The Ersatz Pluriverse*

Theodore Sider *Journal of Philosophy* 99 (2002): 279–315

While many are impressed with the utility of possible worlds in linguistics and philosophy, few can accept the modal realism of David Lewis, who regards possible worlds as *sui generis* entities of a kind with the concrete world we inhabit.[1]

Not all uses of possible worlds require exotic ontology. Consider, for instance, the use of Kripke models to establish formal results in modal logic. These models contain sets often regarded for heuristic reasons as sets of "possible worlds". But the "worlds" in these sets can be anything at all; they can be numbers, or people, or fish. The set of worlds, together with the accessibility relation and the rest of the model, is used as a purely formal structure.[2] One can even go beyond establishing results about formal systems and apply Kripke models to English, as Charles Chihara has recently argued.[3] Chihara shows, for instance, how to use Kripke models (plus primitive modal notions) to give an account of validity for English modal sentences. In other cases worlds are not really needed at all. It is often vivid to give a counterexample thus: "There is a possible world in which P. Since your theory implies that in all worlds, not-P, your theory is wrong." But the counterexample could just as easily be given using modal operators: "Possibly, P. Since your theory implies that it is necessary that not-P, your theory is wrong."

*Since writing this paper I have learned that Daniel Nolan has independently developed a somewhat similar account; see chapter 5 of his *Topics in the Philosophy of Possible Worlds*. Previous versions of this paper were presented at the University of Rochester, the 1997 Pacific Division APA, the 1997 Notre Dame Mighty Midwestern Metaphysician's conference, Syracuse University, Princeton University, and Yale University. I would like to thank David Braun, Phillip Bricker, Cian Dorr, Tamar Szabó Gendler, Jeff Goodman, John Hawthorne, Mark Johnston, Tom McKay, John Mouracade, Mark Moyer, Brent Mundy, Daniel Nolan, Alvin Plantinga, Gideon Rosen, Ed Wierenga, referees, and, especially, Kit Fine, for helpful comments. Two further notes of gratitude. The main idea behind the pluriverse theory occurred to me while thinking about an early draft of Joseph Melia's "Reducing Possibilities to Language" (*Analysis* 61 (2001): 19–29), although he is not responsible for shortcomings of the present paper. Finally, I am grateful to David Lewis for his comments on this paper, and for his writings throughout the years, without which this paper and many others like it would not exist.

[1] *On the Plurality of Worlds* (Oxford: Basil Blackwell, 1986).

[2] See G. E. Hughes and M. J. Cresswell. *A New Introduction to Modal Logic*. (London: Routledge, 1996).

[3] *The Worlds of Possibility* (Oxford: Clarendon Press, 1998).

Unfortunately for those who yearn for desert landscapes, many interesting applications require more ontological seriousness. In *The Conscious Mind*, David Chalmers uses worlds to (among other things) set up his two-dimensional modal framework.[4] In defining the primary intension of 'water' as a function that assigns to any (centered) world, w, the class of things that would count as water in w when w is "considered as actual", Chalmers is not thinking of the set of worlds as a purely mathematical structure. He means to refer to a *particular* function defined on a space of genuine possible worlds.

Or consider Lewis's formulation of materialism as the claim that no two possible worlds lacking "alien natural properties" differ without differing physically.[5] This is no claim about the formal structure of Kripke models; it is a claim about a particular class of worlds, the class of worlds that lack alien natural properties.

Other examples abound. Linguists and philosophers of language utilize set-theoretic constructions out of possible worlds and individuals as semantic values. Philosophers use possible worlds to define probability functions, properties and propositions. Worlds are used to state theories, make claims, formulate supervenience theses and illuminate distinctions. Worlds may be used to give the truth conditions for ordinary quantification over possibilities, for example "there are at least five ways to win this chess match". The journals and books published in the last forty years contain hundreds of uses of possible worlds, few of which could be reconstructed as purely formal uses of Kripke models or as Chihara-style applications. Moreover, these invocations of possible worlds talk usually go far beyond using possible worlds language as a kind of vivid shorthand for sentences containing modal operators: it is well-known that a language employing quantification over possible worlds and individuals has more expressive power than the language of modal predicate logic.[6]

Possible worlds semantics and metaphysics appear to require genuine worlds. And yet, who can believe Lewis's modal realism?

Like many, I turn to reduction. An adequate reduction of talk of possibilia

[4] Oxford: Oxford University Press, 1996; chapter 2.

[5] "New Work for a Theory of Universals", *Australasian Journal of Philosophy* 61: 343–377; especially pp. 361–365.

[6] See Allen Hazen, "Expressive Completeness in Modal Languages", *Journal of Philosophical Logic* 5 (1976): 25–46, especially section 3. Even if one enriches modal predicate logic in various ways, for example with an actuality operator, or even the bolder enhancements of pp. 90–93 of Graeme Forbes, *The Metaphysics of Modality* (Oxford: Clarendon Press, 1985), it is far from clear how to simulate full quantification over set-theoretic constructions of possibilia.

— worlds and their inhabitants — must be materially adequate, but must not appeal to objectionable entities (like Lewisian worlds). Given such a reduction, the contemporary possible-worlds theorist's tools may be freely used with a clear ontological conscience.

A good starting point is what Lewis calls "linguistic ersatzism", which identifies possible worlds with maximal consistent sets of sentences. But linguistic ersatzism faces an apparently devastating problem, the problem of descriptive power. The solution to this problem will lead to an unfamiliar but attractive theory.

Because of my own views (which I will not defend here), the theory to be developed will assume world-bound individuals and counterpart theory.[7] But the theory can be developed under other assumptions.

Section I describes the problem of descriptive power for linguistic ersatzism. Section II introduces the theory to be defended and describes its solution to the problem of descriptive power. Section III then gives an extensive formal development of the theory. Section IV replies to objections, and section V compares the theory with modal fictionalism.

1. Linguistic ersatzism and the problem of descriptive power

The story is a familiar one. We find ourselves apparently quantifying over things we regard as ontologically objectionable. We therefore reinterpret this quantification as *really* being over things more easily accepted. The instance of the story at hand is the reduction of possible worlds to linguistic ersatz worlds.

A linguistic ersatz possible world is a maximal consistent set of sentences. To construct a possible world in which a donkey talks, we need only include the sentence 'A donkey talks' in a set, along with enough other sentences to insure that the set is not silent about any matter (maximality), but not so many that sentences in the set contradict each other (consistency). The notion of consistency here must be modal: a set is consistent iff it is possible that all the members of the set be true together. For mere logical consistency will not do: sets of sentences asserting the existence of married bachelors and round squares may be logically consistent but will not correspond to possible worlds. Unless modal consistency can be reduced in some way, linguistic ersatz worlds

[7] See David Lewis, "Counterpart Theory and Quantified Modal Logic," *Journal of Philosophy* 65 (1968): 113–126. I defend counterpart theory (both modal and temporal) in my *Four-Dimensionalism: An Ontology of Persistence and Time* (Oxford: Clarendon Press, 2001).

cannot be used in a reductive analysis of modality, on pain of circularity. But the linguistic ersatzist can accept this limitation. The reduction of worlds to language still has a point, for it allows us to reduce all talk of worlds — which runs far beyond that which can be said utilizing merely the modal operators — to talk of possibility and necessity. As for these, they may one day be reduced in some way that does not involve worlds, or they may remain primitive.[8]

Certain objections are quickly answered by taking a broad view of what counts in the present context as a sentence. Let us allow infinitely long sentences;[9] and let the language of those sentences be "Lagadonian", in that objects, properties and relations count as names of themselves, so that every actually existing object, property or relation has a name.[10] Natural languages are neither infinitary nor Lagadonian, but there is no need to take 'sentence' or 'language' very seriously. The worldmaking language need not be learnable or speakable (though the difficulty of writing Lagadonian sentences will require me to revert to English in examples). All that is needed is that its "sentences" have well-defined meanings and be ontologically unobjectionable. These sentences can be understood as mathematical objects — infinite sequences of Lagadonian names, quantifiers, variables, and so on.

There are various ways to fill in this sketch of linguistic ersatzism[11], but all face the following *problem of descriptive power*.[12] An actual language apparently cannot fully describe possibilities involving things that do not actually exist. The ersatzist can attempt to construct these possibilities using qualitative sentences describing what non-actual entities would be like. But such attempts conflate distinct possibilities. The problem comes in two forms, one involving non-actual individuals, the other involving non-actual fundamental properties and relations.

A possible world in which I am six feet tall may be constructed by embedding

[8]See my "Reductive Theories of Modality", in Michael J. Loux and Dean W. Zimmerman, eds., *The Oxford Handbook of Metaphysics* (Oxford: Oxford University Press, forthcoming).

[9]See Phillip Bricker, "Reducing Possible Worlds to Language", *Philosophical Studies* 52 (1987): 331–355, especially pp. 343–349; and M. A. Dickmann, *Large Infinitary Languages: Model Theory* (Amsterdam: North-Holland Publishing Corp., 1975).

[10]See Lewis, *On the Plurality of Worlds*, pp. 145–146.

[11]See Lewis, *ibid.*, chapter 3.2.

[12]See Lewis, *ibid.*, pp. 158–165; Bricker, *op. cit.* pp. 349–353; Alan McMichael, "A Problem For Actualism About Possible Worlds", *Philosophical Review* 92 (1983): 49–66. For excellent survey and discussion see Allen Hazen, "Actualism Again", *Philosophical Studies* 84 (1996): 155–181, and chapter 5 of Daniel Nolan, *Topics in the Philosophy of Possible Worlds* (New York: Routledge, 2002).

the sentence 'Ted is six feet tall' in a maximal consistent set of sentences. But one cannot thus construct an ersatz world in which some individual that does not actually exist is six feet tall, since we have no names for non-actual individuals. Instead, one must use a descriptive sentence:

> There is some individual distinct from Ted, Tom, Dick, Harry, …,
> who is 6 feet tall.

where the list 'Ted, Tom, Dick, Harry, …' includes a name for each actual individual. Worlds with non-actual individuals are thereby constructed from *actualist* descriptions of what those individuals would be like — descriptions that name only actual individuals and qualitative, general characteristics. But this procedure will never distinguish between worlds that differ solely over which non-actual individuals play which qualitative roles. We could construct a world in which one man wears a mask of a bat, drives a car and fights crime with a younger masked chum named 'Robin'; but what of a distinct world in which the person who in the first world wears the bat-mask is now named 'Robin', wears the Robin mask, and so on, and who is replaced in the Batman role by the person who wore the Robin mask in the first world? No sentences in an actual worldmaking language will distinguish these worlds, so the current proposal identifies the apparently distinct worlds with a single set of sentences, and hence with each other.

The objection assumes there can be qualitatively identical worlds differing in which qualitative roles are played by which individuals. This assumption is a controversial doctrine sometimes called haecceitism. Its opponents, anti-haecceitists, say that worlds vary only in their qualitative descriptions; therefore the limitation of the linguist ersatzist to actualist descriptions is no limitation at all. Given my acceptance of counterpart theory, I accept anti-haecceitism and thus do not mind this limitation; but it is better to have an account of worlds that is independent of this controversial doctrine. There is, moreover, an analog to the objection that does not depend on haecceitism.

Possibilities involving non-actual *properties*, as well as individuals, must be accounted for. As before, the ersatzer can use descriptions of the roles such properties would play; but as before, such descriptions will not distinguish possible worlds in which non-actual properties swap roles.

There could have existed two fundamental types of matter, call them A-matter and B-matter, which do not actually exist, playing a certain nomic role which may be partially described as follows: A-matter attracts both negatively and positively charged things, whereas B-matter repels each. Let us understand "fundamental" so that neither supervenes on, or may be constructed in any way from, properties and relations instantiated in the actual world. How to construct an ersatz world corresponding to this possibility? We have no names for the properties since they do not in fact exist. Our best attempt will be to use descriptions, perhaps describing the roles in the laws of nature that A- and B-matter would play, or perhaps describing the pattern of distribution throughout spacetime A- and B-matter would have. Choosing the former course, let us embed the following sentence in a maximal consistent set:

> (1) There are two fundamental properties, P and Q, such that i) neither P nor Q is identical to charge, charm, ..., and ii) it is a law of nature that objects with P attract both negatively charged things and positively charged things, while objects with Q repel negatively and positively charged things

But we can also imagine a distinct world in which A-matter and B-matter have swapped nomic roles: in this new world it is B-matter rather than A-matter that attracts charged particles. Since neither sentence (1) nor any other actual sentence will distinguish these worlds, linguistic ersatzism fails as a general reduction of possible worlds talk.

Some will reply that since properties are abstract objects they exist necessarily (or better, that necessarily, every property exists necessarily). If the reply were correct then uninstantiated properties of *being made up of A-matter* and *being made up of B-matter* would exist in the actual world, could serve as Lagadonian names of themselves, and could be included in sentences that would distinguish the worlds in question. One could construct, for example, a sentence saying that objects made up of A-matter attract charged things, which would be true of the first world but not the second.[13] But I cannot accept the reply, for two reasons.[14] First, it requires an ontology on which the existence of properties is radically independent of the goings-on of the concrete

[13] An analogous reply may be made to the objects version of the problem of descriptive power, based on uninstantiated individual essences, as in section 5 of Alvin Plantinga "Actualism and Possible Worlds", *Theoria* 42 (1976): 139–160. The criticisms in the text apply to this reply as well; but there are additional problems: see section 2 of Robert Merrihew Adams, "Actualism and Thisness", *Synthese* 49 (1981): 3–41, and Kit Fine, "Plantinga on the Reduction of Possibilist Discourse", in J. Tomberlin and P. van Inwagen (eds.) *Alvin Plantinga* (Dordrecht: D. Reidel, 1985): 145–186, especially sections 2 and 3.

[14] See also McMichael, *op. cit.*, pp. 59–61 and Lewis, *On the Plurality of Worlds*, pp. 160–161.

world. The theory of worlds should not be thus held hostage to the theory of properties. The properties required for the reply could not be sets of their instances (each would need to be the empty set; A-matter and B-matter would then be identified with each other). Nor could they be immanent universals (in the sense of D. M. Armstrong[15]) which are supposed to be "wholly present in their instances" and incapable of existing uninstantiated. Even Michael Tooley, whose "transcendent" universals can exist uninstantiated in certain cases, would not accept uninstantiated universals of A-matter and B-matter. For Tooley, uninstantiated universals are accepted only when they play a role in the laws of nature, and we may stipulate that A-matter and B-matter play no role in the actual laws of nature.[16]

Note further that since the properties of being made up of A-matter and B-matter would be fundamental, they could not be constructed in any way, even by infinitary means, from less problematic properties, on any of the theories of properties just mentioned.

The second problem with the reply is its ontological extravagance. One of the main reasons to reduce worlds is parsimony. As Quine puts it colorfully, the believer in possibles accepts an "overpopulated" "bloated universe" which is "in many ways unlovely", and "offends the aesthetic sense of us who have a taste for desert landscapes".[17] If non-actualized possibles bloat the universe, so would the actual but uninstantiated properties needed for the reply.

Neither reason for resisting uninstantiated fundamental properties is a reason for rejecting uninstantiated properties as such. On some views, complex properties are composed of or constructed from others, so that constructed uninstantiated properties would be no more objectionable than the properties from which they are constructed. My quarrel is with uninstantiated *fundamental* properties, which cannot be constructed from less problematic ones.

Just as the objects-version of the problem of descriptive power can be avoided by accepting a metaphysical thesis about individuals, namely anti-haecceitism, the properties-version of the problem may be avoided by going in for certain metaphysical claims about properties. For example, if Sydney

[15] *A World of States of Affairs* (Cambridge: Cambridge University Press, 1997); *Universals: An Opinionated Introduction* (Boulder: Westview Press, 1989); *Universals and Scientific Realism (vols. I and II)* (Cambridge: Cambridge University Press, 1978).

[16] See section 3.1.4 of *Causation: A Realist Approach* (Oxford: Oxford University Press, 1987), especially pp. 119–120.

[17] "On What There Is", in his *From a Logical Point of View* (Cambridge, MA: Harvard University Press, 1953), pp. 1–19; pp. 3, 4.

Shoemaker and Chris Swoyer are right that properties have their nomic roles essentially, then the problematic case of nomic role swapping is impossible.[18] Alternatively, one could claim that the identity of a property is determined in some way by its pattern of instantiation. One could then safely account for alien fundamental properties without fear of conflating possibilities with descriptions of the patterns of instantiation such properties would have.[19] But a theory of worlds should be independent of these matters.

It seems that the linguistic ersatzer can solve the problem of descriptive power only by accepting both anti-haecceitism and some strong claim about the identity conditions for properties.[20] This is an unhappy situation. Using possible worlds talk should not commit one to controversial metaphysical doctrines.

2. A new approach to reducing worlds

Fortunately, such commitments can be avoided. A theory of worlds may be constructed that distinguishes possibilities in which non-actual individuals or fundamental properties swap roles, without making special assumptions about the metaphysics of individuals or properties.

Return to the worlds where A-matter and B-matter swap nomic roles. We cannot distinguish the worlds from within, by giving them different descriptions. But we can distinguish the worlds from an external perspective, so to speak. We can say: "there are two distinct worlds and a pair of non-actual properties; in one world the pair instantiates certain nomic roles; in the other world, the

[18] Sydney Shoemaker, "Causality and Properties", in Peter van Inwagen (ed.), *Time and Cause: Essays Presented to Richard Taylor* (Dordrecht: D. Reidel, 1980), pp. 109–135; Chris Swoyer, "The Nature of Natural Laws", *Australasian Journal of Philosophy* 60 (1982): 203–223. For an objection see Lewis, *On the Plurality of Worlds*, pp. 162–163.

[19] See Mark Heller, "Property Counterparts in Ersatz Worlds", *Journal of Philosophy* 95 (1998): 293–316.

[20] Our survey of solutions has not been completely exhaustive. In "Actualism and Thisness", p. 36, Robert Adams, a haecceitist, paraphrases the sentence (RS): "there exist two possible worlds in which a pair of non-actual individuals swap roles" as, roughly, the assertion that there exists a possible world, w, according to which the following is true: there exist two non-actual individuals, x and y, and two ersatz possible worlds, w' and w'', that are alike except for an exchange of roles by x and y. But Adams paraphrases other sentences that quantify over worlds, for example "there exist two worlds, one in which a donkey talks, another in which a bluefish walks" as genuine quantifications over ersatz worlds (for Adams, maximal consistent sets of propositions). A uniform treatment would be preferable.

nomic roles of the properties are reversed".

The theory I propose departs radically from linguistic ersatzism. I do away with ersatz worlds and individuals in reducing worlds talk, and instead use a single "ersatz pluriverse", a single abstract entity that represents the totality of possible worlds and individuals all at once. Two ways of carrying out this idea will be considered below; here I consider the version in which the surrogate is a *pluriverse sentence*, which looks roughly as follows:

> THERE ARE worlds w_1, w_2, …and THERE ARE properties and relations P_1, P_2…that are distinct from the following actual properties and relations: …, and THERE ARE possible individuals x_1, x_2, …that are distinct from the following actual individuals: …, SUCH THAT: …w_1…and …w_2…and …

At the end of the existentially quantified pluriverse sentence, there are conjoined open formulas, one for each possible world. One of these "world-conjuncts" might look like this: "x_1 is in w_1, and has property P_2, and …". The language of the pluriverse sentence must be an infinitary language, since there will need to be infinitely many existential quantifiers and infinitely many world conjuncts, and also some infinitely long world conjuncts.

This requires a departure from the usual way of reducing possible worlds talk. Most familiar reductions are "entity-for-entity" reductions, in that they provide a surrogate entity for each entity to be reduced, an abstract surrogate for each possible world.[21] I propose instead a "holistic" reduction. Instead of individual world-surrogates I offer a single pluriverse-surrogate. Worlds talk cannot, therefore, be talk about possible world surrogates; it must instead be talk about the single surrogate. Suppose we want to say that there is a possible world in which a donkey talks. The usual reductionist paraphrase is that there is a possible world surrogate of some sort according to which a donkey talks. (The linguistic ersatzist's surrogate is a maximal consistent set of sentences; other theories provide other surrogates.) On the pluriverse view we say instead: according to the pluriverse sentence, there is a possible world in which a donkey talks. Thus, quantification over possibles is interpreted, not as quantification over surrogates, but rather as truth of a quantified sentence according to a single surrogate.

[21] See Robert Merrihew Adams, "Theories of Actuality", *Noûs* 8 (1974): 211–231; Plantinga, *op. cit.*; and Robert Stalnaker, "Possible Worlds", *Noûs* 10 (1976): 65–75. Fine ("Plantinga on the Reduction of Possibilist Discourse", pp. 180–183), and Gideon Rosen ("Modal Fictionalism", *Mind* 99 (1990): 327–354) are exceptions.

When I speak of what is true *according to* the ersatz pluriverse, I am not proposing that we take 'according to' as a mysterious or novel primitive. (This is in contrast to Gideon Rosen's modal fictionalism, which I discuss below.) "According to" just means logical entailment of a sort to be explained. (The entailment must not be strict implication, in the sense of the necessity of the material conditional, for it may well be impossible that the pluriverse sentence is true, in which case the pluriverse sentence would strictly imply every sentence.) Thus, a sentence S quantifying over possibilia is reinterpreted as the claim that the pluriverse sentence logically entails S.

A "strict and philosophical" interpretation of quantification over worlds must be distinguished from a more everyday interpretation. I use the words "there are no possible worlds" to deny modal realism, and yet I quantify over worlds when such quantification is reduced via the ersatz pluriverse. The former denial is not intended to receive the ersatz pluriverse analysis. What does get analyzed is quantification over possibilia by philosophers and linguists, as well as ordinary quantification over possibilities, for example when a chess master remarks that there are at least five possibilities for winning a certain chess match.

This holistic approach to reducing worlds is attractive. The object of the reduction is the entire pluriverse, so it is natural to produce a surrogate for that entity. Moreover, the method of representation of this surrogate is parallel to that of linguistic ersatzism. A maximal consistent set of sentences represents a state of affairs iff that set contains sentences that entail that the state of affairs obtains. The pluriverse sentence also represents by entailment: it represents there being a certain sort of world or possible individual by (logically) entailing that such a world or individual exists.

Even before it is spelled out in detail, the pluriverse view can be seen to solve the problem of descriptive power. Assuming properties have neither their nomic roles nor their patterns of distribution essentially, any adequate theory of worlds should allow a pair of possible worlds in which two non-actual fundamental properties swap nomic roles. On the pluriverse view, the sentence "there are such and such possible worlds" is reinterpreted as: "the pluriverse sentence entails that there are such and such possible worlds". Provided the pluriverse sentence is spelled out appropriately, it will logically entail the following sentence asserting the existence of the desired pair of worlds:

> (*) there are two worlds, w_1 and w_2, and two non-actual, fundamental properties, P_1 and P_2, such that: P_1 plays nomic role

R_1 in w_1 and P_2 plays nomic role R_2 in w_1, and P_1 plays nomic role R_2 in w_2 and P_2 plays nomic role R_1 in w_2.

Haecceitists could give a similar response to the problem of alien individuals — the pluriverse sentence could be spelled out so that it entails the sentence:

(**) there are two worlds, w_1 and w_2, and two non-actual individuals x_1 and x_2, such that: x_1 plays qualitative role R_1 in w_1 and x_2 plays qualitative role R_2 in w_1, and x_1 plays qualitative role R_2 in w_2 and x_2 plays qualitative role R_1 in w_2.

(Since I am no haecceitist, I choose not to spell out the pluriverse sentence in this way; instead I reject the existence of worlds differing only by what individuals play what qualitative roles.)

The linguistic ersatzer had to represent non-actual individuals and properties by including existentially quantified sentences within ersatz worlds:

$$\boxed{\exists P \phi(P)} \qquad\qquad \boxed{\exists Q \psi(Q)}$$

But then one ersatz world could not make any assertions about a particular non-actual individual represented within another, since the variables for non-actual individuals were bound to quantifiers that occurred only *within* ersatz worlds. The solution is to move the existential quantifiers outside of the individual world surrogates, to bind occurrences of variables within multiple world surrogates:

$$\exists P \exists Q \left[\quad \boxed{\phi(P,Q)} \qquad\qquad \boxed{\psi(P,Q)} \quad \right]$$

Provided that the details of the pluriverse theory can be adequately filled in, the problem of descriptive power has been solved. The next section concerns those details.

I will develop the pluriverse view within a broadly linguistic ersatzist framework, but my solution to the problem of descriptive power is quite general. Other theories face the problem, and can benefit from the pluriverse strategy. Robert Merrihew Adams, Alvin Plantinga, or Robert Stalnaker could convert their non-linguistic abstract possible worlds into a single non-linguistic abstract pluriverse. Likewise, D. M. Armstrong could incorporate my methods into his combinatorial theory.[22]

[22] See Adams, *ibid*.; Plantinga, *op. cit*.; Stalnaker, *ibid*.; Armstrong, *A Combinatorial Theory of Possibility* (Cambridge: Cambridge University Press, 1989).

3. The construction of the ersatz pluriverse

In the present section I provide a rigorous construction of the ersatz pluriverse. In outline, the construction will run as follows. I will introduce *modal models*, which are similar to Kripke models. Modal models will be called *realistic* when they are faithful to the modal facts. The modal facts will be stated in a *modal language* (containing the $\Box$ and $\Diamond$); truths in this language will determine which modal models are realistic. A second language, the *possibilist language*, contains the sentences to be reduced — sentences about possibilia. The proposed reduction comes in two versions. The version alluded to in the previous section utilizes pluriverse sentences, which are constructed in the possibilist language as maximal descriptions of realistic models. At the end I introduce a non-linguistic version that uses realistic models directly: a sentence, S, in the possibilist language is reinterpreted as meaning that S is true in all realistic models.

3.1 Two languages

I begin with a specification of a *possibilist language*, a language with the resources to describe possible worlds and individuals. This language is in essence the world-making language used by the linguistic ersatzer. Pluriverse sentences will be sentences of this language. Additionally, the possibilist language contains the target sentences of the reduction, those sentences about possibilia that are to be analyzed.

Lewis distinguishes between "rich" and "poor" worldmaking languages. A poor language names only a select few properties and relations, perhaps those of fundamental physics, and therefore will remain silent about many matters. I will use instead a rich language, a language with a Lagadonian name for every (actual, concrete) individual, and a Lagadonian name for every (actual) property and relation (from now on, just "property").[23] The Lagadonian names of properties should include higher-order properties of properties (so that sentences like (*) from the previous section may be formulated). Note the

[23]"Concrete" rules out at least sets, and any other objects that would introduce cardinality problems. Likewise, the properties and relations should be restricted to properties and relations of non-sets. I will speak as if the Lagadonian name of x is just x itself, but officially let it be $\langle 0, x\rangle$. Then we can distinguish between an occurrence of an entity as a variable and an occurrence of a Lagadonian name for that variable. I presuppose an "abundant" construal of properties (see *On the Plurality of Worlds*, pp. 59–69).

use of properties as Lagadonian *names* rather than *predicates*. Given this, the language must also include an instantiation predicate. It will be convenient to make the instantiation predicate multigrade; thus, 'x instantiates P' and 'x and y instantiate R' are both well-formed. The sentences of this language are just sequences of the bits of primitive vocabulary, for example Lagadonian names, the instantiation predicate, quantifiers, Boolean connectives, and so on. The pluriverse theory thus requires set theory, plus the existence of properties and relations (which are the Lagadonian names of themselves) — defending nominalism is not my concern here. The language is infinitary both in allowing infinite sentences and also infinite blocks of quantifiers, in each case of arbitrarily long infinite length (I discuss cardinality worries in section 4.1 below.) The language contains no modal operators. Instead, it contains resources to speak explicitly about worlds: a syntactically distinct category of world-variables w_i, an individual constant @ for the actual world, and a two-place predicate 'exists in'.

Quantification in the possibilist language is first-order. But since the pluriverse theory is to be combined with counterpart theory, there must be certain differences between quantification over individuals and over properties, since the ersatz pluriverse must represent the former but not the latter as being worldbound. Variables for properties are syntactically distinct from variables for individuals; and individual constants for (actual) individuals are syntactically distinct from constants naming (actual) properties.[24] Since individuals are to be represented as worldbound, the instantiation of properties by individuals may be represented as being instantiation *simpliciter*. However, the properties will be represented as recurring in different worlds (as in sentence (*) above); hence the ersatz pluriverse must represent higher order properties as being instantiated *at worlds*. Thus, a distinct predicate for higher-order instantiation must be introduced, multigrade as before, but with an extra place for worlds "$P_1, \ldots, P_n$ *instantiate* Q *at* w". Within the class of property constants and variables, let there also be a syntactic distinction between those for properties of individuals (type 0), and higher order (type 1) properties of properties; and let there be a syntactically associated number of places with each property constant, namely, the number of places of that Lagadonian constant itself. (Syntactically, the constants are names, not predicates, so the number of places does not affect

[24]Since the constants are Lagadonian, this is no stipulative claim about the shapes of the constants! It simply amounts to differential treatment of different sorts of constants at various points in the definitions that will follow.

grammar; rather it will constrain the interpretation of these constants within models.)

The second language is the *modal language*. This language differs from the possibilist language by dropping the world variables, @, and 'exists in', and including in their place the modal operators $\Box$ and $\Diamond$ and a sentential actuality operator. Moreover, the predicate for higher-order instantiation in the modal language no longer has an argument place for worlds. Otherwise, the modal language is like the possibilist language: it is an infinitary first-order language with Lagadonian names for all actual concrete individuals and properties, and syntactic distinctions between constants and variables for individuals and properties.

3.2 Modal models

My modal models are familiar: they are based on a structure in which "possible individuals" instantiate "possible properties and relations" in different "possible worlds". I use scare-quotes here because these elements of modal models are not to be taken with ontological seriousness: the set of "worlds" in a modal model may contain anything at all — numbers, people, fish, pure sets, and so on. That this will not limit the ersatz pluriverse account to purely formal uses of worlds will be made clear in section 3.8 below.

Leaving out accessibility for simplicity, a *modal structure* will be understood as a 6-tuple $\langle \mathbf{W}, \mathbf{r}, \mathbf{D}, \mathbf{P}, \mathbf{Q}, \mathbf{I} \rangle$ where:

- $\mathbf{W}$ is a non-empty set (the set of "worlds")
- $\mathbf{r}$ is some member of $\mathbf{W}$ ("the actual world")
- $\mathbf{D}$ is a set (the set of "individuals")
- $\mathbf{P}$ is a set (the set of "properties"). Each member of $\mathbf{P}$ is to have two associated integers. The first integer (≥ 1) is the "number of places" of the property; the second (0 or 1) is the "type" of the property: type 0 for properties of individuals, type 1 for properties of properties.
- No two of $\mathbf{W}$, $\mathbf{D}$, and $\mathbf{P}$ overlap
- $\mathbf{Q}$ is a function that assigns to each world w an ordered pair $\langle \mathbf{D}_w, \mathbf{P}_w \rangle$, where $\mathbf{D}_w \subseteq \mathbf{D}$ and $\mathbf{P}_w \subseteq \mathbf{P}$. ($\mathbf{D}_w$ and $\mathbf{P}_w$ are to be thought of as the individuals and properties existing in w.) We impose the requirement of worldbound individuals: that if

w$\neq$w' then $\mathbf{D}_w$ and $\mathbf{D}_{w'}$ do not overlap. Overlap of distinct $\mathbf{P}_w$'s is permitted — thus transworld identity of properties is represented.

- **I** is a function that assigns to each n-place p $\in$ **P** an n-place intension. If p is type 0 then the intension of p is a function that assigns to each world w a set of n-tuples drawn from $\mathbf{D}_w$ (the "extension of p in w"). If p is type 1 then its intension assigns to each w a set of n-tuples of members of $\mathbf{P}_w$. The requirement that if p $\notin \mathbf{P}_w$ then the extension of p in w is empty should presumably be imposed.

Both the modal language and the possibilist language may be interpreted in these modal structures. The Lagadonian names are common to both languages, and thus each language may be given the same definition of a *modal model*, which may be understood as a modal structure plus an interpretation function **F** that assigns *denotations* to these names:

- to each individual constant, **F** assigns some member of $\mathbf{D}_r$
- to each Lagadonian name for a n-place property of type m (m = 0 or 1), **F** assigns some n-place member of $\mathbf{P}_r$ of type m

The definition of *truth in* a given modal model must be different for our two languages. For the modal language, the definition is more or less the usual one, subject to the following remarks. As usual, one defines truth-in-a-world recursively; truth in a model is then truth in the actual world of that model. $\ulcorner\alpha_1, \ldots\alpha_n$ instantiates $\Pi\urcorner$ (relative to an assignment to the variables; this will be suppressed from now on) is true at w iff the ordered n-tuple of the referents of the terms $\alpha_1, \ldots\alpha_n$ is in the extension, at w, of the property denoted by Π. (Thus, in a sense, the meaning of the instantiation predicate is "hard-wired" into the definition of truth in a model.) Necessity is truth in all worlds; possibility, truth in some world; actuality, truth in the actual world (**r**). The language is infinitary, so the usual truth conditions for infinite conjunctions and infinite blocks of quantifiers must be adapted to the modal case in the natural way.[25] The quantifiers are actualist; thus, the individual variables range, at a world w,

[25] See Dickmann, *op. cit.*

over $\mathbf{D}_w$, the lower order property variables over the type 0 members of $\mathbf{P}_w$, and the higher order property variables over the type 1 members of $\mathbf{P}_w$.[26]

For the possibilist language, truth-in-a-model must be understood differently. Since the possibilist language contains no modal operators, we do not define truth in worlds, but rather truth simpliciter. The modal model is now treated much like an ordinary non-modal model, with domains for three sorts of variables: the world variables range over **W**, the individual variables over the whole of **D**, the property variables over all of the members of **P** with the appropriate type. When $\alpha_1, \ldots, \alpha_n$ are terms for individuals, $\ulcorner \alpha_1, \ldots, \alpha_n$ instantiate $\Pi \urcorner$ is true iff the denotations of $\alpha_1, \ldots, \alpha_n$ are all in the same world, w, and the n-tuple of these denotations is in the extension in w of the denotation of Π. If $\Pi_1, \ldots, \Pi_n$, and Π are terms for properties, and ω is a term for a world (either a world variable or @), then $\ulcorner \Pi_1, \ldots, \Pi_n$ instantiate Π at $\omega \urcorner$ is true iff the n-tuple of the denotations of the Π_i's is in the extension at the denotation of ω of the denotation of Π.The predicate 'exists in' and the constant @ must have "hard-wired" meanings, in the sense that they will always be given the same interpretation in the models:

- The denotation of @ is **r**
- $\ulcorner \alpha$ exists in $\omega \urcorner$, where α is an individual (property) term and ω is a world term, is true iff a$\in\mathbf{D}_w$ (iff a$\in\mathbf{P}_w$), where a and w are the referents of α and ω, respectively

We have, then, a single conception of a modal model, but two definitions of truth in a modal model, call them "truth$_p$" and "truth$_m$", for the possibilist and modal languages, respectively. There are two model-theoretic notions of entailment (and thus of equivalence) corresponding to these notions: a sentence in the possibilist (modal) language entails$_p$ (entails$_m$) another iff the latter is true$_p$ (true$_m$) in every modal model in which the first is true$_p$ (true$_m$).

The construction has assumed worldbound individuals, to make room for counterpart theory (section III. F.). If this assumption were given up, various changes would need to be made, including adding a place for worlds to the instantiation predicate for properties of individuals, making appropriate

[26]We may ignore the fact that the truth definition for sentences with individual quantification across modal operators would ordinarily require counterpart theory. The only modal sentences I will be examining for truth in modal models are those in CONSTRAINTS, which have no such quantification. See 3.3.

revisions to the definition of truth$_m$, and allowing domains of distinct worlds in modal models to overlap.

3.3 Realistic modal models

Since the ersatz pluriverse is intended to represent the space of *possible* worlds, it had better not imply the existence of worlds with married bachelors, round squares and the like. Just as linguistic ersatzers make use of a modal notion of consistency, I will make use of the modal notions of possibility and necessity in constructing the ersatz pluriverse. More specifically, I assume the notion of sentences in the modal language being *true* under a Lagadonian interpretation, or "L-true" for short. On this interpretation, the Lagadonian names (of both individuals and properties and relations) are interpreted as denoting themselves, the instantiation predicate is interpreted as meaning instantiation, and the modal operators $\Box$ and $\Diamond$ are interpreted as meaning necessity and possibility, respectively. L-truth must be sharply distinguished from both truth$_p$ and truth$_m$, each of which holds only relative to modal models. L-truth does not concern modal models, but rather the "real live modal facts".

A certain class, CONSTRAINTS, of L-true sentences in the modal language will be used to construct the ersatz pluriverse. Which sentences, exactly, should be included in CONSTRAINTS? The truth of the members of CONSTRAINTS will be, in essence, "built-into" the ersatz pluriverse. Since I give a counterpart-theoretic account of *de re* modal sentences below, I do not want truth values for these sentences built into the ersatz pluriverse from the start. Thus, CONSTRAINTS will consist of exactly the L-true *de dicto* sentences in the modal language — i.e., the L-true sentences in the modal language that contain neither i) individual variables in modal contexts bound by quantifiers outside those contexts, nor ii) Lagadonian names of actual individuals in modal contexts. (If the theory is not to be coupled with counterpart theory then CONSTRAINTS may include all L-true sentences of the modal language.)

Call any modal model $\mathbf{M}$ in which every member of CONSTRAINTS is true$_m$ a *realistic* modal model. The modal language allows arbitrarily long formulas, so strictly speaking CONSTRAINTS cannot exist as a set: understand the definition as saying that a realistic model is one in which every sentence in the modal language satisfying the membership condition for CONSTRAINTS is true$_m$. Thus, a realistic modal model is a modal model $\mathbf{M}$ such that for every *de dicto* sentence, S, in the modal language, if S is L-true then S is true$_m$ *in* $\mathbf{M}$. A realistic model is a model of logical space that is as accurate as possible, given that we have only truths expressible in the modal language to guide us in its construction.

3.4 Pluriverse sentences

Define a *pluriverse sentence* as any maximal description of any realistic model, where a maximal description of **M** is a sentence in the possibilist language that is true$_p$ in **M** and which entails$_p$ any other sentence in the possibilist language that is true$_p$ in **M**. More than one sentence will satisfy this definition, so I must cease the pretense of uniqueness.[27]

We need to know that pluriverse sentences do indeed exist. The remainder of this section sketches a proof that every realistic model has a maximal description. So assuming that realistic models exist, pluriverse sentences exist.[28] This section may be skipped with little danger to understanding the rest of the paper.

We first define the concept of *isomorphic* modal models in the obvious way:

> $\mathbf{M}$ (= $\langle \mathbf{W},\mathbf{r},\mathbf{D},\mathbf{P},\mathbf{Q},\mathbf{I}\rangle$ + interpretation $\mathbf{F}$) is isomorphic to $\mathbf{M}'$ (= $\langle \mathbf{W}',\mathbf{r}',\mathbf{D}',\mathbf{P}',\mathbf{Q}',\mathbf{I}'\rangle$ + interpretation $\mathbf{F}'$) iff there exists a one-one mapping, f, such that:
>
> i) f maps $\mathbf{W}$ onto $\mathbf{W}'$, $\mathbf{D}$ onto $\mathbf{D}'$, and $\mathbf{P}$ onto $\mathbf{P}'$
>
> ii) $f(\mathbf{r}) = \mathbf{r}'$
>
> iii) $i \in \mathbf{D}_w$ iff $f(i) \in \mathbf{D}'_{f(w)}$, for any $i \in \mathbf{D}$ and any $w \in \mathbf{W}$
>
> iv) $i \in \mathbf{P}_w$ iff $f(i) \in \mathbf{P}'_{f(w)}$, for any $i \in \mathbf{P}$ and any $w \in \mathbf{W}$
>
> v) $\langle i_1, \ldots, i_n\rangle \in \mathbf{I}(j)(w)$ iff $\langle f(i_1), \ldots, f(i_n)\rangle \in \mathbf{I}'(f(j))(f(w))$, for any $w \in \mathbf{W}$, any n-place $j \in \mathbf{P}$ of type 0, and any $i_1, \ldots i_n$ in $\mathbf{D}_w$
>
> vi) $\langle i_1, \ldots, i_n\rangle \in \mathbf{I}(j)(w)$ iff $\langle f(i_1), \ldots, f(i_n)\rangle \in \mathbf{I}'(f(j))(f(w))$, for any $w \in \mathbf{W}$, any n-place $j \in \mathbf{P}$ of type 1, and any $i_1, \ldots i_n$ in $\mathbf{P}_w$

[27] Distinct logically equivalent pluriverse sentences are harmless. It is unclear to me at present whether a harmful multiplicity arises from CONSTRAINTS failing to constrain realistic models up to isomorphism.

[28] One challenge to the existence of realistic models is discussed in section 4.1. A separate challenge: realistic models will not exist, given the current definitions, if the logic of the sentences in CONSTRAINTS is not S5. In that case, an accessibility relation would need to be introduced into realistic models, a predicate for accessibility added to the possibilist language, and the definitions of truth$_m$ and truth$_p$ adjusted accordingly.

vii) $\mathbf{F}'(\beta) = \mathrm{f}(\mathbf{F}(\beta))$, for every property or individual constant β

The following will be proved below:

> **Theorem** for any realistic modal model **M**, there exists a sentence ϕ of the possibilist language (the "canonical pluriverse sentence for **M**") that is true_p in **M**, and which is such that if it is true_p in an arbitrary modal model **M′** then **M′** is isomorphic to **M**

It is obvious that the same sentences in the possibilist language will be true_p in isomorphic models. Hence, it follows from the theorem that in any modal model there is a true_p sentence, in "canonical form", that entails_p every other sentence in the possibilist language that is true_p in that model — a pluriverse sentence for that model. Given the definition of a pluriverse sentence, it then follows immediately that any pluriverse sentence is equivalent_p to some pluriverse sentence in canonical form.

To prove the theorem, consider any modal model **M** (= $\langle \mathbf{W},\mathbf{r},\mathbf{D},\mathbf{P},\mathbf{Q},\mathbf{I}\rangle$ + interpretation **F**). The following notation will be used:

- $\exists(\mathrm{w}_i\text{: } \ldots \mathrm{w}_i \ldots)$: the (possibly infinitary) existential quantification of the class of variables w_i satisfying condition $\ldots w_i \ldots$. Similarly for variables of other types, and for universal quantification.
- $\bigwedge(\phi\text{: } \ldots\phi\ldots)$: the (possibly infinitary) conjunction of the class of formulas ϕ satisfying condition $\ldots\phi\ldots$. Similarly for disjunction: $\bigvee(\phi\text{: } \ldots\phi\ldots)$. Let the disjunction of the empty set be some logically false sentence, such as $\exists x\, x \neq x$; let the conjunction of the empty set be some logically true sentence.
- w, w_i, etc., are world variables; x, y, x_i, etc. are individual variables; P, P_i, Q, etc., are property variables (of either higher or lower type).

Every member of $\mathbf{D}_r$ is denoted by (i.e., is assigned by **F** to) some individual constant: since we have an individual Lagadonian constant for *every* (actual) individual, the following sentence in the modal language is L-true, and so must be true_m in **M**, and so must be true_m at **r**:

$$\forall x \bigvee(x\text{=a: a is an individual constant})$$

Since the constants are Lagadonian, the sentence ⌜a$\neq$b⌝ is L-true and so true$_m$ in **M** whenever a and b are distinct individual constants. Thus, every member of $\mathbf{D}_r$ is denoted in **M** by a unique individual constant. It can be shown similarly that every member of $\mathbf{P}_r$ is denoted by a unique property constant. Next:

> For each i$\in$**W** other than **r**, introduce a distinct world variable w_i.
>
> For each i$\in$**D** but not in $\mathbf{D}_r$, introduce a distinct individual variable x_i.
>
> For each i$\in$**P** but not in $\mathbf{P}_r$, introduce a distinct property variable P_i.
>
> For each i$\in\mathbf{D}_r$, call the individual constant denoting i in **M** "a_i"
>
> For each i$\in\mathbf{P}_r$, call the individual constant denoting i in **M** "p_i"

To simplify the various special cases that arise in connection with the actual world **r**, let us introduce the following abbreviations:

> For any i$\in$**P**, let "Π_i" denote either the constant p_i or the variable P_i, depending on whether i is, or is not, in $\mathbf{P}_r$
>
> For any i$\in$**D**, let "α_i" denote either the constant a_i or the variable x_i, depending on whether i is, or is not, in $\mathbf{D}_r$.
>
> For any i$\in$**W**, let "ω_i" denote either the constant @ or the variable w_i, depending on whether i is, or is not, **r**.

We now construct the required pluriverse sentence ϕ as follows:

$$\phi = \ulcorner \exists(w_i\!: \ \mathrm{i}\in\mathbf{W} \text{ but } \mathrm{i}\neq\mathbf{r})\exists(x_i\!: \ \mathrm{i}\in\mathbf{D} \text{ but } \mathrm{i}\notin\mathbf{D}_r)\exists(P_i\!: \ \mathrm{i}\in\mathbf{P} \text{ but } \mathrm{i}\notin\mathbf{P}_r)\ [\text{DISTINCTNESS \& COMPLETENESS \& } \textstyle\bigwedge(W_i\!: \mathrm{i}\in\mathbf{W})]\urcorner$$

where

$$\text{DISTINCTNESS} = \ulcorner \textstyle\bigwedge(\omega_i \neq \omega_j\!: \mathrm{i, j} \in \mathbf{W} \text{ and } \mathrm{i}\neq\mathrm{j})\ \&\ \bigwedge(\alpha_i \neq \alpha_j\!: \mathrm{i, j} \in \mathbf{D}\ \&\ \mathrm{i}\neq\mathrm{j})\ \&\ \bigwedge(\Pi_i \neq \Pi_j\!: \mathrm{i, j} \in \mathbf{P}\ \&\ \mathrm{i}\neq\mathrm{j})\urcorner$$

and

$$\text{COMPLETENESS} = \ulcorner \forall w \bigvee(w = \omega_i\text{: i}\in\mathbf{W})\ \&\ \forall x \bigvee(x = \alpha_i\text{: i}\in\mathbf{D})\ \&\ \forall P \bigvee(P = \Pi_i\text{: i}\in\mathbf{P})\urcorner$$

and

$$W_i = \ulcorner \forall x[x \text{ exists in } \omega_i \leftrightarrow \bigvee(x = \alpha_j\text{: j}\in\mathbf{D}_i)]\ \&\ \forall P[P \text{ exists in } \omega_i \leftrightarrow \bigvee(P = \Pi_j\text{: j}\in\mathbf{P}_i)]$$

$$\&$$

$$\bigwedge[\forall y_1 \ldots \forall y_n((y_1 \text{ exists in } \omega_i\ \&\ \ldots\&\ y_n \text{ exists in } \omega_i\ \&\ y_1, \ldots, y_n \text{ instantiate } \Pi_j) \leftrightarrow \bigvee(\text{y}_1 = \alpha_{j1}\ \&\ \ldots\&\ \text{y}_n = \alpha_{jn} : \langle \text{j}_1, \ldots, \text{j}_n\rangle \in \text{I(j)(i)}) : \text{j is an n-place type 0 member of } \mathbf{P}]$$

$$\&$$

$$\bigwedge[\forall Q_1 \ldots \forall Q_n((Q_1 \ldots Q_n \text{ instantiate } \Pi_j \text{ at } \omega_i \leftrightarrow \bigvee(Q_1 = \Pi_{j1}\&\ldots\&Q_n = \Pi_{jn} : \langle \text{j}_1, \ldots, \text{j}_n\rangle \in \mathbf{I}\text{(j)(i)}) : \text{j is an n-place type 1 member of } \mathbf{P}]\urcorner$$

Remarks: DISTINCTNESS ensures that no two property or individual terms, whether variables or constants, denote the same thing. COMPLETENESS says that there are no worlds other than those denoted by @ and the variables w_i, and no individuals or properties other than those denoted by the Lagadonian constants and the variables x_i and p_i. The first component of the world conjunct W_i says that the individuals and properties in world i are exactly those denoted by the terms α_j and Π_j. The second component contains an infinite conjunction, one conjunct for each lower order property; each of these conjuncts says, for the property in question, that the individuals that satisfy the property in i are exactly such and such. The third does the same thing for all of the higher-order properties.

It is tedious but straightforward to verify that any model in which ϕ is true$_p$ is isomorphic to **M**.

3.5 The proposal: linguistic version

Let S be any sentence of the possibilist language, and think of S under the Lagadonian interpretation, as a sentence concerning reality, not any modal model. S might be the sort of sentence that proves so useful in philosophy or linguistics, quantifying over non-actual worlds, individuals and properties. Only a modal realist like Lewis could admit the truth of such a sentence — hence the need for a reduction. So:

> **Proposal (linguistic version)** Reinterpret S as the assertion that S is entailed$_p$ by all pluriverse sentences.

Since the proposal only generates truth conditions for sentences stated within the Lagadonian possibilist language, English sentences about possibilia must be regimented in that language. For example, an English subject-predicate atomic sentence $\ulcorner Fa \urcorner$ may be translated as the Lagadonian sentence $\ulcorner \alpha$ instantiates $\pi \urcorner$, where α is a Lagadonian name for the referent of a, and π is a Lagadonian name for the property expressed by F.

Let us examine pluriverse sentences in more depth.

First, pluriverse sentences need not have the logical form introduced in section II:

> THERE ARE worlds w_1, w_2, …and THERE ARE properties and relations P_1, P_2…that are distinct from the following actual properties and relations: …, and THERE ARE possible individuals x_1, x_2, …that are distinct from the following actual individuals: …, SUCH THAT: …w_1…and …w_2…and ….

beginning with quantifiers for the worlds, properties, and individuals, and concluding with an infinite conjunction of open sentences — world conjuncts — one for each world variable. Maximal descriptions of modal models need not have this logical form. However, as was shown in the previous section, every pluriverse sentence is equivalent$_p$ to some such sentence, which I call a *canonical pluriverse sentence*. It is harmless, therefore, to think of pluriverse sentences as having this canonical form.

Second, no pluriverse sentence can assert the existence of a world in which impossibilities occur. Suppose a world conjunct in some pluriverse sentence contained 'x exists at w' and 'x is a married bachelor'. Any pluriverse sentence is true$_p$ in some realistic model; so some individual in the domain of some world of some realistic model would be in the extension of 'married bachelor' in that world. But this is impossible, since the sentence ' there exist no married bachelors' is L-true, and so is a member of CONSTRAINTS, and so must be true$_m$ in any realistic model.

Third, every possibility will be represented as holding in some world, by any pluriverse sentence. Where **M** is any realistic model, if $\ulcorner\Diamond\phi\urcorner$ is an L-truth of the modal language then it must be true$_m$ in **M**, in which case ϕ will be true$_m$ in some world of **M**. But then any pluriverse sentence based on **M** will entail$_p$ a sentence asserting that ϕ holds in some world. Notice that ϕ might be an infinitary sentence since the modal and possibilist languages allow such sentences; thus, there is no restriction to finitely stateable possibilities. Relatedly, pluriverse sentences will represent necessary truths as holding in every possible world. Thus, on the pluriverse view, truths about possibility and necessity in the modal language "mesh" with their worlds translations.

Fourth, the possible worlds represented by the world conjuncts in pluriverse sentences are "completely specific". Suppose a pluriverse sentence ϕ entails$_p$ the following:

> There exists a world, an individual x existing at that world, and properties $P_1 \ldots P_n$ and Q existing at that world, such that x has properties $P_1 \ldots P_n$.

It must then entail$_p$ one of the following:

> There exists a world, an individual x existing at that world, and properties $P_1 \ldots P_n$ and Q existing at that world, such that x has properties $P_1 \ldots P_n$ as well as Q

> There exists a world, an individual x existing at that world, and properties $P_1 \ldots P_n$ and Q existing at that world, such that x has properties $P_1 \ldots P_n$, but not Q

For if ϕ is a pluriverse sentence then it is a maximal description of some modal model **M**; if ϕ entails$_p$ the first sentence then that sentence is true$_p$ in **M**; but then either the second or third sentence must be true$_p$ in **M**, and hence one must be entailed$_p$ by ϕ.[29]

Fifth, as constructed in the previous section, each pluriverse sentence in canonical form contains a clause COMPLETENESS asserting that its inventory of worlds, individuals, and properties is complete:

[29] I ignore vagueness throughout.

> THERE ARE worlds w_1, w_2, …and THERE ARE properties and relations P_1, P_2…that are distinct from the following actual properties and relations: …, and THERE ARE possible individuals x_1, x_2, …that are distinct from the following actual individuals: …, SUCH THAT: COMPLETENESS and …w_1…and …w_2…and …

The clause COMPLETENESS looks like this:

> COMPLETENESS: there are no worlds other than @, w_1, w_2, …, and there are no properties or relations other than charge, charm, [list all actual properties and relations], and P_1, P_2…, and there are no possible individuals other than Socrates, Aristotle, [list all actual individuals], and x_1, x_2, …

If a pluriverse sentence entailed$_p$ no such clause it could not be a *maximal* description of any realistic model, for it would fail to entail$_p$ many of the universally quantified sentences true$_p$ in that realistic model.

Sixth, each pluriverse sentence will include information about the actual world. For any n-place property Π and any n objects $\alpha_1 \ldots \alpha_n$ that, in fact, instantiate Π, the atomic sentence in the modal language, ⌜$\alpha_1 \ldots \alpha_n$ instantiate Π⌝, is L-true, and so is in CONSTRAINTS, and so is true$_m$ in any realistic modal model. Let **M** be any realistic model, and let **r** be the actual world of **M**. Given the definition of a modal model, $\mathbf{D}_r$ contains the denotations $a_1 \ldots a_n$ of $\alpha_1 \ldots \alpha_n$, and $\mathbf{P}_r$ contains p, the denotation of Π. By the definition of truth$_m$, $\langle a_1, \ldots, a_n \rangle$ is in the extension of p at **r** in **M**. Thus, in any realistic modal model, the actual world **r** will encode all the non-modal truths about the (real) actual world. This information finds its way into pluriverse sentences as follows. Given what we have said, the following sentence of the possibilist language must be true$_p$ in any realistic model:

> (@) Π exists in @, $\alpha_1 \ldots \alpha_n$ exist in @, and $\alpha_1 \ldots \alpha_n$ instantiate Π

It must, therefore, be entailed$_p$ by any pluriverse sentence. Therefore: L-true non-modal subject-predicate propositions are reported by pluriverse sentences as holding in the actual world. (Each pluriverse sentence ϕ in canonical form will collect all the information contained in sentences like (@) into a single *actual world conjunct* — the one of the world conjuncts that completely describes **r**, the actual world of the model on which ϕ is based.)

Seventh, pluriverse sentences say more about worlds than they say about them individually; they do not merely entail$_p$ sentences of the form $\ulcorner$there exists a world such that $\phi\urcorner$. Suppose that property variable P occurs in world conjuncts W_1 and W_2 (which in turn correspond to world variables w_1 and w_2) in some pluriverse sentence. It will then entail$_p$:

> there exists a property P, and worlds w_1 and w_2, such that W_1 and W_2

This sort of consequence, a claim about two worlds "from an external perspective", is what makes possible the solution to the problem of descriptive power.

3.6 Inter-world facts

Though the bare-bones pluriverse theory is now in place, several additions are needed to yield a complete theory of worlds. The present section will sketch, in considerably less detail than the preceding sections, how some of this might be carried out.

To gain the full benefits of talk of possibilia, "paradise on the cheap" as Lewis says, we want to speak, not only of possibilia, but also of set-theoretic constructions of possibilia. Chalmers's two-dimensional framework, for example, requires functions defined over possible worlds and individuals. Ersatz pluriverses must therefore represent *sets* of worlds, individuals, and properties in addition to the worlds, individuals and properties themselves. To accomplish this, the language of set theory could be added to the possibilist language, and the definition of truth$_p$ in a modal model modified so that sentences containing set-theoretic vocabulary are true$_p$ in **M** iff they accurately describe **M**'s set-theoretic structure, in the following sense. **M**'s set of "worlds", **W** is a certain set, and thus has a certain set-theoretic structure: it has certain subsets, which themselves have subsets, and so on. Similarly, **M**'s set of individuals, **D**, has a certain set-theoretic structure. **W** and **D** do not need to be augmented to represent their set-theoretic structures; they simply *have* their structures. The definition of truth$_p$ must be modified so that a set-theoretic sentence in the possibilist language turns out true$_p$ in **M** iff it correctly represents the set-theoretic structure that **M** *has*. Thus, the very same pluriverse sentences defined above will now entail$_p$ set-theoretic sentences, despite containing no set-theoretic vocabulary. Sentences about sets of possibilia will be entailed$_p$ by any pluriverse sentence, and so will turn out true, on the pluriverse view.

Facts about sets of possibilia are "inter-world" facts — facts not from the perspective of any one world, but rather from an external perspective. There are other examples of such facts: i) the pluriverse should, perhaps, represent the existence of transworld mereological sums; ii) the pluriverse should, perhaps, represent relations holding between inhabitants of worlds and transworld sets or sums, for example belief relations between persons and transworld sets thought of as propositions; iii) the pluriverse should represent the holding of a counterpart relation between objects from different worlds. To achieve these goals the possibilist language must be enriched, modal models must be augmented to represent the desired facts, and the definition of truth$_p$ must be adjusted accordingly.

The addition of counterpart theory to the pluriverse view will illustrate some of the complications that can arise with these additions. Where **M** is any modal model and **C** is a binary relation over **M**'s domain, **D**, call $\langle \mathbf{M},\mathbf{C} \rangle$ a *C-model*. **C** is to be thought of as the counterpart relation for **M**. (A refinement would be to allow multiple counterpart relations.)[30]

We now need the notion of a *realistic C-model*. Since the counterpart relation is a similarity relation (although not necessarily a relation of intrinsic similarity), whether it holds depends on what properties are instantiated by its relata and what relations its relata bear to their worldmates. Thus, for $\langle \mathbf{M},\mathbf{C} \rangle$ to be realistic, it is not enough for **M** to be realistic; in addition the holding of **C** must "mesh" with the properties instantiated by the objects in that model. This is not specific to counterpart theory: *inter*-world relations will usually be constrained by the *intra*-world facts about their relata.

In fact, the counterpart relation seems *determined* by intra-world facts. (Whether this is true in general, or whether some interesting inter-world relations are constrained without being determined by intra-world facts, is debatable.) Thus, there should be a function, *c-determination*, which when applied to a realistic modal model **M** yields the appropriate counterpart relation, **C**, over the individuals of **M**. Given this function we can then define the notion of a realistic C-model: $\langle \mathbf{M},\mathbf{C} \rangle$ is realistic iff **M** is realistic and **C** = c-determination(**M**). Thus, our problem of defining 'realistic C-model' reduces to the problem of defining 'c-determination'.

[30] See Lewis, "Counterparts of Persons and Their Bodies," *Journal of Philosophy* 68 (1971): 203–211.

The latter would be easy if we had a particular *theory* of the counterpart relation. Suppose, for example, that x is a counterpart of y iff x and y both instantiate a certain property P. Then we could define c-determination(**M**) as the relation that holds between members of **M**'s domain x and y iff both x and y are in the extension of P in their respective worlds. But the counterpart relation is often taken as an undefined primitive (although explained intuitively in terms of similarity).

One *could* replace the primitive counterpart relation with a new primitive, that of c-determination. C-determination could be explained informally in terms of similarity and so would count as a kind of similarity primitive. But C-determination would be an unwieldy primitive. It would be better to define it in terms of (the relevant sort of) similarity itself, and then take the latter as primitive instead. Under certain assumptions this is in fact possible.

Here is a sketch of a definition of 'c-determination' in terms of a similarity predicate 'c-similar'. Consider any modal model, **M**, containing individuals x_1 and x_2, from worlds w_1 and w_2, respectively. Form a pair of open formulas W_1 and W_2 of the possibilist language describing w_1 and w_2 as follows. Represent each of the individuals and properties in the two worlds with distinct terms, letting any properties from the actual world of the model be represented by the Lagadonian constants which denote those properties in the model, and letting all individuals (whether in the actual world of the model or no) and all non-actual properties be represented by variables. Let x_1 and x_2 be the variables corresponding to individuals x_1 and x_2, respectively; let $\alpha_1,\ldots$ be a list of all the variables other than x_1 used to represent things from w_1, and let $\beta_1,\ldots$ be a list of all variables other than x_2 used for w_2. Next, in the manner of section 3.4, using these variables construct (possibly infinite) open formulas W_1 and W_2 that completely describe the pattern of instantiation of properties within worlds w_1 and w_2. We now say that c-determination(**M**) holds between x_1 and x_2 iff the following is an L-true sentence of the modal language:

> Necessarily, for all x_1, $\alpha_1,\ldots$, and for all distinct x_2, $\beta_1,\ldots$, if W_1 and W_2, then x_1 and x_2 are c-similar

("any objects that are as W_1 and W_2 say that x_1 and x_2 are would necessarily be c-similar"). A potential problem, however, is that W_1 and W_2 may not be satisfiable by distinct sets of objects — perhaps there is no one possible world containing wholly distinct parts that are as described by W_1 and W_2. In this case the definition will hold vacuously, in virtue of the antecedent of the conditional,

⌜if W_1 and W_2, then…⌝, being necessarily false. W_1 and W_2 express massive conditions that are satisfied by all the objects in some possible world of some modal model.[31] The following principle would ensure that the definition never holds vacuously: for any two possible worlds, there exists another possible world containing distinct duplicates of the first two worlds. If the principle is true then the conditionals in the definition will not be vacuous, for even though W_1 and W_2 describe two entire possible worlds, there will *also* be a third world containing two parts, one with objects that satisfy W_1, the other with objects that satisfy W_2. But this principle is controversial.

Reliance on this principle could be avoided if we were willing to accept non-vacuous counterfactual conditionals with metaphysically impossible antecedents[32]: we could then say that c-determination(**M**) holds between x_1 and x_2 iff the following is L-true:

> If there had existed x_1, α_1,…, and distinct x_2, β_1,…, such that W_1 and W_2, then it would have been the case that x_1 and x_2 are (relevantly) similar

The cost here would be accepting this counterfactual connective as a new modal primitive.

Whether 'c-determination' is a primitive or is defined in terms of 'c-similar', realistic models can represent the holding of a counterpart relation over possibilia.

3.7 The proposal, non-linguistic version

Realistic models were used to define pluriverse sentences, which then were used to reinterpret sentences about possibilia. But one could bypass the pluriverse sentences and utilize realistic models directly:[33]

> **Proposal (non-linguistic version)** Reinterpret S as the claim that S is true$_p$ in every realistic model.

[31]Compare Lewis, *On the Plurality of Worlds*, p. 238.

[32]I think we do make non-vacuous sense of at least some counterpossibles, though it is an open question whether there are counterpossible rich enough to ground the proposed definition. See Jeffrey Goodman, "Extended Ersatz Realism", forthcoming in *Pacific Philosophical Quarterly*; William Lycan, *Modality and Meaning* (Dordrecht: Kluwer Academic Publishers, 1994), pp. 38–39; and Takashi Yagasawa, "Beyond Possible Worlds", *Philosophical Studies* 53 (1987): 175–204.

[33]This formulation is due to the extremely valuable suggestions of Kit Fine, to whom I am greatly indebted in my search for an adequate formal construction of the ersatz pluriverse.

Given the theorem of section 3.4, the linguistic and non-linguistic versions of the proposal are equivalent.[34] The non-linguistic version is simpler, but I find the linguistic version more intuitive. The linguistic version also facilitates comparison with other theories, for instance linguistic ersatzism and modal fictionalism (to be discussed below).

3.8 Taking stock

Our inquiry into the ontology of worlds has been driven by the dilemma: Lewisian modal realism is unacceptable, yet possibilia are useful in semantics and ontology. My solution is the proposal sketched in the preceding sections. The proposal is superficially similar to the use of Kripke models as "mere structures" since the particular entities in the sets of "worlds", "individuals" and "properties" in modal models are insignificant; these sets may, for example,

[34]To prove: S is true$_p$ in all realistic models iff S is entailed$_p$ by all pluriverse sentences. First consider the proposal before the introduction of set-theoretic vocabulary discussed in 3.6.

Let S be true$_p$ in all realistic models, let PS be a pluriverse sentence, and suppose that PS is true$_p$ in some modal model $\mathbf{M}$. Since PS is a pluriverse sentence, it is a maximal description of some realistic model $\mathbf{M}'$. S is true$_p$ in $\mathbf{M}'$, and so is entailed$_p$ by PS, and so is true$_p$ in $\mathbf{M}$.

On the other hand, suppose S is entailed$_p$ by all pluriverse sentences, and let $\mathbf{M}$ be a realistic model. By the theorem of III. D., some pluriverse sentence is true$_p$ in $\mathbf{M}$; it entails$_p$ S, and so S is true$_p$ in $\mathbf{M}$ as well.

Now consider the introduction of set-theoretic vocabulary from 3.6. The proof just given relies only on i) definitions which have not changed, and ii) the theorem of III. D., that in each realistic modal model there is a true$_p$ pluriverse sentence. Thus, all we must show is that the theorem holds under the new definition of truth$_p$. That theorem established the existence of a sentence ϕ that characterizes a given modal model up to isomorphism. But isomorphic models have the same set-theoretic structure, and hence make the same sentences of the new, set-theoretic possibilist language true$_p$. Thus, ϕ remains a pluriverse sentence for the model, under the new definition.

Notice that ϕ now has new entailments$_p$ (namely, set-theoretic sentences) without having anything added to it, given the new definition of entailment$_p$. Pluriverse sentences could not have characterized all of the set-theoretic facts by name, via Lagadonian names of all the sets in the hierarchy generated by a given model, for then pluriverse sentences could not exist.

Pluriverse sentences still do not entail$_p$ statements about sets existing *in* worlds; how then can the L-truth of '□(the empty set exists)' be accommodated? Probably the best solution is to follow Lewis in regarding the worlds translation of such a sentence as involving a quantifier over sets that exist "from the point of a world" rather *in* the world; see postscript A to "Counterpart Theory and Quantified Modal Logic", in Lewis's *Philosophical Papers, Vol. 1* (Oxford: Oxford University Press, 1983), pp. 39–40.

contain numbers, persons or fish. But as mentioned at the outset, the purely formal use of Kripke models is severely limited. More interesting uses for possible worlds require the ability to speak of particular worlds, classes of worlds, functions defined on worlds, and so on. The proposal I have given allows this, by providing a reduction of sentences like:

- there is a world in which a donkey talks
- there exists a pair of worlds and a pair of non-actual fundamental properties such that the worlds differ only by the properties swapping nomic roles
- no worlds that lack alien fundamental properties differ without differing physically
- there exists a primary intension of the term 'water' — a function that assigns to any world, w, the set of things that would count as water in w when w is "considered as actual"

The reduction does indeed assign truth conditions to these sentences *via* modal models whose classes of "worlds", "individuals" and "properties" may contain numbers, persons, fish, and so on, but the important thing is that the sentences are indeed assigned truth conditions. Thus, my reduction allows one to partake fully of modal metaphysics and semantics.

The construction of realistic modal models assumes the notion of truth (under the Lagadonian interpretation) for *de dicto* sentences in the modal language, and hence assumes *de dicto* necessity and possibility (which are of course interdefineable). These modal notions remain primitive in this paper. Thus, my account of worlds cannot be employed in a reductive account of these modal notions themselves. As explained at the outset, a reduction of talk of possibilia that employs primitive possibility and necessity is nevertheless valuable since talk of possibilia runs beyond what can be said in the language of quantified modal logic. The pluriverse account thus reduces talk about possibilia to *de dicto* sentences in the modal language.

Moreover, as explained in section 3.6, given a similarity primitive (whether c-determination or a similarity relation over individuals), the holding of a counterpart relation can be built into realistic models. The pluriverse account thus yields a theory of the truth conditions for statements about a counterpart relation holding over possibilia. Given the counterpart-theoretic account of *de re* modality in terms of such statements, truth conditions for *de re* modal sentences can be given. The usual counterpart-theoretic reduction of *de re* modality to *de dicto* modality plus similarity is thereby achieved.

4. Problems for the pluriverse view

4.1 Cardinality problems[35]

The pluriverse view runs into trouble if there are too many possible worlds or individuals to form a set. Suppose, for example, that there is no upper bound, not even an infinite upper bound, to how many (concrete) individuals may possibly exist. Then there can be no realistic models, for any modal model **M** has a domain of individuals, **D**, which is a set with some cardinality; but where $\varkappa$ is some larger cardinal, there will be an L-true sentence in the modal language saying that it is possible that there are least $\varkappa$ objects[36]; such a sentence would be in CONSTRAINTS, but would not be true$_m$ in **M**. Any modal model, no matter how large, would misrepresent the pluriverse as containing an upper bound on the size of worlds. A similar problem would arise if there were no upper bound to the number of properties that could exist.

The very first point to make about the cardinality problem is that it is not particular to the pluriverse view. *Every* theory of worlds encounters trouble in this area. The linguistic ersatzist, for example, may admit arbitrarily large worlds, but cannot admit worlds with so many individuals that they cannot all be members of a set (except in special cases where the objects display symmetries allowing simpler description), for a linguistic ersatz world is a maximal consistent *set* of sentences, and sentences themselves are also sets. That is a limitation on possibility, although a bit less severe than an upper bound on world size.[37] Similar problems confront other views that identify possible worlds with abstract entities other than sets of sentences. This is sometimes less easy to see than with linguistic ersatzism, since defenders of these views do not always

[35] I thank Daniel Nolan for helpful comments on this section.

[36] Given the notation of section III. D., this sentence could be the following, where S is a set of $\varkappa$ many variables: $\exists(x : x \in S) \bigwedge(x \neq y : x \in S, y \in S, x \text{ and } y \text{ are distinct variables})$.

[37] Moreover, some plausible-looking recombination principles would generate proper-class sized worlds from the claim that there are proper-class many possible individuals. One presupposes "trans-world identity": for any possible individuals (perhaps drawn from different possible worlds), there is a possible world containing all of those individuals. Another drops that presupposition: for any possible individuals (again perhaps from different worlds), there is a possible world containing distinct duplicates of those individuals.

supply as rigorous assumptions about the behavior of these abstract entities as do the axioms of set theory, but the problem is no less real for that.[38]

Moreover, anyone who believes in possible worlds and individuals as entities, even a Lewisian modal realist, faces the problem of reconstructing applications of possible worlds talk if there are too many possible worlds or individuals to form a set. If there is no upper bound on the size of possible worlds then there can be no set of all the possible individuals and no set of all the possible worlds (ersatz or genuine). Therefore, applications of worlds talk (e.g., defining semantic values) that require functions defined on the space of possible worlds and individuals are in trouble. Everyone faces the cardinality problem. Below I sketch a few potential lines of response; but the question is complex and calls for further study.[39]

A flat-footed response would be to simply deny that there can be arbitrarily many individuals, as Lewis did at one point.[40] Though theoretically simple, this response requires an unattractive restriction on what is possible. A theme of this paper has been that the theory of worlds should not put controversial constraints on what is possible.

Alternatively, one could try "technical tricks" of various sorts, in various combinations. A drastic move would be to invoke a non-standard logic or set theory on which Cantorian paradoxes with the universal set do not arise.[41] Less drastically, one might invoke class theory[42], and allow **W**, **D**, and **P**, in modal models to be proper classes. Since proper classes are usually not allowed to be members of further classes, some further dancing will be needed. For example, if **W**, **D**, and **P** are proper classes then they cannot be members of

[38] See Chihara, *op. cit.*, pp. 120–141.

[39] See chapter 6 of Nolan's *Topics in the Philosophy of Possible Worlds* for more work on this topic. Does the modal fictionalist face the cardinality problem? Answering this requires deciding whether the fiction must be logically coherent; if not then it could include the claim that there exists a set of all possible worlds and individuals, despite the resulting contradictions.

[40] *On the Plurality of Worlds*, pp. 102–104. For contrary arguments see Phillip Bricker, "Plenitude of Possible Structures", *Journal of Philosophy* 88 (1991): 607–619; and section IV of Daniel Nolan, "Recombination Unbound", *Philosophical Studies* 84 (1996): 239–262.

[41] See, for example, W. V. O. Quine, "New Foundations for Mathematical Logic", in his *From a Logical Point of View* (Cambridge, MA: Harvard University Press, 1953); and Greg Restall, "A Note on Naive Set Theory in LP", *Notre Dame Journal of Formal Logic* 33 (1992): 422–432.

[42] See Paul Bernays, "A System of Axiomatic Set Theory", in G. Muller, ed., *Sets and Classes* (Amsterdam: North-Holland Publishing Company, 1976); and K. Gödel, *The Consistency of the Axiom of Choice and of the Generalized Continuum-Hypothesis with the Axioms of Set Theory* (Princeton: Princeton University Press, 1940).

a modal structure ⟨**W**,**r**,**D**,**P**,**Q**,**I**⟩. One could give up on modal structures as entities, and cash out talk of modal models as *plural* talk of **W**, **r**, **D**, **P**, **Q** and **I**, plus an interpretation **F**. While this would not rescue pluriverse sentences, it would at least allow for a redefinition of "true$_p$ in all modal models" and thus would rescue the non-linguistic version of the theory. However, if **W**, **D**, and **P** are proper classes then the existence of **Q** and **I** is in jeopardy since they are functions defined on **W**, **D** and **P**. There is no problem if the domain of a world can never be proper-class-sized, since then there would be no need for the values of **Q** and **I** to be classes; but if in addition to arbitrarily large set-sized worlds we wish to admit (as we probably should) proper-class-sized worlds then **Q** and **I** cannot be functions construed as classes of pairs. We might now pull mereology from the bag of tricks. **Q** originally was a function that assigns to any w∈**W** a pair ⟨$\mathbf{D}_w$,$\mathbf{P}_w$⟩ of subsets of **D** and **P**. Suppose we accept a mereology of classes on which the following is coherent to require of realistic models: the members of **W**, the subclasses of **D** and the subclasses of **P** never overlap mereologically. We could then define **Q** as a class, each member of which is the mereological sum of a member of **W**, a subclass of **D**, and a subclass of **P**. For a given w∈**W**, $\mathbf{D}_w$ could then be recovered — it would be the subclass of **D** such that it and w are both part of X, for some X∈**Q**; similarly for $\mathbf{P}_w$. The case of **I** is more complicated. Whether the entire pluriverse theory can be reconstructed along these lines is an open question.

Yet another response leaves the definition of a modal model intact, but alters the truth conditions for statements about possibilia. Admitting that no modal model is realistic, one might characterize a transfinite hierarchy of "near-realistic" models of increasingly large size and claim that a sentence about worlds is true iff at some point in this hierarchy the sentence is true and remains true from that point onward. However, this translation procedure would misrepresent the modal facts: the sentence 'there is an upper limit to the infinite number of individuals that exist in any one possible world' would turn out true, despite the fact that it is possible that there exist arbitrarily many individuals.

4.2 "We're not talking about *that*!"

A less technical objection attacks my claim that talk about possibilities concerns infinitary sentences and modal models that no one until now has bothered to characterize. This objection is not particular to the pluriverse view. One might similarly object to the claim that, all along, we have been talking about linguistic ersatz worlds, complex set-theoretic constructions out of bits of language; or spatiotemporally isolated concrete worlds; or maximal consistent states of affairs; or a fiction about possible worlds; or any other proposed conception of possibilia.

The answer to all these worries is the same. Various modal concepts in the neighborhood of possible worlds form an interrelated cluster: possible worlds, possible individuals, possible states of affairs, necessity, possibility, and so on. These modal concepts are concepts of properties and relations that play certain roles *vis a vis* each other and *vis a vis* other notions. But this exhausts the nature of the modal concepts; they do not (much) constrain the intrinsic properties of candidate possible worlds and individuals. This is a sort of structuralism, if you like. Modal concepts lay down a structural requirement: our talk of possible worlds and the rest is about any structure that is suited to play the relevant role. I say that pluriverse sentences and realistic models are best suited to play this role, and that is all it takes for talk of worlds to be talk about them.[43]

[43] I favor an analogous response to Kripke's Humphrey objection to counterpart theory. See *Naming and Necessity* (Cambridge, MA: Harvard University Press, 1972), p. 45, and my *Four-Dimensionalism*, pp. 194–196. For other responses to Kripke see Lewis, *On the Plurality of Worlds*, p. 196, and Hazen, "Counterpart-Theoretic Semantics for Modal Logic", *Journal of Philosophy* 76 (1979): 319–338.

5. The Pluriverse view compared with modal fictionalism[44]

My pluriverse view, particularly in its linguistic form, is similar in some ways to Gideon Rosen's modal fictionalism.[45] There are, however, important differences between the views that favor the ersatz pluriverse view.

According to Rosen's fictionalism, talk about possible worlds is like ordinary talk about fictional characters. Just as "there was a detective named 'Holmes"' is elliptical for "according to the Conan-Doyle stories, there was a detective named 'Holmes"', "There is a possible world in which donkeys talk" is elliptical for "According to PW, the fiction of possible worlds, there is a possible world in which donkeys talk". PW consists of seven postulates, which informally capture Lewis's theory of worlds, together with "an *encyclopaedia*: a list of the non-modal truths about the intrinsic character of this universe."[46] The notion of truth according to, or in, a fiction is an undefined primitive of the theory. Rosen goes on to propose an analysis of modality: P is necessary iff, according to PW, P's worlds translation P* is true in all worlds; P is possible iff, according to PW, P* is true in some world.

Both fictionalism and the pluriverse theory reinterpret assertions about

[44] Two other views, to which the pluriverse view bears some similarity, are worth mentioning briefly. According to Tony Roy ("In Defense of Linguistic Ersatzism", *Philosophical Studies* 80 (1995): 217–242), worlds in which non-actual individuals swap qualitative roles may be distinguished by representing those individuals with "arbitrary names", for example 'Batman' and 'Robin' in 'Batman plays role R_1 and Robin plays role R_2' and 'Robin plays role R_1 and Batman plays role R_2'. But either i) these names mean nothing (since they refer to nothing) and so the sentences represent nothing, or ii) they are implicitly existentially quantified variables (as is suggested by Roy's remarks on p. 226). In case ii), either iia) the variables are bound to quantifiers at the beginning of the sentences, in which case the sentences mean the same thing, namely $\exists x \exists y(x$ plays role R_1 and y plays role $R_2)$ and the worlds have not been distinguished, or iib) the variables are bound to quantifiers outside the two sentences, in which case the view is starting to look like the pluriverse view. The other view is Fine's reduction of possible worlds talk, from his postscript to Kit Fine and Arthur Prior, *Worlds, Times and Selves*. (London: Duckworth, 1977); see also his "First-Order Modal Theories, Part II — Propositions", *Studia Logica* 39 (1980): 159–202; "First-Order Modal Theories, Part I — Sets", *Noûs* 15 (1981): 117–206; and "First-Order Modal Theories, Part III — Facts", *Synthese* 53 (1982): 43–122. Roughly, Fine interprets an existential sentence about possibilia like $\ulcorner$There is some possible object that ϕs$\urcorner$ as meaning $\ulcorner$possibly, there is an object that ϕs$\urcorner$, where the existential quantifier in the latter sentence is an actualist quantifier. Fine's view is attractive, but handles quantification over sets of possibilia (see *Worlds, Times and Selves*, pp. 145–148) and sentences attributing cross-world relations less smoothly than the pluriverse view.

[45] "Modal Fictionalism", *Mind* 99 (1990): 327–354.

[46] *ibid*, p. 335.

worlds as assertions about what is true according to a *single* entity describing the entire pluriverse. Rosen could therefore offer something like the pluriverse theory's solution to the problem of descriptive power. This, it seems to me, is fictionalism's most distinctive benefit.[47]

Rosen's fiction differs superficially from pluriverse sentences by being "generational": it describes the pluriverse by i) describing *one* world (in the encyclopaedia), and ii) supplying a principle that generates new worlds from old. The principle of generation is one of the seven postulates, specifically, the principle of recombination:[48]

> (6e) The totality of universes is closed under a principle of recombination. Roughly: for any collection of objects from any number of universes, there is a single universe containing any number of duplicates of each, provided there is a spacetime large enough to hold them.

A more important difference concerns the content of the fiction, not just its form. The pluriverse theory entails$_p$ everything about the pluriverse stateable from our vantage point in the actual world, whereas Rosen's fiction entails much less. Instead, Rosen relies on the fact that much more is generally true in a fiction than what is explicitly stated.[49] Though never explicitly mentioned, it is presumably true in the Sherlock Holmes fiction that Holmes has ten toes.

My reason for preferring the pluriverse view to Rosen's fictionalism is dissatisfaction with the notion of truth in fiction. If Rosen's fictionalism is to be materially adequate, that notion will need to be very different from the ordinary notion of truth in fiction, and hence will be an obscure, unexplained primitive.

Rosen needs truth-in-fiction because his fiction does not describe the pluriverse exhaustively. But it would be bogus to take this strategy to an extreme. Imagine an extremely thin fiction, consisting of a single sentence "there are other possible worlds", and a fictionalist who claimed to reduce worlds-talk to

[47] Other than the solution to the problem of descriptive power, fictionalism has no real advantage over linguistic ersatzism. Each appeals to truth according to linguistic entities. The linguistic ersatzer needs an infinitary Lagadonian worldmaking language, but Rosen does as well, to construct his encyclopaedia. Ersatzism requires modality, whereas Rosen requires his notion of truth in fiction. But below I argue that the notion of truth in fiction must be at least as powerful as modality if fictionalism is to be materially adequate.

[48] *op. cit.*, p. 333.

[49] *ibid*, p. 347.

truth in this thin fiction. The view seems materially inadequate, for the fiction says too little; it seems not, for example, to imply the existence of a world containing blue swans. The thin fictionalist might reply that it is *implicitly* true in the fiction that there is a world with blue swans. But then the notion of truth in fiction becomes unexplained and obscure, for it is not true in any *ordinary* sense that statements about blue swans are true in the thin fiction. Even if we grant the fictionalist *some* notion of implicit truth in fiction, we should not grant him just any such notion.

Though more extreme, my thin fictionalist is a bit like Rosen himself. Rosen claims that:[50]

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

But the truth of (8f) is far from clear. Rosen's idea, presumably, is that (8f) is true in virtue of the encyclopaedia and the principle of recombination (principle (6e) above). The encyclopaedia will indeed report the existence of swans and blue things in the actual world. But the principle of recombination will not then *logically* imply the existence of a universe with blue swans. There is no way to patch together duplicates of swans and blue things (blue skies, blue shirts, etc.) to arrive at a blue swan. The only hope would be to patch together duplicates of very small actual things — cells, molecules, or maybe even atoms — to form what would be a blue swan. Rosen's fiction PW thus logically entails[51] a certain sentence, BS, about microscopic objects, but it does not logically entail that there is a blue swan.

Consider the conditional statement "If BS is true, then there exists a blue swan". Call such conditionals "microreduction laws"; they express necessary truths[52] connecting the macro and micro realms. The problem under consideration would be solved if microreduction laws are true in PW. But since neither Rosen nor anyone else knows exactly what the microreduction laws are, they cannot be built into the fiction explicitly.[53] So Rosen must either admit that it is *not* true in PW that there is a universe with blue swans, in which case modal

[50] *ibid.*, pp. 335–336.

[51] Even this is not quite true, for various reasons. One is that (6e) is an incomplete statement of the principle of recombination, for it does not say that the duplicates may be patched together in arbitrary spatial arrangements.

[52] On some views BS should be augmented with relational facts or facts about the laws of nature.

[53] Compare Lewis, *On the Plurality of Worlds*, pp. 155–156.

fictionalism is materially inadequate, or he must say that microreduction laws are implicitly true in PW.

Other necessary truths will need to be implicitly true in PW as well. PW says nothing about the multitude of conceptual, mathematical, logical, and philosophical truths that are usually taken to be necessary. In certain cases where PW is silent (for example over whether something could be both positively and negatively charged) Rosen argues with some plausibility that there is no modal fact of the matter[54], but it would be unacceptable to claim this for *all* of these necessary propositions about which PW is silent. The pluriverse theory has no analogous problem: as noted in section III. E., necessary truths automatically turn out true in every world; this is of course because the modal notion of necessity was assumed in constructing realistic models.

PW is silent, then, about many necessary truths. Therefore, either Rosen's theory is materially inadequate, or, if he insists that these truths *do* hold in all the worlds of the fiction, then 'truth in fiction' has moved far from any ordinary notion of truth in fiction. Given the usual understanding of truth in fiction, surely the Goldbach conjecture (if it is indeed true) is not true in the Holmes stories; but Rosen's fiction PW says no more about mathematics than does the Holmes stories. Perhaps *some* elementary necessary truths could be said to be true in all the worlds of PW in virtue of being common knowledge to the writers and readers of typical fictions. We say, for example, that Sherlock Holmes has a liver, not because this is explicitly mentioned in the Conan Doyle stories but because it is common knowledge that detectives have livers. But many necessary truths are not common knowledge.

In evaluating this objection it is important to distinguish the following two claims:

> It is true in the fiction PW that every necessary truth is true in all worlds

> For every necessary truth, N, it is true in the fiction PW that N is true in all worlds

The first may well be unobjectionable, but it is the second that is required by Rosen's theory, and it is the second that is implausible, if the operative notion of truth in fiction is the ordinary one.

[54] *op. cit.*, section 7.

This difficulty for fictionalism emerges in another way, in strengthened form. The encyclopaedia and principle of recombination generate many worlds, but Rosen rightly refrains from claiming that all worlds may be thus generated, because of the possibility of fundamental properties that do not actually exist. Indeed, the fiction contains (in postulate (6g)[55]) a claim to the effect that there are worlds beyond those that can be generated in this way. But this means that Rosen's fiction contains no analog of COMPLETENESS (section III. E.), which says that the worlds a pluriverse sentence mentions are *all* of the worlds that exist. Thus, PW does not logically entail any (non-trivial) claims of the form "there is no world of type T" or "every world is of type T". The point holds for more complex assertions with universal quantifiers over possibilia, for example: "every possible object of type T_1 bears relation R to some possible property of type T_2" or "every function defined on possible individuals of type T_3 is also of type T_4". These claims must be accounted for solely by the primitive of truth in fiction, since the fiction is silent on these matters. This goes far beyond the previous point that all necessary truths must be true in every world in the fiction. As mentioned above, there are truths about the pluriverse that are not stateable in the language of necessity and possibility; few of these with universal quantifiers will be logically entailed by PW.

One final problem for fictionalism arises from the fact, just mentioned, that Rosen builds the possibility of properties beyond those that exist in the actual world into his fiction PW, via postulate (6g). Thus it follows from Rosen's analysis of modal concepts that the actual world does not contain all possible properties. This fact about the actual world is likely true, but it should not be built into an analysis of modality, since there may exist possible worlds with all possible properties, and it seems an open epistemic possibility that ours is such a world. Nor should Rosen build into PW the claim that *no* non-actual properties are possible, for then the fiction would be overly bold in the opposite direction: we would have a conceptual guarantee that this *is* the "richest" possible world. Human fictionalists do not know the modal facts, and so do not know what to build into the fiction.

The pluriverse view has no corresponding problem, since pluriverse sentences are constructed as a function of the modal facts. If 'Necessarily, every property actually exists' is an L-true sentence of the modal language then every pluriverse sentence will entail$_p$ that @ is the richest possible world; if not, not. Rosen cannot analogously construct PW conditional on the modal facts unless

[55] *ibid.*, p. 333.

he gives up on the reductive analysis of modality.

Some fictionalists might be happy to give up on reductive analysis, and utilize modal notions in constructing the fiction.[56] If implicit truth-in-fiction were still needed then fictionalism would require more commitments in ideology than the pluriverse view, which requires only modality. If, on the other hand, the fiction could then be spelled out in perfect detail, fictionalism would become the pluriverse view. Unless the difference between the two views collapses, then, there are reasons to prefer the ersatz pluriverse theory over fictionalism.

Theodore Sider
Syracuse University

[56] Rosen admits that in some sense 'according to' is a modal primitive (*ibid.*, section 8); the question is whether to admit possibility and necessity in addition.