one that Lewis can give. The problem, however, is in the question of *why* any action described in this way is a good action. In the case of an action where one simply causes it to be the case that x lives are saved one can say on Lewisian ethical grounds that what one has done is a virtuous thing, since what one has done is saved a life, thereby making oneself into the sort of person who saves lives, and this is good. However, in the case at hand such simple reasoning does not work, since the action likewise involves causing it to be the case that some lives are lost that otherwise would be saved, or at least it involves preventing

the saving of lives that could have been saved, and this part of the action description is not one that in any obvious way contributes to having a virtuous character. Rather the only reason to think the action described as above is a good one is that the description entails that the action was one in which one intentionally and permissibly acts in a manner which results in the *overall* saving of five more lives. But this consideration is once again partially a consequentialistic one.

One might try for a different kind of account. By suggesting that the stranger donate to the EHH, "one has caused it to be the case that the stranger performs a better action than the one he had intended to perform," and that is a good thing. But this description is likely to get the reasons for the action wrong. Instead of being motivated by making the stranger more meritorious, one is likely to be motivated by the welfare of the existent starving people that are helped by HH and EHH. Besides, because the stranger does not care about numbers of people saved, and one has not even told him that EHH saves more lives (because you knew there would be no point to this as the man does not care about numbers), he derives little or no moral benefit from the fact that his action was the saving of more lives.

It thus does not thus appear that Lewis can give a satisfactory account of why directing the donation to the EHH is good, because it seems that a satisfactory account will have to involve some consequentialistic considerations.

Section 9 Induction and actuality

9.1 *Introduction*

We can say that howsoever much regularity we might have observed in the world, if Lewis is right we have no reason to think it more likely than not that the regularity will continue. For corresponding to any one world that continues to be regular (i.e. governed by the inductively observed natural laws) after this moment, there are infinitely many logically possible worlds that *were* regular but will no longer be so, and thus we have on balance no reason to think that we're in a world that will continue to be regular. To the one world where the apple dropped tomorrow will fall straight down, there correspond infinitely many different ones with the same past but in which tomorrow the dropped apple will take some tortuous path in an unexpected direction (cf. the argument against Lewis given by Forrest 1982).

Call a world that shares our world's past[32] a "continuant world." There are highly irregular continuants, and Lewis is committed to their existence by his principle of recombination according to which anything could co-exist with anything, spatio-temporal layout permitting. Indeed, on his view, any possible past could co-exist with any possible future.

Adapting Lewis's response to Forrest, we can say that the set of continuant worlds that will behave regularly in the future is infinite and has cardinality equal to that of the set of irregular continuant worlds, so that in fact one can pair up worlds in such a way that to each irregular continuant there correspond infinitely many different regular continuants (cf. Lewis 1986a: 119–20). Lewis's second reply is that Forrest-type problems persist on non-Lewisian views of possible worlds, too, and hence the problem should not count against Lewis (Lewis 1986a: 117–18).

However, from the fact that the cardinalities of the sets of regular and irregular continuant worlds are infinite it does not follow that the cardinalities are *equal*. Intuitively, the cardinality of irregular continuants should be larger, because irregularity is compatible with large cardinalities of photons or other bosons suddenly coming into existence while regularity is not compatible with this. Indeed, for any cardinality n, Lewis should say that there is an irregular continuant world where tomorrow n photons come into existence. Thus, the collection of irregular continuant worlds is beyond cardinality. But if the regular worlds are fully governed by the laws of nature, even if these laws are deterministic, there will be an upper cardinality bound on regular continuant worlds (e.g. for any wave-function, there is an upper cardinality bound on the number of ways it can collapse). Of course, if naturalism is false, our world may have aspects that are not fully governed by the laws of nature, but Lewis accepts naturalism, so the argument may work at least as an *ad hominem*.

But I shall allow Lewis his dubious assumption of the equality of the

---

32 By "shares our world's past" I mean, roughly (and I doubt this can be defined more than roughly), that any propositions about what was going on at a past time true in one world are true in the other. There is, of course, a difficulty in defining precisely what one means by "about what was going on at a past time." One wants to exclude such propositions as "Patricia gave birth to the 48th president of the United States" which, even when true, depends not just on the past (the giving of birth) but also on future events (the decision by Patricia's son or daughter whether to run for president and the electoral choices of American voters). I assume that the notion of "what was going on at a past time" has an intuitively clear meaning. Of course, if we were reductive physicalists, this would be easier to define: a continuant world would just be one that has the property of duplicating the spatio-temporal region of our past light cone.

cardinalities between regular and irregular continuant worlds, and argue that even so, a Lewisian still cannot *know*, or even have reasons that make it more probable than not, that the world will continue approximately regularly — the Lewisian in fact does not even have reason to think it more likely than not that gravity will function tomorrow. Assuming that, as seems obvious, there are at least as many irregular possible continuant worlds as regular ones, I shall argue that the problem is with Lewis's indexical notion of actuality. Moreover, I shall argue that the question of the theory of actuality is essential to discussion of the problem of induction. That induction holds places restrictions on what theories of actuality are tenable.

The structure of my argument will be as follows. First, I shall lay ground that will be neutral between various realist theories of possibility as well as between the Humean sceptic about induction and the defender of induction. Next, I shall show how a Humean might attempt to mount a sceptical argument from this ground. The sceptical argument will be seen to fail in general, but, unfortunately for Lewis, works if actuality is merely indexical — i.e. if actuality is merely a matter of relation to us. Therefore, since we *can* inductively know that tomorrow things dropped will fall, Lewis's indexical account of actuality fails.

I am not claiming here that I will give an account of *how* inductive generalizations are to be justified. But I take it that we all think, and that Lewis did as well, that particular cases of induction *are* justified.

9.2 *Neutral ground*

For the moment abstract from the question of what theory of worlds and actuality is right, and assume only that we are going to be realists about worlds so that we agree that there exist such things as logically possible worlds. If it helps, we can even imagine possible worlds as physical books (of infinite size) in some heavenly library, each giving a maximal consistent description.

The notion of "a complete description of a possible world w up to time t" will be needed. This is a maximal set D of propositions true at w such that: (a) the propositions in D do not jointly imply any proposition reporting what happens after time t at w; (b) every proposition true at w that reports a physical state at a time prior to or equal to t is in D; and (c) every proposition true at w that reports the occurrence of a phenomenal state in an enmattered sentient being at a time prior to or equal to t is in D.

Suppose now that we are given a complete description D of a possible

world w_0 up to time t_0. What w_0 is, e.g. a book in the heavenly library, or a concretely existing world, or some abstract entity, will depend on which theory of possible worlds is right. Moreover, we suppose that the propositions in D entail that gravity has been operating up to time t_0 and that induction has worked as well as it has.

Let S_D be the collection of all possible world w of which D is a complete description up to t_0. Now, because of the principle of recombination, Lewis will have to grant that there are infinitely many possible worlds in S_D, and that the number of these at which induction fails in the near future of t_0 is at least as great as the number at which it does continue to work.

Remember now that all we know about w_0 is that it is a logically possible world at which all propositions in D hold, with D being a complete description of w_0 up to t_0, and that the propositions in D entail that induction has worked up to t_0 and gravity has been operational. Does this by itself justify us in inferring, or even give us reason to think, that gravity will continue to operate at w_0 in the near future after t_0? The answer surely has to be negative. All we know about w_0 is that it is a member of S_D and we do not even know that w_0 is actual. The knowledge that w_0 is in S_D is surely insufficient to give us any reason to think that induction will hold at w_0 in the near future of t_0.

The intuition here is particularly clear in the "heavenly library" view of possible worlds. We've picked up a book from the shelf, and read pages one through 1000, which pages gave us a complete description up to t_0. Given our knowledge that there are at least as many books whose first 1000 pages are the same and yet in which the pages after page 1000 describe gravity failing in the near future of t_0 as there are books with the same first 1000 pages and which then tell us of gravity continuing to operate, we certainly are not justified in inferring that our book is one of those which will make it out that induction continues to hold, and we should not think it more likely than not.

The exact same intuition should hold on Lewis's EMR. We are talking of a concretely existing world w_0 about which all we know is that it is a member of S_D. Since we know that in S_D there are at least as many worlds at which induction fails in the near future of t_0 as ones at which it holds, we should not think it more likely than not that w_0 is one of the "nice" worlds.

The above should not be controversial, being merely an observation about the structure of the collection of *possible* worlds, whatever these possible worlds's ontological status might be. This examination of *logical* space is neutral territory on which both the contemporary defender of induction and the Humean sceptic must meet if they are to engage at all.

An over-eager defender of induction who wishes to dispute this might

respond that the argument so far relies on the fallacy of thinking that in general if all we know about an entity x is that it is a member of a set C that can be partitioned into two disjoint subsets, A and B, where A has at least as many members as B, then we have no reason to think that x is in B. The fallacy lies in the fact that the sets A and B might have inner structure that would militate against this conclusion. Suppose, for instance, that A is the set of all real numbers y satisfying $0 \leq y < 1$ and B is the set of all numbers y satisfying $1 \leq y \leq 1,000,000$. Then, A has just as many members as B, namely continuum many, but if all we know about x is that it is a real number between 0 and 1,000,000, then it is *ceteris paribus* more likely that x is in B. It must, thus, be admitted that sometimes we should think it more likely that x is in B than in A when A has just as many members. But that requires special structure, in this case provided by the mathematical structure of the reals. In the possible worlds case, there is no similar story to be told about the not-nice and nice members of S_D that favors the nice ones.

9.3 *The Humean argument and a reply Lewis cannot give*

Thus, we do not have reason to think that the possible world w_0 of which D is a complete description up to t_0 is going to have gravity in the near future after t_0. Someone might think that in fact this justifies scepticism about induction. For suppose we learn two more pieces of information: w_0 is actual and t_0 is now. Then, the sceptic says, if *before* learning this we did not have reason to think it more likely than not that gravity would operate at w_0 after t_0, surely we still have no such reason. For, after all, when we learned that w_0 is actual and t_0 is now, we did not learn anything relevant about *what is true at* w_0. And so if before learning this we would not have reason to expect gravity to continue to operate, we should not after learning it. And *a fortiori* in our actual epistemic situation we should not expect gravity to continue to operate.

I take it that everything the sceptic says is basically correct here, except for the claim:

    (63)   when we learned that w_0 is actual and t_0 is now, we did not
               learn anything relevant about w_0.

For once this claim is granted, the sceptical conclusion that we have no reason to think gravity will operate in the near future follows. We can imagine being given a complete description of our world up to now, and the above

argument shows that were we given that description, we would not have reason to think that gravity will continue to operate even in the near future given (63). And it is plausible say that if we would have no reason to think this given the complete description, likewise we have no reason to think it given the subset of that description actually available to us.

The right response to the Humean, then, is to assert that we *have* learned something relevant when we learned that w_0 is actual and t_0 is now. When we make inductive inferences about the future states of the actual world, our sample space consists of the known states of the *actual* world. States of non-actual possible worlds, such as of the other members of S_p, are largely beside the point. When we have learned that w_0 is actual, the problem of what will be true after t_0 at w_0 was transferred from a question about logical space, to a question about *actual* events. *Pace* the sceptic, we do have reason to suppose that induction will work at least in the near future in the actual world (otherwise we would not be justified in living as we do), and so by learning that w_0 is actual and t_0 is now, we have learned that induction is a good guide to what w_0 is like after t_0. Of course few if any of us can explain *how* it is that we have reason to trust induction, but nonetheless we do have reason; one does not need to be an epistemologist to have knowledge.

However, as we shall shortly see, Lewis's indexical theory of actuality precludes him from being able to make this natural response to the sceptic. Since the sceptical argument succeeds as soon as it is granted that by learning that w_0 is actual and t_0 is now one has not learned anything relevant to the problem of induction, it follows that if Lewis is forced to grant this, Lewis's view does entail scepticism about induction. But since we *know*, or at the very least have reason to think, that gravity will continue to operate in the near future, we therefore know that Lewis's theory of actuality is false.

To see that Lewis cannot make the natural response to the Humean, observe that on Lewis's indexical theory of actuality, in learning that w_0 is actual and t_0 is now, one has only learned an indexical fact, one akin to learning that something is *here*. However, learning a merely indexical fact cannot by itself give one any additional information about non-indexical claims such as "induction will continue to hold at w_0 in the near future of t_0." The reason for this is the Principle of Impartiality of Reason (PIR) that says that whether a belief is epistemically justified does not depend on whether *we* are persons concerned in the matter and whether a time involved in the content of the belief is *now*. More precisely:

(PIR)   Let *K* be a body of knowledge that is known to entail the
        existence of a finite person *x* existing at a time t_0 (where

x and t_0 are definite descriptions or proper names). Let p be the proposition that I am x and that the present time is t_0. Let q be a proposition that does not involve indexicals. Then, K gives one reason to think that q if and only if K & p gives one reason to think that q.

Now, some may dispute PIR in the case of normative propositions, arguing that the fact that I am x can by itself (and not just by telling me other non-indexical facts about x) give me reason to make different normative claims about what x should do. With Kant, I take this to be wrong-headed, but for the purposes of this paper, I am willing to restrict PIR to non-normative claims.

PIR has also been formulated to avoid the following kind of counterexample. Let K be the proposition that a man x alive at t_0 is going to win the lottery. Then, when I additionally learn that I am x, I learn the non-indexical fact that a philosopher is going to win the lottery, a fact I could not know on the basis of K itself. However, this is not a counterexample to PIR, because it is not just the proposition that K&(I am x and t_0 is now) which gives reason to think a philosopher is going to win the latter, but rather this proposition conjoined with the additional proposition that I am a philosopher.

The application of PIR to the case at hand is follows. We let K consist of D and the claim that D is a complete description of a world up to t_0.[33] We let q be the proposition that gravity will function at w_0 for a while after t_0, say until $t_0+\delta$. Then, by the indexical theory of actuality, my learning that w_0 is actual is merely my learning that I am one of those persons who are asserted to exist by a proposition true at w_0. Moreover, since I know some uniquely identifying properties of myself, I know which person x described in D is the one I am identical with. Since it has been shown to be true, in a way neutral between the views of the sceptic and the defender of induction, that K did not give me reason to think that q, by PIR neither does K & p give me such reason. Therefore, if learning that w_0 is actual is nothing but learning that I am one of the persons said to exist by a proposition true at w_0, i.e. if Lewis's indexical theory of actual is correct, it follows indeed that learning that w_0 is actual and t_0 is now does not give us any relevant information with regard to the question of whether w_0 will continue to have induction holding in the near future of t_0. Hence Lewis cannot make the only possible response to the sceptical argument.

---

33 It might well be that D itself implies that it is a complete description of a world up to t_0, in which case K will coincide with D. But I leave the issue open.

Note that the non-Lewisian response to the sceptical argument that I am envisioning is not an argument showing how induction is to be justified. Rather the response shows which premise in the sceptical argument can be rejected by the defender of induction.

9.4 *Conclusions and a formalization*

A reasonable principle of recombination together with Lewis's theory of actuality entails inductive scepticism. Recall now that Lewis has attempted to counter Forrest's argument by saying that exactly the same problem recurs on other theories of possible worlds. But as we have seen, the problem is specifically locatable in Lewis's theory of actuality, and so formally speaking it need not be present once that theory is dropped.

Our argument does, however, show something more general about the problem of induction. In order for induction to be justified, it must be the case that learning that a possible world w_0 is actual and that t_0 is now provides information relevant for figuring out what happens after t_0 at w_0. Since the temporal impartiality of reason shows that the determination of whether t_0 is now or not should not by itself have a significant role in that, it follows that specifically learning that w_0 is actual must be the crucial additional piece of information. This places a restriction on what theories of actuality are tenable. Theories of actuality on which the assertion that w_0 is actual cannot provide any such information are untenable given that we do know induction holds.

Lewis's *tu quoque* reply to Forrest will thus have value if and only if there are no theories of actuality on which learning that w_0 is actual gives us information relevant to the problem of induction. But there are such theories. For instance, Strawson (1959: 126–9) in his discussion of Leibniz considers, though ultimately rejects, an ontology on which all things that exist are maximally specific concepts of individuals, so specific that any given concept could only be instantiated in one possible world. Then, such a concept C is said to be actual if and only if C is a concept that is instantiated in the best of all possible worlds. In other words, Strawson is entertaining a theory of actuality on which to be actual is to be in the optimal world. On *this* theory of actuality, learning that w_0 is actual certainly gives one reason to think that w_0 will continue to be regular after t_0, since surely the best world is regular.

Or take the infinite library account of possible worlds, married with this account of actuality: a world (i.e. one of the books in the library) is *actual* provided that every sentence in it is *true*. Then something substantive

has indeed been learned when we learned that a world is actual: we have learned that every one of its sentences is true, while previously all we knew in that direction was that the sentences (or, more precisely, the propositions expressed by them) were compossible. This is akin to learning whether an ordinary book someone has shown me is an accurate history book or a piece of fantastical fiction. If it is an accurate history book, I can make certain inferences about unread parts of the book on the basis of parts that I have read that I couldn't make if the book were to describe a merely possible world.

We can then formalize the argument as follows:

(64)   Let D be a complete non-indexical description of the actual world up to the present (t_0) in temporally pure terms. (Definition.)

(65) D contains the claim that the law of gravitation has always held prior to t_0. (Premise.)

(66)   Knowing that gravity has always actually functioned prior to t_0 justifies one in believing it will continue to hold for a minute after t_0. (Premise.)

(67)   There are at least as many worlds satisfying D in which the law of gravitation fails during the minute after t_0 as there are worlds in which it continues to hold. (Premise.)

(68)   * Therefore, knowing that an entity w is a world satisfying D does not by itself epistemically justify inferring that w is a world at which gravity functions for a minute after t_0. (Premise, justified intuitively by appeal to (67).)

(69)   * Theoretical reason is impartial with respect to merely indexical facts: If knowing that x is F (where F is purely non-indexical and x is a definite description or proper name) does not epistemically justify inferring that x is G (where G is purely non-indexical), then neither does knowing x is F and that x is I (or now, here, etc.: any pure indexical will do) justify inferring that x is G. (Premise.)

(70)   * Actuality is indexical. (Premise.)

(71)   Therefore, knowing that an entity w is a world satisfying D and w is actual does not epistemically justify inferring that w is a world at which gravity works for a minute after t_0. (By (68)–(70).)

(72)   * But knowing that the actual world satisfies D and w is

actual epistemically justifies inferring that gravity works for a minute in *w* after t_0. (By (65) and (66).)

(73)   Therefore, knowing that the actual world satisfies *D* and *w* is actual both does and does not epistemically justify inferring that gravity works in *w* for a minute after t_0, which is absurd. (By (71) and (72).)

Here, the premises marked with an asterisk form an inconsistent quadruple, all members of which are highly plausible except (70). Therefore, (70) is false, *pace* Lewis.

Section 10   The epistemological objection

But even if all the above problems were solved by Lewis, some think (e.g. Richards 1975) we still have a pressing problem of how we are supposed to know what goes on in the other worlds. The other worlds are causally isolated from us. To know modal propositions is to know what happens in other worlds. But how can we have knowledge of these entities that are causally isolated from us? How can we know, e.g. that it is possible that there exist unicorns and that it is impossible that there exist square circles, without employing an impossible telescope for gazing at worlds other than ours and finding that some of them contain unicorns but none have square circles in them?

Lewis thinks that this objection stems from thinking that all knowledge needs to be causal — the object of knowledge must give rise to the knowledge. But, Lewis insists, the same problem arises for mathematics (Lewis 1986a; Section 2.4). The equality of the angles in a triangle cannot cause anything, and certainly cannot cause me to believe it. Hence, the causal account of knowledge is wrong, as it cannot account for mathematical knowledge.

But the case of EMR seems disanalogous to the mathematical case for two reasons, the first of which is more speculative. This is that even if the truthmakers of mathematical truths do not cause our beliefs about them, what was central about the causal intuition is the notion that in central cases, the object known enters into the *explanation* of our believing. In the case of perceptually based beliefs, the object enters into the explanation of the believing by *causing* the believing. But there are other ways to enter into an explanation than by being the cause. In a Spherical Astronomy class, it is explained why the planets move in elliptical orbits. This explanation does

not, however, consist merely in a statement of Newtonian laws and initial conditions. Rather the answer centrally includes a mathematical derivation of the ellipticity of the orbits. Only when one has understood the mathematics does one really understand why the orbits are elliptical. This is the norm in the sciences. Thus, mathematical facts are part and parcel of scientific explanations, and so the idea that a mathematical truth enters into an explanation of why one believes it is not absurd. It is, of course, implausible that this should happen in every case of mathematical knowledge. But likewise I know that tomorrow the sun will set even though tomorrow's sunset doesn't enter into an explanation of my believing that tomorrow the sun will set, and the causalist should only claim that in *central cases*, ones that anchor (via deductive or inductive logic, say) our knowledge of a subject area, the object known causes the believing. Similarly, it need only be the case that for central cases of mathematical knowledge the believing is explained by the truth known, for instance because one's neural structures are an instance of the mathematical structure that one has the belief about. It is, for instance, plausible that we might have neural structure whose input–output functional relation instantiates addition: its output represents the sum of the numbers represented by the inputs, and then a part of the explanation of our belief that 2+2=4 would be that in fact 2+2=4.

Now, on Lewis's view, happenings on other worlds are not causally efficacious in respect of our beliefs about them (except in weird cases like those we discussed in Sections 8.3.2 and 8.3.3). But insofar as the happenings on other worlds help ground counterfactual truths, and Lewis analyses causation in terms of counterfactuals, facts about happenings on other worlds will enter into explanations if causation does. However, it does not appear that *the relevant* facts about happenings on other worlds enter into the explanations of believings. Thus, my belief that possibly there are unicorns may have some causal explanation in terms of some event *E*, with the causation grounded in the counterfactual <had *E* not happened, I wouldn't have believed that there are unicorns>. However, the facts about other worlds that make *this* counterfactual true do not include unicorns. Rather, they include counterfactual neural happenings, counterfactual social conditionings, etc. Now this example does not establish that some central cases of knowledge of other worlds might not have the facts known enter into the explanations of the beliefs, but I think it makes it implausible that there should be such a story.

A second disanalogy between the EMR case and the mathematical case is that mathematical knowledge is knowledge of abstracta whereas modal knowledge, according to EMR, is knowledge of things that have

the same empirical status as the things we are familiar with. If we can only know whether there are unicorns in this world by empirical observation, why should we be able to confidently affirm that in *some* world there are unicorns, without having made any observation?

Now, there is a Lewisian answer to this. To know whether there are unicorns in this world is to know whether unicorns are spatio-temporally related to us or not. If without any observation we had believed that actually there are no unicorns, this would not count as knowledge because this belief would have been wrong had we been living in another possible world where there were unicorns, and we would have been right only by chance. But there is no world at which it is false that unicorns are possible. Thus it is impossible for someone to be wrong in thinking that unicorns are possible, and hence the belief that unicorns are possible is not due to luck (cf. Lewis 1986a: 113).

But the impossibility of being wrong about something does not suffice for knowledge. Otherwise, the mathematical cranks who believed that they had proofs of Fermat's Last Theorem would necessarily count as *knowing* that Fermat's Last Theorem is true, which is absurd. However, Lewis can always bring in an externalism according to which we count as having modal knowledge providing we have some truth-directed faculty by which we generate modal beliefs, which faculty is in fact right most of the time. Thus the epistemological objection fails on some externalist epistemologies.

However, a view on which the grounds of modal truths contribute causally or explanatorily to our modal beliefs would be intuitively preferable. This is not a fatal objection, but it will be a point in favor of the alternative account I shall eventually sketch that the ontology of possibility on that account will be such that the grounds *could* contribute to the knowledge.[34]

---

34 Richards (1975) goes even further than just saying that the Lewisian account gives no epistemology of possible worlds. If we could work out on the basis of things in this world what things in other worlds are like, the Lewisian theory would have "no explanatory value as a theory of meaning" (1975: 110); indeed, Richards urges that since we would then in the next step want to move from what happens in other worlds to modal claims about this world, our account would be viciously circular. However, this objection confuses truth conditions with assertibility conditions. The move from this world to other worlds is a move justified by *assertibility* conditions: this-worldly situations *epistemologically* ground the truth or falsity of non-modal propositions at other worlds. But the move from other worlds to this world deals in *truth* conditions: other-worldly situations, on Lewis's view, *ontologically* ground this-worldly modal facts.

Section 11 Explaining the actual in terms of the necessary

Consider this claim:

(74)   There are actually no square circles because square circles
        are impossible

This seems a perfectly good case of explanation (if we are worried — as we might be after reading about the Less Radical Theory in Part V — about the intelligibility of the phrase "square circle," we can talk about non-mammalian horses or counterexamples to some complex theorem). However the Lewisian may need to reject it. For on his view, this is equivalent to:

(75)   There are no square circles in this universe because in no
        universe are there any square circles.

And this seems to be no more explanatory than saying that

(76)   our solar system does not have any golden mountains
        because no solar system has any golden mountains.

Now, if it were a law of nature that there cannot be a golden mountain in a solar system, then (76) might be thought to be an explanation. I actually think that is mistaken. If it were a law of nature that there cannot be a golden mountain in a solar system, then the correct explanation would be:

(77)   our solar system does not have any golden mountains
        because *it is a law* that no solar system has any golden
        mountains.

But be that as it may, the point remains that unless there is something nomic-like about the fact that no universe has a square circle, that no universe has a square circle does not explain why our universe has no square circle. And on Lewis's view, such universal generalizations about worlds appear to be simply brute facts about the structure of the space of possible worlds. It seems, thus, that Lewis must reject the explanatory claim in (74), and that should count against his theory.

Section 12 A final assessment of EMR

Lewis notes that the most common "objection" to his view is the "incredulous stare." How could such a crazily rich ontology be the case? I submit that one reason for an incredulous stare may be an intuitive recognition that such an enlargement of what we think reality contains is going to be revisionary to *many* ordinary ideas. We have seen that there *are* in fact such revisions entailed by EMR, and the revision is too expensive.

It would be nice to have a realistic discourse about possibilia. But if that is going to deny us some standard moral notions, remove the possibility of having inductive knowledge, and force a revision of the notion that concrete substances in a kind can always be arranged into sets, then the cost is too large. For instance, without inductive knowledge of things in *this* world, we would have little use for possibilia at all, and so that alone is reason not to accept the theory. The cost of EMR is unacceptable, but perhaps there are other realistic theories of modality where the cost is not as good.

Part IV

Platonic ersatz ontologies

Section 1 The general strategy

Lewis's theory has a number of useful consequences. If the price were lower, one might adopt it. Many people have tried to find a cheaper replacement for EMR, with fewer paradoxes and less ontological extravagance. In all the replacements, possible worlds are not concretely existing worlds ontologically on par with ours. Rather, in Lewis's terminology, they are "ersatz worlds." These may be mathematical constructions, sets of sentences, collections of propositions and the like. All the ersatzists whose views I consider take their ersatz worlds to be abstract. Their ideal would be to find a place for possible worlds within a Platonic ontology they already accept. Then, the hope is, one could talk about possible worlds without any additional ontological commitment.

There are two levels of ambition that an ersatzist may have. The Ambitious Ersatzist claims that her worlds are things that we are really talking about in making modal claims. When we make a true modal claim, the grounds of the claim really involve one or more of the ersatz worlds. The Ambitious Ersatzist sets herself the task of answering the extended Parmenidean problem of what modal language is about. This is, after all, what Lewis was doing. The Ambitious Ersatzist cannot hope to avoid serious metaphysics. Even if the ersatz worlds are things like propositions we would believe in anyway, the claim that *these* things are the grounds of true modal propositions will be an ontological claim we would not have otherwise made.

The Unambitious Ersatzist, on the other hand, merely introduces surrogates for Lewisian worlds which are helpful logical constructions or theoretical tools. While the Unambitious Ersatzist's ersatz worlds had better exist in some way, else in talking of them she is not be speaking truly, nonetheless they need not be at all related objectively to the grounds of modal propositions, if there are any such grounds. The Unambitious

Ersatzist makes no metaphysical claim beyond the fact that her ersatz worlds exist. She need not even deny that EMR is true, though if it were true, the motivation for the ersatzism would be gone. Unambitious Ersatzism thus strives to provide a relatively ontologically neutral way of talking of possible "worlds," but one that makes it clear that such talk can coherently be made sense of.

I shall not raise many criticisms against Unambitious Ersatzism. After all, every basic modal realist, of whatever variety, will take it as plausible that the Unambitious Ersatzist's project can succeed, where "basic modal realism" is the view that there *are*, not necessarily Lewisian, possible worlds. *Whatever* we think possible worlds *really* are, whether concretely existent worlds, or books in a heavenly library, it is likely that somewhere in Cantor's paradise ersatz versions of them can be found. For instance, suppose we accept EMR. Then, it is plausible that one can find some abstract proper class that stands in one-to-one correspondence to these worlds, as well as enough abstracta of any kind sufficient to correspond to individuals and propositions, and then we can find abstract relations T and I such that if p corresponds to a proposition, x to a possible individual and w is an ersatz world, then pTw holds if and only if the proposition corresponding to p is true in the Lewisian world that our ersatz world w corresponds to and xIw holds if and only if the individual corresponding to x exists in the Lewisian world that w corresponds to. Then, T and I can be used to define the *true-in* and *exists-in* relations between propositions and individuals, respectively, and ersatz worlds. And then we have an ersatz theory. In particular if EMR is true and there are enough sets (or other collections), the project of Unambitious Ersatzism is doable, and the same will be true for other ambitious alternatives to EMR.

The Unambitious Ersatzist's task is less trivial if basic modal realism cannot be taken for granted. I will argue that the task succeeds when worlds formed as classes of propositions are brought in. However, I shall argue that contemporary ersatzist approaches fail to conform to the more stringent criteria that Ambitious Ersatzism is to be judged by.

Note that Leibniz is an ersatzist, in that he believes he can make do with worlds that do not exist ontologically on par with ours, but unlike the ersatzists' I shall consider here, Leibniz's worlds are not Platonic self-subsistent abstracta, but concrete thoughts in the mind of God. Therefore, discussion of Leibniz's ersatzist approach will be postponed until Part VI.

There is an accepted distinction between actualist and concretist theories of modality (Lycan 2002). The *actualist* accounts reject any non-actual entities, any entities not found in the actual world, and thus must provide

an account of the truth of modal claims in terms of this-worldly actual entities. The *concretist* accounts, on the other hand, say that there are concrete non-actual entities, such as unicorns existing concretely in concrete physical worlds different from ours, which serve as the truthmakers of modal propositions. EMR is a paradigm case of concretism, while the Ambitious Ersatzist accounts are actualist accounts. There is, however, a subdivision within actualism. One can ground modality in concreta or in abstracta, and the Platonic Ersatzist falls on the side of abstracta, while Leibniz, or my version of him, will work in terms of divine thinkings which are concreta.

1.1 *A cheap ersatzism*

There is a cheap way of getting ersatz possible worlds by producing a Kripke model. Consider the set S of all true sentences in some language with box and diamond operators, and suppose for simplicity that the logic is S5. Suppose S is a countable set, which it will be if the vocabulary is countable and the sentences have finite length. We can then construct a *model* for S as follows. Say that a subset w of S is a "world" providing:

(78) s is a member of w if and only if $\sim s$ is not a member of w;

(79)   if s is a member of w then $\Diamond s$ is true.

Then, let W be the set of worlds. For w in W, say that a sentence s is true at w if and only if s is a member of w. Observe that W is non-empty, because the set of all true sentences forms a world.

What we have is now sufficient for box and diamond work: $\Diamond s$ is true if and only if s is true at some world w. To see this, suppose first that s is true at some world. Then by (79), $\Diamond s$ is true. The converse is a little bit more difficult. Suppose $\Diamond s$ is true. Inductively define a sequence of subsets of S. Let $s_1, s_2, \ldots$ be an enumeration of all the sentences of S such that $s_1 = s$. Let $U_1 = \{s_1\}$, and define the sequence of subsets $U_2, U_3, \ldots$ of S inductively as follows. Suppose U_k has already been defined. Then, let $s_k{}^*$ be a conjunction of all the sentences in U_k (in any order). Then, if $\Diamond(s_k{}^* \,\&\, s_{k+1})$ is true, let $U_{k+1} = U_k \cup \{s_{k+1}\}$. If $\Diamond(s_k{}^* \,\&\, s_{k+1})$ is false, let $U_{k+1} = U_k$. Observe that if we proceed this way, then we will have $\Diamond s_k{}^*$ true for every k (the case $k=1$ follows from the fact that $\Diamond s$ is true).

Now, define w to be the union of all the sets U_k. Observe first that s is a member of w because s is a member of U_1. Suppose that t is a member of w. Then, t is a member of U_k for some k, and hence is a conjunct in $s_k{}^*$. But

$\Diamond s_k{}^*$ is true, and if a conjunction is possible, every conjunct is possible. Thus, $\Diamond t$ is true. Hence, (79) is satisfied by w. It remains to check (78). Suppose t is a member of S. Then, $t=s_k$ for some k, since the s_k form an enumeration of all of S. Likewise, $\sim t=s_n$ for some n. Then, either $k<n$ or $n<k$. I shall only consider the case where $k<n$ — the other case can be handled in exactly the same way after we swap the roles of t and $\sim t$ and of k and n. So suppose $k<n$. Then, t is a member of w if and only if $\Diamond(s_{k-1}{}^* \,\&\, s_k)$ is true, where we may stipulate that $s_0{}^*$ is some tautology like $1=1$. Suppose now t is in fact a member of w. Then, since $n>k$, it follows that t is a conjunct in $s_{n-1}{}^*$. Thus, $\Diamond(s_{n-1}{}^* \,\&\, \sim t)$ is false, and hence $\sim t$ (which, recall, is s_n) is not a member of w.

Suppose now that t is not a member of w. We must show that $\sim t$ is a member of w. To obtain a contradiction, suppose neither t nor $\sim t$ is a member of w. Then, $\Diamond(s_{n-1}{}^* \,\&\, \sim t)$ and $\Diamond(s_{k-1}{}^* \,\&\, t)$ are both false. Since s_{n-1} has as conjuncts all the conjuncts of s_{k-1}, it follows from $\Diamond(s_{k-1}{}^* \,\&\, t)$ that $\Diamond(s_{n-1}{}^* \,\&\, t)$ is also false. But if p is a proposition such that $\Diamond(p\&t)$ and $\Diamond(p\&\sim t)$ are both false for some proposition t, then $\Diamond p$ is false. Letting $p=s_{n-1}$, it follows from this that $\Diamond s_{n-1}{}^*$ is false. But given our construction, $\Diamond s_{n-1}{}^*$ was always true. Hence, a contradiction ensues, and so we cannot have both t and $\sim t$ fail to be members of w.

Thus (78) holds, and w is a world. Since s was a member of w, we have indeed shown that $\Diamond s$ holds if and only if s is true at some world.

So, we have a possible worlds theory on the cheap. All we need are sets of sentences. We can even be cheaper if we so wish by using Gödel numbering, which is a scheme by which a sentence can be assigned a unique Gödel number. For instance, if the vocabulary involves items $x_1, x_2, \ldots$, then the sequence of vocabulary items $x_{733} x_{19991} x_{1001223} x_{13}$ can be assigned the Gödel number $p_{733} p_{19991}{}^2 p_{1001223}{}^3 p_{13}{}^4$, where p_n is the nth prime number. Sentences then can be considered as positive integers. Worlds then become sets of positive integers, with the understanding that a sentence s is true at a world w if the Gödel number of s is a member of w. Our ontological commitment then is only to sets of positive integers.

But we can perhaps do even better. Any set of positive integers can be encoded as a real number between 0 and 1. Here is one simple way to do it. If S is a positive number, let S^* be the decimal number with digits "$0.s_1 s_2 s_3 \ldots$," where s_i is 1 if i is a member of S and s_i is 0 otherwise.[1] The set of possible worlds is then just a subset of the real numbers between 0 and 1. Thus, we

---

1   It might seem more natural to make this a binary number, given that the digits are 0 and 1. But that might run into trouble with uniqueness, because, say, the binary numbers 0.011111... and 0.100000... are in fact equal, just as the decimal numbers 0.099999... and 0.100000... are.

need only be committed to the real numbers, and to relations between them and sentences. Nor do we need to be committed to the real numbers *qua* abstracta if we think there is a continuum of points in space or time; we can map these numbers onto points in time or on a spatial line. Our theory of possible worlds is thus ontologically very cheap. Of course, this very cheapness shows that it is going to be of no help in giving an ontological grounding to alethic modal claims. If possible worlds are just points in space-time or numbers, related in appropriate ways to sentences via some *truth-at* relation, they will not be plausibly thought a foundation for alethic modality.

What if we take our language to have an uncountably infinite set of sentences, for instance because it may have an uncountably infinite vocabulary? If gestures are included as part of the language and space is infinitely subdivisible, then we can utter uncountably many sentences of the form "The fish was *this* [holding hands apart by distance x] big." To make the previous construction continue to work in the uncountable case, we need to fill in some technical details by using the set-theoretic AC in the guise of Zorn's Lemma, and I leave this to the interested reader.[2] Alas, we will no longer be able to reduce worlds to real numbers or sequences of integers, but we can still reduce worlds to sets of sentences.

Unfortunately, attractively economical as the above theory is, whether in its countable or uncountable version, more is required of a theory of possible worlds. While the above works fine for first order sentences, it only works for *some* second order sentences. Second order sentences come in two varieties: They may quantify over linguistic entities such as sentences or over non-linguistic entities such as propositions or properties.

Quantification over sentences works just fine. For instance:

(80)   "There is a sentence which is such that if it is possibly true then it is necessarily true"

comes out correct on a semantics based on the above construction of possible worlds if and only if it is in fact true. The above semantics would translate (80) into:

(81)   "There is a sentence which is such that if it is true at some world then it is true at all worlds."

For if (80) is true, then there is such a sentence. Let this sentence be *s*. Then, it is true of *s* that if it is possibly true then it is necessarily true. But we have

---

2   An explicit argument is given in Pruss (2001).

already proved that *s* is possibly true if and only if it is true at some world, and that it is necessarily true if and only if it is true at all worlds. Thus, (80) implies (81), and the converse is equally easy.

However, it is quantification over propositions that is more problematic. Consider for instance:

>   (82)   "There is a proposition *p* which is contingent and cannot
>   be expressed in *L*."

where *L* was the language used in the construction.

The problem is that the above possible worlds semantics only applies to sentences and derivatively to propositions expressible by sentences. But it is plausible that not all sentences can be expressed in any one language *L*. In fact, even if the cardinality of the set of sentences is uncountable, say because of the use of gestures, there just will not be enough sentences to go around to describe all sets. This may seem to contradict the fact that the theory seemed to work for all sentences. But the criterion for working then was just that the theory should make all and only the possibly true sentences be the ones that are true at some world. We are now using a stricter standard, that of seeing if sentences involving modal operators within the scope of quantifiers (i.e. express *de re* modal claims) can be translated into sentences that do not have any modal operators but only quantifiers.

But the above account does show that if all we want is an ontologically cheap Kripke model account of possible worlds for limited purposes that do not require quantification over propositions, our ontological commitments are low. Arguments explained in terms of quantification over possible worlds are frequently easier to follow than arguments explained in terms of *de dicto* modal operators, and the cheap account above justifies the practice of a philosopher who engages in possible worlds talk without making any controversial ontological commitment to possible worlds as such.

Section 2 Linguistic approaches

2.1 *The basic idea*

On the first ersatzist approach, worlds are maximal consistent collections of sentences in something we can call a "worldmaking language" (cf. Roper 1982; Jeffrey 1983). As discussed in the introduction (Section 4.3 of Part I), we cannot take these to be actual utterances of sentences. Nor can we take

them to be possible utterances of sentences, since if we are entitled to help ourselves to the possibilia that possible utterances are, we might as well help ourselves to the possibilia that possible universes are. Rather, we must take "sentence" to indicate type, not token, and hence the approach is a Platonic one. Presumably, the "type" can be some sort of set-theoretic construction, e.g. a finite sequence of symbols arranged according to some grammar.

A proposition is then true in a world w if and only if it is jointly implied by the propositions expressed by the members of w. Some propositions are not only jointly implied in this way, but are actually expressed by some sentence that is a member of the world. These fortunate propositions Lewis (1986b; Section 3.2) calls "explicitly represented" by the ersatz world; all the rest are implicitly represented. One might think that the fact that the worlds are maximal collections of sentences ensures that all propositions true in a world are explicitly represented by it, since otherwise we could just add to the collection sentences representing the implicitly represented propositions. But that is not the case, because the language in question may lack the resources for explicitly expressing all propositions. Thus, unless each sentence of our language can be translated into a sentence of physics, the language of physics used as a worldmaking language will have to include implicit representation. But it is conceivable that propositions expressed by many sentences that cannot be translated into sentences of physics would be *implicitly* represented by a world made with the language of physics and corresponding to our universe. Any proposition whose truth value supervenes on what can be described by physics would be thus implicitly represented.

A language is *propositionally complete* if and only if for every world made out of maximal consistent collections of sentences of that language and every proposition p, either p is true at that world or not-p is true at it. For instance, if it is possible for there to be properties which do not supervene on physical properties, then the language of physics is propositionally incomplete. Unless we are willing to make do with partial worlds, the worldmaking language had better be propositionally complete.

Note that while I shall continually talk of "propositions" in connection with linguistically based approaches to modality, this usage does not presuppose any Platonic claims about there *being* any abstract propositions out there. Perhaps propositions are just a useful fiction, a useful way of expressing certain facts about language — e.g. facts about whether the language is complete or not. This issue is open. One of the advantages of the linguistic approaches is that they do not close this question in the way that proposition-based approaches do.

Some of the criticisms we shall observe of propositional approaches will

apply to the linguistic approaches, and so now I shall only make the criticisms that do not apply to the propositional approaches as well.

2.2 Objection from alien properties

None of the languages available to us is propositionally complete. For, as Lewis (1986a: 159) notes, surely it is possible for there to be properties that cannot be reduced to the properties actually instantiated in our world. Just imagine the finite denizens of a world with elementary particles — none of which exhibit electric charge. Electric charge being a *basic* property, none of their languages would be propositionally complete since no proposition reporting something's having that basic property would be implied by propositions expressed by sentences of their languages. But surely it is plausible that we are in the same position as the finite persons of that impoverished world, just *vis-à-vis* some property other than charge. At the very least, clearly we cannot *know* that we are not in such a position, and consequently we cannot know that the linguistic approach as based on one of our languages is correct.

But perhaps the denizens of the impoverished world *can* specify electric charge. We bestow the name "electric charge" on that property which causes certain kinds of overt behavior of objects, paradigmatically the attraction with amber (Greek: *êlektron*) that has been rubbed with pieces of fluff. There are two views of how this naming works. First, we could be picking out (whether by ostension or definite description — at this point, it does not matter) a particular complex of patterns of behavior of objects, and defining "electric charge" by the definite description "that property which causes those patterns." This we shall call the "Frege–Russell view" of electric charge. Or, we could be picking out a particular complex of patterns (again, it does not matter how), and then making "electric charge" as a rigid designator or name for that property which in our world causes those patterns. This will be the "Kripke view" of electric charge.

On the Frege–Russell view, necessarily, in any world in which those patterns of behavior that we have used to identify electric charge occur, that property in that world which causes these patterns of behavior will be electric charge. On the Kripke view, it is at conceivable that there be a possible world with the same patterns of behavior exhibited but with a property different from electric charge responsible for them, just as it is possible for there to be a world with a liquid that behaves just like water, but is not water, because it is not H_2O.

If we adopt the Frege–Russell view of electric charge, then the counterexample to Linguistic Ersatzism may fail. For the denizens of the impoverished world could perhaps describe the patterns of behavior that electric charge causes, and since, necessarily, wherever these patterns are exhibited, there electric charge is to be found, worlds described with their language have the resources for implicit representation of predications of electric charge, even if the patterns are not exemplified in their world.

On the Kripke view, the language in the non-electric world is surely propositionally incomplete. For it is then conceivable that there are two worlds exhibiting the same patterns of behavior, but in one world this behavior is caused by electric charge and in the other world by some other property.

But even if we adopt the Frege–Russell view of electric charge, can we do this for *all* properties? This seems implausible. Imagine a world impoverished by a total lack of electromagnetic radiation. This is a world without light. Is it plausible to suppose that finite denizens of that world could *both* express propositions that entail the existence of light *and* propositions that entail the existence of visual experiences? Or consider finite beings in a world where there is no space. Could worlds made of their language represent spatial properties? Probably not. But we cannot know that there is no property to which we are related as those beings are to spatiality. Indeed, it is plausible to suppose there are such properties. It is highly implausible that the Frege–Russell view, even if it works for the special case of electric charge, should work for all physical and phenomenal properties. But if it ever fails, then it becomes highly plausible that there are certain properties which no worlds made from our languages can represent.

But even if none of *our* languages are up to the task, perhaps there are languages that are. Lewis discusses "Lagadonian" languages in which some words represent themselves. Thus, a stone might represent that very stone it is. A real number might represent itself. If there are universals and basic properties are universals, then the world-making language might use these universals for representing themselves. It could even be an extension of our language. Suppose we represent types of sentences as set-theoretic sequences of characters (with each character being represented by a number, if we like, but I omit this for simplicity), so that "There is snow" is set-theoretically represented as the set { $\langle 1, "T" \rangle, \langle 2, "h" \rangle, \langle 3, "e" \rangle, \langle 4, "r" \rangle, \langle 5, "e" \rangle, \langle 6, " " \rangle,$ $\langle 7, "i" \rangle, \langle 8, "s" \rangle, \langle 9, " " \rangle, \langle 10, "s" \rangle, \langle 11, "n" \rangle, \langle 12, "o" \rangle, \langle 13, "w" \rangle$ }. Then a sentence saying that something has an alien property can be represented as the set { $\langle 1, "S" \rangle, \langle 2, "o" \rangle, \langle 3, "m" \rangle, \langle 4, "e" \rangle, \langle 5, "t" \rangle, \langle 6, "h" \rangle, \langle 7, "i" \rangle, \langle 8,$ $"n" \rangle, \langle 9, "g" \rangle, \langle 10, " " \rangle, \langle 11, "h" \rangle, \langle 12, "a" \rangle, \langle 13, "s" \rangle, \langle 14, " " \rangle, \langle 15, F \rangle$ }, where F is the alien property itself.

This approach *does* answer the alien properties difficulty, but only at the cost of an ontological commitment to universals. And if one is willing to countenance *some* abstracta like universals, maybe it is better to go all the way to propositions.

2.3 *Primitive modality*

As Lewis (1986a: 151) observes, the linguistic approach does nothing to clarify the grounds of modality, since it presupposes modality in two places: (1) in the requirement that the worlds be *consistent* collections of sentences, and (2) in the notion of implicit representation. For to say that a collection of sentences is consistent is to say that these sentences, or the propositions expressed by them, are com*possible*. And to say that some collection of sentences implicitly represents a proposition is to say that the propositions expressed by these sentences jointly entail the latter proposition. But of course entailment involves modality: entailment is necessary implication, and indeed a proposition is necessary if and only if it is entailed by a tautology. The above, however, are only objections against Ambitious Linguistic Ersatzism: the Unambitious Ersatzist will not mind.

Moreover, Lewis (1986a; Section 3.2) argues that one cannot simultaneously reduce compossibility to purely formal truth-theoretic notions *and* solve the implicit representation issues. It is plausible that there are certain arrangements of elementary particles such that when they are present, necessarily a donkey is present. But it is implausible to think that the word "donkey" is to be understood as an abbreviation for a description in terms of arrangements of elementary particles. There is then no syntactic way of inferring from a sentence of the form "The joint wave-function of a bunch of particles here is Ψ" that "There is a donkey here." If the worldmaking language is rich enough to include both elementary particles and donkeys, this shows that syntactic definitions of consistency will not do, as there is no syntactic way of gauging whether "The joint wave-function of a bunch of particles here is Ψ" and "There are no donkeys here" are consistent or not. If, on the other hand, the worldmaking language is not rich enough to include both, then there will have to be implicit representation which, for the same reasons, cannot be expressed syntactically.

There are two possible responses. The first, which Lewis considers, is that perhaps there are a number of additional axioms such as "Necessarily, when the joint wave-function of a bunch of particles is Ψ, there is a donkey there" whose use is permitted in the syntactically based proofs. But the

axioms are not mere definitions because the strong reductionistic thesis that one can *define* macroscopic objects in terms of microscopic ones is highly implausible, and so they are substantial. As such, these axioms participate in an unanalyzed primitive modality. The axioms are *automatically* necessary, and their necessity has not been grounded (cf. Section 2 of Part I).

The second response is to complain that these Lewisian examples rest on a weak reductionistic thesis that the presence of macroscopic objects like donkeys supervenes on the presence of elementary particles. One might take an Aristotelian view according to which the presence of a substantial form is a necessary condition for the existence of a donkey. It is not enough that there be elementary particles arranged in a certain way: it is necessary that there be a certain objective telos to their arrangement, a telos bestowed by the form.

However, even this controversial objection will not save the Ambitious Linguistic Ersatzist. It only relocates the problem. Call an aggregate of particles a "donkey*" if the arrangement of particles in this aggregate is such that a normal donkey could be composed of these particles. Admittedly, not every donkey* is a donkey. In fact, no donkey* is a donkey because the criteria of identity for aggregates of particles are different from those for donkeys. However, Lewis's argument can now be rephrased in terms of donkey*s. For it is still true that there being a certain arrangement of particles entails that there is a donkey* there. And the same problems as before ensue.

The second objection against the Ambitious Linguistic Ersatzist taking a syntactic approach to compossibility is similar to an objection levied against EMR in Section 2.2.2 of Part III. One might well wonder why the grounds of <Smith can perform *A*> should have to involve syntactic claims about sentences. It seems natural to say that this proposition is just about Smith. Why should its grounds involve *sentences* at all? This leads to the next objection.

2.4 *The arbitrariness objection*

The Ambitious Linguistic Ersatzist claims that sentences in a language are involved in the grounds of modal propositions. But sentences in *which* language? Any particular human language we choose seems implausibly arbitrary. This objection is much like the one made in Section 3.2 of Part I against Lewis's quasi-linguistic structured propositions. It is clearly absurd to think that facts about English sentences are involved in the truthmaker

of a Frenchman's, or a space alien's, modal claim. Nor will it do to speak of an idealized language, for there are many ways of idealizing a language, and any particular way of doing this will be just as arbitrary as any other.

One might attempt to consider some super-language which involves all actual languages. I have argued in Pruss (1999) that in contexts in which the interlocutors know more than one language, there is in effect a super-language whose grammatical rules govern contextually how each utterance is to be understood. For example, in the case of speakers who know German and English, this super-language governs whether "Gift!" is to be understood as signifying a present given (as in English) or poison (as in German). Extending this, we can consider the totality of all communicative utterances ever made, by humans or others, as part of a single super-language. True, no human knows all the words and grammar of that super-language, but then perhaps no human knows all the words of English. It is less implausible to suppose that such a language should be *the* world-making language. However, this, too, is parochial and arbitrary. For surely, in the spirit of S5, what the possible worlds are should not itself vary across possible worlds, even though what languages are actual does.

Thus, perhaps we should extend our story to the consideration of a totality of all actual and possible languages. But, of course, there is a circularity problem then since one is appealing to class of possibilia, namely possible languages. The same objection can be made against a variant view according to which the world-making language has as its sentences equivalence classes of sentences with the same content in all possible languages.[3]

2.5 Mathematical constructions

One might also try to model possible worlds mathematically (see, e.g. Heller 1998). For a simple Democritean model, suppose S is a set of basic physical types of objects, microscopic or macroscopic, and suppose, apparently contrary to facts about bosons, that it is impossible for two objects to occupy the same point in space. One might then model a world as a function f defined on $\mathbf{R}^4$, where $\mathbf{R}$ is the real line, that assigns to each member of this space an ordered triple consisting of an object type from S, an ordinal acting as a

---

3   Or, more precisely, equivalence classes of pairs $\langle L,s \rangle$ where L is a language and s is a sentence, where the equivalence relation is defined by $\langle L,s \rangle \sim \langle M,t \rangle$ if and only if s in L has the same content as t in M.

"serial number" of that object (just in case there is more than one instance of that type of object) and an orientation.

A simple model of that sort will not do, of course. For one, different possible worlds will have different space-time structures, and using a Euclidean space will be inadequate — even for our world. Perhaps, though, an additional part of our models can pick out the dimensionality of the space and the metric. But even so, the more serious objection is that not all properties are positional. First of all, it is surely at least logically possible that there exist immaterial entities, and we have no idea if these can be modeled mathematically. And even if such entities were impossible, nonetheless there would be non-positional properties, such as charge, mass, and the like. Considerations similar to those applied in Section 2.2 make it implausible to think that we can find a model in which all possible properties find their place.

2.6 *Combinatorialism and Armstrong*

A somewhat similar approach is to construct possible worlds as recombinations of actual-world entities (Quine 1968; Cresswell 1972, 1973; Armstrong 1989; see the discussion in Lycan 1994; Section 3.2). For instance, Armstrong constructs possible worlds as rearrangements of a subset of the objects of the actual world and of arbitrarily many duplicates of the objects of the actual world. Consequently, it is impossible for a property to be instantiated that cannot be reduced to properties found in the actual world. One way to object to Armstrong is to say that therefore S5 is violated, whereas S5 has intuitive plausibility (see Section 3 of Part I). For it is logically possible, on Armstrong's account, that there exists a world w_1 in which there are no electrically charged entities. Assuming electric charge is a property that cannot be reduced to other properties found in w_1, were w_1 actual, no rearrangement of the entities in w_1 and/or of their duplicates would yield a world with electric charge. Thus, were w_1 actual, the existence of a charged entity would be impossible. But it is possible for w_1 to be actual, and hence it is possible for it to be impossible that there is a charged entity. But there is in fact a charged entity, and so the Brouwer axiom, and *a fortiori* S5, is violated.

However this does not worry Armstrong. After all, it *is* Armstrong's basic view that possible worlds are to be built out of the ingredients of the actual world: "The Combinatorialist . . . is an actual-world chauvinist" (Armstrong 1989: 56), and on this intuition, as Gregory Fitch as noted in conversation, it becomes plausible to think that what is possible depends on what is *in fact* actual, and hence that S5 is false.

However, there is a more decisive alien-properties argument against combinatorialism. According to current physics, two photons can collide and produce two charged electrons, and this is a matter of law. (If it's metaphysically necessary that electrons are charged, we can replace "charged electrons" with "electrons.") Now there is a possible world w_0 with the same indeterministic laws L as our world, and initial conditions I such that:

(83)   There are initially no charged particles at w_0

(84)   Two photons collide and produce a charged electron at some point at w_0:

(85)   At w_0 it is necessarily the case[4] that the physical probability given that the laws are L and the initial conditions are I that a charged particle ever exists is non-zero.

(86)   At w_0 the laws L and initial conditions I made it highly physically probable that there would be no charged particles.

In other words, at w_0 it was unlikely that charged particles should be produced, but that unlikely event did in fact happen.

Now, consider the following plausible principle which should hold at every possible world:

(87)   If a physical state of affairs A is made highly probable by the actual laws L and some initial conditions I compatible with L, then it is metaphysically possible that A is explained by the laws being L and the initial conditions being I.

Intuitively, after all, if A were to obtain given I and L, as it is likely to, it at least *might* be explained by the initial conditions being I and the laws being L. We could even try to strengthen the "might" to a "would," but that would rule out the possibility of a supernatural being intervenes to make A happen independently of the laws, which possibility we may wish to keep intact.

Add two more observations:

---

4   If one thinks conditional probabilities given a full set of laws and initial conditions cannot differ between possible or even possibly possible worlds, the "At w_0 it is necessarily the case" will be otiose.

(88)   Necessarily, an absence of charged particles is a physical
       state of affairs.
(89)   Necessarily, only true propositions can explain anything.

From (86), (87) and (88), it follows that at w_0 it is possible that no charged particle ever existed and that this was explained by the laws being L and the initial conditions being I. By (89), it follows that:

(90)   At w_0 it is possible that: No charged particle ever existed
       and I are the initial conditions and L are the laws.

Then, let w_1 be a world that is possible at w_0 and at which no charged particle ever exists, I are the initial conditions and L are the laws. By (85), the physical probability of there being a charged particle, given the laws L and the initial conditions I is non-zero at w_1. The principle is obvious and had better hold at every possible or even possibly possible[5] world:

(91)   Whatever has a non-zero physical probability given the
       laws and the initial conditions is metaphysically possible.

It follows that at w_1, it is metaphysically possible that there is a charged particle, even though there aren't any. Hence, combinatorialism is false at w_1. But if combinatorialism is a correct account of the nature of modality, it is true at every possibly possible world, including w_1. Thus, combinatorialism is not a correct account of the nature of modality.

2.7 *Fictionalist approaches*

2.7.1 Strong fictionalism
Strong fictionalism as developed by Rosen (1990, 1995) interprets modal claims as claims about what is true in a fictional scenario where something like Lewis's EMR is true. One way to look at fictionalist approaches is as providing a grounding for modality through biconditionals like:

---

5   The combinatorialist can and I think should accept S4, but I shall not assume S4 in
    my argument, and hence need to distinguish between the possible and the possibly
    possible.

(92)   Possibly *s* if and only if according to the fictional scenario,
        s is true at some possible world.

Considered as such an ambitious grounding for modal claims, fictionalism is not particularly plausible.

After all, it seems that possibility enters into the right-hand-side of (92) in two places. First, the fictional scenario seems to be the scenario that all *possible* universes exist. Rosen's (1990: 333) way of handling this problem is by specifying that our world is not special, there are no arbitrary limits on the plurality of worlds, and that there is unrestricted recombination of objects, space-time permitting. But these theses are surely not sufficient to specify what universes exist according to the fiction. We have good reason to think that a finite number of bosons can occupy the same spatio-temporal location. Among the universes in Rosen's fictional scenario, is there a universe where there are infinitely many bosons in the same location? Being told that our world is not special, that there are no arbitrary limits, and that recombination of objects, space-time permitting, is unrestricted does not tell us either way. The difference between the finite and the infinite is not arbitrary, and whether space-time permits infinite energy in one place seems precisely to be a modal question. Maybe the lack of arbitrary limits and the non-specialness of our world implies that for any finite *n*, in the fiction there will be a world whose space-time has dimension *n*. But is there an infinite-dimensional world? Is there a world whose dimensionality is beyond cardinality? Surely the answer to each question should be: "Yes, if it is possible."

Second, the "according to" operator seems to have to be taken to be a modal relation (Nolan 1997), presumably either entailment ("the fictional scenario entails that *s* is true at some possible world") or a subjunctive conditional ("were the fictional scenario to hold, *s* would be true at some possible world"). But if we have entailment available, we have much simpler accounts of possibility, such as:

(93)   *p* is possible if and only if *p* does not entail not-*p*.

And, depending on the logic of counterfactuals, we might very well be able to come up with accounts of modality in terms of subjunctive conditionals directly, without any fictional scenarios, such as:

(94)   *p* is necessary if and only if were not-*p* to hold, *p* would
        hold

(95) *p* is necessary if and only if for all *q*, were *q* to hold, *p* would hold.

2.7.2 Timid factionalism

But even if fictionalism is not satisfactory as an ambitious account of the grounds of modality, it may succeed in a more modest role of giving us a semantics for worlds-talk without the ontological costs of EMR. For instance, if we had a promising EMR-based analysis of some concept, we could then reinterpret this using the fictionalist semantics. This could work as follows. Let *s* be a sentence about possible worlds. Then:

(96) *s* is true if and only if it is the case that: were EMR true, *s* would be true.

Since we can now help ourselves to primitive modality, EMR can be the claim that all *possible* universes exist, and we do not need to worry about the modality in the conditional. Then, the sentence "Unicorns exist in all possible worlds" is false, since were EMR true, certainly it would not be true that unicorns exist in all possible worlds. Anything that can be said using possible worlds talk given EMR can be said in this way, without ontological commitment to EMR. Call this "Counterfactual EMR" (CEMR).

One catch is that the counterfactual "Were EMR true, *s* would be true" cannot be understood in a Lewisian way, since then circularity would ensue. Thus, CEMR requires that counterfactuals make sense independently of possible worlds. Note, too, that the counterfactual in question is a *per impossibile* one. For if EMR is false, it is presumably *necessarily* false, as it is surely a conceptual question whether actuality is essentially indexical or not, and EMR is true if and only if actuality is essentially indexical.

Now, *per impossibile* counterfactuals are rather troubling. Unless we take it that trivially every subjunctive conditional with an impossible antecedent is true (as Lewis's account of counterfactuals implies), it is not clear that they have well-defined truth values. If there were a square circle, would circles fail to be round or squares fail to be quadrilateral or both? Nonetheless, some *per impossibile* counterfactuals make perfect sense. Let "Gnosticism" stand for the thesis that there necessarily exists a necessarily perfectly morally good God and a necessarily perfectly morally evil God. Let us agree that Gnosticism is necessarily false. Then, the following *per impossibile* counterfactuals are non-trivially true:

(97)   Were Gnosticism true, then it would be a necessary truth
        that there are at least two persons.

(98)   Were Gnosticism true, then there would be a supernatural
        being.

(99)   Were Gnosticism true, then Christianity would be
        necessarily false.

We have no difficulty seeing the truth of these *per impossibile* counterfactuals. But what do they *mean*? Suppose someone asks what we mean by (97)? I think we would reply somewhat as follows: "Well, only persons can be morally good or evil, and the same person cannot be both perfectly good and perfectly evil, so if Gnosticism is true, there must be at least two persons." What is interesting about this reply is that it explains what (97) means by giving an argument for it. Perhaps in fact the premises of the argument exhaust the meaning of (97). What we have done in affirming (97) is to have summarized several facts, and which facts we have summarized is determined contextually. We could imagine that in a different context, someone who is committed to the necessary truth of monotheism might say:

(100)   Were Gnosticism true, then it would be possible for the
        same deity to be both perfectly morally good and perfectly
        morally evil at the same time.

When asked what (100) means, she would respond: "Well, if Gnosticism is true, there is a God who is perfectly morally evil and there is a God who is perfectly morally good. But, necessarily, there is but one God. And so perfect moral goodness and perfect moral evil would be compatible." It is true that this is more awkward, because for many speakers the context would determine for us that in the impossible world, monotheism is not retained but the incompatibility between perfect moral goodness and perfect moral evil is retained. But only the context determines this: the literal wording of the antecedent "Were Gnosticism true . . ." does not.

If this is right, then it makes no sense to assert a *per impossibile* counterfactual without having some further story at least vaguely in mind. But then the *per impossibile* counterfactuals of fictionalism at least can't be the whole story. And it is not clear whether such a highly contextual beast as the *per impossibile* counterfactual is going to help us in the metaphysics of modality.

However, we can define a *modified* version of CEMR where the counterfactual has an antecedent that at least has a chance at possibility. Let EMR* be the thesis that every possible universe in Lewis's sense of "universe," i.e.

maximal aggregate of spatio-temporally connected entities, is in fact *actual*: thus, EMR* says that reality and actuality is a maximally diverse aggregate of island universes. Assume that in fact, *pace* EMR and EMR*, only one island universe is actual and real and let *u* rigidly designate it. Observe that it makes sense to talk of *u* coexisting with some other universe: there is nothing absurd about two universes coexisting. Let *s* be a sentence using possible worlds language and no other modal operators. Form a new sentence *s** as follows: Replace every quantification over worlds by a quantification over universes and every occurrence of the word "world" by the word "universe." Replace every occurrence of "actual" (and cognates) that has wide scope by "at universe *u*" (and cognates). Now, take every subsentence of the form "*F* at *w*" (and do similarly for cognates like "it is the case at *w* that *F*") and replace all the quantifiers in *F* that are not in the scope of a further ". . . at . . ." (or cognate) operator and that quantify over worldly objects (rather than about propositions or worlds or universes) with quantifiers restricted to objects in *w*. Thus, "$\exists x(x$ is a unicorn) at *w*" is replaced with "$\exists x(x$ is an object of *w* and *x* is a unicorn)."

We can now give a semantics for Lewisian sentences. Sentence *s* is true if and only if were EMR* true, *s** would be true. This purports to give us a version of CEMR with a counterfactual whose antecedent is in fact possible. Of course, one might well think it fails, because it is impossible that all possible universes should co-exist since there are too many of them — they do not even form a set (recall Section 7 of Part III). But this is an objection a Lewisian cannot give since it is equally an objection to EMR. In fact, any conceptual argument that EMR* is impossible seems to also be an argument that EMR is impossible. For any conceptual objection to the possibility of all possible universes being actual is also going to be a conceptual objection to the possibility of all possible universes existing or being real. Admittedly, Lewis distinguishes reality from actuality, but this distinction is an innovation, and an argument starting from our ordinary pre-Lewisian conceptual framework is not going to distinguish the two. Of course if Lewis could give a conceptual argument showing that all actual items are spatio-temporally interconnected, then it would follow that it is impossible for two universes to be co-actual. But Lewis has no such argument, apart from the full cost–benefit argument for EMR which he is not entitled to use as part of an argument against CEMR.

CEMR seems to offer us an ontologically cheap alternative to EMR. It has some of the same faults as EMR. It does not allow for the possibility of non-spatial non-temporal beings, for instance. Within the CEMR semantics, the sentence "It is necessarily true that every actual entity is spatio-temporally

connected to every other" holds. But it seems superior to EMR in that its ontological commitments are weaker.

One might object that counterfactuals presuppose possible worlds and hence CEMR is viciously circular. That would indeed be problematic, but perhaps can modify CEMR. A sentence *s* is true in modified CEMR semantics provided that EMR* *entails* <*s**>.

There is a serious problem, here, however. Given classical logic, we want it to be the case that for every sentence *s*, either *s* or not-*s* is true. But it need not be true that for every sentence *s*, either EMR* entails *s** or EMR* entails not-*s**. And the problem persists on the counterfactual version. For it need not be true that for every sentence *s*, either it is true that were EMR* to hold, *s** would hold or it is true that were EMR* to hold, not-*s** would hold. The latter claim would be a special case of the Subjunctive Conditional Law of Excluded Middle (SCLEM), which is, in general, false: It is neither the case that were the moon made of a single kind of a single kind of French cheese, it would be made of Brie, nor that were it made of single kind of French cheese, it would not be made of Brie.

Of course the fact that some rule, like SCLEM, is false in general does not imply it is false in a specific case, but the onus of proof is on the person claiming it holds in the specific case. Can the CEMR adherent solve this difficulty? The answer, on both the counterfactual and the entailment account, will be affirmative if the thesis EMR* is a *maximally determinate* proposition, i.e. a proposition such that for every proposition *p* either *p* is entailed by EMR* or not-*p* is. A different way to put it is that a maximally determinate proposition is one such that were it true, it would entail all and only the true propositions. There are maximally determinate propositions if one is any kind of a realist about worlds. Let *w* be a rigid designator of some world. Then "*w* is actual" is maximally determinate. If, on the other hand, EMR* is not maximally determinate, then there is little reason to think that CEMR can escape from the bivalence objection.

But there is reason to think that EMR* *is* maximally determinate, at least if one assumes non-temporal non-spatial entities to be impossible.[6] For if EMR* is not maximally determinate, then there are two different ways that EMR* could be made true. Now any way that EMR* could be made true would correspond to a mereological sum of all possible universes. But there is only one mereological sum of all possible universes, surely. If EMR* is maximally determinate, then in fact there is no difference between the

---

6   Otherwise, EMR* is compatible both with *p* and not-*p* where *p* is the proposition
     that there exist exactly 17 non-spatial, non-temporal substances.

entailment and counterfactual versions of CEMR, since it is clear that if p entails q, then were p to hold, q would hold.

However, there is a more serious problem with CEMR than the lack of determinacy on the part of EMR*. EMR* is the thesis that every possible universe is actual. Thus, EMR* quantifies over possible universes, and hence presupposes that possible universes are there to be quantified over. But it is CEMR, which requires the prior meaningfulness of EMR*, that is supposed to make meaningful the talk of possible worlds, and on the assumption that all possible beings are spatial or temporal there is little difference between quantifying over possible worlds and possible universes.

However, EMR* can be rephrased in a way that does the same work. The rephrasing is a little complicated. We need to remember that the word "actual" and its cognates can both take wide scope and be non-rigid. I will call "actual" with this wide-scope non-rigid usage "actual$_2$." Thus, consider the following sentence:

(101) "Possibly there exists a person who is taller than the actual$_2$ height of every person."

Obviously, if we interpret "actual$_2$" as within the scope of the "possibly," then we get an absurdity: possibly there is a person taller than every person (including herself!) But "actual$_2$" has wide scope, and unless (101) is being tokened in a world that contains a tallest possible person — if there is such a thing — then (101) is true. Now, let p be the proposition expressed by (101). We can now make a stronger claim:

(102) "Necessarily p holds."

This is intended to express that necessarily there is no tallest possible person — for any possible person, there could be someone taller than her. Compare this with:

(103) "Necessarily, possibly there exists a person who is taller than the actual$_2$ height of every person."

Because "actual$_2$" takes wide scope, (103) merely says it is necessarily true that there is a person who is taller than every person of w_0, where w_0 rigidly designates the actual world. However, (102) says more strongly that the proposition p holds at every possible world, and thus that at every world w, possibly there is a person taller than the height everybody has at w. It is

essential for this that while "actual$_2$" takes wide scope, it does not designate rigidly. For if it designated rigidly, then p would be the proposition that there is a person who is taller than every person of w_0. But that is not what p says.

We can explain the usage of "actual$_2$" also in the following way. Let p be proposition expressed by a sentence s containing "actually$_2$," or cognates, and without any intensional contexts other than alethic modal ones. Then, p is true at a world w if and only if the result of substituting "w," "at w," or a cognate for every occurrence of "actually$_2$," or its cognates is a sentence that would be true as tokened at w. It is somewhat unfortunate for purposes of CEMR that to explain how "actual$_2$" works, we have had to use possible worlds talk, but perhaps this is only needed to help the reader imaginatively grasp this item of vocabulary that ordinary language is not sufficiently clear on.

With this understanding, let emr* be the following sentence:

(104)   "Necessarily $\forall u(u$ is a universe $\supset$ actually$_2$ $\exists u^*$ (u^* is an actual$_2$ universe and u is exactly like u^* actually$_2$ is))."

Two terms need explanation here. We need here a special wide-scope quantifier "actually$_2$ $\exists x$" which quantifies not over all items, but over all actual$_2$ items with "actual$_2$" taking wide scope. We also require an interworld version of the "exactly like" or "duplicate of" relation that can relate an actual to a non-actual item. Neither of these two terms is a part of every box–diamond modal logic, but both items seem to be needed to express many modal claims. For instance, barring some clever alternative, we need the interworld qualitative identity relation to express the claim that possibly there is a person qualitatively different from everybody who actually exists. And likewise it seems we need the "actually$_2$ $\exists x$" quantifier to express a claim such as that possibly there is a human being whom no actual$_2$ human being has conceived of:

(105)   Possibly $\exists x(x$ is a human being and not actually$_2$ $\exists y(y$ has actually$_2$ conceived of x)).

Now let EMR* be the proposition that the sentence emr* expresses. Then, EMR* is true at a world w if and only if it is true at w that necessarily every universe is qualitatively identical with one of the universes in w. The advantage of the way that EMR* has been defined via emr* has been, however, that we did not make use of the notion of a possible world while defining it, though for clarity possible worlds talk did get used to help the

reader grasp the notion of "actual$_2$." There are two quantifications in (104). Neither is a quantification over possibilia. The outer quantification is supposed to hold necessarily, but this fact does not make it a quantification over possibilia, any more than the claim that necessarily every horse is a mammal is a quantification over possibilia (we should not analyze "Necessarily every horse is a mammal" as "Every possible-horse is a mammal"). What we have, rather, is an innocent sentence of first order quantified logic that every horse is a mammal, embedded within a necessity operator.

CEMR with a thus-modified EMR* appears to give us an alternative to EMR for interpreting modal language. This does not let us do absolutely everything EMR does. We cannot, for instance, blithely define a proposition as a set of worlds. However we can use CEMR's semantics to interpret modal propositions as in EMR in terms of quantification. CEMR is a form of fictionalism: it lets us talk as if there were worlds. Likewise, it lets us talk as if there were Lewisian propositions and properties. If we can rest satisfied with this, then CEMR gives us everything EMR gives us, but seemingly at less cost.

Unfortunately, the metaphysical cost of CEMR may not be that low. Consider the sentence emr*, defined by (104). To understand emr*, one must understand quantification over items in a world different from the one at which the quantification is evaluated. The "actual$_2$ $\exists x$" quantifier is evaluated from within a subsentence of emr* within the scope of "necessarily," i.e. evaluated from the point of view of worlds other than the actual one. Moreover, there needs to be some sort of accessibility of the items in one world from another world if the "exactly like" is to work. While these things do not posit possible worlds as such, they do presuppose quite a bit of realism about non-actual entities.

And of course CEMR is still limited to spatio-temporal items, in the same objectionable way that Lewis's theory was, since the fiction it is based on is Lewis's story. The possibility of non-spatio-temporal entities by itself makes CEMR is inadequate. Thus, CEMR both is inadequate to handle all possibilities, and isn't that cheap metaphysically.

2.7.3 Sider's pluriverse

The limitation of CEMR to spatio-temporal items, together with the complications about defining emr*, might make one opt for a different kind of fictionalism. As the best fictionalism, at least broadly construed, currently available to us is probably that of Sider (2002), it will be instructive to examine it in detail. Sider's fictionalism is a reductive approach. The approach involves two languages. There is the ontologically expensive

possibilist language (PL) which allows one to quantify over worlds and over the properties and individuals in other worlds. We want to get an account of truth of sentences in possibilist language. And we have a less problematic *modal language* (ML), which has the box and diamond operators, but that is all in terms of modal expressiveness. Ideally, we would want to reduce PL to ML, but since it is known that ML lacks the expressive power of PL (Melia 1992; and recall our discussion in Section 1 of Part II), this is not going to work. Instead, we give a fictionalist semantics of PL that correctly lines up with the modal truths of ML.

We need more precision. Both languages are Lagadonian in the sense that the actual individuals and properties existent at our world[7] stand for themselves. We also have an infinity of individual variables such as x_i and y_j, property variables such as P_i and Q, and we allow infinite blocks of quantifiers. PL additionally has world variables, such as w_j, and one special world-constant, @, whose intended interpretation is the actual world. Moreover, both PL and ML have an "instantiates" operator which can be multi-grade: we can say that a sequence of variables $x_1,...,x_n$ instantiates an n-ary property (i.e. relation) P just as we can say that x_1 instantiates a unary property Q. The difference is that in PL the *instantiates* operator requires qualification with a world-variable when the sequence on the left of the operator consists solely of variables; e.g. "P_1,P_2 instantiates Q at w." Let us assume for simplicity that a property is either solely of individuals ("first order property") or solely of properties ("higher order property") — the general case would only complicate the notation.

There is no need for an "at w" qualifier for instantiation by individuals in PL because Sider is a counterpart-theorist so that an individual can only exist at one world and it is at that world that we are interested in what it instantiates. Furthermore, PL contains an "exists in" operator so we can talk of an individual or property as existing in a world.

ML does not have any world variables. Instead, it has ◊ and □ operators which PL does not need given its ability to quantify over worlds. And of course we include the standard logical connectives in both languages.

There is a natural semantic interpretation of ML, which can be defined in the usual way recursively: "u instantiates Q" is *true* provided the individual u in fact instantiates the property Q, ◊p is *true* if in fact possibly p is true, and so on. Let CONSTRAINTS be the subclass of sentences of ML which are

---

7   Which could include all possible properties if there can be uninstantiated properties, though in that case other forms of ersatzism such as Lagadonian linguistic reduction-ism become more plausible than they would otherwise be.

de dicto and true, where a sentence is *de dicto* provided that it has no modal operators within the scope of a quantification over individuals. Sider limits himself to *de dicto* sentences because he is a counterpart theorist. Modal operators within the scope of quantifications over properties are apparently allowed, however: "There is a property which is such that it is impossible that there exist two distinct individuals that instantiate it" is a permissible sentence.

There are two ways of building up Sider's fictionalism, which Sider shows are equivalent. The model theoretic approach is the one I will follow. A modal model **M** is a hextuple $\langle \mathbf{W}, \mathbf{r}, \mathbf{D}, \mathbf{P}, \mathbf{Q}, \mathbf{I} \rangle$, where **W** is a non-empty set ("worlds"), **r** is a distinguished member of **W** ("the actual world"), **D** is a set ("individuals"), **P** is also a set ("properties"), **Q** is a function that assigns to each world an ordered pair of subsets of **D** and **P**, respectively ("the individuals and properties of that world"), and **I** assigns to each n-ary property P its intension, i.e. a function mapping each world to some set of n-tuples of objects (individuals or properties, with types matching those associated with property P so as to avoid category mistakes) in that world. In the intended interpretation, then $x_1, \ldots, x_n$ instantiate P at w if and only if $I(P)(w)$ contains $\langle x_1, \ldots, x_n \rangle$.

From now on, we will restrict the interpretations of PL and ML to what Sider calls "Lagadonian" interpretations, where a property and an individual is always interpreted as *itself*.

We can then define the truth of a PL sentence in a model relative to such an interpretation. An interpretation assigns to each individual constant of PL a member of **D**, to each property constant of PL a member of **P** and to @ the special member **r**. Basically, then, we just check whether under this interpretation the PL sentence is true of the model or not. Thus, for instance, if u is an individual, then "u instantiates Q at @" is true in **M** provided that $I(Q^*)(\mathbf{r})$ contains u^*, where Q^* and u^* are the interpretations of Q and u, respectively.

It is also easy to define the truth of an ML sentence in a model relative to an interpretation. Just take the ML sentence and translate it in the natural way into a PL sentence, replacing modal operators by quantification over worlds, and check if that sentence is true in that model relative to that interpretation.

Finally, a model **M** is said to be *realistic*, relative to some interpretation, provided that every member of CONSTRAINTS is true in **M** relative to that interpretation.

We can now give our reductive account of the truth of PL sentences. A PL sentence s is true provided that it is true in every realistic model under

every interpretation under which the model is realistic. For brevity, I will simply omit mention of interpretations henceforth and just talk of truth in a model — if one wishes, one can consider an interpretation as built into a model. So a PL sentence is true if and only if it is true in every realistic model.

We have thus started with modal truths in ML, and built our way up to giving truth conditions for modal truths in PL. There is, however, a very unfortunate problem here, which we were already worried about in the case of CEMR. If s is a sentence of PL, let $T(s)$ be the proposition that s is true in every realistic model. Observe that $T(\sim s)$ is then a different proposition from $\sim T(s)$ and that it is not obvious whether in general it is the case that $T(s)$ or $T(\sim s)$ holds, i.e. whether Sider's fictionalism supports bivalence in PL. It is clear that at most one of $T(s)$ and $T(\sim s)$ holds. But do we have at least one of them holding? The answer is positive if and only if s is a proposition that has the property of either being true in all realistic models or being false in all realistic models.

However, as Sider knows (cf. Section III.D and footnote 27 in Sider 2002), not all sentences are like that, and bivalence fails. Before we show this, let us reflect on whether this is important. Bivalence is intuitively right, except maybe in cases of vagueness, which our cases will not be — the loss of it is a serious problem. Moreover, let s be a sentence of PL for which bivalence fails so that neither s nor $\sim s$ is true under the fictionalist interpretation of PL. On the other hand, the disjunction "s or $\sim s$" will count as true under the fictionalist story, since in every realistic model either s will be true or s will not be true, assuming excluded middle in ML. Thus, Sider's fictionalism will commit one to excluded middle and the denial of bivalence. And this, in turns, forces one to deny Tarski's T-schema:

(106)   "..." is true if and only if ....

To see this, suppose that "xyzzy" abbreviates a sentence of PL which is not true and whose negation, "~xyzzy" is also not true, and assume the T-schema. By the T-schema:

(107)   "xyzzy" is true if and only if xyzzy.
(108)   "~xyzzy" is true if and only if ~xyzzy.

It follows that:

(109)   ("xyzzy" is true or "~xyzzy" is true) if and only if (xyzzy or ~xyzzy).

But by "xyzzy or ~xyzzy" is true by excluded middle, and so by another application of the T-schema, it follows that "xyzzy" is true or "~xyzzy" is true. Denying the T-schema is costly.

It may be that fictionalism is stuck with the denial of bivalence just by virtue of being fictionalism. According to Rescher (1999), one of the central differences between fiction and reality is precisely that reality is fully determinate while fiction is not. Thus while it is true in the Sherlock Holmes corpus that Holmes had an even or odd number of hairs on his head at 9 a.m. on October 17, 1894, it is neither true in the corpus that he had an even number then, nor is true in the corpus that he had an odd number then. Thus, the lack of bivalence emphasizes that we are dealing here with a fiction.

I shall now show that in fact there is a non-bivalent sentence of PL given Sider's fictionalism. The sentence in question will say that there are at least two indiscernible worlds. Constructing the sentence will take a while. I am going to assume for simplicity that all possible properties are at most N-ary for some possibly infinite N and that the number of individuals and of properties in a given world also has at most cardinality N. (That there is an upper cardinality bound on the arity of properties follows from the fact that there is a set of all properties in Sider's approach.) I will also assume that by definition a sequence of k objects does not instantiate an n-ary property if $n \neq k$.

Suppose $x_1,...,x_N$ and $y_1,...,y_N$ are individual variables and P is a property variable. Let F be a map from the set $\{1,...,N\}$ to the set $\{*,1,...,N\}$. Write $I_F(x_1,...,x_N;y_1,...,y_N;P)$ for the formula:

(110) $x_{F(1)},...,x_{F(k)}$ instantiates P if and only if $y_{F(1)},...,y_{F(k)}$ instantiates P,

with the convention that in a list a variable of the form $x*$ is always to be omitted, so that "$x_1,x*,x_2$" is a fancy way of writing "x_1,x_2."

Let $I(x_1,...,x_N;y_1,...,y_N;P)$ be the conjunction of all the formulae $I_F(x_1,...,x_N;y_1,...,y_N;P)$ as F ranges over all maps from $\{1,...,n\}$ to $\{*,1,...,n\}$ such that for some i we have $F(i) \neq *$. For instance, $I(x_1, x_2;y_1,y_2;P)$ is:

(111)   (x_1 instantiates P if and only if y_1 instantiates P) & (x_2 instantiates P if and only if y_2 instantiates P) & (x_1,x_1 instantiates P if and only if y_1,y_1 instantiates P) & (x_2,x_2 instantiates P if and only if y_2,y_2 instantiates P) & (x_1,x_2 instantiates P if and only if y_1,y_2 instantiates P) & (x_2,x_1 instantiates P if and only if y_2,y_1 instantiates P).

Now, let $\text{IND}_1(w_1,w_2)$ be the sentence:

(112) $\forall x_1...\forall x_N[(x_1$ exists in w_1 & ... & x_N exists in $w_1) \supset$
  $\exists y_1...\exists y_N(y_1$ exists in w_2 & ... & y_N exists in w_2 &
  $\forall P(I(x_1,...,x_N;y_1,...,y_N;P))]$.

This sentence says that w_1 and w_2 are indiscernible with respect to first order properties.

We also need a sentence to say that the worlds are indiscernible with respect to higher order properties. This is reasonably simple. Let A be a subset of $\{1,...,N\}$. Define $J_A(P_1,...,P_N;Q)$ as the formula:

(113) $P_{F(1)},...,P_{F(k)}$ instantiates Q at w_1 if and only if $P_{F(1)},...,P_{F(k)}$ instantiates Q at w_2,

where $F(1),...,F(k)$ is a list of the members of A ordered in ascending order. Let $J(P_1,...,P_N;Q)$ be the conjunction of the formulae $J_A(P_1,...,P_N;Q)$ as A ranges over the subsets of $\{1,...,N\}$. Finally, let $\text{IND}_2(w_1,w_2)$ be the sentence:

(114) $\forall P_1...\forall P_N\forall Q(I(P_1,...,P_N;Q))$.

We can now let IND be the sentence:

(115) $\exists w_1\exists w_2(w_1 \neq w_2$ & $\text{IND}_1(w_1,w_2)$ & $\text{IND}_2(w_1,w_2))$

This sentence says that there are at least two indiscernible worlds.

As it turns out, neither IND nor ~IND comes out true in Sider's fictionalism, unless it should happen that there are no realistic models at all, which would be a worse kind of disaster for Sider. I will only show that ~IND does not come out true. The argument for the other part of the claim is left as an exercise to the reader, though I will sketch the central idea.

To show that ~IND does not come out true, I need only show there is a realistic model in which IND holds. Take any realistic model $\mathbf{M}=\langle\mathbf{W},\mathbf{r},\mathbf{D},\mathbf{P},\mathbf{Q},\mathbf{I}\rangle$, and fix an interpretation. Let u be *any* world in that model ($\mathbf{r}$ for definiteness if one wishes). Let $\mathbf{M}^*$ be a new model produced by starting with $\mathbf{M}$ and adding a new world u^*. This is done as follows. Let u^* be a new set theoretic entity which is not a member of any of the sets making up $\mathbf{M}$. Let $\mathbf{W}^*=\mathbf{W}\cup\{u^*\}$, $\mathbf{r}^*=\mathbf{r}$ and $\mathbf{P}^*=\mathbf{P}$. Let $\mathbf{Q}^*(w)=\mathbf{Q}(w)$ for w other than u^*. Let $\mathbf{Q}^*(u^*)=\langle A^*,B\rangle$ where $\mathbf{Q}(u)=\langle A,B\rangle$, i.e. where A is the set of objects of u and B is the set of the properties of u, and where A^* is a set of new objects, with

A^* having the same cardinality as A, and where the objects do not coincide with any other set-theoretic entity mentioned so far. If x is in A then we write x^* for the corresponding member of A^* under some fixed one-to-one correspondence between all of A and all of A^*. Start defining $\mathbf{I}^*$ by making it coincide with $\mathbf{I}$ on all worlds other than u^*. Finally define $\mathbf{I}(u^*)$ by making it coincide with $\mathbf{I}(u)$ for higher-order properties, and for first-order properties define $\mathbf{I}^*(u^*)(P)$ to be the set of n-tuples $\langle x_1^*,...,x_n^* \rangle$ where $\langle x_1,...,x_n \rangle$ is a member of $\mathbf{I}(u)(P)$. Then, $\mathbf{M}^* = \langle \mathbf{W}^*,\mathbf{r},\mathbf{D}^*,\mathbf{P}^*,\mathbf{Q}^*,\mathbf{I}^* \rangle$. Basically $\mathbf{M}^*$ consists of taking $\mathbf{M}$ and adding an indiscernible copy u^* of the world u. Extend the interpretation of $\mathbf{M}$ in the natural way to $\mathbf{M}^*$.

It is easy to see that IND holds at $\mathbf{M}^*$. Just use u and u^* to instantiate the existential quantifiers in IND and instantiate $y_1,...,y_N$ with $x_1^*,...,x_N^*$ in the definition of $\text{IND}_1(u,u^*)$. The only question is whether $\mathbf{M}^*$ is a realistic model. But this, too, is not difficult to see. For in fact if s is a *de dicto* sentence of ML, then s is true at $\mathbf{M}$ if and only if it is true at $\mathbf{M}^*$. From this it follows that CONSTRAINTS holds at $\mathbf{M}^*$.

An easy way to see that s is true at $\mathbf{M}$ if and only if it is true at $\mathbf{M}^*$ is as follows. Take first a sentence s of ML with no modal operators. It is easy to define a notion of a sentence s holding at a world w of $\mathbf{M}$: just modify the interpretation in such a way that the actual world is mapped to w instead of to $\mathbf{r}$ or, alternately, define a model $\mathbf{M}_w$ which is the same as $\mathbf{M}$ except that instead of $\mathbf{r}$ we have w, and say that s holds at world w of $\mathbf{M}$ if and only if s holds at $\mathbf{M}_w$. Then, it is clear that:

(116) $\forall w$ (s holds at world w of $\mathbf{M}^*$ if and only if s holds at world w of $\mathbf{M}$).

It is also clear that:

(117) s holds at world u^* of $\mathbf{M}^*$ if and only if s holds at world u of $\mathbf{M}$.

Next, note that if s is any sentence satisfying (116) and (117), a sentence formed by prefixing a box or diamond in front of s also satisfies them, as does any sentence formed truth-functionally out of sentences satisfying (116) and (117). But any *de dicto* sentence of ML can be built up in this way. Thus, any *de dicto* sentence of ML satisfies (116) and (117). In particular, any *de dicto* sentence of ML holds at world $\mathbf{r}$ of $\mathbf{M}^*$ if and only if it holds at world $\mathbf{r}$ of $\mathbf{M}$ (since $\mathbf{r}$ is not u^*), and hence any *de dicto* sentence of ML holds in $\mathbf{M}^*$ if and only if it holds in $\mathbf{M}$.

Thus, IND is true in some realistic models. Showing that IND is false in some realistic models proceeds as follows. Take a model **M** at which IND is true. Excise all but one world in each class of indiscernible worlds of **M**, being careful not to excise **r** (this construction implicitly uses the Axiom of Choice). This gives a new model **M***. It is not difficult then to prove that any sentence of ML holds in **M*** if and only if it holds in **M**, using similar methods to those used before.

Hence, Sider's account neither makes IND true, nor makes IND false. A realism about possible worlds as well as our natural pull towards bivalence suggests this is significant evidence against Sider's account. Questions like "Could someone other than me have taken my place in a world just like this one?" seem to make perfect sense. But an affirmative answer to this question entails IND, which makes it very plausible that IND makes sense.

Of course, Sider might claim that sentences like IND are insignificant. But it is interesting to note that it is precisely sentences like IND that are the best examples of the expressive shortfall of ML — and in fact the above argument is one way of showing this expressive shortfall, since within ML, as we have seen, we cannot tell the difference between models in which IND holds and ones in which it doesn't. To the extent that Sider's story cannot handle such sentences well, the advantages over plain ML are significantly reduced, though not completely eliminated *if* there are also Sider-bivalent examples of the expressive advantages of PL.

We could consider an extension of Sider's view that might have a chance of doing better with regard to IND. Suppose we allow true *de re* sentences in CONSTRAINTS. Then, we might have a sentence in CONSTRAINTS that entails IND. For instance, suppose that in the actual world there is only one individual, Fred, and he has only one property, mimsiness. Then in CONSTRAINTS we may have the sentence:

(118)   ◊[mimsiness exists & $\exists x(x \neq$ Fred & $\forall y(y=x))$ &
          $\forall P(P=$mimsiness$)$].

And if we do, then CONSTRAINTS constrains models to ones that contain a world that has only one property, mimsiness, and one entity, who is not Fred, who has that property. Any such world will be indiscernible from the actual world, and so IND will hold in all realistic models, relative to these new constraints. The cost of this extension is that it makes the ersatzism rather less ambitious than it was before by building in more modal constraints — namely the *de re* ones — and hence the ersatzism does even less grounding of modal truth.

In summary, Sider's fictionalism does allow alien individuals and properties, and has an admirable precision and clarity, in its standard form it makes some apparently meaningful modal questions, such as whether IND nor ~IND, fail to have an answer. It can be supplemented to include *de re* constraints, but then it does not do as much work in grounding modal truths.

Section 3 Platonism

3.1 *Platonic approaches to modality*

On a number of ersatzist accounts, possible worlds are Platonic entities, such as collections of propositions (Adams 1974), maximal states of affairs (Plantinga 1974), constructions from properties (Castañeda 1974, 1989; Parsons 1974, 1980) or *sui generis* Platonic entities (Stalnaker 1976). I will specifically focus on the Adams propositional account, though most of the criticisms will apply to all theories that explain possible worlds in terms of Platonic abstracta.

3.2 *What are propositions?*

One can introduce, though not define, the theoretical entities called "propositions" and the relation of "expressing" between assertings and propositions as well as between thinkings and propositions through the following plausible claims:

(119) our assertings and thinkings express propositions and precisely one proposition is expressed by any one asserting or thinking;

(120) assertions have the same content precisely when they express the same proposition;

(121) one says what one thinks precisely when one's asserting expresses the same proposition as one's thinking; and

(122) the intuitive notion of the "content" of a given asserting or a thinking is the same as the notion of the proposition being expressed by the asserting or thinking.

Moreover, to further expand the theoretical account which introduces "propositions," one can tell a story about a relation between propositional

attitudes, such as that between the attitude of *desiring that p* and the proposition *that p*, analogous to the relation of "expressing." One can further say that an asserting is correct if and only if the expressed proposition is true.

Not all propositions are expressed by one of our actual assertions or acts of thinking. However, given the intuition that there are things that we *could* mean or express, but, in fact, which we do not mean or express, and given the further Platonic intuition that these things are not different in kind from those things we in fact mean or express, we conclude that the class of propositions is wider than the class of things that our assertions in fact express. Moreover, it is plausible to suppose that there are some things that none of us is capable of expressing: after all, it would be surprising to think that our expressive capacities should be capable of covering the collection of all possible things that can be expressed. These considerations lead us to conclude that there are propositions that we cannot even express.

It is plausible to suppose, then, that propositions are *necessary* entities, since they do not appear to depend on our existence or on the existence of other contingent expressive beings, and it is not at all clear what else they could depend on.

The collection of propositions then far outstrips us. We *introduced* them as theoretical entities which are the things we express by our claimings and the things that claimings with the same content have in common. But this introduction does not exhaust the collection, any more than the fact that we may have introduced the collection of electrons for explaining a bunch of phenomena in this galaxy means that we have no right to suppose electrons in other cases. Note finally that the collection of propositions is so large as not to be a set (see Grim 1986; Chihara 1998: 125ff; this also follows from the fact that the collection of possible worlds is too large to be a set, since, plausibly, to every world there corresponds the proposition that that world is actual).

It is reasonable to suppose that all propositions stand in logical relations to one another, since only then are they a kind of thing that can be meant or expressed. The ersatzist will want to say: "Let us consider maximal compossible collections of propositions and call them 'worlds'." But it must first be argued that there are such collections.

Now, first of all, it seems we know there is at least one maximal compossible collection of propositions, namely the collection of all propositions that are actually true. This collection is maximal, because if p is a proposition, then either p is actually true or actually false. If p is actually true, then p will be a member of the collection. If p is actually false, then not-p will be a

member of the collection, and hence p could not be added to the collection without introducing inconsistency.

But perhaps there is only one such collection? This would not make the ersatzist happy. In order for modality to be definable in terms of quantification over possible worlds, the ersatzist needs to claim that for any possible proposition p there is a world, i.e. maximal compossible collection of propositions, containing p as a member. This is a nontrivial claim.

One could for the existence of a maximal collection using the SCLEM. For if for every p, it either is the case that were p true, q would be true, or it is the case that were p true, q would not be true, given a possible p, we can define S_p as the collection of propositions q that would be true were p true. It is plausible that S_p would have to be consistent if p is possible. Unfortunately, SCLEM is implausible.

There is, fortunately, another construction that intuitively at least appears to work. Let P be the collection of all propositions, and let F be the collection of all mappings (not functions, since P is not a set) from P to the set { *true*, *false* }. Let V be the predicate that applies to a member f of F if and only if $\forall p(f(p) = true$ if and only if p is true). Now, plausibly, it is a necessary truth that some mapping in F satisfies V, since it seems to be a necessary truth that there is a mapping in F that sends all and only the true members to *true*. Therefore:

(123) $\quad \Box \exists f$ (f is in F and Vf).

But now suppose p is possible. Then, since the conjunction of a possible truth with a necessary truth is also possible, it follows from (123) that:

(124) $\quad \Diamond(p$ and $\exists f$ (f is in F and Vf)).

Hence:

(125) $\quad \Diamond(\exists f$ (p and f is in F and Vf)).

Thus, since p and Vf entail that $f(p) = true$ by definition of the predicate V:

(126) $\quad \Diamond(\exists f$ ($f(p) = true$ and f is in F and Vf)).

But now mappings are necessary entities and one can interchange the order of quantifiers over necessary entities (like mappings) and modal operators (recall the discussion of formula (10) in Section 3 of Part I) so that:

(127) $\exists f\, (\Diamond(f$ is in F and $f(p) = true$ and $Vf))$.

Now let f_0 be a mapping such that it is possible that f_0 is in F and $f_0(p) = true$ and Vf_0 holds. Then, in fact $f_0(p) = true$ since mappings have their values necessarily, and also in fact f_0 is in F since collection membership is necessary when dealing with necessary entities. Thus, $f_0(p)$ is *true* while Vf_0 is possible. Now, let C be the collection of all propositions q such that $f_0(q)$ is *true*. Then, because Vf_0 is possible, C must be compossible, since, necessarily, Vf_0 holds if and only if every member of C is true. And, because $f_0(p)$ is *true*, p is a member of C. Moreover, C is maximal, because for every proposition q if q is not in C, then $f_0(q) = false$ and hence $f_0(not\text{-}q) = true$, and hence not-q is a member of C.

Thus, the Propositional Ersatzist seems to be entitled to her maximal compossible collections of propositions. Since, moreover, we have shown that any possible proposition is contained in a maximal compossible collection of propositions and evidently no impossible proposition is, it is true that a proposition p is possible if and only if it is a member of some world. But of course this cannot be intended as an analysis of possibility, at the pain of vicious circularity, as the concept of possibility is presupposed by the concept of compossibility.

One can also vary the above construction slightly and define worlds as conjunctions of the propositions in a maximal compossible collection of propositions.[8]

While neither construction yields an analysis of possibility, the ersatz propositional worlds are very useful for conceptual purposes. At the very least, they are useful for making complex modal arguments clearer. But, furthermore, they do provide a conceptual tool that mere modal operators do not, at least not in a straightforward way: these ersatz worlds can be used for analysis of supervenience, world-similarity and counterfactuals just as Lewis's worlds can.

3.3 Objections to Platonic Ersatzism

3.3.1 The set-theoretic objection
One might worry that the collection of all propositions is an impossible collection. After all, it seems it is not even a proper class. For, there is no proper class of all proper classes, while there is a different proposition

---

8   cf. Gale and Pruss (1999).

corresponding to every proper class, viz. the proposition that asserts that the class is self-identical (cf. Grim 1986). However, this objection is not fatal. One can distinguish theoretical levels. First-level propositions aren't allowed to be about proper classes or propositions. Second-level propositions are allowed to be about proper classes or first-level propositions, but not higher-level entities. And so on. Moreover, for purposes of analyzing modality, little is lost in restricting ourselves to, say, second-level or third-level propositions, because intuitively contingency is rooted at the first level.

Alternately, one can try to find a different set-theoretic axiomatization for which things would work out. While some collections like the Russellian set of all non-self-containing sets are malformed, a collection is arguably innocent until proven guilty, and so one might have a collection of all sets, without a collection of all non-self-containing sets, if one finds a clever replacement for the Axiom of Separation.

3.3.2 Alien properties

If a proposition predicating a property only exists when that property exists, and if alien properties do not exist, then obviously the Propositional Ersatzist is in the same kind of trouble that the combinatorialist was. But the Propositional Ersatzist can dig her heels in. While a linguistic ersatzist may be limited to propositions we can express, the Platonist's abstract propositions are not the things (i.e. contents) that we can express, but things of the same *kind* as the things we can express. There are many kinds of things that we cannot express.

3.3.3 How do propositions represent?

David Lewis has a very shrewd way of introducing Propositional Ersatzism (see Lewis 1986; Section 3.4). First he describes a variant of Propositional Ersatzism without ever using the word "proposition." There just is some collection of entities with certain formal properties and relations, namely those that propositions have but without their usual names like "reporting," "being true," etc. He then wonders why this collection should have anything to do with possibility, and asks whether calling the members "propositions" and bestowing the usual names on their properties makes it any less mysterious what these entities are and why they have anything to do with modality. A special case of this puzzlement is the wonder of how it is that the propositions represent. What is it that makes some entity *p* represent there being unicorns as opposed to representing there being horses?

However, as I suggested, propositions should be thought of as theoretical entities analogous to those brought in by science. Forget what physics you

know for the moment, while remembering mathematics. Imagine that I gave an account of the following sort. There is a set S of entities each member e of which has a quadruple of quantifiable properties, which we shall call $u_1(e)$, $u_2(e)$, $u_3(e)$ and $u_4(e)$, and binary relations between quadruples, relations which I shall neutrally call "binrels," that ensure that the mathematical relations

$$\frac{\partial^2 u_i(e)}{\partial t^2} = \sum_{e' \in S-\{e\}} \frac{u_4(e')(u_i(e)-u_i(e'))}{((u_1(e)-u_1(e'))^2 + (u_2(e)-u_2(e'))^2 + (u_3(e)-u_3(e'))^2)^{3/2}}$$

are satisfied for $i=1,2,3$, and $u_4(e)$ is constant over time.

Suppose I then said that this explains various astronomical phenomena. Obviously, as far as the account went, my claim would be crazy (if you do not think it is crazy, this is only because you recognize what interpretation the quantities should be given). The account as given explains nothing because it is not sufficiently interpreted: there are many collections of entities that one could imagine singling out that satisfy this relation, and few of them explain astronomical phenomena. It is only when I further specify that $u_1(e)$, $u_2(e)$ and $u_3(e)$ are position coordinates and $u_4(e)$ a mass, all expressed in appropriate units, that this account becomes at all useful. (I can then go on and say that the "binrels" are gravitational forces, though this will not be helpful to my interlocutor if my interlocutor has never before met the concept of a "force.")

Suppose I now make a claim like that which Lewis makes about Propositional Ersatzism in this case: to say that $u_1(e)$, $u_2(e)$ and $u_3(e)$ are "position coordinates" and $u_4(e)$ a "mass" does not make the account in any way more explanatory or clear. The objection will be inappropriate, because prior to my giving the account we had a certain grasp of positions and masses. Likewise, claiming that it is something magical about some property that makes it a "mass" is out of line.

But much the same defense can be made of Propositional Ersatzism as of the Newtonian theory of gravity. Prior to introducing worlds built out of propositions we had a concept of a "proposition," as a generalization from the concept of that entity which is what our language expresses. Just like "mass" and "position" had a prior role in our language game, and were introduced for good explanatory reasons, so too "proposition" had a prior role in our language game and was introduced for a good explanatory reason. Thinking that propositions represent "magically" is just like claiming that gravity is magic (which is what Leibniz's famous criticisms of Newtonian gravity amounted to).

This does not mean that there is no real point to asking what it is about propositions that makes them represent certain states of affairs. There *is* a substantial question there, just as there is a substantial question as to what it is that makes mass and position interact in ways that satisfy the laws of gravitation. However, given how propositions were introduced, there is no problem with thinking that they *do* represent. This is something that the defender of propositions is entitled to, and there is no requirement that she give an account of how propositions represent prior to her being entitled to propositions, just as Newton does not need to explain just *how* gravity manages to "move" things.

This reply to Lewis is why it was essential that I first introduce propositions as fulfilling a different theoretical role from the one that the Propositional Ersatzist needs them for, just as it was essential that notions of mass and position had a role to play prior to the advent of the Newtonian theory of gravitation (mass and position both play a role in Newton's Second Law, for instance).

Furthermore, there is an *ad hominem* answer to Lewis's concern (van Inwagen 1986; Jubien 1991). Lewis's own theory admits abstract entities such as sets. But the relation between a set of concreta and the concreta it is a set of is no less mysterious than the relation which would hold between a proposition and the proposition's truthmaker, were the proposition true with a truthmaker. To the question of why such and such an abstract entity represents there being horses rather than there being dogs, one may respond with the question of why such and such an abstract entity is related in the set-to-members way to cows rather than to chairs.

And if Lewis insists that on his view there is only one magical relation, set membership, while on the propositionalist's view there is both set membership and representation, the friend of propositions can reduce sets to propositions. Plausibly, for any *x*s there is a proposition that those *x*s exist. One can then identify *collections* with those propositions that affirm the existence of one or more things. Then *u* is a *member* of a collection *c* if and only if *c* represents, perhaps *inter alia*, *u*'s existing. We can then pick out our favorite axioms of set theory, and stipulate that any collection of collections that satisfies these axioms counts as a universe of *sets*. Of course, we will have to decide whether to permit propositions that affirm the existences of non-existent things when defining "collections." If we do permit them, and if propositions are all necessary beings, then we will have sets of nonexistent things. If they are not, then we will not.

3.3.4 The unmet extended Parmenidean challenge

However, the Ambitious Propositional Ersatzist ought not rest satisfied with the above answers to Lewis. For there is still Parmenides's challenge to be faced. What real thing are we talking about when we make modal affirmations? The answer our Ambitious Ersatzist offers is: propositions. But what is it about these propositions to make them suitable for being the truthmakers of modal claims?

This is only a question for the Ambitious Ersatzist. The Unambitious Ersatzist can shrug her shoulders and reply that the ersatz worlds constructed out of propositions *model* the grounds of modal propositions, but that's all. But the Ambitious Ersatzist insists that these ersatz worlds, or at least their propositional ingredients, are involved in the truthmakers of modal propositions, and hence provide an answer to Parmenides.

Lewis raises this problem under the head of "primitive modality." Our ersatzist has presupposed modality in the very definition of her worlds, when talking of compossible collections of propositions. It is Lewis's contention that the EMR-theorist can do without *any* primitive modality, and if this is right, then a view that presupposes primitive modality is less theoretically desirable. However, Lewis's account also presupposes some primitive modality. For Lewis's counterpart relation, while contextually determined, surely cannot be *just any* relation satisfying formal constraints. A counterpart relation that makes me a counterpart to a galaxy in w_7 and that makes nothing else a counterpart to anything else is just not the right counterpart relation for *any* notion of *de re* modality. The constraints on what counterpart relations yield a viable notion of *de re* modality will be a kind of primitive modality. Thus, in order to make a "primitive modality" objection against the ersatzist without opening oneself for a *tu quoque* reply, one needs to explain what is particularly objectionable about the ersatzist's primitive modality.

The answer to this is that the ersatzist fails to give any illuminating answer to Parmenides. If the Parmenidean argument against modality is to be escaped, something must be pointed out that can plausibly be said to be the ground of modal claims. The Ambitious Ersatzist has pointed out the necessarily existent entities known as "propositions." When we say it is possible that there are unicorns, we are predicating the property of being possible of the proposition that that there are unicorns. Perhaps this is all that can be said, but if so, this state of affairs is disappointing. It is like explaining mind by saying that there is a primitive property of *mindedness* and any entity that has it is a mind. An account that could say more would be preferable theoretically.

A related worry is as follows. We can deduce that it is possible that there are horses from the fact that there are horses. Let p be the proposition that there are horses. Then, the claim is that possibility is a basic property that p has (it might be that one takes necessity to be a basic property, with possibility defined in terms of it; but *mutatis mutandis* the same objections can be made, so this is an option I shall not explore further). But how is p's having this property connected with p's being true, i.e. with the horses that make p true? In general, necessarily, whenever a proposition p has a truthmaker, then p has the property of possibility. One might think therefore that there should be some connection between the truthmaker and the property of possibility. But it is inexplicable on the "basic property" view of possibility what this connection is. The truthmaker of the proposition is, often, a concrete contingent entity while the proposition itself is an abstract Platonic entity that on standard Platonic views cannot causally interact with the truthmaker.

The fact that propositions cannot be actually true without being possible but can be possible without being actually true suggests that a proposition's having the property of actual truth is ontologically dependent on something related to what makes it have the property of possibility. But what makes it the case that a proposition is actually true, when the proposition has a truthmaker and is true for no other reason,[9] is a truthmaker. Now, plausibly, the truthmaker of <There are horses> is not dependent on abstract propositions and their properties; it is dependent only on its causes (such as some evolutionary process). Now, if the property of possibility is a primitive property of propositions, and hence if it is an abstract fact about something abstract that a proposition is possible, it is plausible that no concrete worldly entity such as a horse will *depend* on some proposition's having this property. But then how can actual truth of p depend on the possibility of p, given that the actual truth of p depends only on horses and their causes?

One place to attack in this argument is the claim that a proposition's being true should ontologically depend on its being possible. I brought this dependence in as an explanation of the fact that no true proposition can fail to be actual. Perhaps an alternate explanation could be given. But until it is given, the above argument gives one reason to doubt whether the Propositional Ersatzism is a satisfactory account of what modality consists in.

All of the difficulties I am now raising illustrate that the "basic property

---

9   Suppose that p is true because of a truthmaker a and q is true without a truthmaker. If r is the disjunction $(p$ or $q)$, then a probably should count as a truthmaker of r. But r is not true only because of a — it is also true because q is true.

of propositions" account of possibility fails to be an illuminating answer to Parmenidean worries. A yet different way of giving voice to this concern is as follows. Many ordinary statements of possibility, if they are statements about anything, seem to be statements about *concrete* aspects of this world. When I say Hitler might never have been born, I am making a claim about the contingencies in the reproductive process, the potentialities for the miscarriage of children, and the like. When I say Smith could have done something, I am speaking about Smith's own abilities. To claim that a proper analysis of these assertions will make them be ultimately statements about abstracta is to raise the worry: So why do we *care* about these statements, if they are merely about abstracta? That Smith could have refrained from an action can be normatively important — but only if it is a statement about concrete things, not if it is a statement about abstract properties of abstracta.[10] There is a force, then, that pulls us in the direction of thinking that assertions of possibility are assertions about concrete things in this world. But at the same time, that line of thought seems to run up against Parmenides's objection, leaving us in aporia.

Roy (1993) offered an account that attempts to explain what a Platonist can ground modality in, and this account will be a paradigm of Platonist attempts that do not make modal properties primitive. Roy thinks that properties have inner structure so that "for example, we might suppose BEING COLORED is BEING RED OR BEING BLUE OR ... and/or that BEING WATER has among its constituents BEING HYDROGEN and BEING OXYGEN" (1993: 342). Thus, necessarily, if BEING WATER is instantiated, likewise BEING HYDROGEN and BEING OXYGEN are instantiated, and necessarily anything that is red is colored.

One difficulty with Roy's view is with *de re* modality. Suppose that *x* has BEING WATER *essentially*. We want to say that *x* could not fail to contain hydrogen. On Roy's account it seems this is to be analyzed by pulling out another Platonic entity, *the essence of x*, which *x* essentially has but nothing else has, and saying this entity is structurally related in an appropriate way to BEING WATER, which in turn is related in another appropriate way to BEING HYDROGEN.

Now, *the essence of x* is some property F. But why, one can ask, is it the case that *x* has F *essentially*, as opposed to having it merely non-essentially? One is tempted to say that the fact that *x* has F essentially is evident if

---

10 The same objection applies to the account of Jubien (2009) in terms of property entailment. For it is puzzling why it is that just because *triangularity* and *polygonality* stand in relation E ("entailment"), then every triangle is a polygon.

we spell out what F is: We could just say that F is the property of BEING x — indeed that is just a paraphrase of "the essence of x." But this property appears to have no inner structure, and the structural account above seems to fail then.

In fact, we have a dilemma. Either the essence F has as one of its ingredients BEING WATER or it does not. If it does not, then perhaps F is something structureless like *the property of being x* and it is unmysterious why x has F essentially. But then Roy's account fails for the *de re* modal claim that x could not fail to be water. But if F does have BEING WATER as an ingredient, then presumably F is a property of the form: BEING WATER AND G, where G contains all the other ingredients of F. But now it has once again become quite mysterious why x has F essentially. Roy wants to explain all modalities in terms of structural facts about the Platonic realm. It is quite unclear how it is that having BEING WATER as an ingredient could explain F's being had by x essentially. Nor does it seem promising to say that G is some property of a sort that makes it clear why x has F essentially; for while it might be that having G would explain why x has G essentially (e.g. if G were the property OF BEING x), it is quite mysterious how it is that having G would explain why x has both G *and* BEING WATER essentially. And it is mysterious how putting both BEING WATER and G together and saying x has both would explain why x has these two properties *essentially*.

A further problem with accounting for modality by using ingredients of properties, or by using more general entailment relations between properties as Jubien (2009) does, is in accounting for the necessary existence of properties. We could suppose that BEING A PROPERTY has as an ingredient or entails EXISTING, but that would only show that anything that has the property of BEING A PROPERTY exists, which is compatible with all properties being contingent. After all, anything that has the property of BEING WATER exists, but water is not a necessary being.

Perhaps one could suppose that BEING A PROPERTY has as an ingredient or entails EXISTING NECESSARILY. But a serious difficulty here is that necessity was supposed to be analyzed in terms of ingredient or entailment relations between properties, and yet here a primitive modal property of necessary existence appears to be posited.

3.3.5 Actuality and possibility

Imagine, *per impossibile* according to the Platonist, that all and only the entities from the Platonic heaven disappeared. Then, so would all the facts which they ontologically ground. There would no longer be any mathematical truths, if mathematical entities are Platonic abstracta. Likewise, if the

Platonic heaven is what grounds possibility, literally no state of affairs would be possible any more.

But actually this is not so. For the things that are actual would surely remain possible. The things that are actual are thereby *automatically* also possible. There is no need to investigate any Platonic heaven to see the possibility of dogs — it is enough to see a dog. And, I submit, it is not the case that the dog's actuality serves as evidence for the Platonic fact that it is possible. Rather, the dog's actuality grounds its possibility. Should a committed, but sane, Platonist come to conclude that there are no Platonic entities and yet continue to believe that they are needed for grounding possibilities, she should not conclude that dogs are impossible.

Hence there is at least one class of states of affairs, namely the actual ones, whose possibility is not in need of a Platonic grounding. This is not a refutation of Platonism, however, because a proposition can have two or more truthmakers. Any dog makes it be true that there is a dog. Perhaps, then, the possibility of the actual is grounded both in the actuality of the actual and in the Platonic realm. This still leaves unexplained, however, the apparent coincidence: Why is it that never does something actual lack a possibilifier in the Platonic heaven? Moreover, it is strange that <There could be a dog> should have two radically ontologically different groundings: one grounding in any dog, and another in the Platonic realm. We might expect such a thing from a disjunctive proposition such <There is a horse or a number> which claim might be grounded both in Secretariat and in the number nine, but <There could be a dog> here does not seem to be such.

But perhaps instead of taking as basic the possible–impossible dichotomy, we should work with the actual–merely possible–impossible trichotomy? If we do that, then a claim that something is possible becomes a disjunctive claim: it is actual or merely possible. However, even though this is not an unattractive position, and in fact my final account will accept it, nonetheless it is not a position an Ambitious Platonist can accept.

For let us consider how to ground mere possibility. It would presumably be grounded in abstract facts about the Platonic realm. But the Platonic realm is there regardless of what actually transpires. But whether <There are dogs> has the property of mere possibility depends on whether there are dogs — it is merely possible if and only if there are no dogs. Thus, while we were working with the possible–impossible dichotomy as primitive we had the puzzle of why the actual was always possible, with the actual–merely possible–impossible trichotomy as primitive, it is puzzling why the merely possible is never actual.

There is, however, a more satisfying escape route for the Platonist. The

arguments above presupposed a positive ontological grounding for possibility in the Platonic realm. But perhaps claims of possibility do not have truthmakers. The true ones simply lack falsemakers, and it is the true claims of necessity that have truthmakers. A proposition is possible providing there is no falsemaker for the claim that it is possible, i.e. no truthmaker for the claim that it is necessarily false. And it is claims of necessity that are ontologically basic.

If this is right, then were there no Platonic heaven, everything would be possible because nothing would be necessary. In particular, the actual would remain possible.

But now we have the problem of the opposite thought experiment. Instead of imagining the Platonic heaven emptying, imagine that the proposition that there are no dogs had the property of necessity. Remember that this is just a fact about the Platonic heaven and should no more affect what happens under that heaven than a dog's acquiring the property of blueness should affect what goes on in the heaven. We are to imagine, thus, a change in the Platonic heaven, but not under it. But any such change notwithstanding, dogs will remain possible. They will remain possible because they will continue to exist. Hence the possibility of the existence of dogs does not flow from facts about the Platonic heaven.

To save the Platonic account from the problem that removing or adding modal properties in the Platonic heaven is conceptually independent of the possibility of existent dogs, we would have to make the existence of dogs depend on the Platonic facts about the possibility property being had by some entities in the Platonic realm. Now the Platonist already believes in some dependence of the concrete on the Platonic. Properties, including doghood, are Platonic entities, and without doghood there are no dogs. Thus we were wrong in the initial gedankenexperiment: if the Platonic heaven disappeared, so would dogs. But we did not need such a drastic destruction of the realm of abstracta to make our point. If we simply imagined modal properties disappearing, there would still be dogs.

Orthodox Platonism allows for one central dependence relation of the concrete on the abstract: participation. Now think about what would happen if only dogness were removed from the Platonic heaven. As long as everything else in the Platonic heaven were undisturbed, the Platonic heaven would still contain poodleness, though it would no longer entail dogness as there would be no such thing as dogness. Thus Fido might still exist as a poodle, if not as a dog. By analogy, if one only removes the modal property of possibility, Fido retains his non-modal properties, including existence.

But perhaps this is based on a "straw man" version of the Platonic realm,

where the various properties are self-standing independent things: here is poodleness, there is doghood, and yonder is possibility. But the ontology of properties may be more intricate. For instance, poodleness might essentially contain within itself doghood, so that, *per impossibile*, once doghood is removed, we are no longer dealing with the same property. By analogy, perhaps existence contains possible existence, and so once we remove possible existence, we no longer have the property of existence.

However, there is a disanalogy here if Kant is right that existence is not a property. For if it is not a property, then there is no reason to think it contains possible existence in the way one property contains another. But let us grant for the sake of argument that existence is a property. The claim now is that existence contains possible existence. But presumably it contains more than just possible existence. Just like poodlehood might be doghood conjoined with some subdividing property, what the Aristotelians called a "differentia," such as being curly haired and having a certain kind of DNA, so too it would seem that existence would be possible existence conjoined with some differentia. It is this differentia that makes for the difference between merely possible existence and possible existence that is actual. But while in the case of poodlehood and doghood, we can with some plausibility point to some property that could be a differentia, we are very hard-pressed to point out any property that is a differentia here. The only plausible candidate is actuality, but existence is not actuality *plus* possible existence, since saying of an object that it is actual just *is* saying that it exists. To say that existence is actuality plus possible existence is like saying that doghood is canineness plus mammality.

The philosophical method of this argument may seem unfamiliar and open to interpretation. Here is one fairly detailed interpretation. What we are doing is using intuitive, non-formalizable *per impossibile* counterfactuals to probe questions about whether some allegedly necessary fact is the ground for another necessary fact, in this case whether some fact about the Platonic realm is the ground for necessary truths of the form $\lozenge p$ (which are necessary whenever they are true, by S5). I am granting that perhaps there really necessarily are propositions, and that the possible ones really and necessarily have the abstract property of possibility, but questioning whether any fact of this sort can be the ultimate ground for possibility claims. The basic structure of the argument, then, is this:

(128)   That the actual is possible *holds no matter what, in the strongest sense of "no matter what."* (Premise)

(129)   Thus, if, *per impossibile*, the modal properties of

propositions residing in the Platonic realm disappeared or radically changed, then those non-modal propositions about concreta that would remain true would nonetheless be possible, and this *per impossibile* counterfactual is non-trivially true. (Premise, intuitively justified with reference to (128).)

(130)   But if *A*-type properties of entities in realm *B* ground *C*-type truths, then a *per impossibile* counterfactual of the form <were the *A*-type properties of entities in realm *B* removed, some *C*-type truth would still be true> would be false or at best only trivially true. (Premise, following from the notion of grounding.)

(131)   Hence, the modal properties of propositions residing in the Platonic realm are not what grounds the truth of propositions of the form *r is possible*, where *r* is actually true. (From (129) and (130).)

The argument is formulated to allow for two views of counterfactuals. On an orthodox Lewisian view, every counterfactual with necessarily false antecedent is trivially true. On that view, there should still be a distinction between those counterfactuals that are *merely trivially true*, i.e. true only because their antecedents are necessarily false, and those that are true for a deeper reason. For instance, <If horses were reptiles, our galaxy would be a star> is merely trivially true. But <If horses were reptiles, then some reptiles would have hooves> would, plausibly, be true for a reason over and beyond the necessary falsity of the antecedent. On a view of counterfactuals that takes the *per impossibile* ones more seriously, on the other hand, <If horses were reptiles, our galaxy would be a star> is simply false, while <If horses were reptiles, then some reptiles would have hooves> is true.

The "no matter what, in the strongest sense of 'no matter what'" seems to indicate a particularly strong kind of necessity, like a conceptual or strictly logical necessity. We have already critiqued the notion of this sort of necessity in Section 2 of Part I. However, that critique assumed that we were to take the notion to be one that is objective and applies to propositions. But we did observe that one can make sense of the notion as applied to sentences in a fixed language with fixed axioms and inference scheme. We can then apply this to the language and inferential structure that we humans find ourselves stuck with (which inferential structure must be taken to be correct if scepticism is to be avoided). Moreover, if we are willing to accept with the Less Radical Theory of Part V, below, that some propositions are necessary

in such a strong sense that they do not have negations, then we will also perhaps be able to make a useful distinction here between levels of necessity.

Now, if a proposition q is necessary with respect to some modality, and p is a proposition that is not necessary in the same sense, then plausibly the counterfactual <Were p not the case, q would still be the case> is true. Thus, to derive (129) from (128), we need the antecedents of *per impossibile* counterfactuals to be negations of claims that have a lower order of necessity than are had by the claims in (128). But this is actually plausible. For that the actual is possible is surely a conceptual truth in the strongest sense. But it does not seem to be a conceptual truth in the strongest sense that there is a Platonic realm at all, much less that certain facts about it hold.

A different way to take the argument against Platonism is to take it as trying to make the reader vividly see that for grounding the possibility of the actual, nothing beyond the actuality of the actual is needed. A parallel to this would be, say, an Aristotelian argument against divine command theory that holds that all one needs to ground the fact that the virtues are morally required is that they are *virtues*.

A final way to take the argument is epistemological. Were we to come to believe that there is no Platonic realm or that it lacks modal properties or that these are very differently arranged from how we thought, we would still not doubt that actual beings are possible. The epistemological version of the argument is the weakest, however. For it might be that the ontological order in reality just does not match the order of our thought.

To see the difficulty of the epistemological argument, consider a parallel argument against materialism. Were I to come to believe that there is no matter in the universe, I certainly would not conclude that my mind does not exist. Thus, my mind's existence cannot be grounded ontologically in the existence of some bunch of matter. I think it is clear that this is a weak argument, and that is why, some textual resemblances to the contrary, it is not given by Descartes, who uses the additional premise that my mind is transparent to myself.

Nonetheless, the simple anti-materialist argument is suggestive as it shows that at least in our conceptual scheme, materiality is not prior to mindedness. And this might well be enough undercut any conviction that we have a *conceptual* argument for our mind's existence being grounded in the existence of a bunch of matter. Moreover, even if the general argument form is not deductively valid, nonetheless we may have some hope that in privileged cases, at least, our conceptual scheme is likely to match the ontological order of reality, and hence the argument may offer some evidence at least.

All of the above considerations strongly suggest that if x is an actually

existing entity, then x itself is a truthmaker of the proposition that it is possible for x to exist. For if x is not a truthmaker of this proposition and the truthmaker is anything else, or if there is some truthmaker/falsemaker account involving things other than x, then arguments similar to the above apply. Likewise, if p is a true proposition, then whatever grounds the truth of p also thereby grounds the possibility of p.

3.3.6 Powers

An Aristotelian can argue that in fact there *are* capabilities and dispositions sufficient to ground the truth of at least *some* possibility claims. That I could have been a biologist is very plausibly made true by my capacities and dispositions and those of various persons and things in my environment. These capacities and dispositions are concrete real-worldly things, albeit ones having modal force. Hence, in fact, we do not need a Platonic realm to make at least some possibility claims true. Indeed, the facts about the Platonic realm — about propositions having or not having some primitive property — are interlopers here. Just as the statement that I could have been a biologist was not made true by what my Lewisian counterparts in other worlds do, so too it is not made true by abstract properties of Platonic abstracta. The common intuition behind both cases is that it is something in me and my concrete environment that makes the statement true.

This would create a major problem for the Platonic approach. On the Platonic approach, what makes it possible that I was a biologist is that the abstract entity <I am a biologist> in the Platonic heaven has the abstract property of possibility. But if there are concrete capacities and dispositions in the universe that are by themselves sufficient to make it possible that I was a biologist, we have two different ways of characterizing possibility: one is via concrete this-worldly Aristotelian properties of concreta which really do exist — the Platonist should not deny their existence — and the other is via some abstract Platonic primitive properties of abstracta. Moreover, anything that is possible on the Aristotelian grounds will be physically possible, and hence also logically possible, and hence possible on Platonist grounds, though perhaps not conversely. But now we can ask: Why is this so? Why is there this apparent coincidence that anything made possible by this-worldly powers and capacities and dispositions happens to correspond to a proposition in the Platonic realm that has the abstract property of possibility?

The problem of coincidence is much more acute when it is a matter of a coincidence between a realm of contingent beings and a realm of necessary beings. If there is a causal relation between the two, which orthodox Platonism rejects, there is no such issue. But without such a connection, the

issue can be acute. Somehow the things here know how to obey the Platonic entities, it seems, without any causal influence being propagated in between.

The Platonist can of course say that the proposition that something has an impossible property is itself impossible. But that will not help here as it is yet another fact about the Platonic realm, and so exactly the same question can be raised: Why is it that this impossible proposition is never in fact true of the changing, shadowy world of contingent things?

Nor will it do just to admit the correlation between the Platonic world and the world of concreta to be a brute fact, lacking in any explanation. For if the connection between what happens around us and what happens in the Platonic realm *just is* a coincidence, then Platonism does not provide an analysis of modal concepts. For surely our modal concepts are such that it is *not* a coincidence that impossible propositions are also false, and that nothing has a power to bring about an impossible effect.

3.4 *Avoiding grounding?*

However, some Platonists[11] are not in the business of providing an ontological grounding for modality. While they believe that propositions or properties exist, and that they have certain properties such as possibility, they do not take these properties to *ground* modality, because they reject the truthmaker principle or any similar demand for grounds.

Now, one cannot reject grounding in *all* cases. There plainly are propositions which are true in virtue of their truthmakers. That there exists a horse is true in virtue of Bucephalus — or Secretariat, for that matter. That this book is rectangular is true in virtue of the rectangular arrangement of its molecules.

The Platonist we are now interested in believes that although the proposition, p, that there are unicorns, say, has the property, P, of possibility, and although it is a conceptual truth that it is possible that there are unicorns if and only if p has P, nonetheless it is not in virtue of p having P that we say that it is possible that there are unicorns. Consider the following general claim of which this is a special case:

(132)   It is a conceptual truth that (q if and only if item a has property F), but nonetheless q does not hold in virtue of a having F.

---

11   For example, Merricks (2009) and Plantinga (personal communication).

There are cases where we do want to say this. For instance, it is a conceptual truth that Fred is intelligent if and only if Fred has the property of being such that were God to exist, God would believe Fred to be intelligent. This is a conceptual truth since it is a conceptual truth that God believes only true propositions. But Fred's subjunctive God-involving property may not be what grounds Fred's intelligence. Rather, Fred might have that subjunctive property precisely because he is intelligent.

One way, thus, of disputing an *in virtue of* claim where one agrees that there is a conceptually necessary relation is to say that the relation goes in the other direction. A different way that could be imagined would be if both propositions were true in virtue of something else. Fred is married to Jane if and only if Jane is married to Fred, and this is true of conceptual necessity. But both of the claims are true in virtue of Fred and Jane having made certain vows under appropriate circumstances and nothing having impeded the continuation of the force of these vows. A final way to dispute a grounding claim is just to say that both sides of the biconditional are actually synonymous. Interestingly, however, we do not dispute a grounding claim by saying that although the two sides of the biconditional are, of conceptual necessity, either both true or both false, nonetheless both sides *just* hold, not in virtue of anything.

It seems, thus, that if the Platonist is to dispute the claim that it is possible that there are unicorns holds in virtue of p having P, she will have to claim (a) that it is in virtue of its being possible that there are unicorns that p has P; or that (b) both hold in virtue of some third item; or that (c) the claims "It is possible that there are unicorns" and "p has P" come to the same thing.

In case (b), the search for an ontological ground for possibility can go on as the objection that there are no grounds has failed.

On the view (c) that the two claims come to the same, surely there *is* a grounding. For if F is an intrinsic property of x and x's having F is not itself further grounded or reducible, then surely the ground (and maybe even truthmaker) of $<x$ has $F>$ just is x's having F. Or so a Platonist should say.

So, can the Platonist say that that p has P is synonymous with its being possible that there are unicorns? On the face of it, no. For one can believe that it is possible that there are unicorns without believing that there are any Platonic entities. That there is a Platonic universe, if there is one, is a discovery that humankind made *after* discovering modality, no doubt, or at least it might have been so. On the synonymy view, the opponent of the Platonist is horribly confused: she does not understand everyday modal propositions. This is implausible.

The most promising solution for the Platonist is to take option (a). The abstract entity *p* has property *P* in virtue of its being possible that there are unicorns. This is problematic for classical Platonism on which participation in a property is primitive and not further analyzable into anything else. But perhaps our Platonist's view of the relationship between *p* and *P* and its being possible that there are unicorns is that it is like the sensible view of the relations between *Fred* and the property *being such that Napoleon lost at Waterloo* (or *being in a world at which Napoleon lost at Waterloo*) and the state of affairs of its being the case that Napoleon lost at Waterloo. Fred has the property of being such that Napoleon lost at Waterloo precisely in virtue of its being the case that Napoleon lost at Waterloo and that Fred exists. Fred's participation in *being such that Napoleon lost at Waterloo* is clearly not primitive. A sane Platonist will have to allow that in such cases, at least, participation in a property is not ontologically primitive but reduces to something else, and that something analogous could be said about possibility claims.

But I suspect that in fact the right way of understanding how it is that Fred has the property of being such that Napoleon lost at Waterloo is to deny that in this case there *is* any such property. Talk of Fred's having such a "property" is to be understood as a paraphrase of the claim that Napoleon lost at Waterloo and that Fred exists. This is a controversial claim, but can be justified as follows. If we say that it is true in virtue of some state of affairs *q* that *a* has *F*, and if *q* does not actually involve the attribution of any property to *a*, then we are saying that <*a* has *F*> is subject to a *reductive* analysis that does not ontologically commit us to the existence of *F*. But if <*a* has *F*> does not commit us to *F*, it is not plausible to think there is such a thing as *F*, since <*a* has *F*> is a paradigm case of a proposition about *F*, if there really is such a thing as *F*.

Let us illustrate this by a different example. Suppose that the proposition that a knife is hot is made true by certain molecules having high kinetic energy. In particular, the facts about these molecules are then also sufficient to ground the existence of the knife and heat, since that the knife is hot entails that the knife and heat exist. Thus, the knife and the heat exist only in a sense which reduces to the existence of something else, certain molecules. If one denies this, then one will presumably likewise deny that that a knife is hot is made true by certain molecules having high kinetic energy. For instance, if one is an anti-reductivist constitution-theorist about knives, one will say that what makes it true that a knife is hot is that "the molecules that constitute the knife have high kinetic energy." But now the grounding

proposition still mentions the knife, and the knife's own existence is not reduced away, though the heat is.

Thus, if <*p* has *P*> reduces to its being possible that there are unicorns, then in fact there really is no property *P* in the Platonic heaven. In other words, this way of denying that <It is possible that there are unicorns> is grounded in *p* having *P* requires our Platonist to deny the existence of a property of possibility.[12]

Can this Platonism at least maintain that there are propositions? Perhaps. But just as the property of possibility, and by the same token the property of necessity, no longer enter into the theory of modality in an ontologically committive way, neither need our story include propositions. We had allowed *p*'s having *P* to be grounded in its being possible that there are unicorns, and for exactly the reason the grounds do not ontologically involve the existence of a property of possibility, neither do they involve the proposition that there are unicorns, in both cases notwithstanding surface grammar. For if the proposition is involved, it is surely involved precisely through having something attributed to it, namely the property of possibility. Admittedly, our Platonist could turn into an off-and-on nominalist, and say that <*p* has *P*> commits one to the existence of *p* without committing one to the existence of *P*. But this will not do. For even the nominalist should say that <*p* has *P*> is a proposition that attributes something to *p*, while we should *not* say that the proposition that possibly there is a unicorn attributes anything to the proposition that there is a unicorn. For if it attributes something to it, it is possibility which it attributes to it. And then the claim that *p* has *P* becomes *synonymous* with the claim that possibly there are unicorns (since I take this to be just a rephrasing of the claim that it is possible that there are unicorns), and we have seen that that option is not available to the Platonist who believes that modal claims are ungrounded.

Thus, while such an Unambitious Platonist may believe in propositions, she should not bring them into the analysis of modal talk. Modal talk does not involve attributing anything to propositions, *if* the Platonic theory is not supposed to yield a grounding for modality. But while this Platonist should not bring propositions into the analysis of modal claims, she should go further. She should reject the real existence of the properties of *possibility* and *necessity* as applied to propositions. For the having of these properties would surely have some conceptually necessary relation to modal claims, such as that it is possible that there are unicorns, and unless it is a complete

---

12   For an illuminating discussion of reduction and ontological commitment, see Daniel Johnson (unpublished).

coincidence that (a) there can be unicorns; and (b) the proposition <There are unicorns> has the property of possibility (or that its negation lacks the property of necessity), surely (b) would, if true, ground (a), if (b) did not in some way reduce to (a).

Now, a Platonism on which *possibility* and *necessity* do not exist may not be a particularly popular Platonism but it is not obviously incoherent. It is a particular species of sparse Platonism.

What this leaves is a story on which it *just is true* that possibly there are unicorns. There is no truthmaker. No claim is made about some proposition having some property. There is no further analysis. It's just that it's possible that there are unicorns. That is all: this is rock bottom.

Such a story should, I take it, be a last philosophical resort. Euthyphro could just say: "Some actions are pious and others are impious, and that is all there is to it." But it would be unphilosophical to say this *if* there were a more informative account of the grounding of piety attributions. Since we shall see that there *is* a more informative account of modality, we need not take this desperate resort.

Section 4 Conclusions

Linguistic Ersatzism fails even in an Unambitious form unless the language is really expansive, for instance having Platonic properties as linguistic components. The main objection against standard linguistic ersatzism is that it cannot take account of alien properties. This does not apply to Sider's version, but Sider's version does not yield bivalence.

The main objections against Unambitious Platonic Propositional Ersatzism seem to fail. That propositions are abstract is no reason to disbelieve in them. That it is not clear *how* they represent the states of affairs they represent is also no objection, just as it is no objection to various scientific views that they fail to explain how it is that certain objects "manage" to have properties like mass.

However, Ambitious Propositional Ersatzism does not succeed, because it fails to give a satisfactory account of what we are talking about when we are making modal claims. On one alternative, it is completely unilluminating in connection with the grounding problem, simply moving the bump in the carpet to the problem of what it is about a proposition that makes it have the property of "possibility" or "necessity," with this being exactly the original problem in different terminology. On the second alternative, if our ersatzist more ambitiously insists that *possibility* is a basic unanalyzable property of

abstracta, the account fails to explain how *this* particular basic property has anything to do with our ordinary claims about possibility, what its connection to actuality is, and how it is relevant to concrete facts about the world.

PART V

SKETCHES TOWARDS A SPINOZISTIC–TRACTARIAN ACCOUNT OF MODALITY

There is an account of modality that has received little attention in the contemporary literature, even though it gives an elegant answer to the ontological grounding problem, without adding anything to the ontology. It does this by *subtracting* items, namely impossible propositions, from our ontology. This account is inspired by Spinoza's *Ethics* and Wittgenstein's *Tractatus*, though both philosophers go further than is needed for this account. The view is highly counterintuitive, and we will need to develop it in greater detail to see why it in fact has some plausibility. We shall see that the view does not work as a full solution to the metaphysical problems of modality, but it may work in a more limited way. At several points in Part VI we shall see that invoking a part of the Spinozistic–Tractarian view can help solve difficulties with the Aristotelian–Leibnizian view of modality.

Section 1 Asserting, naming, and infallibility

In Plato's *Theaetetus* we find a discussion of the problem of intentionality couched in terms of a debate about how error could arise. The difficulty stems from the fact that an erroneous claim needs to have some *meaning*, needs to be about some worldly state of affairs, but since it is erroneous, that state of affairs does not in fact exist. How can a claim be about something that does not exist? Does it stand in relation to something non-existent, in a mysterious one-sided relation? In other words: if claims are about their truthmakers, a false claim, lacking a truthmaker, is not about anything.

The problem arises on a pre-philosophical level in the Greek language. Where we say that a claim is *true*, the Greek might just say that the claim *is*. A semantic theory behind what might otherwise seem like a linguistic oddity has apparently deeply influenced Plato's thinking in the *Theaetetus*. This theory states that a sentence is a *name* of something worldly. The sentence

is true provided the named object *exists*. Presumably, the kind of object we are talking about here is a state of affairs — so-and-so's being thus-and-such, say. These are Armstrong-style rather than Plantinga-style states of affairs: they exist if and only if they obtain, rather than being abstract and necessarily existing Platonic objects.

Thus, when I am asserting something, I am speaking the name of a state of affairs; I am verbally pointing out this state. In the Ugaritic language it was possible just to say "The dwelling of 'Il [*mt̠b il*]" and that would count as an assertion that 'Il's dwelling exists or is present (Gordon 1965: 111). The idea here is that all assertions are like the Ugaritic "The dwelling of 'Il."

Now this saying-is-naming semantic theory is all fine and good until we get to false statements. We know how difficult it is to come up with a semantics for non-referring names. But on the theory, a false statement names a non-existent concrete state of affairs. But there are no non-existence concrete states of affairs. Hence the false statement does not name anything. But a statement is on the saying-is-naming theory a name. Thus, a statement that fails to name anything appears to be a contradiction in terms. A string of noises that fails to name anything is not a statement, but a string of noises. "Mimsy were the borogroves" is not a statement despite the appearance of grammaticalness, and according to the theory neither is "Socrates was a warrior-king."

What we are looking at, then, is a theory on which the inscriptions:

(133) "Mimsy were the borogroves."
(134) "Colorless green ideas sleep furiously."
(135) "Akjlfjiudfoiwefaljksjlk
       dk𝑒𝑟●·𝑒𝑟&;●♒⚹𝑒𝑟●☺⚹sdiwealiflij."
(136) "Elvis is alive."
(137) "Napoleon won a victory at Waterloo."
(138) "5+7=11."

are semantically on par. Not one of them has any meaning.

But this seems absurd. For surely the person who says Napoleon won a victory at Waterloo commits a different kind of error from that committed by the person who says that 5+7=11, and both are in a very different state from that of the person who inscribes one of the first three examples. Furthermore, the person who *denies* that Napoleon won a victory at Waterloo speaks truly. But the person who denies one of the first three examples speaks nonsense. One can only deny a meaningful sentence. One cannot deny a nonsensical inscription.

In the *Sophist*, Plato famously abandons the theory on which there is no error in favor of introducing a grammatico-logical distinction between *subject* and *predicate* and defining error in terms of the mismatch of the two. A theory on which there is no such thing as error is not likely to be successful, we might think. And yet a philosophical line of thought on which at least certain kinds of thinkings are innately immune from error reappears in Aristotle and Aquinas, and culminates in Spinoza who thinks that *all* thinkings are immune from error, having come full-circle back to the pre-*Sophist* semantics.

In Aristotle and Aquinas, the immunity from error occurs at the level of the knowledge of essence. To know the essence of a horse is to become a horse, though not in the material mode of being a horse but *in esse intentionale*. If one has the essence of the horse in one's mind, then one's contemplating of this essence is without error. But if one does not have the essence of the horse in one's mind, then one is not making a mistake about horses. For to make a mistake about the essence of the horse requires that one should *refer* to this essence, and that in turn requires the presence of that essence. You either see the essence of the horse or you see something else, and in neither case are you making a mistake about the essence of the horse.

In the Aristotelians this special kind of knowledge that is immune from error is knowledge of only certain kinds of facts — necessary facts about essences. But more to our point, Thomas Aquinas (1948: I, 25, 3) further claims that an apparent assertion of a necessary falsehood is not a "word." He does this in the context of defending the compatibility of God's being able to anything (literally "any word") with God's inability to do something self-contradictory. The self-contradictory item is not a thing, not a word, and hence God's inability to do "it" does not limit God.

Section 2 Spinoza

Neither Aristotle, nor Aquinas, dreams of making the theory apply to *all* believings. That needed the mad genius of Spinoza's *Ethics* and *Treatise on the Emendation of the Intellect*. On Spinoza's dual-aspect theory, all of reality can be seen as *extension* or as *idea*. All of reality falls under both of the attributes of Extension and Thought. Now, each idea is an item under the attribute of Thought that corresponds to an *ideatum*, which is that which it is *of*. The *ideatum* of my idea of x is nothing but x itself, except that if x is material then the *ideatum* is considered under the attribute of Extension while my idea is under the attribute of Thought. Every idea then has an

ideatum, and this is just a simple consequence of Spinoza's metaphysics, since every idea is something under the attribute of Extension which then must be identical to something under the attribute of Thought — and that is its *ideatum.* Something that lacks an *ideatum* is then a non-idea, a non-thought. Every thought, then, is true, since it has an *ideatum.*

Spinoza then needs to give us a rational reconstruction of what is happening in those everyday cases where we are apt to say that someone is thinking falsely. There is only one thing Spinoza can say. Sometimes it is as if we were thinking *p* when in fact we are not thinking *p.* In those cases, we are either not thinking anything or we are thinking something other than *p.* It is obvious that there is nothing else that Spinoza can say, given his basic assumption of the coextensiveness of Extension and Thought and his account of intentionality in terms of identity. The question is whether he can make the story plausible.

One way to flesh out Spinoza's position is that sometimes when we *think* we are thinking that *p,* we are not in fact thinking that *p.* That would be unhelpful: it would just push the problem back, from an error about who the dark figure in the fog is to an error about our minds. If this is the only way to flesh out Spinoza's position, he suffers from a regress problem. Let us bracket this for the moment.

To help make the story plausible, Spinoza brings in the central concept of *confusion.* When it seems that I think there is a broken stick in the water — this is of course the sceptics' famous stick stuck in the water and appearing broken — I am confusing together two different ideas. I am confusing the idea of my retinal imprint, which we now know to consist of a combination of electrical states, with the idea of the stick. Both ideas are true, but I confuse them together. I project the idea of the broken stick onto the external world. But what I am *really* thinking of are two true ideas, that of the retinal imprint and that of the stick, though I possess them in an incomplete manner. The *ideata* of both ideas exist: there is a retinal imprint and there is a stick.

What if we have a more radical case? I have drunk much and now I think there is a purple tiger in the corner. Unlike in the stick case, I am not just wrong about the tiger's properties but about the tiger's existence. But there is still a blip in my nervous system that corresponds to my hallucinatory image of the purple tiger. This blip is something "in the imagination," Spinoza insists, and imagination is something physical. Spinoza theorizes it is something in the back of the eye. Nowadays, we are more likely to think that in ordinary hallucinations it is something in the brain. It does not matter: it is some physical event in me. My idea, though seemingly of a purple tiger, is in fact a true idea of that physical event in me. But I have the idea quite

inadequately. I am ignorant of many aspects of the idea I have, such as of its being an idea of something internal to me.

All error is just ignorance and confusion for Spinoza. We can now come back to the problem that had been put aside. Does not the account simply shift error further back? Instead of my being wrong in thinking there is a purple tiger, I am wrong in thinking the purple tiger image is an image of something external. But I am still *wrong*. When I confuse the stick in the world with the stick image in the mind, am I not erroneously thinking that they are the same? Do I not then have an idea of the identity of the stick image with the stick, a false idea?

Spinoza's own answer here seems troubling at first sight. He appears willing to admit that indeed you wrongly think the broken stick to be outside the mind, which leads to the above problems. In fact, it appears he has a metaphysics of thought on which by default an idea is associated with an image from the imagination. The imagination is the faculty of images, some of which come from the outside world by the operation of the senses and some of which are internally produced. Whether an image in fact ultimately originates in the external world or is produced by the imagination (a "figment of the imagination"), it seems to be an idea of an external thing, and one would need to add a *qualification* to that idea in order to make it adequately refer only to the blip in the brain or the event at the back of the eye that constitutes the image as found in the imagination. Like Wilfrid Sellars (1997, Sections 14–18) in his Myth of Jones, Spinoza here appears to take it that being is prior to seeming. According to Sellars, we first have the concept of a tiger visibly existing in external reality, and then we *add* to this concept in order to form the concept of the appearance of the tiger: the appearance of the tiger is nothing but its being *as if* there were a tiger visually presented to one.

It seems that according to Spinoza, when I confuse together the ideas of the stick image and the stick in the world, what is happening is this. Normally my possession of two ideas counts as a thought that these two ideas are ideas of one and the same thing, unless the ideas are qualified through something that distinguishes them as ideas of different things. If in my mind there occurs an unqualified idea of A and an unqualified idea of B, and if these two ideas are such that it is logically possible for one thing to instantiate both, then this pair of occurrences constitutes me as believing that there is an item which is both A and B. For the occurrence of these two ideas to constitute me as believing that a *different* item is an A from the one that is a B, I would need to have these ideas be logically incompatible, perhaps through a qualification built into A that states that the A is not a B.

And to believe the weaker claim that some item is an *A* and some item is a *B*, without any commitment whether these are the same or different items, would require a different sort of qualification. Spinoza does not say what that qualification is, but perhaps it involves tacking onto *A* the qualifier "which is perhaps not a *B*."

This account lets Spinoza save a notion dear to him, namely that error is always metaphysically a lack. My error in thinking that there is a purple tiger visually presenting itself to me is constituted by the *lack* of a qualifier in my mind — the "as if" qualifier. Unfortunately, if we take this to be the basic account of error in Spinoza, we get a system that is inconsistent both with what Spinoza says and with the notion that ideas cannot be wrong. Spinoza insists that in fact *all* our knowledge of macroscopic objects in the external world is tainted by our perception. We know the objects only insofar as they affect our bodies, including our senses. Thus our senses, as in Descartes, have a certain priority. It seems thus that Spinoza does *not* give the Sellarsian account.

Moreover, on this account my idea of the purple tiger *is* wrong. It is wrong through a lack: a component of my thinking that the purple tiger exists *imaginarily* is missing, and without this component I am constituted as merely thinking that the purple tiger exists. Moreover, the lack here is one of a different sort from that which ordinary ignorance involves. Compare two assertions:

(139) Napoleon was a great general.
(140) Napoleon was a great general and did win at Waterloo.

Now, the first assertion, if it represents the sum total of someone's beliefs about Napoleon, is lacking: the agent does not know the highly relevant fact that Napoleon was eventually vanquished. Consider now the second assertion. It too is lacking: it is lacking the word "not" between "did" and "win." But this sort of lack is a very different kind of lack. It is a linguistic lack, but one that does not correspond to a *conceptual* lack. Likewise, the Spinozistic account discussed above involves a quasi-linguistic lack, where we are talking about the language of ideas. A qualifier is missing. But yet the idea, given that the qualifier is missing, seems to be false.

The above is an account, thus, that betrays Spinoza's basic intuitions, though Spinoza's text occasionally may give the impression that it is his view. There is a better way to read Spinoza, and this is by using the notion of a *pragmatic* confusion or error, a notion for which I am grateful to James Conant in the context of his discussion of Frege and the early Wittgenstein.

Take me in my hallucinatory state. I have an idea of a purple tiger. This really *is* an idea of something in my imagination. In having this idea what I am believing just is that there is a purple tiger image in my imagination. However, I *act*, both linguistically and practically, as if the purple tiger image were not merely in my imagination. But I do not have any belief to the effect that the purple tiger really exists in the external world, even if I utter the words "There is a purple tiger here in the external world." It is a basic Spinozistic contention that grammatically correct strings of words need not convey *any* meaning, much less the "meaning" that as it were they seem to convey. Thus, it is quite open to Spinoza to claim that when I utter these words, I am really claiming that there is a purple tiger in my imagination — or if my confusion is bad enough, not claiming anything at all.

At the same time, Spinoza needs to give us an explanation of our behavior. I act exactly as a person who really believes a purple tiger to be present would act. It is plausible that the same thing explains the activity of both me and the non-hallucinator. Traditionally, we would say that this common explanans is the belief that a purple tiger is present, a belief that in my case is false and in her case true. But Spinoza can make a radical move here. What explains the activity of both me and the non-hallucinator is the belief that one has a vivid image of a purple tiger in the imagination (remember that images in the imagination can be figments of the imagination, but they can also be received by the imagination from the senses). And the natural causal consequent of having such a belief is making the noise "There is a purple tiger here in the external world!" and doubtless running or sneaking away in fear.

There seems to be something perverse here. After all, surely, in the non-hallucinatory case at least my *reason* for running away is not that I have an *image* of a purple tiger but that there really *is* a purple tiger there. Jonathan Dancy (1993) would insist on this criticism. A mere image of a purple tiger is no reason to run away, except insofar as such an image might be correlated with the actual presence of a purple tiger. However, it seems a very natural evolutionary adaptation to run instinctively as soon as an image of a tiger — of whatever hue — should occur in one's imagination. There is no time to check whether the image is *merely* imaginary or whether it arrived in the imagination from the eyes — remember that Spinoza's imagination can both generate images as well as be receptive of visual images. It is only when we already know that the image is a *mere* image, maybe because we remember mentally painting it, that we have reason to stay still.

Now it is true that in this case we do not want to say that the tiger image was the *reason* for running. It was the *cause* of running, and surely we should

not confuse causes with reasons. Indeed, on this evolutionary account, we act in a way that is mechanistically determined rather than freely rational. This, however, will not worry a Spinoza who explicitly affirms that we have no free will. Moreover, Spinoza can insist that there *is* a reason for one's activity. Reasons and actions operate in parallel under the respective attributes of Thought and Extension. My having the *idea* of the tiger image causes me to have the *idea* of my running and is a conclusive reason for it. The *idea* of the tiger image is something under the attribute of Thought that is identical with a blip in the brain or event in the back of the eye under the attribute of Extension. That blip or event then *causes* my running, which is an event under the attribute of Extension identical with the idea of my running. Conclusive reasons under the attribute of Thought then are identical with deterministic causes under the attribute of Extension.

In fact, Spinoza can go even further. If all you and I have is the image of the purple tiger, then even if your image came from external reality and mine was a figment of my imagination, we both are thinking the very same thought: *There is a purple-tiger image*. Appearances to the contrary, you do *not* have the idea of a purple tiger simply by virtue of your image's having come from a purple tiger. As long as your belief has exactly the same subjective warrant as mine, it is the same belief. Fully to gain the belief that there is a purple tiger in reality would require further work on your part — it would, perhaps, require scientific knowledge of your sensory system and of the way it receives information from the external world, which is why Spinoza insists on the importance of scientific investigation for knowledge.

Recall now that I suggested that according to Spinoza *by default* an idea of a purple-tiger image, if not further qualified, constitutes one as thinking that there is a purple tiger in reality. That can't be right, at least not if we are to make Spinoza self-consistent. But there is something to be recovered from this, as we indeed need to recover it to make sense of a number of Spinoza's claims. What is to be recovered is that ideas of purple-tiger images, when they are not adequately qualified as purely internal phenomena, make one *pragmatically* act the way the *hoi polloi* do when presented with tigers, namely tremble, run and exclaim "Tiger!," and more specifically the way the *hoi polloi* would when presented with a *purple* tiger, say by exclaiming "What a weirdly colored tiger!" One does not thereby *think* that one is presented with a purple tiger, but one *acts* in that way. It takes a further qualification of one's idea, say in the way that the word "merely" can qualify the word "image," to make it be an idea that makes one act as if one were presented with a mere mental image.

It is very difficult to talk in a way consistent with Spinoza's theory. Take

the sentence I wrote earlier: "It is a basic Spinozistic contention that grammatically correct strings of words need not convey *any* meaning, much less the 'meaning' that as it were they seem to convey." There is a reason why the second occurrence of "meaning" has to be in scare quotes and why there is an "as it were." Consider the sentence, as uttered by the hallucinator: "There is a purple tiger here." We want to say: "The hallucinator is acting and speaking as if he believed a purple tiger to be present." But what is the content of the belief that the hallucinator acts as if he had it? The content of a belief is a proposition or, in Spinoza's setting, an idea. But on Spinoza's account there are no false ideas, and hence no false proposition. Thus, since there is no purple tiger present to the hallucinator, the phrase "acting as if he believed a purple tiger to be present" is meaningless. We cannot consistently say that there is a meaning which the hallucinator is trying to convey by his words "There is a purple tiger here" that is other than that there is a purple-tiger image in his head. For that meaning would be the proposition that there is a purple tiger here, while *ex hypothesi* there is no such proposition.

Thus, in fact, what we should say is that the hallucinator is in the throes of a pragmatic confusion. He is acting the way people do when they are speaking of the presence of surprisingly colored animals. This much we *can* say because presumably people sometimes *are* in the presence of surprisingly colored animals. He is speaking using noises similar to those made by people who speak about surprisingly colored animals in their environment. But we cannot consistently say that there is some meaning that he is aiming at which he fails to convey. His words do not *deceive* as to the meaning: there cannot be cognitive error on Spinoza's view. The meaning of his "There is a purple tiger here" utterance, if it even has one, is that he in fact has a purple-tiger image. This is the only meaning that he could have intended to convey by the words, because it is the only meaning there *is* in the vicinity. Admittedly, the words "There is a purple tiger here" are rather deficient at conveying this.

It may well be that upon hearing the speaker's words you will fail to grasp the meaning. But this, too, Spinoza will say is not a cognitive error. You will simply be *acting* confusedly, and *failing* to have an idea of the speaker's meaning, rather than having a false idea of that meaning. Probably you find yourself uttering the noise: "He told me that there is a purple tiger here." This noise may or may not be meaningful. But if it has a meaning, the meaning is something like: "I have an auditory image of his uttering the words 'There is a purple tiger here'" or, in a more favorable case: "He uttered the words 'There is a purple tiger here.'"

Section 3 A radical theory of modality

The above explanation of how to make sense of a view that there is no error rests on three ingredients:

> (141) A story to the effect that by default we act in certain ways upon being convinced of the presence of an image of the imagination, as long as the image is not qualified by something like "merely."
>
> (142) The notion that all error is pragmatic.
>
> (143) A crucial dichotomy between language, considered as a bunch of noises, and thought.

It is essential that the notion of pragmatic error not become cognitive in some sneaky way. Spinoza can explicate pragmatic error in a simple causal way: pragmatic error is activity that does not tend to contribute (or maybe just: does not contribute) to human happiness.

But the account is implausible. Assume the mental supervenes on the physical (Spinoza certainly thinks so). Suppose you and I have brains that are precisely in the same internal state, but with you perceiving a purple tiger and me hallucinating it. I do not, on Spinoza's view, have a belief about the presence of a purple tiger. But surely you do, though Spinoza might deny it (he is apt to say, most implausibly, that to have a belief about the presence of a purple tiger you would have to have made, say, a scientific investigation of your sensory apparatus and verified the correctness of its functioning). But then we are left with an implausibly radical externalism on which whether you believe there is a purple tiger or not *always* depends on whether your mental states were caused by a purple tiger or not.[1]

Now this argument against Spinoza only worked because it is possible for you to be right about the presence of a purple tiger and for me to be wrong, since the term "presence" is indexical here. A similar argument can be made in the case of non-indexical contingent truths. Suppose that I am in the sort of state we would ordinarily describe as falsely believing that there exists (timelessly) a purple tiger. Of course I do not *really* believe this if Spinoza is right, but only act in a certain confused manner. Then, we can ask counterfactually: Were I to have the same brain state but a purple tiger

---

1 Moderate contemporary externalists who think that you only have the concept of a tiger if *sometimes* your tiger-thoughts are caused by tigers will not accept this radical version.

really existed, would I have a belief in virtue of that state about the external existence of a big predator? Surely the answer is positive (*pace* Spinoza), but this again leads to an implausibly radical externalism.

But the refutation of Spinoza by consideration of duplicates of brain states is specialized to the case of contingent or indexical truths. Thus, prima facie, Spinoza's view might be defensible in the case of necessary truths. For if p is a necessary truth, then we cannot imagine someone who has a brain state like that of a believer in p, but who is in a world at which p is false.

What we thus come back to is the more limited view of Aquinas that there can't be a necessarily false proposition (assuming "proposition" is what he means by "word"). This view, as extended by the addition of a pragmatic theory of error much as the one that I offered to Spinoza above, is found in an even stronger form in Wittgenstein's *Tractatus*, at least on one reading of this work. Observe, too, that the historical Spinoza will not be troubled by a restriction of his theory to non-contingent propositions, since in fact he thought that *no* proposition was contingent. This fatalistic view we will not follow him in, but we have learned *something* from Spinoza. We now have before us the notion that sometimes seemingly sentential noises either are not sentences at all and have no meaning or are sentences with a meaning that does not sit well with the form of the words used. We can strengthen our conviction in this notice by considering the wide variety of Liar paradoxes. "This sentence is not true" is just as grammatical as "This sentence is short," but the former cannot express a proposition while the latter is true.

So now we have the Radical Theory of Possibility: *every* proposition is possible. It follows that a proposition p is contingent if and only if it has a negation $\sim p$.

It is a strange idea that there might be propositions that have no negations, and Wittgenstein's *Tractatus* takes the Radical Theory and goes one step further by excising the necessarily true propositions, and hence yields Even More Radical Theory that every proposition is *contingent*. But while this lets Wittgenstein avoid the seemingly absurd conclusion that we might have a proposition that has no negation, it still has the equally implausible consequence that there are propositions that have no disjunction, say <There are horses> and <There are no horses> (if these had a disjunction, the disjunction would be a necessary truth). Since the existence of a disjunction for every pair of propositions is not particularly less plausible than the existence of a negation for every proposition, there does not appear to be a particular advantage to the Even More Radical Theory. If one can bite the bullet and say that some pairs of propositions have no disjunction, one can bite the bullet and say that some propositions have no negation, and additionally

have the advantage that one does not need to deny the existence of necessary truths. Moreover, Wittgenstein's account in terms of atomic propositions is a kind of combinatorialism and hence subject to the alien-properties objection (cf. Section 2 of Part IV). So I will resist the Even More Radical Wittgensteinian step.

We now have on the table a Radical Theory that seems to render quite unnecessary the whole game of explaining the grounds of possibility. The statement that a proposition p is possible is trivial, like the statement that a bachelor x is unmarried. Just as no further grounds are needed for an explicit tautology of the form an A is an A, we need no further grounds for the claim that a proposition is possible, since possibility is just propositionhood. If we wish to save a truthmaker theory, we just need to say that a tautology asserting that an A is an A is made true by that very A itself, or perhaps by A *qua* A. What makes it true, then, that a proposition p is possible is that very proposition.

It is a mainstay of the Spinozistic–Tractarian views of modality that:

(144)   "There is a square circle"

is just as much nonsense as:

(145)   "Mimsy were the borogoves"

or

(146)   "Xkjfdkj ûüØ2134 ☺✿√‹ frog."

If we do not opt for the Wittgensteinian More Radical Theory, we will, however, be able to say that it is not the case that there is a square circle. This seems to make for a distinction between (144) and (145), since "It is not the case that mimsy were the borogroves" and "It is not the case that xkjfdkj ûüØ2134 ☺✿√‹ frog" are *not* sentences.

It is tempting to say that when we deny that there is a square circle, we understand what we are denying, while when we deny (145) or (146), we do not understand what we are denying. But on the Spinozistic–Tractarian views, in none of the three cases is there a proposition we are denying. However, there *is* a difference that the Spinozistic–Tractarian theorist can acknowledge. The difference is in our interlocutor's attitudes when the interlocutor produces one of the noises or inscriptions. We have some idea of what normally it would be like to try to sincerely "assert (144)." It

would be to make the relevant noise with a desire to speak about circles and geometry, in a context and with intentions that would make one apt not to turn away from squares when looking for a circle, and that would make one apt to assert such further sentences as "There is a circle which has several straight sides," and so on. Our knowledge of this is a knowledge we have of how verbiage gets used and misused by human beings. While there is no proposition that our interlocutor would be attempting to convey through any of the three non-sentences, we have *some* idea of what she would be doing by trying to assert (144), and we would be able to say that she has failed in that. Moreover, we might somewhat confusingly (or confusedly) express that failure by a denial of (144).

However, in the case of (145) and (146), we really do *not* have any idea of what the speaker or writer is trying to accomplish. In fact, it seems likely that she is not even *trying* to make an assertion, as she would most likely be trying to do if she uttered (144) with apparent sincerity. The denial of an asserted proposition is a central sense of "denial." But we might extend the notion of a denial to pseudo-assertions, i.e. noisy or inscriptive acts that are an attempt to make an assertion of a sort that cannot possibly be made. To extend it thus is quite natural, and so we might say: "I deny that!," after someone uttered (144). But we draw the line at talking about denials (at least in a sense related to the relevant one) in the case of activities that are not even attempts at assertion, such as scratching, swimming, or inscribing (146).

Of course, if we did have evidence that our interlocutor was making an attempt at an assertion while producing (145) and (146), then things might be different. But we then would not have enough of a grasp of what she wants to assert about to be able either to deny or affirm. For maybe in fact she *does* mean something sensible by these. Thus, in some way we might extend more charity to the producer of (145) and (146), since she could be speaking a language unknown to us, while the producer of (144) is trying, it seems, to say something in ordinary English, and in doing so is violating the norms of ordinary English.

Let us explore the Spinozistic–Tractarian theory further. It is a commonplace that given meaningful sentences s and u, we can form further meaningful sentences such as:

(147) $s \,\&\, u$
(148) s or u
(149) $\sim s.$

But on the Radical Theory, if $<s>$ is a contingent truth, then "$u \,\&\, \sim u$"

does not express a proposition. If <*u*> is a necessary truth, then there is no proposition expressed by "~*p*." However, unless we take Wittgenstein's More Radical variant, "*s* or *u*" will be meaningful if *s* and *u* are.

But the Radical Theory does not in fact end the problem of explaining the grounds of possibility. Unless we take Wittgenstein's variant, we have a crucial distinction we can make. Some propositions *p* have the property of being *contingent* and others of being *necessary*. The contingent ones are the ones which have negations. But our old problem of grounds is now returning. What is it about a proposition that makes it be the case that it *has* a negation? So the Radical Theory does seem to entirely solve the problem that has been haunting us, though it may be seen as reducing it to another problem, that of what grounds the relation of being-the-negation-of between propositions.

It is worth noting that the Radical Theory has an additional advantage over the *Platonic* account, which the Aristotelian–Leibnizian theory of Part VI will also share. As I argued in Section 3.3.5 of Part IV, the Platonic story does not make clear why it is that all actually true propositions must also be *possibly* true. Why does the world of concreta cooperate so nicely with the Platonic heaven? But on the Radical Theory there is no mystery about the possibility of the actually true. If a proposition is true, then it is a proposition, and hence it is possible.

Section 4 Costs

It would indeed be the simplest theory of possibility to claim that all propositions are possible. But this simplicity has a cost. The first objection we have already touched upon: Logical connectives do not always apply. The second objection is related. Consider a proposition like <~(5+7=11)>. Wittgenstein's theory rejects this as a proposition, but the Radical Theory accepts it. This leads to the following refutation of the Radical Theory:

(150) I grasp the proposition <~(5+7=11)>.
(151) To grasp a negation of *p*, one must grasp *p*.
(152) Therefore <5+7=11> is graspable.
(153) Only propositions are graspable.
(154) But <5+7=11> is not a proposition if the Radical Theory holds.
(155) Therefore, the Radical Theory is false.

Now, the Radical Theorist could introduce a range of pseudopropositions that can be understood, e.g. by saying that anything truthfunctionally composed of propositions, whether it is a proposition or not, counts as a pseudoproposition, and deny (153) on the grounds that <5+7=11> is a graspable pseudoproposition meant by the sentence or pseudosentence "5+7=11." But this destroys the advantage of the Radical Theory, since the question of what it is that makes something be a proposition rather than a pseudoproposition is then just a version of the question of what makes something possible rather than impossible. This will not do.

However, there is a better solution. There is something deceptive about the notation "<~(5+7=11)>." It makes a proposition seem to be a string of symbols with a certain structure. But propositions are abstract objects. Given the Radical Theory, we should abandon the idea that the proposition that <~(5+7=11)> is somehow constructed by applying an operation, *negation*, to an entity <5+7=11>, though the *sentence* "~(5+7=11)" that expresses the proposition is indeed constructed out of "5+7=11" using a negation functor. (There is a further question whether we should call "5+7=11" a meaningless sentence or deny that it is a sentence.) Recall the central Tractarian idea that words only have meanings in sentences. The string "5+7=11" is like the word "sake" — as in "do this for my *sake*" — in that its "meaning" is derivative from the meaning of texts containing it. The string "5+7=11" is not a meaningful sentence on its own. It receives a crucial part of its meaning from the sentence "~(5+7=11)," rather than the other way around. I say only "a part" because the string is embedded in a number of other meaningful sentences, all of which contribute to its meaning.

It is an oddity of the Radical Theory, then, that if *s* and *t* are meaningful sentences, then "*s* ⊃ *t*" need not be a sentence, and likewise it does not follow from "*s* ⊃ *t*" being a meaningful sentence that *s* and *t* are. The apparently compositional linguistic forms of sentences do not reflect a compositional logic of propositions. The *Tractatus* sought to remedy this by creating a language where there would be such a reflection. But the Radical Theorist might not opt for such an ambitious project. Instead, she simply might simply gently remind us that propositions are abstract entities and we should not confuse them with their linguistic representations. Thus, <~(5+7=11)> no more has any object "like <5+7=11>" as a component than an elephant has any object "like a phant" as a component.

Now consider this question: "What does it mean to say that circles are square?" We would not be surprised to be told: "Surely you know what it means to say that circles are square! After all, surely you will agree that circles are *not* square, and when you say that circles are *not* square, you

are simply denying that which is asserted when one says that circles *are* square." This answer is both a challenge and a help to the Radical Theorist. It is a challenge because it insists that there is a meaning to the sentence "Circles are square." It is a help because it suggests that linguistic form notwithstanding, the inscription "Circles are square" is understood by virtue of understanding the sentence "Circles are *not* square."

Our Radical Theorist can accept the help and defend against the challenge. There is no meaning to the phrase "Circles are square" except in the same way that the word "sake" has a meaning. It has meaning only *qua* ingredient — what one may call "derivative meaning." To understand it is just to have an understanding of the genuine, meaningful sentences containing this phrase, and it is these sentences that *primarily* have meaning. But on its own the phrase is indeed meaningless. We might even deny it is a sentence in the relevant philosophical sense.

The above observations also help to solve another problem. There seems to be a difference in the beliefs of the people who believe, respectively, that:

(156) 5+7=11
(157) the evening star is not the morning star
(158) there is a married bachelor.

According to the Radical Theory there is no difference in *beliefs*. But these people are respectively disposed to utter different phrases, and these different phrases while having no primary meaning have different derivative meanings, in the way that "of" and "sake" have different derivative meanings. Such persons labor under pragmatic confusions. They act as if they had made an assertion in uttering, say, "There is a married bachelor" but do not realize that they are just like the madman who shouts "Of!" in a voice resonant with meaning.

And just as the madman might in fact be trying to convey some meaning, so too the person who utters "There is a married bachelor" may be trying to convey some meaning. She might be in the throes of a minor pragmatic confusion where she thinks the word "bachelor" stands for a young man of dissolute habits, which habits, however, are logically compatible with being married. But her confusion could take a different form. For instance, she might have heard from one person that John is a bachelor and from another that John is married. She holds on to both beliefs and seemingly draws the conclusion that there is a married bachelor. Perhaps a few minutes later she will realize that this is self-contradictory and that she needs to reject one of her beliefs, but right now she seems to accept this as a surprising new conclusion.

Here is where we can use the central part of Spinoza's defense of his theory that there is no error. Spinoza shows us how a sentence can convey a meaning that does not sit well with the form of words. He would say that when I hallucinate a purple tiger I do not actually mean by "There is a purple tiger here" anything other than "There is a purple-tiger image in my imagination." Likewise, the person who has heard of John that he is a bachelor and that he is a married man may not be saying anything more than that there is a person of whom *it is said* that he is bachelor and *it is said* that he is a married man. This seems an implausible solution at first sight, but it would beg the question against the Radical Theorist simply to complain that the content of the belief fails to match the form of the words. And there is actually some plausibility to the solution. The sorts of circumstances when one says "There is a married bachelor" surely are circumstances of serious confusion. There is some reason to think that one is indeed not operating with the concepts of *bachelor* and *being married* under such circumstances.

But consider a more complicated case. Take a fourteenth-century mathematician who became convinced that she had a method for trisecting an angle with compass and straightedge. She fully understood what angle trisection is all about, and she fully understood the canons of proof. She was no crank. In the fourteenth century it was not yet known that angle trisection is impossible. Unfortunately, her lengthy proof that the angle trisection procedure works is erroneous. Somewhere in the middle of the proof there is a minor mistake. She says: "I have trisected the angle." Has she really affirmed a necessary falsehood as the form of her words indicates?

Spinoza's reconstruction would presumably be to say that all our mathematician means is: "I have an apparent proof of an angle trisection," where an "apparent proof" is an inscription that looked to one like a proof upon step-by-step examination. Alternately the Radical Theorist can say that the mathematician has the contingently false empirical belief that the manuscript in front of her in fact is a proof. But imagine now that her colleague has come up with an equally long and complicated proof, but this time the proof was of a genuine theorem and was in fact correct. By the same token, should we not say that what her colleague really only means: "I have an apparent proof of my claim" or "There is a proof lying before me" when the colleague says: "I have a proof of my claim"? After all, their respective relationships to their assertions are closely parallel. Both checked their proofs equally well.

Nonetheless, one might insist, there is a difference between the two. For whereas the trisector at some step in the proof had committed a fallacy — she wrote down a formula that failed to follow from previous statements according to the canons of proof — her colleague did not. In fact the Radical

Theorist can say that some step did not even involve the trisector's *making a logical inference*, but only doing something that seemed like one.

The Radical Theory thus is committed to the claim that part of what makes a mental activity into a belief that *p*, at least in the case of non-contingent matters, is that this is an activity that has a certain cognitive history — a history of *proof*. Without the right kind of cognitive history, we do not have the right kind of belief. Perhaps the Radical Theorist could go further and insist that believings in non-contingent matters are a different kind of mental act from ordinary believings, and to count as engaging in a believing in a non-contingent matter one must have gone through the right kind of a process. This would be like Spinoza's insistence that one only adequately knows about the external world through scientific investigation.

Consider now a third mathematician. This one comes up with the same correct conclusion as the second mathematician did. But she has a fallacious proof of it, just as the trisector has of hers. On the above view, we have to assimilate the case of this mathematician to that of the trisector. Neither the third mathematician, nor the trisector, has the right kind of cognitive history to have a belief about a mathematical conclusion. Thus, this third mathematician does not believe "her theorem" to be true — to do that her line of thought would have to have the right kind of cognitive history. Rather, on this account what she believes is that some manuscript before her contains a proof or that she has an apparent proof.

The intuition behind the above judgments is that similar cognitive histories, and *apparently* similar kinds of mental states, make for actually similar kinds of mental states. Since the first and third mathematicians have similar cognitive histories, they are thereby either both believing non-contingent propositions or both believing contingent propositions. The trisector cannot be believing a non-contingent proposition, since there is no proposition expressed by the inscription "One can trisect an angle," and hence neither mathematician believes a non-contingent proposition in this case. The case of the second mathematician is less clear. If in fact she *carelessly* produced her proof so that she did not *see* every step as necessarily valid, the Radical Theorist might have to say that she too only believes some contingent proposition, though in her case at least a contingently *true* proposition.

What about mathematical belief based on testimony? My own conceptual history could be relevantly the same regardless of whether what I believed on testimony were in fact a mathematical truth or not. Suppose I said "Goldbach's conjecture is true" on the basis of testimony, but in fact Goldbach's conjecture was false. Then, on the Radical Theory, I cannot have asserted *that* Goldbach's conjecture is true. I must have asserted something

like: "I have been told that Goldbach's conjecture is true." But by the same token, when I say "Fermat's last theorem is true," also on the basis of testimony, it seems I am also not asserting Fermat's last theorem to be true — only those who have grasped the proof can assert it. There is only one way out of this unwelcome conclusion, and that is by saying that the cognitive history that partially constitutes beliefs is communal rather than individual, and it matters not who proved the theorem as long as someone did and my belief originated from that person or from someone else who understood the proof.

Thus, the Radical Theorist is committed to a view on which to understand a necessary proposition involves communal or individual grasp of the necessitating grounds for it. It will be good to review the above argument, its assumptions, and why the Radical Theorist must accept it.

To do this, let us put the argument in the form of a *reductio*. I will restrict myself to the case of individual grasp, but the communal case follows *mutatis mutandis*. Suppose for a *reductio* Jones understands a mathematical proposition p but does not grasp the necessitating grounds for it. This does not mean she lacks all ground for it. She might, for instance, accept p on the basis of the trustworthy testimony of witnesses or through the use of an electronic calculator or computer, as when by pressing buttons on a calculator or computer we might seem to learn that $328908432089 \times 177217 = 58288165609516313$ (which is in fact the case). But her grasp of the grounds for p is not tantamount to her understanding of a proof from unimpeachable self-evidently necessary premises. Suppose Jones expresses her belief that p in uttering the sentence s which means that p. Moreover, Jones has a full understanding of all the linguistic components of s. We would account Jones a fully competent mathematician in respect of s-type claimings. Of course, Jones does not grasp a proof of p, but that is quite compatible with competence: Many great mathematicians never grasped the proof of Fermat's Last Theorem simply because they died before this proof was first given, though one would not impeach their competence to understand the components of the sentence:

(159)   "$\sim(\exists x \exists y \exists z \exists n(x, y, z$ and n are positive integers & $n>2$ & $x^n+y^n=z^n))$."

We can then imagine a second person, Smith, who utters a similar kind of arithmetical sentence or pseudosentence, s^*. Smith does not see a proof of s^* but he has it on good authority that such a proof exists, or he has used a hand calculator to verify the proposition.

Now, we make a plausible internalist move, though alas some vagueness sets in here. If two people utter a similar sort of apparent sentence apparently about mathematics, on the basis of similar sorts of reasons, and have a similar amount of competence in the area of mathematics, then we do not want to say that the kind of propositions they affirm through uttering these sentences are different. It would be quite implausible, for instance, to say that Smith utters a geological proposition while Jones utters a mathematical one. Their practical activities are very similar, and very plausibly the mental processes in the two cases are very similar. Thus, unless we're dealing with something bizarre like a version of the Liar paradox, it is plausible that either Smith and Jones both grasp a mathematical proposition or neither does.

But now fill out the case of Smith further. Suppose that the sentence or pseudosentence Smith utters is one that people who don't accept the Radical Theory would take to express an arithmetical falsehood (say, "$328908432089 \times 177217 = 58288165619516313$"). Thus, on the Radical Theory, Smith is not affirming a mathematical proposition (unless we can come up with some *other* mathematical proposition for him to affirm, which seems a dubious task). Therefore, neither is Jones, since Jones's relation to the true sentence is parallel to Smith's relation to the pseudosentence.

The alternative for the Radical Theorist is to accept a radical externalism. One cannot tell from the sort of reasoning process by which one comes to affirm a sentence what if anything the sentence is about. The fact that the sentence uttered by Jones expresses a necessary truth while the one uttered by Smith does not is what makes it be the case that Jones utters a mathematical sentence while Smith does not. Granted, except in Gödelian cases, one can in principle tell from the form of the sentence, provided one is deductively omniscient, which of the two it is, and so one does not need to *physically* go outside of Jones and Smith to tell which of them if either is speaking mathematically. But it is still an implausible sort of externalism.

And, even more seriously, this externalist approach destroys the central advantage the Radical Theory has over other theories, the advantage for the sake of which one might be willing to tolerate its counterintuitive aspects. That advantage was that the problem of what makes a proposition possible disappeared as all propositions were possible. But now the problem creeps back in. What is it that makes an utterance be an expression of a necessary truth versus just plain nonsense or an expression of a contingent truth? It is not some fact about the form of the utterance or of the sorts of grounds on which it is adduced. In the mathematical case, the question of what makes for the difference in the utterances is precisely a *mathematical* question. And because of the externalism, we cannot simply say: "Well, in one case

there just *is* a mathematical belief and in another there is *not.*" For on the externalist solution to the problem there really is an issue here: we must, as it were, leave Jones and Smith and visit the universe of mathematical entities, and therein check what is going on. Thus, our externalist Radical Theorist has the implausible view that nobody can falsely believe a mathematical proposition, which every Radical Theorist has, without the compensating philosophical benefits, because the bump under the rug has simply shifted from the question of which propositions are possible to the question of which apparent sentences express a proposition.

The Radical Theorist may thus be stuck with the internalist account on which *neither* Jones nor Smith grasp a mathematical proposition, since one only grasps mathematical propositions by grasping their proofs. There is some difficulty here as to how proofs are to be characterized. Are they chains of self-evident deductions from self-evident premises, as in Descartes? Must one grasp the chain as a whole, or is it enough to grasp it step by step? If one goes for the step-by-step option, then a mathematical proposition can be grasped by anybody with the leisure to go through the steps. But there is some plausibility to the idea that the Radical Theorist needs to go for the holistic view of proofs, despite the unwelcome fact that it entails that many published theorems have never been grasped. For grasping a chain of deductions step by step is not very different from grasping a mathematical proposition on the epistemic authority of testimony. For it is by the "testimony of memory" that one holds on to the results of the preceding steps in the proof as one plunges into the next step. We will come back to this difficulty later.

Some plausibility for a part of the Radical Theorist's position might be thought to reside in the following observations. Even though at first sight it seems that when I used a computer to calculate that $328908432089 \times 177217 = 58288165609516313$, I thereby learned that a certain arithmetical proposition is true, this is not so clear. Do I really *grasp* what is involved in $<328908432089 \times 177217 = 58288165609516313>$? Do I even have an arithmetical concept that we would normally call "the concept of the number 328908432089"? What I have is a string of digits before my eyes. I do not even have the string of digits memorized (in fact, as I write this book, I do a lot of copying and pasting). It is much more plausible that I am in fact operating not with "the concept of the number 328908432089 *simpliciter*" but with the concept of *the number that such-and-such an inscription expresses in the decimal system*, which concept I have in virtue of having a concept of a decimal system. But if that is the concept which I have, then in fact the proposition that I believe, even though I express it

with the inscription "$328908432089 \times 177217 = 58288165609516313$" is at best a proposition that:

(160)   If the number indicated in decimal by the first numerical inscription is multiplied by the number indicated in decimal in the second numerical inscription, then the result is the number indicated by the third numerical inscription.

And this is an empirical proposition rather than an arithmetical one since its truth value depends on what the inscriptions in fact say. And so it is not so obvious that we grasp all the arithmetical propositions that we ordinarily would say we know.

The Radical Theory has now been laid out in somewhat greater detail, and let us take its internalist form, on which neither Smith nor Jones believes a mathematical proposition, but both believe empirical propositions, one of them a true one and the other a false one. We are now in a position to try to evaluate whether the theory is true. Consider the dilemma about proof. Either we understand a mathematical proposition by having sequentially understood every transition in a step-by-step proof or we need a holistic vision of the proof as a whole.

Now, in the step-by-step case, memory is involved, as already noted, which is relevantly similar to testimony, and we have seen already that belief on the basis of testimony is not sufficient for grasp of necessary propositions on the Radical Theory. It seems to matter little when I use a lemma in a proof whether (a) I have proved a lemma myself, and now remember that I have proved it without remembering its details; or whether (b) I take it on another mathematician's testimony that *she* has proved the lemma. Whether I am believing a mathematical or a contingent proposition should not depend on whether ten steps back in my proof process I remembered or misremembered something.

But perhaps there is another approach available to the Radical Theorist who allows that multi-step proofs can lead to the grasping of a necessary truth. It is not the fact of having gone through a proof that makes one grasp a mathematical proposition, but the practical *ability* to produce a proof. This, however, does not work once we think a little more about what "an ability to adduce a proof" is like. This cannot be an *infallible* ability, an absolutely reliable ability. Thus, it must be an ability to engage in an activity whose exercise *reliably* produces proof. But *how* reliably must the exercise of the ability produce proof? Reliability comes in degrees, whereas whether the proposition one is grasping is necessary or contingent does not.

According to the holistic version of the Radical Theory, what makes for the difference between a mathematician who grasps a mathematical proposition and a mathematician in relevantly similar circumstances who merely grasps a contingent proposition is that the successful mathematician insightfully sees of the proof as a whole. This option corresponds to the historical Spinoza's highest "intuitive" form of knowledge: "in the case of the given numbers 1, 2, 3, everybody can see that the fourth proportional is 6 [i.e. the number x such that the ratio 1:2 is equal to the ratio 3:x is 6], and all the more clearly because we infer in one single intuition the fourth number from the ratio we see the first number bears to the second" (*Ethics*, Sch. 2 to Prop. II.40, in Spinoza 1992), a knowledge Spinoza explicitly praises higher than that of the person who learns this answer "from the force of the proof of Proposition 19 of the Seventh Book of Euclid" (ibid.). Interestingly, Spinoza thinks that his views allow for a lower order of knowledge — or even belief — about necessary matters in cases of multi-step proofs, but here the Radical Theorist can say Spinoza errs.

But if this is right, then we have to say that a vast number of competent mathematicians do not believe the theorems that, it is generally said, they have proved. For the proof of a theorem is frequently long and involved, and often a mathematician does not grasp it in a holistic way. The mathematician may have intuitions guiding the proof, but (a) that is not always the case as sometimes the proof is just a matter of slogging through: and (b) often the intuitions are not sufficient to generate all the details of the proof.

Furthermore, a typical published proof tends to rely on intermediate results, theorems the mathematician learned in graduate school, say. Let us suppose, often overly charitably, that the mathematician actually remembers the proofs of all of the theorems she employs in the course of her proof. Even so it is highly implausible that she sees in a holistic way how her theorem follows *from the axioms*: she would holistically have to grasp not only her own proof, but all the proofs of the intermediate results. At best she sees in a holistic way how her theorem follows *from other theorems that she remembers*. And there we have a reliance on memory.

Granted, our Radical Theorist could say that what the mathematician really believes is not that the theorem she proves is true, but that the theorem follows from the intermediate results. However, to claim that mathematicians rarely believe the theorems they have proved would be too revisionary an account. And thus the holistic-grasp variant of the Radical Theory fails as well.

Section 5 *The Less Radical Theory*

But there is a variant of the Radical Theory that survives the above arguments. This is the Less Radical Theory, which accepts the holistic-grasp-of-proof variant of the Radical Theory, but no longer claims to be a complete theory of modality. The Less Radical Theory allows that some propositions have a particularly strong form of necessity, call it R-necessity, where p has R-necessity provided that there is no proposition that is a negation of p.[2] She can further posit that R-necessary propositions can only be grasped through a one-step intuitive seeing of their truth — one cannot wonder whether or not an R-necessary proposition is true. For instance, propositions such as that $1=1$ or that green is a color or that the law of excluded middle is true might be examples. Those who appear to deny these self-evident propositions fail to grasp them while attempting to deny them, and if they deny anything, it is some other proposition.

Unfortunately, R-necessity is a form of necessity that is too strong to be of much use. For instance, it is not the case that every proposition derived in a syntactically valid way from R-necessary propositions is R-necessary. Thus, it is quite possible that p and $p \supset q$ are both R-necessary, but q is not R-necessary. This is just because it might be that the derived proposition is not such that it can only be grasped through a one-step intuitive seeing of its truth.

Since the Less Radical Theory cannot give a complete account of modality, what would be its point? Well, it might answer some epistemological criticisms of other theories of modality. Take for instance the Lewisian theory on which something is possible if and only if it really occurs in some physical universe: How do we know what goes on in other physical universes?

Indeed, how do we know, for instance, that at every world we have $1=1$? The answer for the Less Radical Theorist is easy: At every world, every proposition is true or else its negation is true. We can see that there just is no thought with a content that could be sensibly described as "that $\sim(1=1)$," and yet if $\langle 1=1 \rangle$ were to have a negation, there should be such a thought. At this point, we may be proceeding fallibly. We can see, perhaps, that *we* cannot think a thought that would be a negation of $\langle 1=1 \rangle$, and then we take that as evidence that there is no such thought.

Now, at every world, $\langle 1=1 \rangle$ is true or else the negation of $\langle 1=1 \rangle$ is true. Since there is no negation of $\langle 1=1 \rangle$, it cannot be that the negation of $\langle 1=1 \rangle$

---

2   We may want to make the following stipulation here: A proposition p is a negation of a proposition q provided that either $p = \sim q$ or $q = \sim p$.

is true, and so <1=1> is true at every world. One might worry that when I use the words "the negation of <1=1>," I am talking nonsense according to the Less Radical Theory. To dispel that worry, apply a Russellian analysis of definite descriptions to replace a sentence of the form "S(the negation of <1=1>)" with the sentence "$\exists! p(p$ negates <1=1>) & $\forall p(p$ negates <1=1> $\supset S(p)$)," where "$\exists! x\ Fx$" abbreviates "$\exists x(Fx\ \&\ \forall y(Fy \supset y=x))$."

Furthermore, anything that logically follows from necessary propositions is going to be necessary. This is true because, plausibly, it is R-necessary that anything that logically follows (according to some set of syntactically valid inferences, such as *modus ponens* — we only want to give a sufficient condition here, so we need not catch them all) from a set of necessary propositions is necessary. So we can argue for the necessity of various non-self-evident propositions, such as Pythagoras's Theorem (considered as a conditional: if such-and-such axioms hold, then such-and-such follows) or that it is impossible to have $p\ \&\ {\sim}q\ \&\ (p \supset q)$ holding. For these propositions, although not R-necessary, nonetheless logically follows from R-necessary propositions, and it is R-necessary that any claim that does that is itself necessary, since the R-necessary propositions are also necessary.

Thus, the Less Radical Theory can be of help to someone searching for an account of alethic modality. For it can supply epistemological aid, explaining how we know *some* modal truths — these are truths that we can see not to have negations. It also removes the problem of how we ground *these* truths, since these truths can be grounded in the non-existence of a negation. However, the Less Radical Theory does not give an account of the grounds of modal claims in general. R-necessity just does not cover all the cases, since not all necessary truths are self-evident or even derivable from self-evident truths: just think about Gödelian mathematical truths, say.

One might try to build a more general theory. Let us inductively define R*-necessity as follows.

(161) If p is R-necessary, it is R*-necessary.

(162) If $p_1,...,p_n$ are R*-necessary and <If $p_1\&...\&p_n$ holds, then q holds> is R*-necessary, then q is R*-necessary.

It might, then, be R-necessary that whatever is R*-necessary is necessary. If so, then this is a useful criterion for necessity. But it will fail to work for Gödelian or Kripkean cases (cf. Section 2 of Part I), and so it does not yield a notion of metaphysical necessity. However, it might yield a useful notion of strictly logical or conceptual necessity, and that could be of value, if the kind of self-evidence that is involved in the notion of R-necessity makes sense.

Part VI

ARISTOTELIAN–LEIBNIZIAN ONTOLOGY

Section 1 Leibniz's approach

1.1 Leibniz's argument for the existence of God and the explanation of the nature of abstracta

We have examined a number of ways to account of possibility and possible worlds, and none have been entirely adequate, but two yielded defensible tools. Platonic ersatzism did not answer the grounding problem, but it did provide a way to construct possible worlds out of propositions. And the Less Radical Spinozist–Tractarian Theory gave an account of R^*-necessity, which might yield a narrower notion of necessity that could be epistemologically helpful.

It is time to sketch a positive approach that combines ideas from Leibniz and Aristotle. However, it will be a sketch, leaving details for future research, and how plausible this account will be will depend in large part on how well these details can be worked out.

First we need to discuss the Leibnizian half of the account. After having set out his account of how the possibilia vie with one another for existence, as it were trying to persuade God to actualize them, Leibniz writes:

> But, you say, this comparison between a certain determining metaphysical mechanism and the physical mechanism of heavy bodies, though it seems elegant, is faulty insofar as the heavy bodies striving really exist, while possibilities or essences before, or rather outside of existence, are imaginary or fictional, and therefore, one cannot seek a reason for existence in them. I respond that neither those essences nor the so-called eternal truths pertaining to them are fictitious. Rather, they exist in a certain realm of ideas, so to speak, namely, in God himself, the source of every essence and of the existence of the rest. The very existence of the actual series of things shows that we

seem not to have spoken without grounds. For the reason for things must be sought in metaphysical necessities or in eternal truths, since (as I showed above) it cannot be found in the actual series of things. But existing things cannot derive from anything but existing things, as I already noted above. So it is necessary that eternal truths have their existence in a certain absolute or metaphysically necessary subject, that is, in God, through whom those things which would otherwise be imaginary are realized, to use a barbaric but graphic expression.[1]

This provides Leibniz with yet another argument for the existence of God. Necessary truths must have some kind of a reality, and since, Leibniz thinks, it is inconceivable that ideas should have a self-standing existence outside of a mind, it follows that there must be a mind that contains the necessary truth. Since ideas of the form <A world containing seventeen donkeys, the first of them lame in the right foot, 6.54 feet tall, ... is possible> are also necessary truths, ideas of possibilia are included in this argument. Observe that it is essential to this view that the "ideas" be taken to be divine *thinkings*, rather than merely *thinkables* which are the *objects* of divine thought,[2] since arguably once one countenances thinkables independent of thinkings, one is already a Platonist, and Leibniz's desire for grounding these thinkables in a mind will no longer impress one.

The crucial assumption in this argument is that there cannot be self-standing necessarily true propositions. One way to attempt to argue for this would be along the lines of Lewis's doubts about how propositions could represent. One might hold the controversial view that representation is something that is necessarily dependent on there being minds and that only minds have original intentionality. If X is a non-mental item, one can argue, using Putnam-like shuffling arguments (Putnam 1978: 123–40, 1983: 1–25) that it is completely up to the minds that exist to decide what they want X to represent if X is to represent anything at all — even if X is some sort of a Platonic entity.

Even if there is some relation, like Platonic "participation," between X and the states of affairs that make X true, something like the noetic rays that Putnam (1978) disparaged that are supposed to emanate from a sentence to that which makes the sentence true, it would still be possible to take X to represent something completely different, simply by choosing not to see *this* relation as "the relation of representation" but choosing to see another

---

1   "On the Ultimate Origination of Things," Ariew and Garber (1989: 151–2).
2   I am grateful to Peter van Inwagen for forcing me to clarify this point.

relation as being representation (cf. Jubien 2001). We can reinterpret the noetic rays themselves: perhaps each noetic ray really points to the complement of that to which we first took it to point. Arguably, the only thing we cannot reinterpret in this way is something that has intentionality by virtue of being in someone's mind. For very plausibly *everything else* can be reinterpreted, and realism requires that there be *something* that cannot be re-interpreted. If this theory is right, then at least in the case of clear thought, we must then assume that *our* thoughts are intrinsically tied to meanings through some faculty of intentionality that we have (and if the controversial theory is wrong, then Leibniz's argument fails). But at the same time we need to extend the same courtesy to thoughts in the minds of other persons: thinkings in other people's minds have an intrinsic intentionality that we can understand and cannot re-interpret because these thinkings stand in a relation to their minds and the world objectively analogous to the one that our thinkings stand to our minds and the world.

Brian Leftow (1989) has noted that even if we accept that necessary truths must be found in a mind, it does not follow that the mind must be a necessary being. Perhaps there is a different omniscient mind in every world, each such mind being contingent but with it being logically necessary that some such mind exist. One way to argue against this possibility would be to claim that surely it is logically possible that no contingent person exists, and hence if Leftow's suggestion were true, then there would be worlds without necessary truths, namely those worlds that do not contain omniscient minds, which is absurd.

One might argue that there could be multiple divine minds, each of which has the ideas of some, but not all, necessary truths, and together they have the ideas of them all. A less than completely satisfactory solution to this objection would be just to define as "God" the aggregate of all these minds, leaving it as a topic for further investigation how united God's substance is. A more satisfactory solution is to follow Adams's (1994, chapter II.7) exposition of Leibniz, according to which the unity of the divine mind follows from the fact that the propositions are supposed to be interrelated, and it is mysterious for Leibniz how they could be interrelated if they are in different minds. Indeed, after all, the idea of a possible world will contain the ideas of all other possible worlds since at that world it will be true that they are possible. Likewise, one might also argue that if p and q are necessary truths, so is $p\&q$, and hence $p\&q$ will have to exist in some divine mind. On the assumption that one cannot have the idea that $p\&q$ without having the ideas that p and that q, it follows that if there are finitely many divine

minds, then one of them contains all propositions,[3] and we can reasonably call that one *God*. If there are infinitely many divine minds, and *if* there is a conjunction of all necessary truths (a rather complicated beast, since it will be a conjunct of itself), then the mind that contains that conjunction will contain all necessary truths.

That still leaves open the possibility that while in every world there is a necessarily existing mind that contains all necessary truths, in different worlds it is a different necessarily existing mind that does that. But this seems a needless complication. By Ockham's razor, it is simpler to suppose that the same mind contains all the necessary truths in every world.

1.2 *How propositions and possible worlds represent*

If the notion of divine ideas or thinkings can be made to work coherently, Leibniz can define a proposition as a divine idea, and a possible world as a coherent maximal divine idea. Recall Lewis's question of how it is that the possible worlds built up out of propositions represent worldly states of affairs. Progress has been made on this: they represent in a way analogous to the way *our* ideas represent our world. Of course how *that* happens is itself a difficult question, but one cannot, without pragmatic self-contradiction, doubt that it *does* happen.

Identifying propositions with divine thinkings gives an account of propositions. Now, Jubien (2001) distinguishes between three kinds of theories of propositions. "Mathematical" accounts provide ersatz surrogates for propositions. Typically, as we saw was the case for Lewis's structured propositions,

---

3   The claim follows from the following observation: If $S_1,\ldots,S_n$ are collections of ideas such that every idea is contained in some S_i and if $p\&q$ is a member of S_j, then both p and q are members of S_j, then at least one of the S_i contains all the ideas. The proof of the observation is as follows. First we prove this for the case $n=2$. Suppose for a *reductio* neither S_1 nor S_2 contains all the ideas. Then there is an idea p contained in S_1 but not in S_2 and an idea q contained in S_2 but not in S_1. But $p\&q$ must be in S_1 or in S_2. And whichever of the two collections it is in, that collection must contain both p and q, which violates our assumptions. Now, we proceed by mathematical induction. Suppose the observation has been proved for all $n<m$, and we are supposed to prove it for $n=m$, where $m>2$. Let S_2^* be the union of $S_2,\ldots,S_m$. Then, every idea lies either in S_1 or in S_2^*. By the case $n=2$ of the observation, it follows that every idea lies in S_1 or every idea lies in S_2^*. If every idea lies in S_1, the proof is complete. If every idea lies in S_2^*, then every idea lies in at least one of $S_2,\ldots,S_m$. Applying the $n=m-1$ case of the observation, we conclude that there is an S_i where $2\leq i\leq m$, such that all the ideas are in S_i.

there will not be a unique system of such surrogates. "Ontological" accounts tell us what the propositions *really are*. The system of propositions given by such an account must be unique. And, finally, there are "primitive" accounts that just take the notion of a proposition as a primitive. We could, unambiguously, take the identification between divine thinkings and propositions merely to give a "mathematical" account. But there is some plausibility in supposing that the account is actually ontological — that divine thinkings are what propositions are.

Jubien essentially argues that ontological theories of propositions suffer from a tension between two desiderata. One desideratum is that an account be given of how propositions represent how things are. The second desideratum is that the system of propositions be unique. For if we give a story about how a proposition represents how things are, a different story could also be given — a different "coding scheme" could always be used. One way to see this is that either the propositions do or do not have structure. If they do have structure, a different correspondence between the structure and the world could be given, much as discussed in Section 3 of Part II. If they do not have structure, then some external relation yields their correspondence to the world, and a different relation could be chosen. Moreover, even if there are Platonic entities, they only gain intentionality from how they are treated by *minds*.

However, the Leibnizian theory satisfies both desiderata. We get uniqueness from the fact that (at least according to the theism Leibniz is working with) there is only one necessarily existing thinking being. But the same token thinking cannot be engaged in by a different being from the one who is engaging in it[4] — you can think the same *kind* of thought that I am thinking, but my act of thinking is essentially mine. If propositions are thinkings, then any necessarily existent proposition will have to, thus, be a thinking by a necessary being, and if God is the only necessary being, then it will have to be one of God's thinkings. On orthodox views of propositions, all propositions are necessarily existent, and it is in any case plausible that at least the necessarily true propositions should exist necessarily. Thus, if the propositions are thinkings, they can only be God's thinkings. In Section 7.2.3, we consider the possibility that propositions making *de re* reference to contingent singulars exist only contingently. However, some propositions will still exist necessarily. And if we make simplicity a constraint on a plausible ontological

---

4   If the theism is to be Trinitarian, some careful qualifications will be needed. To ease a Trinitarian reading of the above, I shall use "being" stipulatively in such a way that the doctrine of the Trinity claims that there is one *being* who is three persons.

candidate for a system of propositions, it will be reasonable to take even the contingently existing propositions to be divine thinkings, rather than thinkings by finite beings.

The uniqueness thesis does require that God not have multiple token thinkings of the same proposition. But this is automatically true on the Leibnizian account. For the content of the thinking is a proposition, and that proposition is to be identified with that act of thinking itself. Hence, where there is a multiplicity of divine thinking tokens, there is by definition a multiplicity of propositions being thought.

As for the question of how the propositions represent the world, the answer is that they do so by virtue of the intentionality of the divine mind. But while Jubien took it to be a vice in an ontological account if "our active intentionality is required" for the propositions to represent (Jubien 2001: 53), it should not be a vice if *divine* mental intentionality is required. Moreover, this divine intentionality is not something super-added to the thinkings: the thinkings are only *thinkings* because of this intentionality. Nor should one worry that perhaps the relation between the thinkings and the reality that they represent is external and another could be chosen (e.g. maybe one could instead choose the relation that holds between a thinking T and the negation of T's content). For even if it is external, it is privileged by the fact that it is surely the most relevant natural relation if we are looking for a representation relation and the first relatum is a thinking.

1.3 *The Parmenidean challenge revisited and an assessment of the Leibnizian approach*

However, Leibniz's account does not answer the Parmenidean challenge any more than the propositional account did. It is a necessary truth that it is possible for there to be unicorns. There is an idea in the mind of God which expresses its being possible for there to be unicorns, since there are ideas in the mind of God that express the necessary truths. This idea is true or correct. But what makes it correct? Ideas are correct if they accurately reflect reality. Saying that a necessary truth is an idea in the mind of God does not remove the task of explaining what the truthmakers of these ideas are. If we say that the truthmaker of the idea consists in the idea that there are unicorns having the property of possibility, we are seemingly no further ahead than the Ambitious Propositional Ersatzist was.

Though, actually, we do have a gain. We do not have to worry about how we can meaningfully refer to entities that are of a kind that plays no

causal role in anything. For the divine ideas as aspects of the divine mind presumably can be thought of as playing a causal role in creation. Moreover, if we have a substance-based ontology according to which all things that exist are substances or things that ontologically depend on substances, then the Leibnizian approach is the more tenable: the divine ideas ontologically depend, in some way, on the substance that God is. Furthermore, Lewis's worry of how it is that there can be entities like propositions which somehow manage to represent states of affairs is alleviated. For, although it is a difficult problem to say *how* the thinkings of a person represent the world, it would be self-defeating to doubt that they *can* do it.

Section 2 Aristotelian possibility and causality

2.1 *Proximate matter, proximate cause, and Parmenides*

Now consider a very different theory. Begin by recalling Aristotle's theory of change in the *Physics*. Change involves matter and two forms, a positive form and a privative form. The matter in question is directly capable of having both forms. This matter is termed in the *Metaphysics* "the proximate matter [*hê eschatê hulê*]" (e.g., Aristotle 1984: 1045b18-19). It is "proximate" because with no further prior changes, it can take on either of the two forms. Thus the proximate matter for a bronze statue is not the elemental earth. Nor is it the copper and tin that the elemental earth is worked up into. Rather, the proximate matter is the bronze that is produced from the copper and the tin. The copper and tin are not the proximate matter of the statue because copper and tin require a further process in addition to the process of taking on the form of the statue in order to gain that form, namely the process of alloying.

Aristotle sees the possibility of there being a statue as *materially* grounded in the capabilities of the proximate matter, viz. the bronze. The proximate matter is the matter found at the last stage in a process that can branch into either of the two forms that the matter is capable of. But along with the proximate matter responsible for the material grounding, there is also what one might call the "proximate efficient cause," which is the cause capable of operating on the proximate matter to produce in it one of the forms that the matter is capable of. In the case of the statue, this proximate cause will involve the sculptor's decision whether or not to make a statue. Aristotle does not distinguish the proximate cause in the explicit way he does the proximate matter, but it is arguably implicit in his account.

Armed with this terminology, let us imagine that we have before us a lump of bronze and a sculptor. It is possible that a statue will be there in the future, but it is also possible that a shapeless mass will be there in the future. Suppose that in fact no statue will exist here. Nonetheless, it is possible that a statue will exist here. And this is an assertion of possibility in connection with which we can give an answer to the Parmenidean problem. The assertion that it is possible that a statue will exist is made true by the *actual* capabilities, or more precisely second potentialities or first actualities,[5] of the sculptor and the bronze: the sculptor's actual know-how, the actual strength and dexterity of her limbs, the malleability of the bronze, and the like. All these things are *actually* possessed by the sculptor and the bronze. We are only talking of the actual world here. And being *second* potentialities, these capabilities have in themselves all that is needed for producing the effect.

But is this not a circular answer? The "actual" capabilities of the sculptor and of the bronze are *dispositional* properties, second *potentialities*, and hence modal through and through. However, it is here that the importance of speaking of the *proximate* matter and cause shows up. The proximate matter is the matter which *in its own right* (*per se*, *kath' hauto*) is capable of taking on the various forms. It needs nothing else, except the proximate cause, for this. Thus the items which actually exist in this world, namely the proximate matter and the proximate cause are sufficient to guarantee the possibility of the statue's existence. If the statue were to come into existence, they would yield an explanation of the statue's existence: the sculptor worked the bronze into a statue. And because they are sufficient to guarantee this possibility, it is plausible to say that when we are talking of this possibility we are talking of them. Actuality is prior to possibility in the sense that assertions of possibility are in fact grounded in actual worldly states of affairs.

Of course this is not a completely reductive account of modality. We are analyzing a possibility in terms of things that are themselves through-and-through modal: capabilities, causal powers, dispositions, etc. However, these things are actual and concrete in the paradigmatic way: they enter into causal explanations of actual phenomena. We can be actualists in good standing and yet believe in them.

5　　I am now in first potentiality with respect to speaking German: I can *learn* to speak German, but I don't know how to speak German. Were I to learn to speak German, I would then be in *second* potentiality, also known as first actuality, with respect to speaking German. Were I to actually open my mouth and speak German, I would then be in second actuality with respect to speaking German (and second potentiality with respect to saying something different from what I am saying).

2.2 *The basic Aristotelian view*

However, the account so far is incomplete. For it would still be true to say that it is possible that a statue will exist even if the proximate matter of the statue had not come to be, but, say, had remained as tin and copper, and even if there had never been a sculptor on earth. Moreover, as noted in Section 4.5 of Part I, the Aristotelian view runs into a serious objection in its inability to handle global possibilities.

First suppose that in fact in the course of human history, sculptors had never arisen, while tin and copper had never been alloyed into bronze. We can still give an Aristotelian account of why a bronze statue could have existed. The possibility of the statue is grounded in the possibility of bronze and of a sculptor, while the possibility of bronze is grounded in the capabilities of a metallurgist, of smelting equipment and of tin and copper which could be the proximate material causes of bronze, while the possibility of a sculptor is grounded in more general capabilities of human beings which could be proximate causes of the existence of sculptors. The possibility of the statue is thus seen to be ontologically grounded in a longer chain of possible causes, which chain nonetheless starts with some *actual* items in the world. Moreover, the items at each point in the chain are sufficient for producing the items in the next stage in the chain. The more general capabilities and the free will that Quine had are sufficient for making it possible for Quine to have been a sculptor, for instance.

We can make this into an inductive definition of possibility for propositions. Start by saying that a proposition p is possible$_0$ if and only if p is true. Then for $i>0$, say that p is possible$_i$ if there exists (timelessly, or in the past, present or future) an actual item A (past, present, future or timeless) that has the causal capability for bringing it about that p is possible$_{i-1}$, in such a way that the actuation of this causal capability would explain p. Possibility$_1$ is *proximate* possibility, and we can say that p is possible *simpliciter* if and only if p is possible$_i$ for some i. Note that the causal capabilities will in general not be deterministic in nature, and so it can be that both p and not-p are proximately (or non-proximately) possible.

Equivalently, we can say that it is possible that s if and only if either s, or there is something that has the causal capability to make it be that s, or there is something that has the causal capability to make it be that there is something that has the causal capability to make it be that s, or …. And we can summarize this by saying that a non-actual state of affairs is made possible by something capable of initiating a chain of causes leading up to that state of affairs.

Because each step in the chain is capable of yielding an explanation (albeit perhaps indeterminstic) of the next, all the possibilities in the chain are inchoately present in the first actually existent member of such a chain. This chain-based definition takes care of the difficulty with possibilities whose proximate causes are not in actual existence and gives an account roughly like that of branching modalists like Penelope Mackie (1998).

On Aristotle's own view, the items will presumably be substances and their accidents. However, even then, they will in general be unlikely to be individual substances or individual accidents. What makes a statue proximately possible in Aristotle's story will be the sculptor and the bronze. In the above definition, we seemingly talked of a *single* item A that makes possible a chain of causes. But in doing so, I used "item" in a loose and ontological non-committal way. Thus, the sculptor and the bronze together can count as an "item."

If one is uncomfortable with this loose way of talking, two options are available. The first way is to commit oneself to mereological sums or fusion. What makes the sculpture possible is the causal capabilities of a single item: the fusion of the sculptor and the bronze. This is an un-Aristotelian way of thinking, but it is available.

A better way would be to replace the loose quantification over items with *plural quantification*. Thus, for $i>0$, we say that p is possible$_i$ if there are actual As that have the causal capability for bringing it about that p is possible$_{i-1}$. We can formalize this, if we wish, in terms of a plural-to-singular two-place non-distributive relation $Can(xx, q)$ which holds when the xx have the causal capability for bringing it about that q: if $i>0$, then p is possible$_i$ if and only if $\exists xx(Can(xx, <p$ is possible$_{i-1}>)$. Whether this is necessary, or whether we can make-do with single-agent powers, shall be discussed further in Section 6.6.

It may be that taking the items to be substances and their accidents, or even pluralities of substances and accidents, is untenable. For what makes something immediately causally possible need not merely be the *existence* of one or more substances or accidents, but the existence of one or more substances or accidents combined with the non-existence of others. For instance, what makes it possible for me to blow up a balloon seems to go beyond my and the balloon's powers: the space that the balloon would supposedly expand into has to be *not occupied* by bricks or anything else that would preclude that expansion. Now, it *could* be that in the correct metaphysics, the availability of space for balloon expansion is grounded in the existence of one or more substances and accidents. For instance, it could be that space-time is itself a substance and has the accident of being

occupied only by low-pressure air in region R, or maybe of being vacant in R. But we are right now occupied in giving an account of the metaphysics of modality, rather than a full metaphysics of the world. And so at this point at least, we should leave open the option that negative facts could be involved in the grounding of modality.

To that end, one option that should be left open is that the items are *states of affairs*. Now, the term "state of affairs" is unfortunately ambiguous in the literature. There are Plantinga's (1974) Platonic states of affairs which are all abstract beings, and which exist whether or not they obtain — thus, the there is the state of affairs of Napoleon winning at Waterloo in addition to the state of affairs of his losing. And there are Armstrong's (1997) states of affairs which exist if and only if they obtain. It is the Armstrongian kinds of states of affairs that are of interest here. It is tempting to call these states of affairs "concrete" but some of these of states of affairs are also abstract,[6] like the state of affairs of seven being odd. From now on, when I talk of states of affairs without further qualification, I shall mean the Armstrongian ones. It is these states of affairs that can potentially enter into causal relations, and it is they that are suited to an Aristotelian account of modality.

It is also left open that talk about the items in the account reduces to something else — as long as it does so in a way that does not threaten circularity or triviality for the account of modality. Thus, it may be that while the items in the account are states of affairs, facts about states of affairs reduce to facts about what substances and accidents do or do not exist. If this turns out to be the case, then our account will give *grounds* for modality but not *truthmakers*: a possibility might then end up being grounded in such-and-such substances and accidents existing and such-and-such substances and accidents not existing, as in Lewis's generalization of the truthmaker principle (Lewis 1983a; cf. Section 4.1 of Part I).

With it unsettled as to what exactly the items are, the account may seem to threaten vacuousness. However, without settling what exactly the items are, we can say that the items are the very items that figure in causal explanations. The item whose causal capability that grounds the possibility$_1$ of p is such that the actuation of that very causal capability would explain p if that causal capability were actuated. And this is a non-trivial claim: it is denied by other attempts to ground modality.

---

6   I am grateful to Jonathan Kvanvig for this caution.

2.3 *Global possibilities: a difficulty*

However, the difficulty about global possibilities remains. Surely, the universe could have been radically different at every instant of time in the past. Perhaps our cosmos might not have been temporal at all. The universe might always have been 22-dimensional. Our universe started existing after a Big Bang. But it is difficult to see how these possibilities are to be grounded in an actuality, since in the Aristotelian account the possible world and the actual world in question always have something in common, namely the potential causes in which the possibility is grounded, e.g. the sculptor capable of producing or not producing a statue.

It is here that combining the Aristotelian account with Leibniz's theistic story would solve the difficulties. On this combined account, there would be a God outside the universe, though not outside the cosmos (recall that in our preferred terminology "the cosmos" is the aggregate of all existing entities while "the universe" is the aggregate of all spatio-temporal ones), who has in his mind ideas of all kinds of possible worlds. We get to retain Leibniz's intuition (very much in the Aristotelian spirit) that propositional-type entities can only exist in minds. But what we now add is the fact that this God is not just all-knowing but also all-powerful. He is capable of putting his divine ideas into action. He can produce a 22-dimensional universe, not just think about one. If he himself is atemporal, then he is capable of creating beings that are atemporal. If an infinite past is coherent, he can even make a universe that has always existed — by *creatio continua* if he is temporal and by a timeless act of creating time and a universe that had always existed if he is eternal.

Still, the theistic story is not the only one available to the Aristotelian. The Aristotelian might, with the historical Aristotle who thought that the existence of the actually instantiated natural kinds to be necessary, simply deny the possibility of the universe having been radically different. On this view, though for any time t it might have been that the universe was different at t from the way it actually was at t — this is a possibility we can ground in the powers of things at a time t^* prior to t — nonetheless it is false that it might have been for all t different at t. I take this to be an implausible view, since it seems plausible that if the state of the universe on any time-slice is contingent, it might surely have been that the state of the universe was different on *all* time-slices. But it is indeed a view an Aristotelian might take. Alternately, the Aristotelian might accept something atemporal prior to the course of history which non-deterministically causes the start of our history, but deny that this atemporal being is a

person — perhaps this being is just some event in a supercosmos, it might be said.

The Aristotelian thus has three options: (1) deny that the whole course of history could have been different; (2) accept that it could have been different due to the causal agency of an ahistorical non-person; or (3) accept that it could have been different due to the causal agency of an extra-historical person, a God. We examine an argument similar to the one behind this trichotomy in more detail in Section 2.4.

Observe also that any uncaused being, such as God, will have to be a necessary being on the Aristotelian account. For in fact the Aristotelian account implies that, necessarily, every contingent being has a cause. To see this, suppose for a *reductio* that it is possible that x is a contingent being that has no cause. Let p be the proposition that there is an uncaused contingent being. Then, p is possible. There is then a possible world (e.g. an Unambitious and hence uncontroversial Propositional Ersatz possible world — I am just using possible worlds as an aid to expression) where p is false. But by S5 (which can be argued for on Aristotelian grounds as we shall see in Section 2.4), since what is possible is necessarily possible, p is still possible in that world. But by the Aristotelian account, this possibility would have to be grounded in a cause or causes capable of making p true. But it is logically impossible that the proposition that there is an *uncaused* contingent being should be made to hold by a *cause*. Therefore, there cannot be an item capable of making p true, and hence p is impossible on the Aristotelian account, and yet possible by assumption, which is absurd. Consequently, there cannot be an uncaused contingent being.

2.4 *A direct argument for a necessary being*

While there is good theoretical reason to supplement the Aristotelian account with theism, we can also give a direct argument for a necessary being. Let S be the set of all the contingent beings in existence. It is very plausible that:

(163) Possibly, none of the beings in S exist.

Why is this plausible? Well, first, if (163) is false, then there necessarily are contingent beings. Although this has been defended (e.g. Rundle 2004), it is implausible. For instance, it implies there not being any non-unicorn contingent beings entails that there are unicorns, which is strange. Second, even if there necessarily are contingent beings, it seems implausible that at

least one of these beings must always be from among the contingent beings of our world.[7]

Now, given (163), there must be an item, x, that can initiate a chain of causes leading to its being the case that (163). Now, plausibly, only an item that is or contains a being can have any causal powers. Thus, x must include a being that has causal powers. This being cannot be contingent, since then it would be in S, and no item in S can contribute to initiating a chain of causes leading to none of the beings in S existing. But if this being isn't contingent, it is necessary.

2.5 *S5, the Principle of Sufficient Reason, and the cosmological argument*

The Aristotelian account of modalities gives an argument for both S5 and a Principle of Sufficient Reason (PSR).

First, observe that S4 follows quickly. Suppose that it is possible that it is possible that p holds. Then, there is an item x which had the power of initiating a causal chain that can lead to its being possible that p. In other words, x can initiate a chain that can lead to there being an item, say y, that can initiate a chain that can lead to its being the case that p. Putting these two potential chains together we see that x can initiate a chain that can lead to p's holding. In other words, $\Diamond\Diamond p \supset \Diamond p$, and moreover this argument applies at all possible worlds and so we have S4.

The PSR that I shall now argue for states that every contingently true proposition has a causal explanation, where something counts as having a causal explanation provided it is the actualizing of a causal capability or the result of such an actualizing. To argue for the PSR, however, I shall need a somewhat odd, but very plausible, technical assumption. Say that "the PSR must hold for a contingent proposition p" provided that necessarily, if p is true, then there is a causal explanation of p. Then, the assumption is:

(164)   If every false contingent proposition is such that the PSR must hold for it, then the PSR must hold for all contingent propositions.

---

7   A typical implausible scenario ruled out here would be like this. This world's contingent beings include Smith and Jones and no one else. There are worlds where Smith doesn't exist and worlds where Jones doesn't exist, but no world where neither of them exists.

It would be strange indeed if, say, there actually were an uncaused horse, but dragons had to have causes. Indeed, the standard Humean objection to the PSR is that one can imagine, say, a brick coming into existence *ex nihilo*, not that there actually is such a brick.

But now the antecedent of (164) directly follows from the Aristotelian account of modality. For suppose that *p* is a false contingent proposition, and suppose, for a *reductio*, that it is possible that *p* holds without a causal explanation. Let *p** be the proposition that *p* holds without a causal explanation. Then *p** is possible, and hence there is something that can initiate a chain of causes leading to *p** holding. But then the chain of causes leads to *p* holding without a causal explanation. But that is absurd, since then that chain of causes would causally explain why *p* holds and why it has no causal explanation. So the antecedent of (164) holds.

The reason for the need for the technical assumption is that it rules out scenarios like the following. In the actual world, a brick *B* comes into existence causelessly at spatio-temporal location *u*. This brick, however, is a contingent entity as there is an *A* that had the causal capability of preventing *B* from existing, perhaps by filling up location *u* with something that precludes *B*'s coming to be present there. Moreover, no purely other-worldly brick could be like that, since by the Aristotelian account the possibility of a purely other-worldly a brick would need to be grounded in a potential cause. But the Aristotelian account does not require a ground of the possibility of a this-worldly brick beyond the actuality of the said brick. Hence we need the technical assumption.[8]

Given that the PSR must hold for all contingent propositions, Brouwer follows. For suppose that *p* is actually true. We must argue for $\Box\Diamond p$. So, for a *reductio*, suppose that $\sim\Diamond p$ is possibly true. Then there must be an actual item *x* that has the causal capability of initiating a chain of causes leading to $\sim\Diamond p$ holding. But now, by the PSR, in the actual world, there must be a causal explanation of why both *p* holds and *x* did not initiate any such chain of causes. The item, *y*, whose causal capability grounds this explanation must be prior to any such potential chains, and hence the exercise of its causal capability cannot be prevented by the activity of *x*. But, then, *y* would still exist even if *x* were to initiate a chain of causes leading to $\sim\Diamond p$ holding, and *y* would still ground the possibility of *p*. Hence, *p* could still be possible even if *x* led to $\sim\Diamond p$ holding, which is absurd.

---

8   In some earlier works (e.g. Pruss 2006, Section 19.5), I have argued that the PSR follows from the Aristotelian view of possibility. Those arguments did not use this assumption, but were unfortunately defective.

Assuming Brouwer, one can argue for PSR without (164). Actually, what we need is a slightly weaker version of Brouwer:

(165)   If p holds contingently, then it is possible for p to be both possible and false.

This follows from the full Brouwer axiom, since if p holds contingently, then p possible, and so it is necessarily possible, but since it is contingent it is possibly false, so possibly it is both false and possible. Suppose for a *reductio* that a contingent p has no causal explanation. Let p^* be the state of affairs of p's obtaining without causal explanation. Then p^* is a contingent state of affairs. By the weaker version of the Brouwer axiom, it is possible that p^* does not hold but is nonetheless possible. Let us suppose a possible world w (perhaps an ersatz one) where that happens. Thus, there is an item x in w which could initiate a chain of exercises of causal powers capable of leading to p^*. But that is absurd, since in doing so the chain would lead to p's holding without causal explanation.

The above remarks show that in the Aristotelian setting, the necessity of PSR is closely akin to the Brouwer axiom.

Something like the Brouwer axiom can also be seen as codifying the fact that the Aristotelian account yields a branching structure where taking a branch is compatible with the possibility of not having taken it. Suppose an item x makes a false proposition p possible$_1$. Then, Brouwer says that, were x to have exercised the causal capability to make p hold, the actual world's branch, on which $\sim p$ holds, would still have been possible. The latter possibility would, presumably, be partly grounded in x's ability not to actuate the causal capability for p.

And, of course, S5 follows from Brouwer and S4. Given the intuitive plausibility of S5 (see Section 3 of Part I), the fact that we are able to derive S5 from the Aristotelian view together with the plausible assumption (164) stands in favor of the Aristotelian account, by providing at least a partial explanation of why S5 holds.

Given the PSR, one can run a cosmological argument for the existence of a necessarily existing First Cause (see, e.g. Pruss 2009). For instance, one might demand a causal explanation for the existence of the contingent beings in our world. That causal explanation cannot involve the existence of any contingent being. Hence there must be a causally efficacious necessarily existing First Cause that explains the existence of our world's contingent beings.

2.6 *The Aristotelian–Leibnizian view*

The Aristotelian who accepts the plausible assumptions of the previous section is committed to a necessarily existing First Cause. Does this Aristotelian need to identify this First Cause with the God of theism, or can she allow it to be something in a super-universe atemporally prior to the Big Bang? This will depend on other issues. For instance, Koons (1997) has argued that once one has an argument for the existence of a First Cause, then teleological arguments, say from Fine Tuning or from the unlikelihood of self-reproducing organisms coming into existence, gain in plausibility and so one might have reason to accept that the First Cause of our complex universe is intelligent.[9] The plausibilistic considerations in Gale and Pruss (1999), Gellman (2000), Rasmussen (2009), and Pruss (2009) are also relevant here.

The Aristotelian account by itself does not yield possible *worlds*. One could supplement the account with Platonism. The main objection to Platonic accounts of modality was that they failed to give a satisfactory answer to the grounding problem. But once we have the Aristotelian answer to the grounding problem, we can adopt a Platonic answer to the worlds problem. Possible worlds are, say, maximal collections of compossible propositions, where a collection of propositions is compossible provided there is something that can initiate a causal chain leading to its being the case that they are all true. While this seems to be a coherent account of both possible worlds and of the grounds of modality, it is inelegant in that it posits that there really is no connection between worlds and the grounds of modality. Worlds are just certain collections of propositions. Their ontology as collections of propositions seems completely separate from the grounding of the fact that they are *compossible*. Worlds become an add-on to the system, one that is not directly relevant to the grounding of modality.

However, if one wants the benefits of a theory of possible worlds and finds the Leibnizian approach plausible, then in order to have *both* possible worlds *and* a grounding for alethic modality on the scene, one can posit one God whose powers ground alethic modality and some of whose thinkings are the possible worlds that he can initiate causal chains leading to the actualization of. God, as a rational being, can only act after rationally considering the options, and his considering of the options might well constitute the divine ideas, the thinkings that possible worlds are.

---

9   Gellman (2000) has also argued in a related context that the God argued for by means of a PSR would have to be essentially omnipotent, and perhaps his argument can be adapted to the present context, too, since our account also entails the PSR (see Section 5, below).

This merges the Aristotelian and Leibnizian accounts and this solves the difficulty which was the main objection to Leibniz's account of possibility. Granted, propositions expressing modal truths are ideas in the mind of God. But what, we asked, makes these propositions true? The answer is that they are made true by the capabilities of the God whose action-guiding ideas they are[10] and by those of beings that he might have created. What makes it true to say that it is possible for there to exist a world with unicorns is that the idea of such a world is an idea in the mind of an all-powerful God capable of acting on it and creating such a world.

This account also gives a good answer to Leftow's (1989) concern that perhaps different divine minds can satisfy the Leibnizian need for a contemplator of necessary truths. For the Aristotelian–Leibnizian account entails that the same God must be the First Cause of all things in all possible worlds. For suppose that there was a world *w* where the God who is the First Cause of all things in the actual world was not the First Cause of all things, but instead Smith was. Then it would be possible for God to initiate a chain of causes leading to Smith being the First Cause. But if God initiated a chain of causes leading to Smith being the First Cause, then Smith wouldn't be the First Cause. (cf. also the arguments in Gellman 2000.)

Moreover, there is no problem with our grasping the nature of the modalities involved in this branching account, because this nature is nothing else than the nature of non-deterministic causation, a causation that we ourselves ordinarily take ourselves to exhibit, and which we can introspectively understand. Note that for this grasp it perhaps need not even be required that we *exhibit* such causation, but only that we *experience*, perhaps on occasion non-veridically, the exercise of such causation.[11] For an account of causation that starts with agency in basically this way, see Swinburne (1997).

---

10 One might think that this solution is indicated in Leibniz's words at the end of his necessary truths argument for the existence of God that through this God "those things which would otherwise be imaginary are realized, to use a barbaric but graphic expression" (Ariew and Garber 1989: 152). However, this charitable interpretation of Leibniz would require Leibniz to have said "are *capable of being* realized," since not all the imaginary things are in fact created. The phrasing, at least on this translation, suggests that Leibniz is not talking of a realization of things that are possible but rather of ideas being made real (though not necessary being made *true*) through having an ontological home in God's mind.

11 I am grateful to Richard Gale for this suggestion.

2.7 *Branching and other modalities*

The present theory allows one to have a branching theory of modality, rather like those that one might have in a tensed logic system, except that this theory allows for global possibilities — for the whole temporal sequence to have been different — because it has a God who of logical necessity is the First Cause and can create different cosmic systems. The branch-points are instances of proximate causal powers that can go in more than one direction. And, of course, the branching theory here, unlike the one in the many-worlds interpretation of quantum mechanics, only assigns actual reality to one line — the other lines are not really existent. Moreover, the branching system here is not necessarily temporal: God might not be in time.

   A major advantage of such a branching view is that one can handle different modalities in a closely parallel way. Ordinary language does not concern itself with metaphysical possibility as much as with physical possibility (cf. Place 1997), and logical possibility is a generalization from physical possibility. If we see physical possibility and necessity as grounded in the causal powers of actual things, then it is difficult to see how logical possibility could be a generalization of this, unless one takes the Aristotelian view. For on the Aristotelian view, there is no radical difference between logical and physical possibility. The only difference is that when we talk of the capabilities that ground physical possibility, we restrict the quantifiers in the definition of possibility$_1$ to take into account only the capabilities of physical objects, whereas when we talk of metaphysical possibility this restriction is relaxed and the capabilities of *all* entities are now permitted to be talked about. Similarly, a temporalized notion of possibility is arrived at when we restrict the quantifications in the definition of possibility to the *future* capabilities of actual entities. The Aristotelian view thus brings multiple alethic modal notions under a single framework.

2.8 *Indeterminism*

Consider the following possibility: A twin brother of Napoleon could have freely become a Franciscan friar. This possibility is not grounded in the capabilities of Napoleon's twin brother, since none such exists. Rather, let us say, it is at best grounded in the capabilities of Letizia Buonaparte's ovaries, such as the power of doubly ovulating, and of Carlo Buonapartes' sperm, such as the power to have also fertilized a second egg, as well as

various environmental factors and any other needed causes.[12] Call the state of affairs including all this *S*. We probably do not want to say that these things grounding the possibility of Napoleon's twin freely becoming a friar ground it by having a capability *to make Napoleon's twin be a friar*. For the chain between *S* and Napoleon's twin's becoming a friar is surely indeterministic. Rather, *S* has the capability to produce a state of affairs that has the capability to produce a state of affairs that ... that has the capability to make Napoleon's twin become a friar.

However, the possibility of Napoleon's twin becoming a friar must still in some way be contained in *S*. For instance, it is within *S*'s capabilities to produce the sort of being that would have the capability to become a Franciscan. Therefore, there is something in *S*, some actual property in *S*, in virtue of which *S* has this higher order capability. Becoming a Franciscan friar is thus in an implicit way contained in *S*, though only at a remove. It is *this* sort of grounding in the powers of the cause that the theory yields. And this result is similar to the medieval dictum, made use of by Descartes, that the reality of the effect is found eminently or actually in the cause.

Section 3 Combining with the Spinozistic–Tractarian view

The Radical Spinozostic-Tractarian view denied that are any impossible propositions. This was untenable. However, we also saw that there was a Less Radical Theory (LRT) on which self-evident propositions, such as <1=1>, are so strongly necessary that they have no negations. One can ground the LRT on the Wittgensteinian view that to try to deny certain tautologies is just to make a *grammatical* mistake, i.e. not to say anything. We should not take the LRT to be a complete account of modality, as it is not a grammatical mistake to say that Bucephalus is not a mammal, even if Bucephalus is essentially a mammal. What makes it impossible that Bucephalus is a non-mammal is, perhaps, that nothing can initiate a causal chain leading to Bucephalus being a non-mammal.

The LRT could then answer an epistemological objection to the Aristotelian–Leibnizian theory. We can know that there cannot be any square circles without surveying possible causes for things with a square-circle-producing capability, or a capability to produce something capable of producing square circles, etc. The LRT gives us a story here. <There are no

---

12   For instance, if it is necessarily true that God has to create each soul, then God's
     power will be involved.

square circles> is not only self-evident but such that no proposition aptly expressed by the words "There is a square circle" expresses a proposition, and hence it is grammatical nonsense to try to say "that something can causally lead to there being a square circle." Thus, nothing can causally lead to there being a square circle.

But what if we have more complex kinds of contradictions? Let us say *s* is a very complex formula which is, nonetheless, provably but not self-evidently self-contradictory. Then the proposition <*s*> on the *Less* Radical Theory may well have a negation, <~*s*>. Can we explain why <~*s*>? Maybe. For then it may well be that <~*s*> is R*-necessary (recall Section 5 of Part V), and it might be R-necessary that nothing has the power to initiate a chain capable of leading to the truth of a negation of an R*-necessary proposition. This might give us an explanation why the only powers there are are powers to initiate chains capable of leading to the truths of propositions that are not the negations of R*-necessary propositions, and hence why there is nothing that can initiate a chain leading to its being the case that <*s*>.

The truth of the LRT is not a real limitation on the Aristotelian–Leibnizian theory's applicability to the explication of modality. For the Aristotelian–Leibnizian theory can still be thought of as the only theory that tells us what propositions are possible. The LRT does not tell us what propositions are possible: it tells us what putative sentences express propositions and which propositions lack negations. Whatever theory of modality we take, we need to say which putative sentences express propositions and which do not. The LRT differs, however, from traditional grammatical theories in that it denies that we can always truth-functionally compose sentences to form new sentences. If we can live with this, then the LRT makes a fine supplement to the Aristotelian–Leibnizian account.

Note that the LRT in this role escapes the objections of Section 2 of Part I against a notion of strictly logical necessity. The objections there were that what is strictly logically necessary is relative to a language and a system of axioms. If we think with the Wittgenstein of the *Tractatus* that the system of axioms is embedded in the grammar of the language, then we will not be worried by this relativity in the case of the LRT. The reason we will not be worried is that the LRT is precisely a theory about the meanings of putative sentences, and hence language-relativity is not a problem.

Section 4   Ordinary alethic modal talk

Ordinary language uses the concepts of possibility and necessity in many ways:

(166)   "It is *possible* that one day the Riemann conjecture will be disproved."

(167)   "Smith got into the room, but didn't come in through a door or a window, so he *must* have come in through a skylight."

(168)   "Were the apple dropped, it would *have to* fall."

(169)   "It is *impossible* for human beings to survive without air."

(170)   "It is *impossible* that a human being survive without air."

(171)   "If all men are mortal and Socrates is a man, then Socrates *must* be mortal."

Prima facie, sentence (166) talks of epistemic possibility. The word "must" in (167) can be read either as asserting the epistemic necessity of Smith's coming in through a skylight or, less likely, as taking wide scope and affirming the causal necessity of the whole conditional that if Smith got into the room not through a door or a window, then he came in through a skylight. In (168) the "have to" indicates a causal necessity given certain background conditions, and in (169) and (170) we get nomic or causal necessity. The last example could be read as once again giving an epistemic necessity, with the "must" again taking wide scope, but a more likely reading is that what we have here is an affirmation of the logical necessity of the conditional <If all men are mortal and Socrates is a man, then Socrates is mortal>.

One thing is grammatically striking about these ordinary usages. Not one of the unambiguously non-epistemic cases grammatically takes the paradigmatic alethic modal form so beloved of philosophers: "It is necessary that *s*" or "It is possible that *s*," where *s* is an assertion. The closest we get is (170), but there we have "It is impossible that *s**," where *s** is the subjunctivized version of *s*, and *s* itself does not literally occur. Whether this is of any significance is yet to be examined. But it looks like in English the use of the indicative after a modal operator appears to correlate with epistemic modality.

More significantly, observe how (170) sounds stilted, since one is more likely to use (169). But there is a significant difference between (169) and (170). The "impossible for" version suggests that human beings lack a certain ability, namely the ability to survive without air. The more awkward

"impossible that" version abstracts from any one being's capabilities. It is tempting to say that the "impossible for" version suggests where we should look for the grounding of the impossibility, namely in the human being or her nature, while the more "impossible that" version gives no such hint. Could this be why the "impossible that" version sounds more awkward?

What is also philosophically striking about these sentences is that only one of them, the last, uses a notion of logical as opposed to causal modality, and even that sentence allows for an epistemic rather than alethic reading. The logical necessity in (171) is an epistemological tool: it is used for the evaluation of the validity of an argument. There is a related way that (171) seems to differ from the other non-epistemic cases, though this is more subjective. Sentence (171) sounds like it is a meta-level, while the other cases are at ground level. In (171), we are evaluating our reasoning, holding it up to the standard of apodeictic argumentation. In the other non-epistemic cases, it seems we are talking of things themselves, their abilities and what is possible or not *for them*. The only possible exception here is (170), but its felt awkwardness supports the general point.

This is all quite unsurprising from a pragmatic point of view. We do not *need* logical necessity in daily life apart from the meta-level. One would not benefit in the wild from attaching much significance to the difference between its being causally or logically impossible to survive being eaten by a tiger.

Causal necessity and possibility occur most naturally in contexts where there is an antecedent relative to which we are considering the matter. *Were the apple dropped*, it would have to fall. The natural home of causal necessity is the subjunctive conditional, and this might have something to do with the above-noted fact that alethic "It is impossible that . . ." sentences are naturally completed in the subjunctive, as they are closely related to "Were it the case that . . ., it would have to be the case that . . ."

In any case, we see that there are at least three modalities in common ordinary use: epistemic, causal, and logical. Anything causally possible is logically possible, but the converse is thought by many philosophers to be false. Something can be logically impossible, but epistemically possible, for instance a false mathematical proposition whose falsity is not known. Something can be epistemically impossible while logically necessary, for instance a true mathematical proposition which is ruled out by evidence taken to be conclusive, say an apparent disproof with a very subtle mistake.

The logical form comes in at a meta-level. We now have two choices. We can boldly identify the logical and the causal, as in the unmodified Aristotelian–Leibnizian theory, or we can use the LRT to distinguish a more

strictly logical version of necessity, R*-necessity, which is so strong that some Gödelian arithmetical truths might well come out non-necessary, and the "broadly logical" account given by the Aristotelian theory.

Section 5   The Principle of Sufficient Reason

5.1   *From the PSR to the Aristotelian view of modality*

As we have seen in Section 2.5, the Aristotelian account of modality entails a version of the PSR, which says that, necessarily, if p is a contingently true proposition, then there is a causal explanation of why p holds. We can also argue for the Aristotelian account on the basis of the PSR, assuming S5.

If the PSR is necessarily true, then every contingent state of affairs ultimately has to derive from a necessarily existing First Cause. To see this, suppose that p reports a contingent and occurrent state of affairs that does not thus derive from a necessarily existing First Cause. By the PSR, there is an explanation, q_1, of p. If the state of affairs reported by p does not derive from a necessarily existing First Cause, then q_1 is a contingent proposition. Moreover, q_1 is not self-explanatory. For q_1 could be taken to report the activation of some causal powers which explains p — that is how we ultimately explain contingent states of affairs. These causal powers are *ex hypothesi* not those of a necessary being, and hence the existence of these powers is not self-explanatory. Thus, by the PSR, there is a further explanation q_2. And the chain continues forever. Form the whole chain of explainers of p. By the PSR, there must be an explanation of the chain as a whole, since this contingent chain is not self-explanatory as none of the links in it is.[13] But the chain was that of *all* the explainers. Hence we have a contradiction.

Thus, every contingent true proposition reports a state of affairs deriving from a necessarily existing First Cause which essentially has the power to cause that state of affairs. Moreover, if the PSR is necessarily true, this holds in every possible world. The necessarily existing First Cause from each world must also exist in the actual world by S5, and must have the essentially had causal powers, too. Therefore, every possible proposition that is not actually true is such that were it true, it would have to derive, perhaps through a causal chain, from the activation of the actual causal powers of something that actually exists.

---

13   For a discussion of Hume's objection that a chain of causes is sufficiently explanatory, see Pruss (1998b).

Therefore, if the PSR and S5 hold, all false possible propositions are Aristotelian-possible. Since all true propositions are automatically Aristotelian-possible, it follows that all possible propositions are Aristotelian-possible. The converse is just a conceptual truth. If there is something that can initiate a causal chain that can lead to its being the case that *p*, then *p* is possible, since clearly the causally possible is possible. Indeed, a paradigmatic way of showing a proposition to be possible is to show that something could make it come out true. Thus, if the PSR and S5 are true, Aristotelian possibility is co-extensive with metaphysical possibility. But by Ockham's razor, then, there is good reason to suppose the two are actually identical. Given the intuitive plausibility of S5, this shows that if we had good reason to accept the PSR, we would have reason to accept the Aristotelian account.

But should we accept the PSR?

The PSR has a venerable history. *Ex nihilo nihil fit*, goes the classic adage: nothing comes from nothing. Parmenides used the Principle of Sufficient Reason to argue that there was no such thing as change: If there was change, why did it happen when it happened rather than earlier or later? "Nothing happens in vain, but everything for a reason and under necessitation," claimed Leucippus (Diels and Kranz 1985: 67B2, my translation). Saint Thomas insisted in the *De Ente et Essentia*: "Everything, then, which is such that its act of existing is other than its nature must needs have its act of existing from something else" (Aquinas 1949: chapter 4). Leibniz thought the principle that everything that is the case has a reason why it is so and not otherwise was one of the two central principles of philosophy, the other being the principle of contradiction.

All these claims are closely related and have significant intuitive appeal. Each claim insists that an existent or occurrent thing has an explanation. The ordinary person accepts the claim: it is taken for granted that airplane crashes have causes. Some of the claims, like the *ex nihilo nihil fit* adage, limit themselves to saying that certain kinds of things, such as those that come into existence, have explanations or at least cannot come from nowhere. Others, like the claims of Leucippus and Leibniz are fully general and state the PSR that every true proposition has an explanation for why it is true.

All of the particular claims are special cases of the PSR. It is difficult to see what intuitions could be given to support any one of the particular claims without supporting the full PSR. There is little reason to think, for instance, that a contingent being has any less need for an explanation of its existence if it has existed longer, indeed if it has always existed, than if it has existed for a finite time (cf. Pruss 1998b). Thus, intuitive support lent to any of the versions of the PSR transfers to a significant degree to other versions.

I have given arguments for the PSR elsewhere (Pruss 2006, 2009).[14] The simplest is perhaps this. If the PSR is not true, then it is puzzling why it is that we do not observe objects popping causelessly into existence *ex nihilo* all around us. If we say that the laws of nature somehow manage to preclude that (though I do not know how one prevents a coming into existence *ex nihilo*), then put the question differently. Why don't supernatural, non-physical beings, not governed by the laws of nature, pop causelessly into existence *ex nihilo*, and then for no reason at all cause unexpected physical effects here? Since our laws only say what happens in the absence of external intervention, such things would not be violations of the laws. We cannot say that without the PSR such events have low probability. For while we can assign a probability to events that are produced by stochastic causal processes, assigning probabilities to causeless brute events does not appear to be a promising task. Moreover, the number of ways for such an event to happen is so large as to be beyond cardinality. For, for any cardinality n, it should be possible for exactly n supernatural beings to pop causelessly into existence, and then for no reason at all to produce unexpected physical effects. The PSR provides a simple explanation for why such things do not happen: they do not happen because they *cannot* happen. Thus, by Inference to Best Explanation, we arrive at the PSR.

The illustrious philosophical history and intuitive appeal of the PSR notwithstanding, the PSR is widely denied in analytic philosophy circles. One sociological reason for the denial is that the PSR implies the existence of a First Cause of the universe. This is not a very good argument against the PSR unless one has independent reason to believe that the kind of being that the PSR implies the existence of does not exist. The best argument for the PSR is an argument that the PSR is incompatible with the existence of contingent propositions.

5.2 *The van Inwagen objection*

Take the Big Contingent Conjunctive Fact (BCCF) which is the conjunction of all contingent true propositions. This is itself a contingent proposition. By the PSR, it must have an explanation. This explanation is a true proposition.

---

14 However, in that earlier work I defined the PSR as the principle that every contingent proposition has an explanation, without requiring the explanation to be causal. The other form of explanation that I allowed was reductive. However, I also argued in Pruss (2006, chapter 1) there that reductive explanations can't be all there is, so if the PSR is true, there will also be causal ones.

This proposition is then either contingent or necessary. If it is contingent, then this explanation will itself be a part of the BCCF — remember that the BCCF contains *all* contingent true propositions, and hence it also contains the explanation of the BCCF. But if this explanation both explains the BCCF and is contained in the BCCF, then *inter alia* it explains itself. But no contingent proposition can explain itself. Thus, the explanation of the BCCF must be necessary. But an explanation has to entail that which it explains: the *explanandum* must follow logically from the *explanans*. Otherwise, how does the explanation do any explaining? Thus, the explanation of the BCCF entails the BCCF. But anything that logically follows from a necessary proposition is itself a necessary proposition. Thus, if the explanation of the BCCF is a necessary proposition, so is the BCCF. But the BCCF is contingent. Hence absurdity ensues on this horn of the dilemma, too.

This argument is most clearly given by Peter van Inwagen (1983: 202–4), though it was in effect earlier made by James Ross (1969) and William Rowe (1975, 1984). As I formulated it, it has precisely two controversial premises:

(172) No contingent proposition can explain itself.
(173) No necessary proposition can explain a contingent proposition.

I shall argue that we ought not to accept (173).

The intuition behind (173) is, I take it, that if p is a necessary truth that explains a contingent q, then there are possible worlds where p holds but q does not, and hence it is mysterious why it is that in *our* world p is true but q is not. This mystery appears to depend on the principle:

(174) If p explains q, then any world where p holds is a world where q holds.

In other words, if p explains q, then p entails q. If the motivation behind (173) is (174), then we should not accept (173), since, as we shall see, (174) is false.

To see that (174) is false, observe that most scientific explanations fail to satisfy (174), for multiple reasons. First of all, contemporary science is indeterministic. The initial conditions and laws do not entail the explanandum. At best they make the explanandum *probable*, and sometimes they do not even do that. A standard example in the philosophy of science is the explanation of why the subject got paresis (Salmon 2006: 49). The subject got paresis because the subject had latent untreated syphilis. This

explanation holds even though only a minority of untreated latent syphilitic patients get paresis.

Second, scientific explanations involve *ceteris paribus* laws, so that even when the laws are deterministic, they typically do not entail the explanandum. Why was the rock falling at t_1? It fell because of such-and-such laws, and because at t_0 it was unsupported in a gravitational field. But the explanans does not entail the explanandum. If, for instance, the rock had a high iron content and there was a strong enough magnetic field pulling upward, the rock would be falling at t_1. Perhaps we should add the lack of such a magnetic field to the explanans. But there are other possibilities not yet ruled out. For instance, if there was a sufficiently strong upward current of air, the rock would not be falling at t_1. We could try to list in the explanans the negations of all such conditions. If Cartwright (2002) is right, this is a hopeless task. But hopeless or not, for our purposes all we need to note is that in fact it is a task that we do not in fact perform. We just give scientific explanations that do not exclude such possibilities, and leave it at that. Whether or not there could be a scientific explanation where the explanans entails the explanandum, most and perhaps all scientific explanations violate (174).

We could say that scientific explanations are not *really* explanations. But that is not satisfactory. For then we have to say that we typically misuse the word "explanation," a dubious proposition in light of the tie between meaning and use. But even if scientific explanations are not Real Explanations in some exalted sense, they surely are explanations in some more modest sense. Very well: stipulate the more modest sense to be the sense that the PSR involves. Besides, since the PSR is, to ordinary people, a very intuitive principle, it is charitable to interpret the notion of explanation in the PSR to be one that does not make the PSR incoherent, in the way that an exalted notion that entails (172) and (173) would. Furthermore, ordinary people's illustrations of the PSR will presumably all violate (174). For instance, one might illustrate the PSR by noting that Francine did not get sick for no reason; she got sick because she spent some time talking with someone who had the flu (cf. Anscombe 1999). But, of course, there is no entailment here.

Needless to say, even if (174) is the motivation for believing (173), one might still think that (173) is true even though (174) is false. But without another argument for (173), and given the intuitive plausibility of the PSR as well as the other arguments for the PSR,[15] this is not a very convincing move.

And leaving aside (174), we can argue that it is *not obvious* that (173)

---

15   See, e.g. Pruss (2006, 2009) and Della Rocca (2010), as well as Koons (2008).

is true. Consider a scientific explanation of some effect E. The explanation cites laws L and initial conditions I, and says:

(175) E happened because the laws are L and the initial
        conditions are I.

But, as in the above examples, that the laws are L and the initial conditions are I does not entail that E happened, for instance because the laws are *ceteris paribus* or stochastic. Next, consider a different explanandum, the material conditional $I \supset (E$ happens$)$. To explain that conditional, we do not need to cite I, but only L:

(176) $I \supset (E$ happens$)$, because the laws are L.

Now, it is *not obvious* that the laws are contingent. And if they are necessary, then the explanans is necessary, while the explanandum is still contingent, because the laws, although necessary, are still *ceteris paribus* or stochastic. Hence, (176) would be a counterexample to (173) if the laws were necessary.

Nonetheless, it is unlikely that (176) is what the ultimate explanation of the BCCF looks like. To that end, consider a model of what goes on in free will. Suppose that Smith is choosing, in a libertarian free way, between two incompatible options, A and B. Option A is supported by a set of reasons R while option B is supported by a set of reasons S. The reasons are all good, but the two sets are incommensurable. Thus, Smith can rationally choose A and Smith can rationally choose B. Roughly speaking, if he chooses A, that is explained by R and if he chooses B, that is explained by S.

But some more detail is helpful. Agents are *impressed* by reasons. To be impressed by reasons is to take the reasons into account in deliberation. Being impressed is something that happens to a greater or lesser degree. Then, prior to his choice, we may suppose that Smith is impressed by the reasons in R to degree d_R and by the reasons in S to degree d_S. Then, if Smith chose A:

(177) Smith chose A because he was impressed by the reasons in
        R in favor of A at least to degree d_R and by the reasons in S
        in favor of B at most to degree d_S.

But if Smith chose B instead, then:

(178) Smith chose A because he was impressed by the reasons in

Section 6 Ontology and implications

We have sketched the Aristotelian–Leibnizian account. Now it is time to work out some of the details.

Modalities on the Aristotelian–Leibnizian view are to be grounded in Aristotelian items such as causal capabilities. We have some basic idea of what these Aristotelian items are like, and it is an advantage of the Aristotelian–Leibnizian account that it is compatible with multiple ways of understanding the Aristotelian ontology, which we can leave for important future investigation. Nonetheless, we shall speculate further about the ontology.

Substances are found in various actual states, such as *seated*, *in Washington*, *green*, or *twice as tall as Fred*. I will use "state" as the most general description for what can be said of a substance, including both essential states (such as Fred's being human, perhaps) and accidental ones. For the Aristotelian view of modality to work, substances must have some states that have modal force, states such that their possession takes us beyond mere actuality. The most plausible candidates for such states are natures (or essences), dispositions (or propensities) and capabilities (or powers or abilities). I have previously talked of capabilities, but now it is time to consider other options, with an eye to getting clearer on what the capabilities we are talking about are.

Note that on an austere Aristotelianism, there are only substances, and talk of states reduces to talk of the substances that have them, while on a less austere Aristotelianism, states themselves exist (perhaps analogically and not univocally to substances). But we shall not need to resolve the question of which is the right form of Aristotelianism. Rather, we shall try to address the question of which kinds of Aristotelian states should be used to ground modality.

6.1 *Natures and other necessitators*

The Aristotelian–Leibnizian account is primarily based on *possibilifiers*. What makes a non-actual state of affairs possible is that there is something with a capability for producing a chain leading up to it. There is an alternate broadly Aristotelian account, in terms of *necessitators*. One version of this was in effect suggested to me by David Manley in discussion.[16] What makes

---

16 Actually, Manley's account is in terms of impossibilifiers, but that comes to the same

a state of affairs that is occurrent after a point t in causal history be necessary at t is that something existing at or prior to t has necessitated this state of affairs. States of affairs are always taken to be necessary at t whenever their occurrence is at or prior to t. Points in causal history, here, may be points in time but may include more general loci, such as the locations in the order of explanation of activities by timeless entities, if there are any. Something is necessary *simpliciter* provided it is necessary at every point in causal history, and possible *simpliciter* provided it is possible at some point in causal history. I will call these "constraint-necessity" and "constraint-possibility," respectively, in contrast to the "power-necessity" and "power-possibility" given by my preferred Aristotelian view.

Both the possibilifier and necessitator accounts answer to the Aristotelian call to ground modality in actuality. The two accounts differ, however, in what they take to be the ultimate positive items of reality. On the possibilifier account, making something possible is a positive property of an entity. To claim that something is necessary is to make a negative universal quantification: There is nothing which makes possible the negation. On the necessitator account, it is possibilities that involve negative universal quantification: p is possible provided there is nothing that necessitates $\sim p$, i.e. p is possible provided that nothing constrains things to be such that $\sim p$.

A good reason to prefer the possibilifier account is that there surely really *are* capabilities. I *can* ride a bicycle. Fred *can* sing. Maybe, too, the electron *can* spin up and *can* spin down. And capabilities are positive things that are explanatorily relevant to their exercise. Given the existence of and our epistemic access to capabilities, the possibilifier account is indeed going to be a viable account of modality. But it is harder to see what exactly necessitators are. What kind of a thing is it that makes it impossible, say, for a square to be circular? Moreover, since the necessitator account was supposed to be broadly Aristotelian, the necessitators are not Platonic things, but presumably work in some causal way. But how can something causally prevent squares from being circular? If, on the other hand, we say with the Aristotelian–Leibnizian account that nothing can make a square circle, this is perfectly comprehensible.

Could we not say that the necessitators are the essences of things (cf. O'Connor 2008)? What makes it impossible for Bucephalus to be a frog is then Bucephalus's equine essence. But this only seems to go so far. Consider the impossible proposition, p, that there is a member of a natural kind of essentially seventeen-legged animals which naturally has sixteen legs. This

---

thing: to impossibilify p is to necessitate $\sim p$.

proposition p concerns a non-actual natural kind, I assume, since there are no animals that are naturally 17-legged. Since no natural kind invoked in p actually exists, on Aristotelian grounds we cannot ground in the nature or essence the impossibility of this kind's members having 17 legs. Thus, p, while impossible, does not seem to be ruled out by the essence of anything.

But that was a little too quick. Perhaps there *is* something whose essence rules out p. It is difficult to see what kind of a thing this would be, however. While this is an argument from ignorance, I think the only option available to us is a being, such as God or an initial state of the universe, who is such that any natural kind of 17-leggers would have to be caused by it, and whose essence bars it from producing any such kind that contains a 16-legger that is normal. Now, this being's essence has to not only (a) bar it from producing any naturally 17-legged animal that naturally has 16 legs; but it has to be such as to (b) necessitate its being the case that all naturally 17-legged animals are caused by this being. It is not outrageous to suppose that the being's essence can do job (a). But how can it do job (b)?

It is implausible that there is any item other than God or the initial state of the universe that *must* be in the causal history of any naturally 17-legged animal. Let x be such an item. What is it that makes it necessary, then, that if there is a naturally 17-legged animal, it is caused by x? Given the apparent diversity of possible causes of an animal (evolutionary processes, laboratory design by aliens, random movements of molecules, etc.), it seems that the only way this can be necessary is if x is such that it must be the cause of all contingent beings. But what grounds the "must" in this sentence? The necessity whereby if there is a contingent being, it is caused by x is a constraint-necessity, since the impossibility of the naturally 17-legged animal having naturally 16 legs is a constraint-impossibility. If it were a mere power-impossibility, there would be no mystery: it would be enough for x to actually be the First Cause of everything for it to be the case that everything else that power-could exist power-must come from it.

But in any case, it seems we have come to a First Cause, x (perhaps an aggregative item if monotheism is false), which has the property that all contingent beings, or at least all other contingent beings, must arise from it. And the only plausible candidate for such an item is a being, like God or perhaps an initial state of the universe, that is such that in fact it must be the First Cause of everything. But if x is constraint-necessarily the cause of all constraint-possible things outside of itself, it seems that constraint-possibility is no wider than power-possibility, once we account for x's power.[17]

---

17 It might be thought that x's own existence is constraint-contingent, though it will of

Thus, x's own existence must be constraint-necessary. But if so, then constraint- and power-possibility coincide extensionally. For it is constraint-necessary that all constraint-possible things should come from x, and hence, constraint-necessarily, all constraint-possible things are power-possible. And the converse is, of course, trivial.

Thus, by Ockham's razor it seems we should cut away one of the two kinds of modality. Furthermore, it would be strange if two different kinds of entities respectively ground two different kinds of modality which then happen to coincide extensionally. Such a coincidence would call out for an explanation. And the best explanation would seem to be that claims about one kind of entity can be reduced to claims about the other. Since we have independent reasons to think there are powers — I have empirical evidence that I have an ability to hop, for instance — it is necessitators that should be reduced away, it seems, not powers.

But perhaps the advocate of constraint-necessity can avail herself of the LRT to rule out the existence of such a proposition as that x is a normally 17-legged member of a natural kind of 16-leggers. The difficulty here is that on the *Less* Radical Theory, it is only the self-evident propositions that do not have negations, while non-self-evident ones do. Let "f" be an abbreviation for a complicated mathematical function such that it is true, but not self-evident, that $f(3) \neq 16$. Then, consider the proposition <A member of a natural kind of 16-leggers normally has $f(3)$ legs>. The LRT has no way to rule out the existence of this proposition, since it is not a negation of a self-evident proposition. Hence, the argument above can be run again with this more complex proposition which it is hard to find an Aristotelian necessitator for.

On the possibilifier view, however, there is no difficulty with a normal $f(3)$-legged 16-legger. There just is nothing there that can initiate a chain leading to the production of that kind of a creature. That is all that is needed for ontologically *grounding* the impossibility of such a creature. And the ground itself is necessary, because there is nothing that could make it be the case that there is something that can initiate a chain leading to the production of that weird beast.

The above arguments make it implausible that necessitators on their own

---

course be power-necessary. But that won't do. For if x's own existence is constraint-contingent, then it is constraint-possible that x does not exist. Suppose, then, that we are at a world where x does not exist. Then, at that world it is constraint-possible that p, i.e. that there is a naturally 17-legged animal which naturally has 16 legs. But that world was constraint-possible. Hence, it is constraint-possible that it is constraint-possible that p, and by S4, it is constraint-possible that p.

ground modality. It is compatible with the arguments that necessitators and possibilifiers *combine* to ground modality. But Ockham's razor prefers a possibilifier-only account, unless it could be shown that that is insufficient. After all, working out how necessitators and possibilifiers interact and correlate would be difficult. Can something possibilify something that a necessitator rules out, for instance? This might lead to some questions similar to the ones about the correlation between capabilities and Platonic properties of possibility that helped lead us to the Aristotelian view (see Section 3.3.6 of Part IV).

6.2 *Dispositions*

A dispositional state of a substance directly determines how the substance would or might act if placed in a given non-modal state. Thus, sugar has the dispositional property of being soluble in unsaturated water at room temperature: were it placed in unsaturated water at room temperature, it would dissolve. A fair and indeterministic coin would have a dispositional property of coming up heads with probability ½ when tossed. Dispositions, thus, can be deterministic or indeterministic.

When the disposition is indeterministic, there will sometimes be an objective probability attached, but perhaps not always. For instance, Winston Churchill no doubt had a disposition to tend to reject bribes offered him by Nazis. Supposing this was an indeterministic disposition, so that he might have accepted such a bribe were it to have been offered, there is little reason to think that there is a fact of the matter about what exact probability is to be assigned to the action, unless one thinks on physicalist grounds that his choice would have been governed by quantum mechanical processes in his brain and hence an exact objective probability is assignable. But in any case, there is nothing absurd about something having a strong disposition to act a certain way without the strength of that disposition being precisely quantifiable. Moreover, the strengths of dispositions can be compared absent quantifiability. Churchill was more likely to reject a one hundred pound bribe than a one billion pound bribe. At the same time, it might be that there is incommensurability and Churchill's disposition to reject the bribe cannot even in principle be compared to something numerically quantifiable such as the likelihood that 20 fair indeterministic coins all land heads when tossed simultaneously.

Dispositions would seem to both possibilify and necessitate. Suppose Churchill has a *deterministic* disposition to reject the bribe. Then, this

disposition necessitates his rejection should he be offered the bribe under the circumstances involved in the disposition's antecedent. Presumably, the disposition to reject the bribe when offered is conditional on more than just the offering of the bribe; for instance, it is conditional on a lack of brainwashing. If his disposition is indeterministic, then we cannot say it necessitates the rejection, but at least it seems we can say that it possibilifies it, though still only *if* the conditions are somehow already possibilified.

But in fact it is a mistake to think of dispositions as possibilifiers. Consider an indeterministic disposition for x to tend to do A in circumstances C. What happens here is that were C to obtain, then x would be in such a state that it could, but did not have to, do A. In other words, were C to obtain, then x would have the capability to do A. And it is the capability to do A that is the last step before the actual doing of A, and hence, plausibly, it is the capability to do A that directly possibilifies A. And when C occurs, the disposition does not just possibilify the presence of the capability: it *necessitates* the presence of the capability, though it does not necessitate the exercise of the capability. So it seems that, in fact, if dispositions are going to play a role in grounding necessity, it will always be as *necessitators*.

But now consider one problem. If dispositions are going to be a part of the ground of modality, it does not seem that they can conflict. We cannot have a situation where x has a deterministic disposition to do A in C while also having a deterministic disposition to do non-A in C, where the circumstances C are possible. But *why* can't we have such a situation? After all, it is perfectly possible for x to have a deterministic disposition to do A in C and to do non-A in D, as well as to do A in C and to do B in C. Is it just a brute fact about the way dispositions are in fact coordinated with one another that the deterministic disposition to do A in C never coexists with a deterministic disposition to do non-A (or something more complex than non-A but that is clearly conceptually incompatible with A)[18] in C?

A good solution is to turn to capabilities, and say that the impossibility of warring dispositions is grounded in the fact that, as it happens, there has never been anything causally capable of giving rise to such dispositions. But then it is capabilities that we should be grounding modality in.

Another difficulty concerns modal claims about non-existent entities, such as that a unicorn could not be both a mammal and a non-mammal at the same time. A dispositional account of the impossibility of conflicting dispositions would require that there be entities which collectively have the

---

18 This parenthesis suggests that use of the LRT is not going to help — cf. the discussion of $f(3)$ in Section 6.1 above.

dispositional property of being such as to rule out the existence of a unicorn being both a mammal and a non-mammal. But this is a strange dispositional property, since it rules out not just what the entities that have the property can do, but what other possible entities could do. The only way, I think, of making sense of this is if the entities that have this property are such as to be First Causes of everything else, with its being *necessary* that they are First Causes of everything. Otherwise, they are not sufficient to secure the impossibility that a unicorn, not caused by them, is both a mammal and a non-mammal.

But what, then, is the ground of the necessity of the proposition that all other things are caused by these First Causes? Is it that the First Causes have the dispositional property of being such that in the circumstances of y existing, y had been created by one of these entities? But in a dispositional property, the consequent is posterior in the order of explanation, whereas that y is created by x is prior in the order of explanation to the proposition that y exists.

6.3 *Shoemaker's powers*

Sidney Shoemaker (1984) has also argued for the notion of a power as basic. It will clarify what the Aristotelian notion of a power is if we sketch Shoemaker's alternative. Shoemaker argues (though he weakens the claim in response to a counterexample of Richard Boyd's — see Shoemaker 1984: 232–3) that we need a distinction between a real property and a merely Cambridge property. A merely Cambridge property is the sort of property that acquiring or losing it does not entail a real change, such as the property of being taller than Callias, which one can gain or lose on account of change in Callias rather than in oneself. Shoemaker thinks that an account of the distinction requires one to bring in powers. Real properties are defined by facts about what clusters of them give rise to different causal powers. Thus, being knife-shaped is partially defined by the fact that when clustered together with being knife-sized and made of steel, we get the power to cut. In fact, real properties are second order powers: being knife-shaped is the power to make something that is knife-sized and made of steel be capable of cutting, and to … (where the ellipses stand for a list of all the other clusters that being knife-shaped enters into).

Shoemaker denies that this is a reduction of properties to powers. For the properties are *explanatory* of the powers. Part of the explanation of why a given steel object has the power to cut is that it has the property of

being knife-shaped. But, nonetheless, this appears to provide an independent argument for the notion of a power: powers are helpful for making a distinction we need, that between Cambridge properties and genuine properties. Unfortunately, as Shoemaker notes, things are not so simple: there are also powers that are merely Cambridge powers. For instance, one might observe that the power to win a beauty contest against Helen is a merely Cambridge power, and the Cambridge property of being more beautiful than Helen can give rise to and explain this power. However, the mutual interconnections between real powers and real properties do at least, Shoemaker thinks, help us understand both concepts.

Nonetheless, much as I welcome arguments for the existence of powers, this argument should be rejected. For it follows from Shoemaker's account that properties are causally explanatory and are properties in virtue of this. Consider, however, the property of being a good reason for me to sleep. This property is had by some but not all propositions. That I am tired is a good reason for me to sleep. That I am at an interesting colloquium is not. Now, the possession of the property in question might well be causally efficacious, and hence it is not a direct counterexample to Shoemaker's account. For a proposition's being a good reason to sleep when combined with its being known by me (one might worry that this is a Cambridge property, though) has the power to motivate me to sleep, perhaps. The problem here, however, is that the explanation goes the wrong way around. For on Shoemaker's account, the property of being a good reason to sleep should be *defined* by its powers. But something is not a good reason to sleep *because* when combined with other properties it motivates me to sleep: to say that would be to confuse a desire with a good reason. Rather, it motivates me to sleep *because* I recognize it as a good reason to sleep.

Shoemaker also gives an explicit account of what a power is. "For something to have a power . . . is for it to be such that its presence in circumstances of a particular sort will have certain effects" (1984: 211). Moreover, "it is possible for things to have the same power in virtue of having very different properties," e.g. as when two different poisons have the power to kill by affecting the heart and the nervous system respectively (ibid.). Shoemaker's powers, then, are much like what I have called "dispositions," in that his powers seem to be necessitators rather than possibilifiers. We have already discussed the pitfalls of basing an account of modality on dispositions, and so it does not seem that Shoemaker's powers would alone do the work for grounding modality.

Moreover, Shoemaker's powers are positioned in such a way that it is not the having of the power that is ultimately explanatory, but the having of the

properties that give rise to the power. For it is not a very good explanation of why a poison killed someone that it had the power to kill and that this power was activated: a better explanation is that it was such-and-such a substance which, say, attacks the nervous system and that the substances was activated in such-and-such a way. But the capabilities I wish to invoke are more fundamentally explanatory, since they are causally efficacious. Thus it is Shoemaker's properties, or rather the possession of them by things, that are closer to what I have previously called "capabilities or powers."

6.4 *Explanation*

The Aristotelian need not deny, *pace* Molière's sneer about the *virtus dormitiva* that explains why opium makes one sleepy, that the activation of the capability to φ can explain why something φs. Rather, she needs to distinguish between explanations and informative explanations. It is indeed due to the activation of the capability to φ that something φed. This is an explanation, but not an *informative* one. Why did Obama choose to run for president? Obama chose to run for president because of reasons that he had that favored running for president. This is true, but uninformative. When we were told that he *chose* to run for president, we already were in a position to know that he did so for a reason (randomness is not choice). What we want to know is what the reasons are. Likewise, the possession of the *virtus dormitiva* by the opium does not informatively explain why opium makes one sleepy. Moreover, as we now know, the *virtus dormitiva* of the opium is reductively constituted by lower-level capabilities that opium and nervous systems have. Mentioning *these* capabilities would give an informative explanation.

There may be cases where a non-informative explanation is the only one possible. This seems to be the case for tautologies: 1=1 because 1=1. This may be a genuine explanation because the proposition that 1=1 is self-explanatory, but it is certainly not an informative explanation. Typically, however, when causal capabilities are involved, there is more to be said than just that they were activated. Often, causal capabilities are not fundamental. But even when they are fundamental, typically (though not when this capability is one of divine omnipotence) there will be a further story as to how the capability came to be there.

6.5 Reduction

Both the possession of essential properties and the possession of dispositions necessitate, while powers possibilify. We have already seen that necessitators like dispositions and essences are insufficient for an account of modality: we need possibilifiers, and these seem to be capabilities. So capabilities had better not be reduced away into dispositions and essential properties.

If we are not seeking after a parsimonious theory, we can get away with just talking of the Aristotelian trio of dispositions, essences and capabilities, though even so we will have to say something about why they cannot conflict. Or we can try to reduce two of the items in the trio, or at least their alleged contributions to the grounding of modal truths, to the third item.

We can try to reduce talk of essential properties to talk of dispositions by talking of Bucephalus's essentially having horseness as consisting in Bucephalus's having the disposition to cease existing upon ceasing to be a horse. But this cannot be right. For the impossibility of a non-equine Bucephalus can only be grounded in a disposition *essentially* had by something. Otherwise, one could tell a story where Bucephalus first loses the disposition to cease existing upon ceasing to be a horse, and only then ceases to be a horse. It is only a disposition *essentially* had that can necessitate as it is supposed to, and hence one cannot reduce the essential to the dispositional. Nor will it help to bring in the dispositions of other things, even of God, such as the disposition not to allow Bucephalus to exist without being a horse, since for these dispositions to the job, they, too, would have to be essentially had.

It is also difficult to reduce dispositions to essences, since things can have non-essential dispositional properties, such as Patrick's dispositional property of tending to say what the voters want to hear, a property that Patrick might have lacked had he not gone into politics.

It now seems that if we are to have a reduction, it will be to *powers*. And there things look more promising. We can ground the impossibility of a non-equine Bucephalus in the non-existence of anything that can initiate a causal process leading up to a non-equine Bucephalus. In general, it seems, we can ground the essentiality of essential properties in this way. And there is something plausible about this. After all, a part of the intuition behind the impossibility of a non-horse being Bucephalus is that we cannot imagine how one could produce something non-equine that still counts as numerically identical with Bucephalus.

It is more difficult to reduce claims about dispositional properties to claims about capabilities or powers. To do this, we will need to say more

about what causal powers are like. But the first step of the story is easy. Consider a disposition of x to do E or F, in circumstances C, with precisely specified probabilistic tendencies towards E and towards F, where the disposition is indeterministic in respect of which of E and F is done, but deterministic in respect of the disjunction E-or-F being performed. Suppose that x is in fact in C at a time t. Then, x at t has the power to do E and it has the power to do F, while lacking the power to do anything incompatible with the disjunction. But to have the disposition to do E or F in the present circumstances while refraining from anything incompatible with the disjunction it is both necessary and sufficient that one have the power to do E and the power to do F and that one lack the power to do anything incompatible with the disjunction. Thus, the modal aspects of those dispositions whose antecedents' circumstances are actualized can in fact be reduced to claims about powers.

This leaves the reductionist with two problems. First, our reduction has lost sight of the probabilistic information about how likely x was to do E, information contained within the claim about the dispositional property. That is why it is only the *modal* aspects of the dispositional property that are captured so far. Second, this needs to be extended to cases where the circumstances in the disposition's antecedent are in fact not actualized — say, a case where a sugar cube is soluble but never actually placed in water.

For any dispositional property D, let D^* be a weaker property that drops all probabilistic data. Thus, if D was the dispositional property of doing E with probability 99%, F with probability 1%, and deterministically refraining from doing anything else, all in circumstances C, then D^* is the deterministic dispositional property of doing E or F in C. It is this weaker property that we can capture in terms of powers, though so far only in cases where C is actual.

But what about the lost probabilistic data? First observe that one *can* meaningfully talk, as one did, of a dispositional property having a probabilistic and a modal aspect. The modal aspect of necessitation is captured by D^*, and anything that has D also has D^*. The probabilistic aspect is whatever else there is to D beyond D^*. In fact, we can take the probabilistic aspect of D to be a property $D^\dagger$ described by replacing all claims of what is determined to happen with claims about what will happen with probability one. In the above example, $D^\dagger$ is the dispositional property of doing E with probability 99 percent, F with probability 1 percent, and anything else with probability zero. To see that $D^\dagger$ is weaker than D, note that $D^\dagger$ does not necessitate that E or F happens. It only says that with probability one, E or F happens. Probability zero is not the same as impossibility. For instance,

given a random process that picks a real number between zero and one in an unbiased way, the probability of *any* particular number being picked is exactly zero. But there *are* numbers which it is possible for the process to pick out: indeed, *all* numbers between zero and one are such. Thus, $D^\dagger$ is strictly weaker as a property than D, just as D^* is.

The conjunction of $D^\dagger$ and D^*, however, yields D. We can thus separate out the modal necessitation and the probabilistic aspects of a dispositional property. But if we can do this, then we can think of the probabilistic information as something superadded to a modal dispositional property. And we can do exactly the same thing in the case of powers. For there is no reason why a power, in addition to being such as to be capable of yielding some result, might not also be such as to be 99 percent likely to produce that result. Its having this probabilistic feature is something we do not need to ground modality, but it may have it nonetheless.

I will leave aside the question of what objective probabilistic features are. That would be a subject for another book. Modality is difficult enough. But that powers should turn out in the end to be more than just possibilifiers but also probabilifiers would not be unwelcome. If so, our theory can be extended to accomplish more with the same ontological commitments.

Our second problem with the proposed reduction of dispositions to capabilities was what to do about unactualized dispositions. Suppose that x is a sugar cube with a disposition, D, to dissolve in the circumstances, C, of being placed in unsaturated water at room temperature, together with any other provisos needed to ensure C is a normal set of circumstances for actuating D, and suppose that x never finds itself in C. Presumably, the claim that D grounds is a claim about the impossibility of x's being in C without dissolving. But note that we have not said that x has D essentially. Thus, it possible that x is in C without dissolving, but only when x does not have D. Therefore, it is actually incorrect to say that x's possession of the dispositional property as such grounds the impossibility of x not dissolving in C. (This observation reminds one of standard cases of finkishness; see Martin 1994.)

When we thought of dispositions as necessitators, that was because we thought of the disposition D as offering a grounding for the necessity of x's dissolving in C when having D. But note that here the *actual* possession of D by x is completely irrelevant. Whether x is water-soluble now or not, were it placed in unsaturated water at room temperature *while* being water-soluble, it would have to dissolve. But this claim can be reduced to saying that if x is in a position where it lacks the power to remain solid in unsaturated room temperature water (i.e. where it has D) and it is in unsaturated room

temperature water, then (necessarily) it dissolves. The necessity here comes from the lack of power to remain solid. One can use power-talk here to accomplish exactly the same thing as disposition-talk did.

But perhaps the dispositional account can do more. For not only is our sugar cube disposed to dissolve in water under appropriate conditions, but it has a second-order disposition to *maintain* this first-order disposition. And maybe is this *second*-order disposition that grounds the necessity of its dissolving in C. Moreover, the problem of grounding things in the first-order disposition does not return, because this second-order disposition has *its* conditions actual. It is a disposition to have the first-order disposition in the future given present circumstances being as they are. Presumably, too, in order to avoid a regress, the second-order disposition also contains a disposition to maintain itself. But we have already seen that we can give a power-based account of the necessitation done by dispositions whose conditions are in fact met, and the second-order disposition is one whose conditions are met. So we can just say that the sugar has the power to maintain its state of having no power to resist dissolving and no power to act contrariwise. Or we need not even suppose the power to maintain the state — we can just suppose that nothing has the power to make it swerve from that state.

Let us now consider cases where the disposition to dissolve is essentially had (maybe it is only essentially had dispositions that are necessitators). This *essential* possession of the disposition is indeed sufficient to ground the impossibility of *x*'s remaining solid in C. But now it is the *essentiality* rather than the dispositionality that is doing the modal work. After all, one can just think of *x* as essentially having the non-dispositional disjunctive property of either being dissolved or being outside of C, and that will be sufficient to ground the necessity involved. It may be true that the disjunctive property is there because of the dispositional one, since presumably there are no primitive disjunctive properties. But this only says that the possession of the dispositional property explains the possession of the disjunctive property. Since the disjunctive property as such lacks necessitating force, and since all the modal force needed here is provided by the essential possession of the disjunctive property, it seems that the dispositional property, even if it is what explains the possession of the disjunctive property, need not be thought of as doing any *modal* work in grounding the given impossibility.

And we have already seen that we can account for the necessities involved in essential predication in terms of powers, or, more precisely, the lacks of powers. We can just say that nothing has the power to make *x* lose the dispositional property of dissolving in water in appropriate circumstances.

And this dispositional property is itself to be explicated in terms of powers or the lacks of them. Thus, the whole claim becomes that nothing has the power to make x have the power of remaining solid in water in appropriate circumstances.

6.6 *Some assembly required?*

We have already seen that the possibilifier for a single entity, say a statue, might include several things. There is the marble, the sculptor, her abilities and plans, her tools, and a favorable environment. But how do these different possibilifiers conspire to make possible a single entity, the statue of Hermes? The possibility of the Hermes is not grounded in any one of these alone. Is the power for making the Hermes spread out through them? Is the possible hardness and color of Hermes grounded in the marble, while the possible shape is grounded in the sculptor, her abilities and plans and her tools? But these things cannot be separated. Without hardness and color, there is no shape, either.

One option is to say that the various items *jointly* have a capability.[19] This is a promising suggestion, but let us explore whether there couldn't be a bolder one. The bolder option is to say that each of the ingredients grounds a different aspect of the final result, even though these aspects are interdependent. The better to understand the case of Hermes, consider a simpler case. A piece of paper is divided up into a 1000×1000 grid. Each grid square is assigned to a different one of a million people. Each person is told where her grid square is and is given a device whereby she can either fill the square with whatever color she wishes, and she does this independently of the other people, without being able to find out what they are doing. It is possible, though highly unlikely, that the result of this will be one particular exquisite picture of Hermes in 1000×1000 resolution, call that picture H. And let us look at only one aspect of the problem: The question of how it is that the 1000 possibilifiers, each of which is a person with a tool, conspire to make possible H's showing up on the page.

An obvious proposal is this. For any given square, s, of the grid, H would have some color on s, which we can denote $H(s)$. Then, s's having color $H(s)$ is made possible by the powers of the person and tool assigned to s, and since this is true for every s, H's showing up on the page is made possible by the

---

19   cf. Martin's (1997) talk of "reciprocal disposition partners."

powers of all the persons and tools. After all, it is necessary and sufficient for *H*'s being there on the grid that each square *s* should have color *H*(*s*).

However, this suggestion assumes that if we have possibilified every conjunct, namely every proposition of the form *s* has color *H*(*s*), we have possibilified the conjunction. But it is false that by possibilifying every conjunct one has possibilified the conjunction. After all, it is possible in general for a conjunction of possible propositions to be impossible: if *p* is any contingent proposition then *p*&~*p* is a conjunction of possible propositions. In the example at hand, if there were only one square's worth of paint of each color available, then for any *s* that square's having color *H*(*s*) might have been possibilified, but if there are two or more squares of the same color in *H*, then we would not have thereby possibilified *H*. Of course my formulation above rules this out by saying that the persons and tools act independently. We can now see that this independence assumption is essential to the possibilification. But what this means is that just as in the original case of the statue we needed to take into account the interrelations between the causes, we must do the same here, except now only in a negative way.

Therefore, it seems that the possibilifier must involve the joint capability of persons and tools, in their interrelations, or else the capability of a mereological sum (or arrangement, as in Rasmussen 2010). This is not entirely a happy conclusion. But grounds of propositions often will be pluralities or mereological sums — take, for instance, any ground of <There are at least two horses> — and so this just something one may have to live with.

It does seem, thus, that even in a case where the causes are apparently working independently, they need to be considered as a system that is doing the modal grounding. *A fortiori*, we will need to consider that which does the grounding in the case of the statue to be a system. Aquinas gives this example:

> no one man's strength is sufficient for moving the boat; *while all together are as one mover* [sed omnes simul sunt in loco unius motoris], in so far as their united strengths all combine in producing the one movement. (Aquinas 1948: I, 52, 3, italics added)

The next question, then, is whether we can ontologically ground the powers of a system in the powers of the constituents. Intuitively, at the very least facts about the collection of independently working people should reduce to facts about the individuals. And that appears to be right. The power of the system as a whole to make the picture *H* show up on the page is grounded in the powers of the individuals to do their bits in a way that does not interfere

with the relevant powers of the other individuals. Thus, the powers must be described more carefully.

If one wants to completely reduce modality to the powers of single individuals, one will have to say something more, though, about cases where several individuals cooperate in a non-independent way, say as when the sculptor, the chisel, and the block of marble cooperate. One way to do this is to insist on the Aristotelian distinction between agents and patients.

This may seem to be untenable, given Newton's third law which seems to imply symmetrical cases where there is no distinction between agent and patient — as I push the wall, the wall pushes me, and there is no difference between pushing and being pushed. But the situation is not hopeless once one realizes that two items can both be agent and patient in a given interaction. Moreover, given post-Newtonian science, it is not correct to say that in pushing against the wall, the wall is pushing back on me. Influences from me propagate at some speed not exceeding that of light and there is a gap between my molecules and those of the wall. Thus, the effect I have on the wall at time t is not actually caused by what I am doing at time t, but what I was doing at time $t-\delta_1$. Likewise, the wall's effect on me at t is caused by the wall's activity at time $t-\delta_2$. This means that if we localize agents and patients to specific moments in time, we may not even have to say that we have cases of the agent being acted on by its own patient. But even if we do have to say it, it does not seem a very hard thing to say.

We might, then, take a power to be an ability of a particular agent to produce a particular effect in a particular patient or kind of patient. In many cases of collaboration, the collaborating causes are not in fact brought to bear on the same particular patient if the patient is sufficiently localized in space and time. Rather, they are sequentially brought to bear, though perhaps in very close temporal succession, or brought to bear on different parts of an agent spatially spread. In the case of the sculptor, the sculptor does not directly bring her power to bear upon the block of marble while the chisel is bringing its power to bear upon it. Rather the sculptor directs her power at the chisel, thereby enabling the chisel to strike at the block. Of course, she might with another hand be doing something else to the block. But that activity is not directed at that precise point where the chisel is.

What if two people are simultaneously pushing the same block, and are doing so in such a manner — obviously physically impossible given our laws — that they are pressing the block at the same point? Then, indeed, we would have two agents producing one effect, one movement of the block. Nor can we say that this movement is composed of two smaller movements, since then we would be asked which part of the movement is caused by

which agent, and if they are pushing in the same direction, the question is senseless, just as it is senseless to ask: "Which half of the speed of 60 miles per hour was caused by your additionally pressing down the gas pedal when moving at 30 miles per hour?"

But what we *can* say here is that each agent contributes kinetic energy to the block. We cannot say which "part" of the kinetic energy is contributed by which agent, but we can say that each agent brings it about that the block has such-and-such an amount of kinetic energy over and above what it would otherwise have had. In other words, the power that the agent has is an indefinite power — what effect the power brings about depends on what else is acting, what other agents are involved (this is the point of the central arguments of Cartwright 2001).

One may still ask: What is it that explains the necessity of the fact that when a block is stationary and, in the absence of other causes, each of two agents brings it about that the block has, say, 100 Joules more kinetic energy than it otherwise would, then the block comes to have 200 Joules of kinetic energy? On the plain Aristotelian–Leibnizian account, we could simply say that there is nothing that can make it not be so. Or we could invoke LRT and say that it is just a matter of logical grammar given that the agents are giving the block 100 Joules more energy *than it would otherwise have had*.

On the Aristotelian–Leibnizian account, all modality is grounded in the powers of items. It is a difficult substantial question whether we need to consider the powers of substances acting together in systems as irreducible or as reducible to the powers of individual substances. How we answer this question will not, however, affect the basic account.

6.7 *How we know what is possible*

The Aristotelian–Leibnizian account does a good job, indeed better than competing accounts, of our knowledge of local possibilities, such as that of Hitler never having existed. Possibility is nothing but capability, and we know the capabilities of human beings. Thus, we know that Hitler's parents had the capability of remaining abstinent on the occasion when Hitler was conceived, and that he probably or even necessarily (see Section 7.2.2, below) would not have existed had he not been conceived, and so we know that Hitler could have failed to exist. Likewise, I know that it is possible that I finish this sentence, because I know my capabilities and that finishing this sentence is among them. It is possible for there to be horse-like animals with a single horn because a plausible evolutionary story could be sketched

under which they would evolve. In fact, it seems that in general our best bet for arguing that something is logically possible is to show that it is physically possible according to our best scientific theories, and this technique coheres very well with the present account.

If one insists that knowledge in each domain must have some causal cases, some cases where the object of knowledge causes the knowing, then one will be pleased at the fact that *causal powers* are the sort of thing that can sometimes be known causally: a causal power can cause us to know of it, by being activated and the activation directly or indirectly causing our knowledge.

Moreover, the Aristotelian–Leibnizian account is perhaps one's best hope for an account of knowledge of necessary truths that has a *causal* component, if one should want that. The necessary truths are ultimately grounded in the nature of the power of God, and his power is causally relevant to everything in this universe, and in particular to our thinking about necessary truths. Leibniz had an illuminationist epistemology of necessary truths according to Adams (1994, Section II.7.2.2), and on such an illuminationist epistemology, the ground of the necessity of p can motivate God to cause us to think that p is necessary, and hence the ground of the necessary truth helps to causally explain why we think that p is necessary.

6.8 *The structure of the space of possibilities*

The space of possibilities is intricately structured. Some of the structure of the realm of possibility is revealed in the intricate beauty of the mathematical realm, where, arguably, mathematics is a study of certain kinds of necessity.

This beautifully intricate structure gives rise to an explanatory demand. Why is logical space so regular? It seems most unlikely prima facie to have such a deeply regular structure. A random assignment of *necessary, possible,* and *impossible* to the space of propositions would be unlikely to give rise to such a structure. On the Platonic, Lewisian, and pure Aristotelian views, the order should be a surprise. Prima facie, why should we expect the Platonic realm, the collection of concrete universes, or the arrangement of causal powers in our history to give rise to such an order?

Admittedly, the explanatory demand here seems strange. There are no alternate possibilities for the space of possibilities! But take a parallel case. Imagine that we find on the moon, under the dust, a gigantic golden inscription, which, in fact, spells out all of *Hamlet*, in English. Scientific investigation reveals that it is a basic law of nature that Hamlet be inscribed

in gold on every moon of a third planet of a non-binary stellar system. It would surprise us that such a specific and macroscopic proposition should be a basic law. But so it is: the evidence leaves us no choice but to accept it. Suppose, further, that those philosophers who contend the laws of nature are necessary have conclusively proved their case. Would this lessen our surprise at the inscription or at the law that necessitates it? We still would want an explanation, though perhaps we could not have one. And the demand for explanation is quite intelligible here.

We might have to admit that there is no explanation for the structure of the space possibilities. But this should be a last resort.

Now, one explanation that could be given is a selection effect. We see many beautiful mathematical truths because we *select* for beauty. A logician once showed me a referee's report he got for a mathematical paper. The report said, as best I can remember: "The paper is correct and interesting. Publish." Interest is not just newness. It is easy to generate correct and new mathematical results: just randomly generate two 20-digit numbers, multiply them together, and write down the resulting multiplication equation. But these results are not *interesting*. Mathematicians try to find beautiful and interesting connections, and the space of mathematical entities is so rich that such connections can often be found.

This may or may not be a satisfactory explanation. But even if it is a satisfactory explanation, if the Aristotelian–Leibnizian account, can give *another* explanation compatible with this one, that will most probably result in some "residual confirmation" for the Aristotelian–Leibnizian account (cf. Glass 2009). Let S be the fact that mathematicians select for beauty. Let B be the unified and beautiful structure we have seen in the realm of necessity. If the Aristotelian–Leibnizian account can give an independent explanation fully compatible with S, then given *both* the Aristotelian–Leibnizian account's explanation *and* S, will make B even more likely than S alone: $P(B|S\&AL) > P(B|S\&{\sim}AL)$. And so B will provide evidence for the Aristotelian–Leibnizian account even given S.

And indeed the Aristotelian–Leibnizian account, offers hope for a further explanation of B, fully compatible with S. The ultimate grounding of possibility is in the powers of God. These powers derive from the nature of God. If the nature of God is ontologically simple, then it is plausible that there be a unified explanation of the space of possibilities — though perhaps one only knowable by God. And this unified explanation would make for a deep structure. Moreover, God, in whom modality is grounded, is traditionally said to be supremely beautiful. The picture then would be neo-Platonic: the emanation of the orderly multiplicity of the space of possibility from the

One. This is not very much confirmation for the Aristotelian–Leibnizian account, but is additional evidence for it.

Section 7 The main challenges to Lewisian and Platonic ontologies can be resolved

Observe that the challenges to Lewisian and Platonic ontologies can be answered on the Aristotelian–Leibnizian account. I will discuss some of the major ones below.

7.1 *Primitive modality*

The present view certainly depends on a primitive Aristotelian modality: substances' capabilities for bringing about effects. However, this primitive modality is not objectionable as it can be argued, *pace* Hume, that we have a direct grasp of Aristotelian modality. Leibniz has claimed that because we are substances, we have a grasp of what it is to be a substance. Similarly, because we are, or at least experience ourselves to be, active substances, we have a grasp of what it is to bring about effects (cf. Anscombe 1999; Swinburne 1997). But even more than this, we have a grasp of what it is to be capable of doing something. This grasp is, for instance, in play when we see ourselves as responsible for having neglected a duty. Our knowledge of ourselves as guilty involves a knowledge of ourselves as having neglected something we had a capability of doing, much as Kant had argued in the second *Critique*.

Lewis objected to Platonic accounts of modality as having a *magical* relation of representation. The problem was that there are many potential relations one could take as *the* relation of representation, and we have no way of picking out which one. I have argued against Lewis in Section 3.3 of Part IV, but even if my defense of the Platonists fails, the Aristotelian has no such problem. We can specify what kind of a thing a causal power is by ostension: we meet with many cases of the exercise of causal power, and then ostend to that which lies explanatorily behind them all. If this works, then the "grasp" of the concept of a causal power is not more objectionably magical than our grasp of the concept of water. What would be objectionably magical is the specialness of a property such that we cannot give a story about what distinguishes *this* property from all others: a property that is supposedly special but yet one which we cannot individuate from all other properties.

It might be argued that, on the contrary, by taking the causal chain of modalities back to God, I have brought in the *mysterium tremendum et fascinans*, and there is nothing more mysterious than that! However, the mystery involved in God is of a different kind than that which is found in Platonic views. The objectionable "mystery" would be an inability to pick out which particular relation is the one that is involved in one's theory. In the case of God there is no parallel problem, because we can pick out God uniquely, e.g. as "The First Cause of all contingent beings" or "A personal being that is necessary." Further objections concerned with the mysteriousness of God will be considered in Section 8.4.

7.2 Transworld identity

We no longer need to espouse the paradoxes of counterpart theory, because the only serious objections to identity theory in Section 2.1 of Part III were based on the Lewisian assumption that all possible worlds concretely exist, which is no longer the case. But we still need to sketch what an Aristotelian can say about transworld identity. In fact, we will sketch *three* options.

7.2.1 Identity option 1: haecceities
How exactly on the Aristotelian–Leibnizian view one explains transworld identity will depend on whether one thinks there are haecceities; i.e. individual essences or thisnesses. If for every possible individual there is a haecceity, a property that exists in every possible world, such that the individual has that haecceity essentially and necessarily no one else does, then we can simply define an identity between the individuals in terms of identities between haecceities.

But of course the theory of haecceities is controversial. For instance, I read Adams (1981: 11) as observing that there is a problem about the relation of me to my thisness; i.e. to the property of being me. It is not just that I have my thisness essentially. Rather, I have my thisness essentially and my thisness essentially has the property of not being had by anyone other than me. Thus my thisness has a property that depends on me *de re*. It will not do to specify me *de dicto* here; e.g. by saying that my thisness essentially has the property of being had only by beings that have the property of being me, since that would be just making the trivial claim that my thisness essentially has the property of being had only by beings that have my thisness, and that would not distinguish my thisness from *having green-dyed hair* since *having green-dyed hair* essentially has the property of being had only by beings

that have green-dyed hair. But if my thisness has a property that depends on me *de re*, then it cannot exist if I do not exist. Alternately, my thisness is essentially related to *me* and hence can only exist if I exist since nothing can be related to what is not. Therefore, my thisness cannot be a haecceity, since a given haecceity exists in all possible worlds. But my haecceity, were it to exist, would indeed be my thisness, and if I have no necessarily existing thisness, I have no necessarily existing haecceity.

The haecceitist might try to bite the bullet and say that if *H* is my haecceity, then I simply am *that entity, whatever it may turn out to be, that has H.* And there is no need to further specify that *H* can only be had by me, because that would be tautologous. This would in fact allow a reduction of *de re* modality to *de dicto* modality: the claim that I could not be a horse could become the claim that, necessarily, anything that has *H* is not a horse. Unfortunately, this leaves unanswered the question of what distinguishes haecceities from other properties. Not all properties give rise to a trans-world identifiable entity in the way that my thisness, *H*, was supposed to. For instance, let *F* be the property of being the uniquely tallest woman. Then, in every possible world where there is an instance of *F*, that instance is the only instance of *F*. But we do not want to say that there is some entity which in one world is Smith and in another is Jones, depending on who is the tallest woman in the given world, unlike the case of the haecceity where we *do* want to say that the same entity is picked out in all worlds. The haecceitist will have to bite the bullet and reject the call for an explanation of the difference between the haecceity and other properties, taking it to be primitive.

7.2.2 Identity option 2: the identity of indiscernibles and a causal-historical account of identity

If we reject haecceities we have a major problem with how there can be possible worlds that contain non-actual individuals. The possible worlds of the Aristotelian–Leibnizian theory are supposed to be complete representations. But then the worlds have to specify which individuals are in them. Moreover, since the worlds are divine ideas, God in grasping a world has to have a grasp of the identities of the individuals in them. But how could he, if some of these individuals don't exist and there are no haecceities?

In fact, without haecceities it is difficult to see how on the Aristotelian–Leibnizian view one can avoid a PII that states that there cannot be numerically distinct indiscernible individuals. Here, indiscernibles are individuals that are exact duplicates, being exactly alike in all "purely qualitative" properties, i.e. properties that do not reference the identities of particular individuals. For suppose that some world, w_0, contains two indiscernible individuals, say x_1

and x_2. Then, unless there are haecceities, the two individuals are represented in the same way by the divine idea that w_0 is — if they were represented differently, the difference in representation would make them discernible. Suppose w_1 is another world containing at least one of these two individuals, and let y be one of these two individuals in w_1. Then there must be a fact of the matter as to whether y is identical with w_0's x_1 or with w_0's x_2, a fact that had better be a part of the content of the worlds if they are representationally complete. Yet there cannot be such a fact because w_0 represents x_1 and x_2 in the same way and hence any reason for taking y to be identical with one is also a reason for taking it to be identical with the other. In other words, if we reject haecceities, it seems we must reject the starting assumption that w_0 contains two indiscernible individuals, and hence it seems we should accept the PII. But the PII is very controversial.

If the proposed approach were committed to haecceities or PII, that would in fact count to some degree against it. Haecceities are controversial, while it seems prima facie possible for there to be two distinct individuals that are indiscernible. In regard to the apparent possibility of indiscernibles, one answer is this. On the Aristotelian–Leibnizian account, given that all things ultimately come down to one cause, if there are to be indiscernibles, then somewhere along the line the same cause acting at the same time and in the same manner would have to produce two indiscernible effects. Apart perhaps from some sub-atomic phenomena which can be variously interpreted (we might, for instance, take all particles of the same type to be a single multilocated particle), we do not have cases of the same cause acting at the same time and in the same manner to produce two indiscernible effects. Thus, given a causal account of possibility, the source of confidence that indiscernibles are possible cannot be empirical observation. Rather, it must be from some intuitive notion that such a thing would be possible for an omnipotent God. However, it is not clear that such a thing would indeed be possible for God. For it is not clear whether it would make sense to say that God has instantiated some complete individual concept C (a complete description of a possible individual) more than once as opposed to just instantiating some other concept once (cf. Hacking 1975; Pruss 1998a).

Admittedly, one feels an intuitive plausibility to the possibility of, say, two distinct but indiscernible consciousnesses that are non-embodied, or embodied in a single body, or embodied in two indiscernible bodies. But unless we can give an account of how they could be caused in an indiscernible way by indiscernible causes or by one cause (since otherwise one could distinguish them by their causes), the defender of the Aristotelian–Leibnizian view should simply reject this intuitive plausibility. The intuition may count

against the account, but it is not decisive. (And if one thinks it is decisive, one should go for Identity Options 1 or 3, instead.)

But in fact, even without haecceities, we can try to turn this issue into an asset for the proposed theory by also giving an explicit account of transworld identity that naturally comes out of the spirit of the Aristotelian–Leibnizian account. The usefulness of having such an account will count in favor of the theory, outweighing the somewhat counterintuitive consequence above which needs to be tolerated if we are to have a theory of transworld identity. The proposal here is that what individuates individuals are precisely *their origins*. One formulation is this:

(181)   individual x is identical with individual y if and only if
          x and y are initially exactly alike and the chain of all the
          causes leading up to x is exactly like the chain of all the
          causes leading up to y, both in what the causes are like and
          in how they acted.

Here, we might say that "how [causes] acted" includes what purely qualitative properties (i.e. properties not depending on identities of individuals) they caused their effects to have.

We can affirm (181) whether or not we think of x and y as existing in the same world. Thus, x is in w_1 is the same individual as y in w_2 if and only if the chain of causes of x in w_1 is exactly like the chain of causes of y in w_2. This is a significant strengthening of Kripke's claims about essentiality of origins (Kripke 1980; Mackie 1998). If x is uncaused, we can stipulate "the chain of causes of x" to be empty.

An interesting corollary is that if x is uncaused, then y is identical with x if and only if y is uncaused, and hence there is only one possible uncaused being, which gives another nice argument for the unity of the First Cause. If we do not want to prejudge the latter question, we might add to the right-hand-side of (181) the condition that the *initial* purely qualitative properties of x and y are the same, where a property P of x is initial provided that there is no property Q of such that x's having Q is explanatorily prior to x's having P.

One might of course argue that it *is* logically possible for the same human being to have been conceived by different parents, *pace* this view. I know of no better answer to this than Kripke's (cf. Kripke 1980: 114n56). Suppose in the actual world w_0 I am conceived by TW (my father) and IW (my mother) and a world w_1 where I am conceived by AB (a man) and CD (a woman), who are not identical with TW and IW, respectively. Now, surely there is a possible world w_2 in which TW and IW have a child at the same time as they

do in the actual world and which child has the same genetic makeup as I do, while AB and CD *also* have a child at the same time as they do in w_1. But then, surely, that child that AB and CD have in w_2 is the same child as they have in w_1. After all, the actions of TW and IW should not be relevant to the identity of the child of AB and CD (assuming enough causal isolation). And likewise, surely, that child that TW and IW have in w_2 is the same child as they have in w_0. Hence, in w_2, I am the child of AB–CD and I am the child of TW–IW, since it is I whom AB and CD have in w_1 and whom TW and IW have in w_0. Hence in w_2 I am conceived by two different sets of parents — and that surely is absurd.

The above argument only shows that a person could not have been conceived by a pair of parents having no one in common with the parents that she in fact has. But given S4, this is enough to show that the exact identity of the pair of parents I am descended from is an essential property of me. For suppose that I could have had a different mother. Well, by the same token, that person with the different mother could have had the same mother that he does, but a different father. By S4, it follows that I could have had a different father and mother, contrary to the above argument.

This does not show that *all* the parts of my causal history are essential to me, in the way that (181) claims. But it seems hard to draw any non-arbitrary line as to which parts of my causal history are and are not essential, and hence it is reasonable to take them all to be essential.

Of course, this approach to individuation builds PII in, and hence intuitions against PII are intuitions against this view.

However, the view has three significant theoretical assets. First of all, it is a simple and clear account of the identity of individuals within and between worlds, giving determinate answers to questions about transworld identity.

Second, this historical view account of identity significantly reduces the number of apparently unexplained facts. For instance, if someone else could have had the very same origins as I, then there is a question as to why it was that *I* resulted from these origins instead of that other person. It may be that God's appreciation of the value of one particular haecceity — despite the fact that another haecceity has an exactly similar value — explains why it was that he had me originate from my parents rather than that other person, but a neater and simpler story is that there is nothing here to be explained, as anybody with exactly the same origins as *I* would be identical with me.

Third, the view solves Chisholm's (1967) Adam–Noah problem, the problem that one could imagine a continuum of worlds where an individual occupies a role that in the first world is that of Adam and in the last is that of Noah, and where the roles change slightly in between. The puzzle is that,

plausibly, the first world represents the individual as being Adam and the last as him being Noah, and any two worlds next to each other in the series seem to represent the same person as the roles are very similar. But on the present view, what matters is exact likeness. As long as the different roles include exactly similar histories, we have the same individual, but as soon as anything is changed, we have a new individual. Thus, the first individual is Adam, the last is Noah, and presumably there are many other possible individuals in between. Divers (2002: 263) writes that if one denies that the last individual is Adam, one has "a mystery about representation de re." But if one accepts the historical view, there is no mystery.

7.2.3 Identity option 3: divine ideas as differing between worlds
Alternately, one might deny PII and try to resolve the problem that indiscernibles pose for the proposed account of possibility by denying that the divine ideas that constitute possible worlds encode all the information about possibilities. What the divine ideas sometimes fail to encode is information about transworld identity for non-actual individuals.

To see how an account along these lines would work, consider Max Black's (1952) counterexample to PII, modified to be theistic. We have a world, w_1, at which there are two exactly alike iron spheres with set of intrinsic properties S, and nothing else but God. Now, the following is also possible: we have a world just like w_1 but where only one iron sphere with intrinsic properties S exists. Suppose now that w_1 is actual. Then the two iron spheres are actual, and there will be *three* worlds, w_{11}, w_{12} and w_{13}, corresponding to this possibility: one of w_{11} and w_{12} will have one of the spheres of w_1 and the other, the other, while w_{13} is a one-sphere world with a sphere numerically *different* from either of those of w_1. From the point of view of w_1, there are three possibilities corresponding to the one-sphere hypothesis. God represents these two of these possibilities by means of different thisnesses, which thisnesses relatively unproblematically depend on the identities of the actual individual spheres, and represents the third by means of the negation of the disjunction of these two thisnesses.

Suppose, on the other hand, a different world, w_0, were actual, a world that does not contain either of the spheres. Now, relative to w_0, there is no way to represent the difference we had between w_{11} and w_{12}. Instead, when we consider the possibility of a world containing nothing but one iron sphere with properties S, we have only one possibility, w_{01}, a world consisting of one iron sphere just like either of the iron spheres in w_1. There is no fact of the matter at w_0 as to which of the spheres of w_1 exists at w_{01}, and the divine ideas w_{11}, w_{12} and w_{13} do not exist in w_0. Thus what worlds or divine ideas

there are depends on what world is actual: when w_0 is actual, we have only one one-sphere world, namely w_{01}. But when w_1 is actual, there are three one-sphere worlds: w_{11}, w_{12} and w_{13}.

Nonetheless, both at w_0 and at w_1, all modal propositions involving boxes and diamonds can still be explicated by means of quantification over possible worlds, and when a proposition p exists at w_0 and at w_1, and hence directly refers only to individuals that exist at *both* worlds, then it is possible or necessarily, respectively, at w_0 if and only if it is possible or necessary, respectively, at w_1. For while different worlds may exist at w_0 from those that exist at w_1 (i.e. God would have different ideas were w_0 actual than the ones he would have were w_1 actual), the differences should not be germane to the evaluation of the truth value of a proposition p that directly references only individuals that exist in *both* w_0 and w_1.

This view makes the propositions that directly reference contingent individuals be themselves contingent beings, and it may seem obvious that this contradicts S5. For suppose that p is a contingently existing and possible proposition. Then, by S5, we have $\Box \Diamond p$. Now, plausibly, if p is contingently existing, so is the proposition $\Diamond p$. But $\Diamond p$ is necessary, and surely if a proposition is necessary, then, necessarily, it is true, and hence it necessarily exists, since it surely cannot have the property of truth without existing (cf. Plantinga, 2006).

However, the argument from $\Box \Diamond p$ to the necessary existence of $\Diamond p$ can be challenged. First, one can think of the box as not expressing a property of propositions, but an operator that transforms the proposition *that s* into the proposition *that necessarily s*. If one does that, then one need not grant the inference from $\Box \Diamond p$ to $\Box \text{True}(\Diamond p)$. Granted, the Tarskian T-schema is hard to dispute. But the T-schema as normally stated only lets us move from "$<s>$" to "$<s>$ is true," where "$<s>$" denotes the proposition that s. It does not let us move from "necessarily s" to "necessarily $<s>$ is true." Granted, if one takes every instance of the T-schema to be an axiom, and one takes axioms to be necessarily true, and one's modal logic includes Necessitation, one will be able to make the move. But someone who thinks propositions are contingent should deny that instances of the T-schema are axioms. Instead, she could say that every instance of:

(182)   (s and the proposition $<s>$ exists) if and only if $<s>$ is true

is an axiom. But this only lets us move from "necessarily s" to "necessarily $<s>$ is true if $<s>$ exists," which is not enough for the argument from $\Box \Diamond p$ to the necessary existence of $\Diamond p$.

Second, even if one grants that necessarily ◊p is true, one might deny the serious actualism that lets one move from ◊p being necessarily true to ◊p existing necessarily. After all, one might think that necessarily Socrates is a man, but Socrates doesn't exist necessarily.

Moreover, we can follow Adams's distinction between truth *at* a world and truth *in* a world. A proposition is true *at* a world provided that it correctly describes how it would be if that world were actual. A proposition *p* is true *in* a world provided that <p is true> is true *at* that world. Given serious actualism, a proposition can be true in a world only if it exists in that world, but it can be true at a world without existing there. If we do this, we can grant that ◊p is true at every world, without granting that in every world it is true. Granted, Adams (1981) himself seemed to think S5 was incompatible with his view, but it does not appear that there is good reason to think this.

7.2.4 Conclusions about identity
So there are in fact three possible solutions to the problem of indiscernibles: (1) allow haecceities; (2) bite the bullet and deny the possibility of indiscernibles while presenting an elegantly complete historical-causal account of transworld identity; and (3) allow the collection of divine ideas to differ between worlds.

Each of these solutions has advantages and disadvantages. The second appears to have the greatest theoretical advantages, since it provides an informative and powerful account of transworld identity. On the other hand, the first and third have the advantage that they do better justice to intuitions about indiscernibles. Nonetheless, each of the solutions avoids the challenges facing Lewis's counterpart theory (see Section 2 of Part III).

7.3 *Attributions of ability*

The objection to Lewis's view (Section 2.2.2 of Part III) as well as to the ersatzist approaches (Section 3.3.4 of Part IV) on the grounds that statements of the form "I am able to φ" should only be about me and not about persons in other worlds or abstracta disappears immediately on the present approach, of course, since if I can be the cause of something, those capacities whose exercise would make me the cause are the grounds for the proposition that reports the possibility of that thing's existence. The Aristotelian account is based precisely on the causal powers causally relevant to a possibility.

7.4 No *set of all possible worlds*

I have argued in Section 7 of Part III that on no reasonable account of possibility is there a set of all possible worlds and that this is a serious objection to Lewis's view, since on his view worlds are physical objects at ontological ground level so that there *should* be a set of them. After all, they are paradigmatically the sort of thing that makes up a set (assuming, of course, a set theory with ur-elements). One might think that the very same argument works against the theory I have proposed. After all, the worlds on this theory are in a sense concrete: they are ideas, not of the Platonic variety, but actually existent in a mind. Should there not be a set of them, then?

But while physical-type entities *are* paradigmatic of the sorts of things from which we build sets, mental things are not. It is not clear that there is such a thing as the set of all the pains I have. Indeed, it is not clear how pains are to be individuated. If my legs hurt, is the pain in my left leg something different from the pain in my right leg? If so, what about the pain in the lower 3 inches of my left leg and the pain in the next 3 inches? Similarly, it is not clear that it makes sense to talk of the set of my thoughts. Is my thinking of the bachelor different from my simultaneously thinking of the never married marriageable man? When is an idea that is implicit in another idea to be counted as separate from it? It is far from clear that such questions *always* have determinate answers. On the other hand, the members of sets should always be individuable as sets and their subsets should have well-defined cardinalities.

Thus, ideas and like mental entities are *not* paradigmatic ur-elements of sets (an ur-element of a set *S* is a member *e* of *S* that is not itself a set) in the way that physical objects, or more generally substances, are. Now, the above *specific* reasons why my ideas and my pains do not form a set may not apply to the divine ideas of worlds. Those ideas can be individuated by the logical maximality and mutual incompatibility of their contents. However, having seen that ideas in general are not paradigmatic ur-elements of sets since not all collections of ideas are sets, and having seen that paradigmatic ur-elements are substances, we should at least be very cautious about assuming that there is a set of all worlds if worlds are divine ideas.

7.5 *The ethical objection*

The ethical objections to Lewis's theory have no force here. It is clear that we have no reason to be concerned in the same way about things in worlds that

God has not actualized as about things in the world that God has actualized, since the former things do not exist — only ideas of them do.

7.6 *Inductive reasoning*

The inductive reasoning objection to Lewis's theory in fact turns into a positive asset for the present theory. For it is an asset of a theory of actuality that if the theory is true, one has more reason than not to believe in induction. God is a personal being, a rational being, and so there is reason to think that he would act as one would expect rational beings to do so. One would not expect rational beings to act haphazardly. One would expect a certain order to the states of affairs they bring about. The fact that we have found order in the past gives us reason to believe that God *does* value a certain amount of order in the universe, just as orderliness in any rational being's past activity gives reason to believe that a certain amount of order is valued by the being.

One might try to make a Goodmanian objection. Consider two theories. On one, the universe is always orderly. On another, it is gruesome, i.e. it is orderly *except* that emeralds are grue as opposed to green: they are green before the year 2050 (we need to update the dates, given that the flow of time has disproved the original gruesome hypothesis!) and blue thereafter. One can form such gruesome predicates for all predicates we use for things in time. Now, it is claimed, all the evidence we have about God is just as compatible with God valuing ordinary order as with God valuing gruesome order. Hence we have no more reason to think emeralds are grue than that they are green.

However, this reply neglects the fact that part of what makes a being rational is that the being responds to objective values. Order in the universe *is* an objective value. Gruesomeness is not. It is not the sort of thing to which a rational being is sensitive. Thus we have some *a priori* reason to think that God would be more responsive to order than to gruesomeness as a value. This is not enough to give an *a priori* proof that God would create an orderly universe. After all, God can be responsive to a value but choose to act on an incompatible value, such as the value of an utter diversity of things and events in the universe. However, that *a priori* reason taken together with our past observations of the universe which showed that there were consistently obeyed laws there *does* give one some reason to prefer the hypothesis of the universe being ordinarily orderly over the hypothesis of gruesome order.

7.7 *Explaining with necessity*

In Section 11 of Part III, I argued that Lewis has a problem accepting explanations like:

> (183)   There are actually no square circles because square circles are impossible.

The arguments in Section 3.3 of Part IV about the connection between the concrete and the abstract suggest that Platonists may have similar problems.

However, on the Aristotelian view, that square circles are impossible is simply the claim that nothing has a capability of initiating a chain of causes leading to a square circle, and so (183) is akin to:

> (184)   There are no flying elephants because no elephant has the capability to fly.

But (184) is a genuine explanation, though maybe not as informative as we would like. But it is at least *somewhat* informative. We might, after all, have previously thought that the reason there are no flying elephants is that although they can fly, they all suffer from a terrible fear of heights.

Section 8 *Objections to the Aristotelian–Leibnizian view*

8.1 *Primitive causal modality*

One might argue that the Aristotelian view, whether in its theistic version or in a version that bites the bullet and says that the whole history of the universe could not have been different, has a serious disadvantage over a view like Lewis's, because it presupposes primitive causal powers, whereas Lewis manages to reduce both powers and causality to his possible worlds.

However, to fix Lewis's reduction of causation to similarity relations between possible worlds, we need an independent account of the order of time. And, as I have argued in Section 2.3 of Part II, our best story about the order of time presupposes causal relations and indeed not only actual causality but potential causality. If this is right, then both Lewis and I need potential causality, and neither can reduce it to something non-modal.

8.2 *Evil and the apparent compatibility of all things*

If things are as described, then some worlds that are apparently possible are impossible. For if God is necessarily omnibenevolent, omniscient, and omnipotent, then it is a necessary truth that any evil that exists is *justified*, i.e. there is a morally extenuating reason for an omnibenevolent, omniscient, and omnipotent deity to allow this evil. Richard Gale (1991: 229) has in fact used this observation to argue against the possibility of a necessarily existent, omnibenevolent, omniscient, and omnipotent deity, since, he claims, it is plainly possible for there to be an unjustified evil. Moreover, Gale claims that theists by taking seriously worries about the problem of evil are admitting this logical possibility.

One will be particularly confident about Gale's argument if one accepts a Humean view that anything could co-exist with anything else. For then we could have a universe in one corner of which there is a perfectly innocent person who co-exists with a horrid pain in her mind for all eternity, and with the rest of the universe being set up in such a way that no one can benefit from this pain. But such a pain, surely, would be an unjustified evil — yet the Humean claims it is possible. Hence, there cannot be a necessarily existent, omnibenevolent, omniscient, and omnipotent deity.

One can, however, challenge the Humean intuitions by bringing in the global nature of the Aristotelian–Leibnizian account of modality. To make a claim of the possibility of some non-actual event is to make a claim about how something actual could have brought it about. To be certain that what Gale claims is possible is indeed possible, one would have to give some sort of a causal account of how it could come about. But no such causal account can be given without begging the question against the existence of a necessarily existent, omnibenevolent, omniscient, omnipotent creator of all things. For if there is such a creator, then any full causal account will have to ultimately go back to him, perhaps in a non-deterministic way, and if the evil is unjustified, then no such story going back to him is available. The confidence with which one affirms such possibilities withers when one realizes that in asserting the possibility of a proposition one is asserting something's capability for ultimately bringing this possibility to actuality.

Of course there is a simpler answer if one is not interested in defending all of traditional theism. One could just say that the Aristotelian–Leibnizian account of possibility does not require the assumption that God is necessarily omnibenevolent, and hence Gale's argument is irrelevant as an objection to this account of possible worlds. But the answer above is preferable if one is impressed by the explanatory advantages of classical theism. But in any

case, the advocate of the Aristotelian–Leibnizian account cannot accept all of the Humean combinatorial intuitions since it is an essential part of these intuitions that there are no necessary beings.

8.3 *Is God omnipotent if logical possibility is defined in terms of his power?*

One might worry that if possibility is defined in terms of God's power, then the claim that God is omnipotent is tautologous, since to be a possible state of affairs and to be capable of having God as First Cause are the same.[20] This criticism would appear to presuppose that "*x* is omnipotent" is defined as something like:

(185)   For any possible state of affairs *S*, *x* can initiate a chain of causes that can lead up to *S*.

While I find this a plausible start of an account of omnipotence, this is by no means a standard account of omnipotence. But for the sake of the argument, let us start by granting this account of omnipotence.

Still, even if the claim that God is omnipotent is tautologous on the theory, the claim that God is omnipotent is still contentful. For we have independent epistemic access to what is logically possible, though we may not have a full epistemology of how this access works, so that saying that the metaphysically possible is nothing is the same as the possible-for-God *does* convey some information to us. We have defeasible intuitive reason to believe, for instance, that it is possible for the universe not to have contained any living things, and so upon learning that possibility is nothing but divine ability to be the First Cause of a given state of affairs we learn that it is possible for God to have created a world without any living things.

But in fact the Aristotelian–Leibnizian account of modality does not identify possibility with producibility-by-God (mediate or immediate). Rather, possibility is defined as actuality or producibility-by-any-actual-entities. It is then a *substantive* claim that possibility is extensionally the same as producibility-by-God, a substantive claim that follows from God being necessarily the First Cause. Thus if we accept (185) as a definition of omnipotence, we should look at the Aristotelian–Leibnizian account

---

20   An objection like this was made by Cameron (2008). Also, cf. Loux (1979: 59–60).

as providing us with a derivation of God's omnipotence from God's being necessarily a First Cause, a thesis of independent interest.

Of course we might worry that (185), even if nontrivial, is unsatisfactory. Suppose that God is capable of only two things: he can produce a demiurge or not produce anything. There is only one demiurge God can produce, and God has no control over what the demiurge is like. The demiurge, on the other hand, has a choice what sorts of things to make: unicorns or no unicorns, horses or no horses, donut-geometry space-times or flat space-times, etc. If we identify omnipotence with being the First Cause in every possible world, then this God counts as omnipotent. But that seems mistaken. At least in respect of the *variety* of producibles, the demiurge is more powerful, though no doubt it does show a very great power for God to be able to make such a demiurge.

To rule out scenarios like this one, one might wish to supplement (185) with some additional conditions. For instance, one might say that x is omnipotent provided that:

(186)   For any possible state of affairs S, x can initiate a chain of
         causes that can lead up to S. *and* for any possible state of
         affairs S, x has the power of *directly* producing a duplicate
         of S.

We do not require that x be able to directly produce S itself, because essentiality of origins might preclude that possibility.

There are at least two types of counterexamples, however, for such an account. For the first, consider the state of affairs of

(187)   there being an island not directly produced by God.

A volcano can produce (187), but it does not seem that God can produce any duplicate of (187). For a second counterexample, consider:

(188)   there being an injustice.

A finite being can produce (188). But God cannot directly produce an injustice.

One might try to refine the definition by restricting the quantification over states of affairs said to be directly producible by God to states of affairs consisting of fundamental entities, their possessing positive intrinsic properties and their standing in positive relations only to one another. To refine

this, one would need an account of the positivity of properties that excludes such properties as *injustice* and *not being directly produced by God*. A limitation of the resulting account, however, is that while it guarantees that an omnipotent being can directly make an island with at least 100 coconut trees, it does not seem to guarantee that the being can directly make an island with at most 100 coconut trees. Nonetheless, it is worth saying that at present all accounts of omnipotence have difficulties, and there is at least room to hope that approaches in terms of powers can do at least as well as other approaches — and maybe better.

And, in any case, the objection that the Aristotelian–Leibnizian account makes omnipotence tautologous fails.

8.4 *The wrong way around?*

According to the Aristotelian:

> (189) I could not be a point in space because there is no power capable of producing a chain of causes that could lead up to my existing as a point in space.

But, the objection goes, this is perverse. For, surely:

> (190) It is precisely because it is impossible that I be a point in space that there is no such power.

The reasoning behind this is that if something is impossible, it is likewise impossible for there to be a power capable of leading to its occurrence, and if it is impossible for there to be such a power, then there is no such power. Thus, we have a vicious circularity.

However, "because" is used in different senses in (189) and (190), and one can perhaps have *A* occurring because of *B* while *B* occurs because of *A* if there are two senses of "because." Here, in (189) we have a "because" of grounding, as in the sentence "The iron is hot because the molecules in it have high kinetic energy." The sense of "because" in (190) is not actually completely clear, but it is not likely to be that of grounding: nobody's having a power to make me be a point in space is surely not *grounded* in the impossibility. A very plausible reading is that it is simply an epistemic "because," and if that is the right reading, then the argument dissolves.

But let us linger on the suggestion that (190) is in fact an *explanation*.

But what sort of an explanation is it? On non-Aristotelian views, there is no hope for its being a causal explanation. On Lewisian views, it could, I suppose, be an explanation involving the reduction of causal power to possible worlds. Perhaps once we do that reduction, we get the result that nothing can have a causal power to produce a state of affairs that obtains at no world. But Lewis's reduction of causation to the structure of the space of possible worlds failed, and if causation can not be reduced to possible worlds, it is unlikely that causal powers would be thus reducible. On Platonist views, on the other hand, (190) is one of those puzzling coincidences between the Platonic realm and the concrete realm that were used in Section 3.3 of Part IV as an argument against Platonism. Thus, while we may have an intuition that (190) is a good explanation, it does not appear that the two major alternatives to the Aristotelian–Leibnizian view bear out that intuition. And the Aristotelian–Leibnizian view at least has the advantage of making the two sides of the "because" in (190) not be coincidentally connected — it is just that the explanation goes in the reverse direction.

8.5 *Necessary being*

The God of the Aristotelian–Leibnizian account is a necessary and concrete being, and this may seem absurd (cf. Findlay 1948). But there is no absurdity in a necessary being. For instance, it is difficult to see how the grounds of necessarily true mathematical claims could fail to involve necessary beings. Moreover, there is good reason to believe in propositions, and at least some of these are necessary. There is nothing prima facie absurd about a necessary being.

Of course Hume has argued that any being that can be imagined to exist can also be imagined not to exist. As a general principle this is already refuted by the mathematical cases, though it might be patched up by restriction to cases of concreta. However, on our Aristotelian epistemology of modality, a state of affairs is only possible if there actually is some causal factor that could have brought it about. To claim that the non-existence of some entity is possible is, then, to claim that something actual could have brought it about that this entity did not exist, and there is nothing absurd about the idea of an entity such that nothing could have brought about its non-existence.

Still, imaginability is a guide to possibility. But it is a fallible guide, and one could simply say that in this case it fails. But we can say something more here. It is difficult to really imagine negative states of affairs. Here

is a thought experiment. Tell an ordinary person: "Imagine yourself in an otherwise empty room. Describe your experiences." I surmise that most ordinary people when faced with this will not say: "I am gasping for lack of air and I die in pain." For I suspect that most people when asked to imagine an empty room imagine a room empty of large macroscopic objects, but do not imagine a room empty of air. Much less do they imagine a room empty of gravitational and electromagnetic fields, non-zero components of wave-functions, or space-time points.

But suppose now that one takes care to imagine a *really* empty room. This room would have no macroscopic objects, no atoms, no fields, no wave-function components and no points. But is it something one would *imagine*? I suggest that, rather, one imagines a room without macroscopic objects, and mentally *labels* it as "a room without atoms, fields, wave-function components, points or anything else." But the possibility of mentally labeling an imagined *F* as "a *G*" provides only a little evidence that an *F* could be a *G*.

I suggest that, likewise, when one purports to imagine a world without a necessary being, one is simply imagining a world and *labeling* it "There are no necessary beings here."

One might retort: "But I can imagine a world that consists only of an iron sphere." But if one takes this seriously, one should take it as evidence against Platonism. For if the world consists only of an iron sphere, it has no sphericity in it either. And yet it is not a very compelling argument against Platonism. Moreover, it may be incorrect to take one as having imagined a world without sphericity. Rather, one has imagined a world without imagining sphericity in it. Similarly, I submit that one probably hasn't imagined a world without God. One has imagined a world without imagining God in it, and then one has labeled the imagined world as "A world without God."

There is an opposite objection. Instead of complaining that one has made use of the notion of a necessary being, one may object that we have not done justice to it. For, the objection goes, there is *much more* to the necessity of the existence of God than the merely negative fact that nothing could have brought it about that God never existed, or, alternately, that there is nothing which is causally prior to God. The sort of necessity imputed to God by the Aristotelian–Leibnizian theory is weak. All that the necessity means is that there was nothing that could have prevented God from popping into existence thus. It is generally accepted that a necessary being exists always. But all we get here is that the being exists either timelessly or at least at one instant of time in all possible worlds: not that he *always* exists in all possible worlds.

The Big Bang brings us to the atheological objection. Paul Tillich (1957)

identified God with one's ultimate concern, so that anybody who had an ultimate concern *eo ipso* believed in God. There is an anecdote that once when Tillich was speaking, a questioner objected that he did not want to be made a theist by definition. A theological account on which an atheist counts as a theist simply because, say, he has money or reputation as an ultimate concern makes theism too cheap.

A similar objection can be offered to the Aristotelian–Leibnizian account of necessary beings: it just makes belief in a necessary being too easy. Take, say, an atheist who believes that the Big Bang happened causelessly so that nothing was causally prior to the Big Bang. On the Aristotelian–Leibnizian account of necessity, the atheist believes that the Big Bang was necessary. The atheist can object: "I do not want to be a believer in a necessary being by definition."

The theological and atheological objections are closely related. Both complain that not enough has been said about what makes a being necessary. But while on the Aristotelian account, what makes God a necessary being is that nothing can prevent his existence, the account does not close the door to saying something more about the necessity of the necessary being. For while having nothing be causally prior to God just *is* God's being necessary, nonetheless we can ask why and how God can exist without having anything causally prior to him. It is only if one has the Humean intuition that it is quite easy for beings to exist without causes that one sees the lack of a cause as failing to say anything impressive about a being. But if existence requires explanation, then once we are told that there is nothing prior to God, we leave open the door to a further philosophical investigation: How is it that God has no cause and yet exists? And it is the openness to this kind of a question that makes for the difference between the traditional Western theist's view that God's existence is necessary and a typical atheist's view that there is nothing causally prior to, say, the Big Bang. Admittedly, the atheist *does* count as believing that the Big Bang is necessary on the Aristotelian account — there is no way out of that. But a typical atheist denies the PSR and adds: and that's *all* there is to the Big Bang, with there being nothing special to being a causeless being — any being could be like that.

Of course this leaves completely open the question of what it is that explains how God manages to exist causelessly. Answers to this question are beyond the scope of this book. One suggestion is that God's existence flows from his essence. Another is in the Thomistic account on which some beings require causes precisely because they are such as to have a difference between essence and existence so that the contingent union of the essence and existence is to be explained. But God is *simple*, his essence being

identical with his existence, and it is because of this identity that he exists — or perhaps we might say that this identity is what makes God's existence not problematic and not calling out for explanation like the more shaky existence of creatures. Or perhaps the answer to the question is beyond our knowledge. In any case, the theist can insist that a full understanding of God's nature would give an answer, and does not need to say that there is here just a brute fact.

A different option is to go back to the LRT account. We have various concepts like *the First Cause* and *the omnipotent being* that uniquely pick out God. If *G* is one of these concepts, then perhaps both the proposition <There is a *G*> and the proposition <There are no *G*s> exist, though only the first is true. But, maybe, there is a deep but simple concept *G** of God which is such that <There is a *G**> has no negation, but alas in this life we do not possess *G**. In that case, there is a deeper story to be told about God's non-existence: God falls under a deep but simple concept under which his existence is undeniable. This may be what Aquinas means when he says that God's existence is self-evident in itself but not to us (Aquinas 1948: I, 2, 1).

8.6 *Obscurity*

It does not seem as though the concept of God is one that we are very clear about. After all, although we know enough to individuate God from all other possible beings — God is the only First Cause, for instance — this is not much. What are divine ideas? What are God's capabilities and choices like? All these are invoked in the theory, but they are shrouded in obscurity.

However, they are not in complete obscurity. Analogy plays a crucial role here. The divine ideas are related to God in a way analogous to the way our ideas are related to us. Divine capabilities and choices are related to God in a way analogous to the way our capabilities and choices are to us — and we are capable and do choose. The analogies have their limits, but they render the theory determinate enough to be open to discussion, and indeed more determinate on various points, e.g. the nature of propositions which is now seen to be that of divine ideas, than competing theories. Analogical reasoning in constructing theories is not to be brushed aside in general. Sellars (1968) has observed this in the case of scientific reasoning. When introduced, the particles of the atomic theory of gases were seen as analogous to hard billiard balls. The analogy had its limitations, but gave some determinateness to the content of a theory whose ontology would otherwise be objectionably obscure.

8.7 *Creaturely free will*

The full Aristotelian–Leibnizian account posits libertarian free will in God, which was, to some extent, defended in Section 5.2, but does it leave any room for *our* libertarian free will? It is essential to the account that there be a chain of causes stretching back from Jones's mowing of his lawn last Saturday to God's acts of will, so that even were Jones not to have been alive last Saturday, the proposition that Jones mowed the lawn last Saturday was a possible proposition, because God could have initiated a chain of causes capable of leading to Jones's being alive last Saturday and freely choosing to mow the lawn there. God's power is a possibilifier of Jones' freely mowing the lawn last Saturday.

But this only says that God's power is sufficient to *possibilify* Jones's freely mowing the lawn. It is not said here that God's power can *make* Jones freely mow the lawn. No claim is made that God could have *ensured* that the chain of causes would have this outcome. Rather, God is able to bring it about that there was an entity able to bring it about that there was an entity able to bring it about that … that there was an entity able to bring it about that Jones mowed the lawn on Saturday. This last entity would have to be Jones himself if Jones has libertarian free will. Each link in the chain can ensure that the next link in the chain has the power to produce the link after it (with the power to produce the link after it, and so on), but it need not be able to *ensure* that the next link should exercise that power.

8.8 *Divine simplicity*

Hitherto, I have been blithely talking about ideas, in the plural, found in the mind of God. But it is a tenet of traditional Judaism and Christianity that God is *simple*. If there is no division in God, how can he have multiple ideas?

We have three options here. First, one might abandon divine simplicity. Those who have a theological commitment to divine simplicity or are impressed by the arguments for the simplicity of the First Cause (e.g. that diversity always requires a further explanation, but the First Cause is the ultimate explainer) will not take this route. Second, one might abandon the notion of divine ideas. The Aristotelian part of the Aristotelian–Leibnizian account would still remain, but the Leibnizian would have to go, and we would be left without worlds. There would still be much to be said for the remaining Aristotelian account: it would give an answer to the grounding problem. Or, third, we could try to reconcile divine simplicity with a

multiplicity of divine ideas. If this could be done, it would be preferable, as then one would not need to choose between divine simplicity and the Aristotelian–Leibnizian acocunt.

Thomas Aquinas thought one could have both divine ideas and divine simplicity. Aquinas noted that we can think of ideas in two ways. First, we could say that x has an idea of y provided that there is an image of y in the mind of x. In this sense, there cannot be multiple ideas in the mind of God by divine simplicity. Second, we could say that x has an idea of y provided simply that x cognizes y. Thus, when we say that x has two ideas, this means that there are two things, y_1 and y_2, each of which x cognizes. God can cognize diverse items in and through a single act of cognition, indeed the same act of cognition by which he cognizes himself.

> Inasmuch as He knows His own essence perfectly, He knows it according to every mode in which it can be known. Now it can be known not only as it is in itself, but as it can be participated in by creatures according to some degree of likeness. But every creature has its own proper species, according to which it participates in some degree in likeness to the divine essence. So far, therefore, as God knows His essence as capable of such imitation by any creature, He knows it as the particular type and idea of that creature; and in like manner as regards other creatures. So it is clear that God understands many particular types of things and these are many ideas. (Aquinas 1948: I, 15, 2)

So while I introduced divine ideas as divine *thinkings*, in an important sense there is only one act of divine cognition, though we can consider this act under many aspects, depending on its many objects: the thought of Smith, the thought of Jones, etc.

As it stands, this is only an answer to the problem if God's complete knowledge of himself is compatible with divine simplicity. In the dialectical structure of the *Summa Theologica*, this is not a difficulty. For Thomas has already argued that there is a God, that he is simple, and that he knows everything, and in particular thus that he knows every aspect of himself. Hence, God's complete self-knowledge and divine simplicity have to be compatible, because both in fact are co-instantiated.

But even apart from such an argument based on the success of Thomas's Five Ways, we can build on Thomas's account in Question 15 of the First Part of the *Summa Theologica*. The central point of the account is that a simple cognitive act can be simultaneously directed to more than one known thing. We might say that if I think of something as material, *eo ipso* I think of it as temporal and spatial. But my grasping it as material need not be

a mental act composed of two other acts. Rather, I think of it as material and *thereby* I think of it as temporal and spatial, rather than thinking of it as temporal and inferring that it is extended and spatial. In one sense there are two thinkings, of extension and spatiality, and in another there is one think*ing* and two *objects* of thought. This fits with the phenomenology. My thought that Fred is material does not appear to be *composed* of two sub-thoughts, even though the two thoughts are in some way "contained" in it. Similarly, there is no phenomenological reason to suppose that my thinking of a square has multiple parts to itself, whether parts corresponding to individual sides and individual corners, or whether one part corresponding to sides and another to corners, even though my thinking of a square includes my thinking of something with sides and corners.

Of course the phenomenology could be deceptive. But what appears to us is at least likely to be coherent. This is a weaker version of a principle of the Indian philosopher Śaṃkara, AD 788–820, which says that "Nothing even appears to be like an impossibility."[21] It may be that there is no material world, but the fact that there seems to us to be one makes it likely that it is coherent.

Thus, when we pay attention to our phenomenology, we see something that gives us a model for how a singular and simple act of thought could in another sense involve more than idea, more than one thinking. And presenting ourselves with a vivid model of something is a primary way by which we convince ourselves that it is possible. Thus, even if the phenomenology deceives us and there is a finer ontological structure to my act of thought, there is reason to think it coherent to suppose that such an act of thought might be simple.

On this view, the worlds and other divine ideas are not fundamental entities. Rather, facts about worlds are grounded in the cognition of the simple God.

8.9 *Incompatible powers*

Consider any given item. It has multiple powers. Thus, a dog generates an electromagnetic field with its brain, depresses the ground, stirs the air-molecules, barks, perceives, etc. The idea of one item having multiple powers is entirely familiar. However, as we have seen, in order to avoid modal fatalism, some substances must also have multiple powers in a stronger sense:

---

21   Quoted in Chakrabarti (1995: 319).

some causes must be indeterministic. They must both be able to produce effect *E* and not to produce *E*.

Of course, an indeterministic cause cannot exercise the powers for *E* and for non-*E* simultaneously. This raises a question about the ontology of these indeterministic powers. What prohibits such simultaneous exercise? After all, the dog, in virtue of its multiple powers, can displace air molecules and generate an electromagnetic field. Why can't Curley simultaneously accept the bribe and not accept it, in the same respect, if he has the power to accept it and the power not to accept it? The worry is that non-Aristotelian modality is being smuggled in when we deny Curley the ability to exercise incompatible powers simultaneously.

However, our picture of the dog's capabilities is overly simplistic. For a substance's having the ability to do *E* and the ability to do *F* is not sufficient to ensure its ability to do both *E* and *F* simultaneously. To be able to do both *E* and *F* simultaneously is something more, a greater power. The dog has the capability to *both* displace air molecules and generate an electromagnetic field, rather than just having the power to displace air molecules and the power to generate the electromagnetic field, while a beginning tightrope walker may have the power to walk the tightrope and the power to carry another person, but not the additional power to do the two simultaneously. Likewise, then, Curley lacks the power to simultaneously accept and not accept the bribe, much as he might wish for such a power.

Thus, what bars Curley here from both accepting and not accepting the bribe (at the same time and in the same respect, of course) is his powerlessness to do that, and on the Aristotelian view this appears not different in kind from the beginning tightrope walker's inability to walk the tightrope while carrying someone. Yet, there does seem to be an important difference between the two cases, and it is not clear that the Aristotelian can make sense of this difference.

But the Aristotelian can say this: The beginning tightrope walker might have had the ability to walk the tightrope and carry someone simultaneously. Curley could not have had the ability to both accept and not accept the bribe, notwithstanding how useful it could be to a politician. We can ground both facts. Someone or something had the power to give a beginning tightrope walker the ability to walk the tightrope while carrying someone. For instance, evolutionary processes could have produced aliens who could have given that power to the tightrope walker. But nothing could have initiated a process leading to Curley having the power to accept and not accept a bribe simultaneously. And on the bare Aristotelian or Aristotelian–Leibnizian

view, that is all there is to the impossibility of simultaneously accepting and not accepting the bribe.

If more is wanted, we should opt for the LRT, which may imply that there is no proposition that is a conjunction of <Curley now accepts the bribe> and <Curley does not now accept the bribe>, for $\sim(p \;\&\; \sim p)$ is always R-necessary. (And in the case of more complex contradictions, we can invoke R*-necessity, as in Section 5 of Part V) We have met the same pattern over and over. Whenever the Aristotelian is pressed about something which seems to call out for a deeper explanation of some necessity, she can either bite the bullet or she can take the LRT way out. I find the latter somewhat attractive, but it has the disadvantage that it leads to an unorthodox logic.

PART VII

FINAL CONCLUSIONS

Section 1 Cost–benefit arguments for the Aristotelian–Leibnizian ontology of possible worlds

1.1 Modality

The Aristotelian–Leibnizian theory gives an account of the existent realities that are spoken of when we make modal claims. Moreover, it succeeds in doing this without making any implausible distinction between actuality and existence as Lewis does and without invoking any *objectionable* primitive modality as the ersatz views do. In a precise sense, the possible worlds exist in the actual world: it is true at the actual world that possible worlds exist. But this is not paradoxical: The possible worlds are just divine ideas, one of which is actualized, and the actualized one contains within itself an account of all the other divine ideas. (Obviously, they must be infinite lest the regress prove vicious.)

Moreover, the account squares nicely with the intuition that attributions of the possibility of doing something should be claims precisely about the beings to which this possibility is attributed and their environment, something which neither Lewisian nor ersatzist approaches allow for.

1.2 Explanation of what propositions are

As a non-modal bonus, the Aristotelian–Leibnizian model provides us with a deeper insight into the nature of propositions, which it says are divine ideas. Admittedly, the nature of divine ideas is not perfectly clear, and so this clarification is only partial. However, we do have an analogical grasp of the nature of divine ideas based on our grasp of our own ideas, and this provides us with a better understanding of what propositions are than just

considering them to be bare Platonic abstracta that represent realities in some mysterious way. Divine ideas represent realities analogously to the way that our ideas represent realities.

It is true that David Lewis's EMR also purports to give an account of the nature of propositions. Unfortunately, his reduction of propositions to collections of worlds is unacceptable whereas his "structured propositions" involves too much arbitrariness (see Section 2.3 of Part I). It might be argued that our present view also contains some arbitrariness. Why should we define "propositions" as *God*'s ideas rather than the ideas of someone else? But there is good reason for this. For one, God is the only necessarily omniscient being, and presumably also the only being that necessarily has ideas of all realities. But more importantly, what makes these ideas be ideas of possibilia is the power of God: it is God's choice that is necessarily the first branching in the "history" (in quotes as God's choice might not be in time) of the cosmos, and so it is appropriate that the propositions be found in the God who is the ultimate possibilifier.

1.3 *Not a completely new ontology*

The Aristotelian–Leibnizian model *does* have a rich ontology containing God, an ontology strictly richer than that of the ersatzist. For, all the items in the ersatzist's menagerie exist on the Aristotelian–Leibnizian model: there are propositions (i.e. divine ideas) and there are sentence-types for all possible languages (i.e. divine ideas of sentence-tokens in all possible languages), but there is one more entity required: God. However, this added ontological richness is necessary to do justice to modality. Very briefly, mere propositions or sentences are "descriptive": for the Parmenidean problem to be resolved, there needs to be an "executive power" by virtue of which possibilities are made *possible*, and this is God and other items with causal capabilities.

However rich the Aristotelian–Leibnizian model's ontology may be in positing God, positing God is not revisionary of our ordinary concepts in the way that EMR is. For theism has historically been very widely believed through many centuries of the development of many of our ordinary concepts, and hence it is prima facie likely for theism to be compatible with central ordinary concepts, unlike EMR.

1.4 *Connection with theistic arguments for the existence of God*

If traditional theism is true, then God is the "ground of being," necessarily the First Cause of the existence of all contingent states of affairs. But if this is so, then traditional theism comes very close to the proposed Aristotelian–Leibnizian view. The only additional step that the Aristotelian–Leibnizian view takes is that not only is it the case that God must be the First Cause of the realization of any possibility *S*, but that *S*'s being a possibility should be *analyzed* in terms of God. Theoretical simplicity and Ockham's razor are in favor of taking the two predicates "is something that which God or some other substance can initiate a chain capable of leading to" and "is possible," which according to traditional theists necessarily have the same extension, to express one and the same property. Thus, if there is a God who is the ground of being, it is plausible to suppose him also to be the ground of possibility.

Consequently, arguments for traditional theism also lend plausibility to the Aristotelian–Leibnizian view, and certainly, there are many such arguments.[1] And, conversely, the present view's theoretical virtues in explicating modality lend plausibility to traditional theism.

Observe, too, how the Aristotelian–Leibnizian account of possibility not only gives an account of what possibility is but also gives a partial explanation why what is actual is actual: namely, because of God's creative act. To fill in this partial explanation, one would have to state the values that God was promoting in creating the actual world. We cannot do this fully, though it does appear that a balance of diversity and unity is one of them:

> For He brought things into being in order that His goodness might be communicated to creatures, and be represented by them; and because His goodness could not be adequately represented by one creature alone, He produced many and diverse creatures, that what was wanting to one in the representation of the divine goodness might be supplied by another. For goodness, which in God is simple and uniform, in creatures is manifold and divided and hence the whole universe together participates the divine goodness more perfectly, and represents it better than any single creature whatever. (Aquinas 1948: I, 47, 1)

---

1   See Swinburne (2004) and the articles in Craig and Moreland (2009). Not all the contemporary natural theologians think that God is a necessary being. But the arguments the authors adduce for theism do not show that God is *not* a necessary being, and it is plausible that if God exists, then he is the ground of being in a sense that requires him to be a necessary being. Therefore, even the arguments for theism by authors like Swinburne who deny that God is a necessary being can be invoked here.

Section 2 *Choosing between accounts of possibility*

We have seen a plethora of accounts of possibility. This naturally pushes on us the desperate question: Is it all likely that any one account we have available to us is one that we can know to be true?

The question of how likely any one of the accounts is to be true is one that applies in all philosophical areas, whether in special or general metaphysics, ethics, or aesthetics. On both Aristotelian and theistic grounds, I am confident that we can indeed find philosophical truth. We have a natural inclination for philosophical investigation, and the *telos* of such investigation is truth. But we do not have natural inclinations for largely unattainable *telē* — nor would God have given us supernaturally an unattainable *telos*. Can we say with any confidence and without undue hubris that with the Aristotelian–Leibnizian account we have indeed found in some measure at least the truth of the matter about possibility?

I think so. We have very good reason to think that items in the world have capabilities. I see myself as being morally responsible for certain omissions of mine, and in this I see myself as having had a power to remedy these omissions. When I play chess with a friend I may observe not only that he won a game of chess, but also that he has a capacity for playing chess well.

But if there are such Aristotelian features of the world as capabilities, then we can ground *some* possibilities in them, as the Aristotelian account proposes. We can then be confident that there *is* such a thing as Aristotelian causal possibility. Moreover, most of our day-to-day uses of "possible" and "necessary" involve restricted versions of Aristotelian modality, restricted to consideration of some capabilities of some beings at some times. Thus, we can be confident that Aristotelian possibility is at least a *part* of the full story about possibility.

But is it the whole story? Or is there a deeper sort of possibility, "metaphysical possibility," that escapes the Aristotelian analysis? By Ockham's razor we should like to answer in the negative. Ockham's razor is admittedly a defeasible consideration, but we would need a good reason to posit a deeper sort of possibility. The most plausible such reason would be that Aristotelian possibility does not go widely enough, and hence fails to embrace the whole realm of the conceptually possible.

But note that there are two prominent views, each of which makes it very plausible that Aristotelian and metaphysical possibility coincide extensionally. The first view is a PSR: if every contingent fact has to have a causal explanation, then it is very plausible that the Aristotelian view is extensionally right. The second view is theism. As God is omnipotent, then for any

metaphysical possibility *S*, he can initiate a causal chain capable of leading up to *S*. And both of these views can be backed up with evidence independent of the Aristotelian account of possibility, and so we have reason to think that the Aristotelian view is extensionally correct. Applying Ockham's razor, we may conclude that it is simply *right*.

The Aristotelian–Leibnizian account overcomes objections to other accounts of modality, ranging from their lack of a solution to the Parmenidean problem to paradoxical conclusions such as that there is no such thing as inductive knowledge. The Leibnizian component of the account also clarifies the nature of propositions. Since we need a theory of modality, and would like to have propositions and possible worlds, and the present theory is coherent and avoids egregious paradox, this is good reason to believe the Aristotelian–Leibnizian model to be true. Recall Leibniz's remark:

> (1) it seems that it is quite a considerable thing when a hypothesis manifests itself as *possible*, when none of the others manifests itself so at all, and (2) . . . it is extremely *probable* that such a hypothesis is the true one.[2]

Nevertheless, stepping back from our various arguments, we can find that depending on just how much we demand from our theory of modality, we can make do with other ontologies, or we may want to supplement the Aristotelian–Leibnizian account with a Spinozist–Tractarian story. It is easiest to summarize these results with a table (see Table 1).

Table 1: Summary of conclusions

If all we want from our theory is:	Then we can:	At the price of:
Possible worlds	Take a pure Platonic view of possible worlds as abstracta such as collections of propositions	Not having an explanation of the intentionality of propositions or of the truthmakers of modal claims
Possible worlds and something to say on the problem of the intentionality of propositions	Accept the Leibnizian view of possible worlds as thinkings in the mind of a God˙	The primitive intentionality of a divine mind

(*continued*)

---

2   Gerhardt (1960–61, vol. III: 353) my translation, emphases in original.

If all we want from our theory is:	Then we can:	At the price of:
Grounds of modal claims	Accept the Aristotelian view on which possibilities are grounded in causal powers and dispositions	Not having global possibilities or possible worlds
Global possibilities and grounds of modal claims	Accept the Aristotelian view together with a necessary First Cause without commitment to this First Cause being a theistic one	Not having possible worlds
	Or: accept that there are no impossible propositions	Not having impossible propositions
All of the above and more: grounds of modal claims, global possibilities, possible worlds, something to say on the problem of the intentionality of propositions, and whatever other theoretical benefits the existence of God provides	Accept the Aristotelian–Leibnizian model	The primitive intentionality of a divine mind and any other difficulties involved with the notion of a personal and immaterial First Cause
An account of why it is necessary that 1=1, etc., that it is somehow nonsense to deny this	Accept the Spinozistic–Tractarian view, at least in its Less Radical Version	Having propositions that have no negations
All the benefits of the Aristotelian–Leibnizian model, and justice to be done to the intuition that it is somehow nonsense to deny that 1=1, etc.	Accept the Aristotelian–Leibnizian model combined with the Less Radical Theory	The primitive intentionality of a divine mind, any other difficulties involved with the notion of God, and the existence of propositions without negations

*   This is the view that Plantinga has espoused in discussion.

Appendix: the $\Diamond^*$ operator and S4

Suppose as axioms that:

(191) $\Box(p \supset q) \supset (\Box p \supset \Box q)$

And

(192) $\Box p \supset p$

where "$\Box p$" is short for $\sim\Diamond\sim p$. Suppose also that Necessitation holds: if p is a theorem, so is $\Box p$. And assume as an axiom the Barcan Formula specialized to quantification over positive integers, for every predicate F:

(193) $\Diamond \exists n F(n) \supset \exists n \Diamond F(n)$.

Moreover, suppose we can formalize mathematical induction in the usual way in our logic. Let $\Diamond^n$ and $\Box^n$ be the nth iterates of $\Diamond$ and $\Box$, respectively. Observe that $\Box^n p$ if and only if $\sim\Diamond^n\sim p$ (by (191) and induction). Let "$\Diamond^* p$" be short for $\exists n \Diamond^n p$ and let "$\Box^* p$" be short for $\sim\Diamond^*\sim p$. It is easy to see that $\Box^* p$ if and only if $\forall n \Box^n p$. Moreover, assuming (193) holds for all F is equivalent to assuming for all F:

(194) $\forall n \Box F(n) \supset \Box \forall n F(n)$.

I shall argue (in an informal way, but in a way that can be formalized if one so desires) that $\Diamond^*$ satisfies S4, i.e. (191) and (192) hold with $\Diamond^*$ and $\Box^*$ in place of their unstarred versions, Necessitation holds for the starred operator, and if $\Diamond^*\Diamond^* p$, then $\Diamond^* p$.

The starred version of (192) follows from the unstarred because $\Box^* p$ entails $\Box p$.

Starred Necessitation is only slightly more difficult. Suppose that p is a theorem. Then so is $\Box p$ by unstarred Necessitation. We can iterate this

argument. Or, more precisely, assuming our logic allows mathematical induction, it will be a theorem that $\forall n\square^n p$, and hence that $\square^* p$.

Now let us argue for the starred version of (191). I will begin by showing that:

$$(195) \quad \square^n(p \supset q) \supset (\square^n p \supset \square^n q)$$

by induction, for every n. In the case where $n=1$, this is just unstarred (191). Suppose we have shown that:

$$(196) \quad \square^{n-1}(p \supset q) \supset (\square^{n-1} p \supset \square^{n-1} q)$$

for some $n>1$. Applying unstarred Necessitation we get:

$$(197) \quad \square(\square^{n-1}(p \supset q) \supset (\square^{n-1} p \supset \square^{n-1} q))$$

and then by unstarred (191) we have:

$$(198) \quad \square^n(p \supset q) \supset \square(\square^{n-1} p \supset \square^{n-1} q)).$$

Applying unstarred (191) again, we get:

$$(199) \quad \square(\square^{n-1} p \supset \square^{n-1} q)) \supset (\square^n p \supset \square^n q).$$

Combining (198) and (199), we then get (195), as desired. So by induction we have (195) in general.

Now we can show the starred version of (191). From (195)'s holding for all n, by First Order Logic it follows that:

$$(200) \quad \forall n\square^n(p \supset q) \supset \forall n(\square^n p \supset \square^n q).$$

But since $\forall n(\square^n p \supset \square^n q) \supset (\forall n\square^n p \supset \forall n\square^n q)$ by First Order Logic, it follows that:

$$(201) \quad \forall n\square^n(p \supset q) \supset (\forall n\square^n p \supset \forall n\square^n q).$$

And that is just the starred version of (191).

It remains to argue for S4. To do that, I shall first show by induction on n that:

(202) $\Diamond^n \Diamond^* p \supset \Diamond^* p$.

Start with the case $n=1$. What we need to show is that the following is a theorem:

(203) $\Diamond \exists n \Diamond^n p \supset \exists n \Diamond^n p$.

By the Barcan special case (193), we have:

(204) $\Diamond \exists n \Diamond^n p \supset \exists n \Diamond \Diamond^n p$.

But $\Diamond \Diamond^n p$ is just $\Diamond^{n+1} p$, so if $\exists n \Diamond \Diamond^n p$, then $\exists m \Diamond^m p$ (just take $m=n+1$). So the right hand side of (204) implies the right hand side of (203), and the argument for (203) is complete.

Suppose now that (202) is a theorem for $n=m-1$, where $m>1$, i.e. that it is a theorem that:

(205) $\Diamond^{m-1} \exists n \Diamond^n p \supset \exists n \Diamond^n p$.

For purposes of mathematical induction, we need to show that (202) holds for $n=m$.

Now, by the unstarred version of (191):

(206) $\Box(\sim q \supset \sim r) \supset (\Box \sim q \supset \Box \sim r)$.

Contraposing within the box operator on the left hand side (and using Necessitation and (191)), while using the duality between box and diamond on the right:

(207) $\Box(r \supset q) \supset (\sim \Diamond q \supset \sim \Diamond r)$.

Contraposing on the right hand side we get the standard theorem:

(208) $\Box(r \supset q) \supset (\Diamond r \supset \Diamond q)$.

Next, because (205) is a theorem, by Necessitation and (208) we get:

(209) $\Diamond \Diamond^{m-1} \exists n \Diamond^n p \supset \Diamond \exists n \Diamond^n p$.

Since $\Diamond \Diamond^{m-1}$ is $\Diamond^m$, combining (209) this with the already shown (203), we get

(210) $\Diamond^m \exists n \Diamond^n p \supset \exists n \Diamond^n p$

which was to be shown. By mathematical induction, we get (202) holding for all n. Therefore, if $\exists n \Diamond^n \Diamond^* p$, then $\Diamond^* p$, i.e. if $\Diamond^* \Diamond^* p$, then $\Diamond^* p$, and so S4 holds.

Bibliography and Further Reading

Adams, Robert M. "Theories of actuality," *Noûs*, 8 (1974): 211–31; reprinted in: Loux (1979).

——"Middle knowledge and the problem of evil," *American Philosophical Quarterly*, 14 (1977): 109–117; reprinted in: Gale and Pruss (2003).

——"Primitive thisness and primitive identity," *Journal of Philosophy*, 76 (1979): 5–26.

——"Actualism and thisness," *Synthese*, 49 (1981): 3–42.

——*Leibniz: Determinist, Theist, Idealist*. Oxford: Oxford University Press, 1994.

Alston, William. "The inductive argument from evil and the human cognitive condition," in: J. E. Tomberlin (ed.), *Philosophical Perspectives*, 5 (1991): 29–68; reprinted in: Gale and Pruss (2003).

Anscombe, G. E. M. "Causality and determination," in: P. van Inwagen and D. W. Zimmerman (eds.). *Metaphysics: The Big Questions*. Oxford: Blackwell, 1999: pp. 244–58.

Aquinas, Thomas. *Summa Theologica*. The Fathers of the English Dominican Province (trans.). New York: Benziger Brothers, 1948.

——*On Being and Essence*. Armand Maurer, C.S.B. (trans.). Toronto, ON: Pontifical Institute of Mediaeval Studies, 1949.

Ariew, Roger and Garber, Daniel. *G. W. Leibniz: Philosophical Essays*. Indianapolis, IN: Hackett, 1989.

Aristotle. *The Complete Works of Aristotle in the Revised Oxford Translation*. Jonathan Barnes (ed.). Princeton, NJ: Princeton University Press, 1984.

Armstrong, D. M. *A Combinatorial Theory of Possibility*. Cambridge: Cambridge University Press, 1989.

——*A World of State of Affairs*. Cambridge: Cambridge University Press, 1997.

Augustine of Hippo. *The Confessions*. John K. Ryan (trans.). New York: Doubleday, 1960.

Barker, Stephen. "Counterfactuals, probabilistic counterfactuals and causation," *Mind*, 108 (1999): 427–69.

Beedle, Andrew. "Modal fatalism," *Philosophical Quarterly* 46 (1996): 488–95.

Beer, Michelle. "Temporal indexicals and the passage of time," *Philosophical Quarterly*, 38 (1988): 158–64.

Bennett, Jonathan. "Counterfactuals and temporal direction," *Philosophical Review*, 95 (1984): 509–27.

Bigelow, John. "Possible worlds foundations for probability," *Journal of Philosophical Logic*, 5 (1976): 299–320.

Black, Max. "The identity of indiscernibles," *Mind*, 61 (1952), 153–64.

Boolos, George S. and Jeffrey, Richard C. *Computability and Logic*. 3rd edn. Cambridge: Cambridge University Press, 1995.

Cameron, Ross Paul. "Truthmakers and modality," *Synthese*, 164 (2008): 261–80.

Cartwright, Nancy. *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press, 2001.

——*How the Laws of Physics Lie*. Oxford: Clarendon, 2002.

Castañeda, H.-N. "Thinking and the structure of the world," *Philosophia*, 4 (1974): 3–40.

——"Fictional and reality: their basic connections," *Poetica*, 8 (1989): 31–62.

Chakrabarti, Arindam. "Metaphysics in India," in: Jaegwon Kim and Ernest Sosa (eds.). *Companion to Metaphysics*. Oxford / Cambridge, MA: Blackwell, 1995: pp. 318–23.

Chalmers, David. "Two-dimensional semantics," in: E. Lepore and B. Smith (eds.). *Oxford Handbook of the Philosophy of Language*. Oxford: Oxford University Press, 2006.

Chihara, Charles S. *The Worlds of Possibility: Modal Realism and the Semantics of Modal Logic*, Oxford: Clarendon, 1998.

Chisholm, Roderick. "Identity through possible worlds: some questions," *Noûs*, 1 (1967): 1–8; reprinted in: Loux (1979).

Clarke, Randolph. "Indeterminism and control," *American Philosophical Quarterly*, 32 (1995): 125–38.

Craig, W. L. and Moreland, J. P. (eds.). *The Blackwell Companion to Natural Theology*. Oxford: Blackwell, 2009.

Cresswell, M. J. "The world is everything that is the case," *Australasian Journal of Philosophy*, 50 (1972): 1–13; reprinted in: Loux (1979).

——*Logics and Languages*. London: Methuen, 1973.

Dancy, Jonathan. *Moral Reasons*. Oxford: Blackwell, 1993.

Davis, Wayne A. "Indicative and subjunctive conditionals," *Philosophical Review*, 88 (1979): 544–64.

Della Rocca, Michael. "PSR," *Philosopher's Imprint*, 10 (2010): 1–12.

Diels, Hermann and Kranz, Walther. *Die Fragmente der Vorsokratiker*. Zurich: Weidmann, 1985.

Dirac, P. A. M. *The Principles of Quantum Mechanics*, 4th edn. Oxford: Oxford University Press, 1987.

Divers, John. *Possible Worlds*. London: Routledge, 2002.

Donnellan, Keith S. "Reference and definite descriptions," *Philosophical Review*, 75 (1966): 281–304.

Edgington, Dorothy. "On conditionals," *Mind*, 104 (1995): 235–329.

Elga, Adam. "Self-locating belief and the sleeping beauty problem," *Analysis*, 60 (2000): 143–7.

——"Statistical mechanics and the asymmetry of counterfactual dependence," *Philosophy of Science*, suppl. to vol. 68 (2001), *PSA 2000*: S313–S324.

Feldman, Fred. "Counterparts," *Journal of Philosophy*, 68 (1971): 406–9.

Findlay, J. N. "Can God's existence be disproved?" *Mind*, 57 (1948): 176–83.

Fine, Kit. "Critical notice for *counterfactuals* by David Lewis," *Mind*, 84 (1975): 451–8.

Forbes, Graeme. "Two solutions to Chisholm's paradox," *Philosophical Studies*, 46 (1984): 171–87.

Forrest, Peter. "Occam's razor and possible worlds," *Monist*, 65 (1982): 456–64.

——and Armstrong, D. M. "An argument against David Lewis' theory of possible worlds," *Australasian Journal of Philosophy*, 62 (1984): 164–8.

——"Ways worlds could be," *Australasian Journal of Philosophy*, 64 (1986): 15–24.

Frankfurt, Harry. "Alternate possibilities and moral responsibility," *Journal of Philosophy*, 66 (1969): 829–39.

French, Steven. "Quantum physics and the identity of indiscernibles," *British Journal for the Philosophy of Science*, 39 (1988): 233–46.

Gale, Richard M. *On the Nature and Existence of God*. Cambridge: Cambridge University Press, 1991.

——and Pruss, Alexander R. "A new cosmological argument," *Religious Studies*, 34 (1999): 461–76.

——(ed.) *Blackwell Companion to Metaphysics*. Oxford: Blackwell, 2002.

——(eds.) *The Existence of God*. Aldershot: Ashgate Publishing, 2003.

Gellman, Jerome. "Prospects for a sound stage 3 of cosmological arguments," *Religious Studies*, 36 (2000): 195–201; reprinted in: Gale and Pruss (2003).

Gerhardt, C. I. *Die philosophischen Schriften von Gottfried Wilhelm Leibniz*. 7 vols. Hildesheim: Olms Verlagsbuchhandlung, 1960–61.

——*Mathematische Schriften: G. W. Leibniz*. 7 vols. Hildesheim: Olms Verlagsbuchhandlung, 1962.

Glass, David. "Can evidence for design be explained away?" Presented at the Formal Methods in the Epistemology of Religion Conference, Leuven, June 2009. http://www.infj.ulst.ac.uk/~dvglass/Research/ExpAway_thms.pdf (accessed 9 December 2010).

Gordon, Cyrus H. *Ugaritic Textbook*. 3 vols. Rome: Pontifical Biblical Institute, 1965.

Grim, Patrick. "There is no set of all truths," *Analysis*, 46 (1986): 186–91.

——"Worlds by supervenience: some further problems," *Analysis*, 57 (1997): 146–51.

Grünbaum, Adolf. "A new critique of theological interpretations of physical cosmology," *British Journal for the Philosophy of Science*, 51 (2000): 1–43.

Haack, Susan. "Lewis' ontological slum," *Review of Metaphysics*, 33 (1977): 415–29.

Hacking, Ian. "The identity of indiscernibles," *Journal of Philosophy*, 72 (1975): 249–56.

Hare, R. M. *The Language of Morals*. Oxford: Oxford University Press, 1964.

Heller, Mark. "Five layers of interpretation for possible worlds," *Philosophical Studies*, 90 (1998): 205–14.

Horwich, Paul. *Truth*. Oxford: Blackwell, 2005.

Hunter, Graeme and Seager, William. "The discreet charm of counterpart theory," *Analysis*, 41 (1983): 73–6.

Jackson, Frank. "A causal theory of counterfactuals," *Australasian Journal of Philosophy*, 55 (1977): 3–21.

Jeffrey, Richard C. *The Logic of Decision*. 2nd edn. Chicago: University of Chicago Press, 1983.

Johnson, Daniel. (doctoral dissertation in progress, Baylor University).

Jubien, Michael. "Could this be magic?," *Philosophical Review*, 100 (1991): 249–67.

——"Propositions and the objects of thought," *Philosophical Studies*, 104 (2001): 47–62.

——*Possibility*, Oxford: Oxford University Press, 2009.

Kane, Robert H. *The Significance of Free Will*. New York: Oxford, 1996.

Kant, Immanuel. *Critique of Pure Reason*. Norman Kemp Smith (trans.). New York: St. Martin's Press, 1965.

——*Critique of Practical Reason*. 3rd edn. Lewis White Beck (trans.). New York: Macmillan, 1993.

Kaplan, David. "Dthat," in: P. Cole (ed.). *Syntax and Semantics 9: Pragmatics*. New York: Academic Press, 1991: pp. 211–43.

Kirk, G. S., Raven, J. E., and Schofield, M. *The Presocratic Philosophers: A Critical History with a Selection of Texts*. 2nd edn. Cambridge: Cambridge University Press, 1983.

Klagge, James C. "Supervenience: perspectives versus possible worlds," *Philosophical Quarterly*, 37 (1987): 312–15.

Koons, Robert C. "A new look at the cosmological argument," *Philosophical Quarterly*, 34 (1997): 171–92.

——"Epistemological foundations for the cosmological argument," in: J. Kvanvig (ed.). *Oxford Studies in the Philosophy of Religion*, vol. 1. Oxford: Oxford University Press, 2008: pp. 105–33.

Kripke, Saul. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.

Krips, H. "Irreducible probabilities and indeterminism," *Journal of Philosophical Logic*, 18 (1989): 155–72.

Kuratowski, K. and Mostowski, A. *Set Theory: With an Introduction to Descriptive Set Theory*. New York: North-Holland, 1976.

Kutach, D. N. "The entropy theory of counterfactuals," *Philosophy of Science*, 69 (2002): 82–104.

Ladyman, J., Don Ross, David Spurrett, and John Collier. *Every Thing Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press, 2007.

Lappin, Shalom. "Why one ought to care about identity across possible worlds in moral judgments," *Ratio*, 22 (1980): 167–9.

Leftow, Brian. "A Leibnizian cosmological argument," *Philosophical Studies*, 57 (1989): 135–55.

Leslie, John. "The prerequisites for life in our universe," *Truth*, 3 (1990): 97–119.

——"A neoplatonist's pantheism," *The Monist*, 80 (1997): 218–31.

——*Infinite Minds: A Philosophical Cosmology*. Oxford: Clarendon, 2001.

Lévinas, Emmanuel. *Totality and Infinity*. Dordrecht: Kluwer, 1991.

Lewis, David. "Anselm and actuality," *Noûs*, 4 (1970): 175–88; reprinted in: Gale and Pruss (2003) and, with added postscripts, in: Lewis (1983b), vol. 1.

——*Counterfactuals*. Oxford: Blackwell, 1973.

——"Counterfactual dependence and time's arrow," *Noûs*, 13 (1979a): 455–76.

——"Attitudes *de dicto* and *de se*," *Philosophical Review*, 88 (1979b): 513–43; reprinted, with added postscripts, in: Lewis (1983b), vol. 1.

——"New work for a theory of universals," *Australasian Journal of Philosophy*, 61 (1983a): 343–77.

——*Philosophical Papers*. 2 vols. Oxford: Oxford University Press, 1983b and 1986.

——*On the Plurality of Worlds*. Oxford: Blackwell, 1986a.

——"Against structural universals," *Australasian Journal of Philosophy*, 64 (1986b): 25–46.

——"Truthmaking and difference-making," *Noûs*, 35 (2001): 602–12.

Linsky, Bernard. "Truth makers for modal propositions," *Monist*, 77 (1994): 192–206.

Loux, Michael. *The Possible and the Actual: Readings in the Metaphysics of Modality*. Ithaca, NY: Cornell University Press, 1979.

Lycan, William G. "The trouble with possible worlds," in: Loux (1979).

——"Two —— no, three —— concepts of possible worlds," *Proceedings of the Aristotelian Society*, 91 (1991): 215–27.

——*Modality and Meaning*. Dordrecht: Kluwer, 1994.

——"The metaphysics of possibilia," in: Gale (2002).

Mackie, Penelope. "Identity, time, and necessity," *Proceedings of the Aristotelian Society*, 98 (1998): 59–78.

Martin, C. B. "Dispositions and conditionals," *Philosophical Quarterly*, 44 (1994): 1–8.

——"On the need for properties: the road to Pythagoreanism and back," *Synthese*, 112 (1997): 193–231.

McCall, Storrs. "Counterfactuals based on real possible worlds," *Noûs*, 18 (1984): 463–78.

McTaggart, John E. "The unreality of time," *Mind*, 17 (1908): 456–73.

Melia, Joseph. "Against modalism," *Philosophical Studies*, 68 (1992): 35–56.

Merricks, Trenton. *Truth and Ontology*. Oxford: Oxford University Press, 2009.

Meyer, Robert K. "God exists!," *Noûs*, 21 (1987): 345–61.

Monk, J.D. *Introduction to Set Theory*, New York: McGraw-Hill, 1969.

Mourelatos, Alexander P. D. "Pre-Socratic origins of the principle that there are no origins from nothing," *Journal of Philosophy*, 78 (1981): 649–65.

Nolan, Daniel. "Recombination unbound," *Philosophical Studies*, 84 (1996): 239–62.

——"Three problems for 'strong' modal fictionalism," *Philosophical Studies*, 87 (1997): 259–75.

O'Connor, Timothy. *Theism and Ultimate Explanation*. Malden, MA: Blackwell, 2008.

Parmenides. *Peri Phuseôs*. In: Kirk, Raven, and Schofield (1983).

Parsons, Terence "Prolegomenon to a Meinongian semantics," *Journal of Philosophy*, 71 (1974): 561–81.

——*Nonexistent Objects*. New Haven, NY: Yale University Press, 1980.

Peacocke, Christopher. *Being Known*. Oxford: Clarendon, 1999.

Place, Ullin T. "'*De re*' modality without possible worlds," *Acta Analytica*, 12 (1997): 129–43.

Plantinga, Alvin. "Transworld identity or worldbound individuals?," in: Milton Munitz (ed.). *Logic and Ontology*. New York: New York University Press, 1973; reprinted in: Loux (1979).

——*The Nature of Necessity*. Oxford: Clarendon, 1974.

——"Actualism and possible worlds," *Theoria*, 42 (1976): 139–60; reprinted in: Loux (1979).

——"Two concepts of modality: modal realism and modal reductionism," in: J. E. Tomberlin (ed.). *Philosophical Perspectives, 1: Metaphysics*. Atascadero: Ridgeview Publishing, 1987.

——*God and Other Minds: A Study in the Rational Justification of Belief in God*. Ithaca, NY: Cornell University Press, 1990.

——"Why propositions cannot be concrete," in: M. Davidson (ed.). *Essays in the Metaphysics of Modality*. Oxford: Oxford University Press, 2006: pp. 229–33.

Pruss, Alexander R. "Leibniz's approach to individuation and Strawson's criticisms," *Studia Leibnitiana*, 30 (1998a): 116–23.

——"The Hume-Edwards principle and the cosmological argument," *International Journal for Philosophy of Religion*, 43 (1998b): 149–65; reprinted in: Gale and Pruss (2003).

——"Lying and speaking your interlocutor's language," *The Thomist*, 63 (1999): 439–53.

——*Possible Worlds: What They Are Good For and What They Are*. Doctoral Dissertation, University of Pittsburgh, 2001.

——"Śaṃkara's Principle and two ontomystical arguments," *International Journal for the Philosophy of Religion*, 49 (2001): 111–20.

——"The actual and the possible," in: Gale (2002): 317–33.

——"David Lewis's counterfactual arrow of time," *Noûs*, 37 (2003): 606–37.

——*The Principle of Sufficient Reason: A Reassessment*, Cambridge: Cambridge University Press, 2006.

——"Conjunctions, disjunctions and Lewisian semantics for counterfactuals," *Synthese*, 156 (2007): 33–52.

——"The Leibnizian cosmological argument," in: Craig and Moreland (2009).

——"Probability and the open future view," *Faith and Philosophy*, 27 (2010): 190–6.

——"The A-theory of time and induction," *Philosophical Studies* (forthcoming).

Putnam, Hilary. *Meaning and the Moral Sciences*. London: Routledge and Kegan Paul, 1978.

——*Realism and Reason*. Cambridge: Cambridge University Press, 1983.

Quine, W. V. O. *Word and Object*, Cambridge, MA: M.I.T. Press, 1960.

——"Ontological relativity," *Journal of Philosophy*, 65 (1968): 185–212.

Rasmussen, Joshua L. "From a necessary being to God," *International Journal of Philosophy of Religion*, 66 (2009): 1–13.

——*What Propositions May Correspond To and How They Do It*. Doctoral Dissertation, University of Notre Dame, 2010.

Rescher, Nicholas. "How many possible worlds are there?" *Philosophy and Phenomenological Research*, 59 (1999): 403–20.

——*Nature and Understanding: The Metaphysics and Method of Science*. Oxford: Clarendon, 2000.

Rhoda, Alan. "Probability, truth, and the openness of the future: a reply to Pruss," *Faith and Philosophy*, 27 (2010): 197–204.

Richards, Tom. "The worlds of David Lewis," *Australasian Journal of Philosophy*, 53 (1975): 105–18.

Roper, Andrew. "Toward an eliminative reduction of possible worlds," *The Philosophical Quarterly*, 32 (1982): 45–59.

Rosen, Gideon. "Modal fictionalism," *Mind*, 99 (1990): 327–54.

——"Modal fictionalism fixed," *Analysis*, 55 (1995): 67–73.

Ross, James F. *Philosophical Theology*. Indianapolis, IN: Bobbs-Merrill, 1969.

Rowe, William L. *The Cosmological Argument*. Princeton, NJ: Princeton University Press, 1975.

——"Rationalistic theology and some principles of explanation," *Faith and Philosophy*, 1 (1984): 357–69.

Roy, Tony. "Worlds and modality," *Philosophical Review*, 102 (1993): 335–62.

Rundle, Bede. *Why there is Something rather than Nothing*. Oxford: Oxford University Press, 2004.

Salmon, Wesley. *Four Decades of Scientific Explanation*. Pittsburgh, PA: University of Pittsburgh Press, 2006.

Sellars, Wilfrid. *Science and Metaphysics*. London: Routledge and Kegan Paul, 1968.

——*Empiricism and the Philosophy of Mind*. Cambridge, MA: Harvard University Press, 1997.

Shalkowski, Scott A. "The ontological ground of the alethic modality," *Philosophical Review*, 103 (1994): 669–88.

Shoemaker, Sydney (ed.). "Causality and properties," in: *Identity, Cause and Mind*, Cambridge: Cambridge University Press, 1984: 206–33.

Sider, Theodore. "The ersatz pluriverse," *Journal of Philosophy*, 99 (2002): 279–315.

——"Reductive Theories of Modality," in: M. J. Loux and D. W. Zimmerman (eds.). *The Oxford Handbook of Metaphysics*. Oxford: Oxford University Press, 2003: 180–208.

Skyrms, Brian. "Possible worlds, physics and metaphysics," *Philosophical Studies*, 30 (1976): 323–32.

Spinoza, Benedict de. *The Ethics; Treatise on the Emendation of the Intellect; Selected Letters*. Samuel Shirley (trans.). Indianapolis / Cambridge: Hackett, 1992.

Stalnaker, Robert. "A theory of conditionals," in: Nicholas Rescher (ed.). *Studies in Logical Theory*. Oxford: Blackwell, 1968.

——"Possible worlds," *Noûs*, 10 (1976): 65–75; reprinted in: Loux (1979).

——"Possible worlds and situations," *Journal of Philosophical Logic*, 15 (1986): 109–23.

Strawson, P. F. *Individuals: An Essay in Descriptive Metaphysics*. London: Methuen, 1959.

Swinburne, Richard. "The irreducibility of causation," *Dialectica*, 51 (1997): 79–91.

——*The Existence of God*. 2nd edn. Oxford: Oxford University Press, 2004.

Takeuti, G. and Zaring, W. M. *Introduction to Axiomatic Set Theory*. New York: Springer-Verlag, 1971.

Tillich, Paul. *Dynamics of Faith*. New York: Harper, 1957.

Turner, Donald A. "The many-universes solution to the problem of evil," in: Gale and Pruss (2003): 143–59.

van Inwagen, Peter. "Indexicality and actuality," *Philosophical Review*, 89 (1980): 403–26.

——*An Essay on Free Will*. Oxford: Oxford University Press, 1983.

——"Two concepts of possible worlds," *Midwest Studies in Philosophy*, 11 (1986): 185–213.

Wittgenstein, Ludwig. *Tractatus Logico-Philosophicus*. Transl. C. K. Ogden, Trowbridge UK: Redwood Books.

Index