The Nature of Epistemic Space

David J. Chalmers

1 Ways things might be[1]

There are many ways things might be, for all I know. For all I know, it might be that there is life on Jupiter, and it might be that there is not. It might be that Australia will win the next Ashes series, and it might be that they will not. It might be that my great-grandfather was my great-grandmother's second cousin, and it might be that he was not. It might be that brass is a compound, and it might be that it is not.

There are even more ways things might be, for all I know with certainty. It might be that there are three chairs in this room, and it might be that there are not. It might be that water is H_2O, and it might be that it is not. It might be that my father was born in Egypt, and it might be that he was not. It might be that I have a body, and it might be that I do not.

We normally say that it is *epistemically possible* for a subject that *p*, when it might be that *p* for all the subject knows. So it is epistemically possible for me that there is life on Jupiter, or that brass is a compound. One can define various different standards of epistemic possibility, corresponding to various different standards for knowledge. For example, one might say that it is *epistemically possible in the Cartesian sense* (for a subject) that *p* when it might be that *p*, for all a subject knows with certainty. So in the Cartesian sense, it is epistemically possible for me that water is not H_2O, and it is epistemically possible for me that I do not have a body.

A natural way to think about epistemic possibility is as follows. When it is epistemically possible (for a subject) that *p*, there is an epistemically possible *scenario* (for that subject) in

[1]This paper is forthcoming in *Epistemic Modality*, edited by Andy Egan and Brian Weatherson (Oxford University Press, 2010). Thanks to the many audiences to which I have presented versions of this paper: the 2001 Central APA, the 2003 Modality workshop at St. Andrews, the 2003 Northwest Philosophy conference, the 2004 Aspects of Knowing conference at UNSW, the 2006 Epistemic Modality conference at ANU, and talks at the Institut Jean Nicod, Leeds, Melbourne, Nebraska, Oxford, Sheffield, Texas, UCLA, and Vermont. Thanks also to Jens Christian Bjerring, Berit Brogaard, Kit Fine, Gregg Rosenberg, Wolfgang Schwarz, Brian Weatherson, and Bruno Whittle for comments on earlier versions of this paper.

which *p*. A scenario is a maximally specific way things might be: a sort of epistemically possible world, in a loose and intuitive sense. On this picture, corresponding to the epistemic possibility that Australia will win the next Ashes series are various epistemically possible scenarios in which they win in all sorts of different ways. And corresponding to the Cartesian epistemic possibility that I have no body are various scenarios in which I am disembodied, each epistemically possible by the Cartesian standard: e.g. scenarios in which I am a brain in a vat, or in which I am a disembodied Cartesian mind.

To fill out this picture, we might imagine that there is an overarching space of scenarios. These scenarios constitute *epistemic space*. If a subject did not know anything, all scenarios would be epistemically possible for the subject. When a subject knows something, some scenarios are excluded. Every piece of substantive knowledge corresponds to a division in epistemic space: some scenarios are excluded out as epistemically impossible for the subject, while others are left open. More specifically, it is natural to hold that for a given *p*, there may be scenarios in which *p* is the case, and scenarios in which *p* is not the case. Then when a subject knows that *p*, scenarios in which *p* is not the case are excluded, while others are left open. The scenarios that are epistemically possible for a subject are those that are not excluded by any knowledge of the subject.

One can naturally suppose that the space of scenarios is equally divided by *belief*, and perhaps that the division by belief underlies the division by knowledge. Every substantive belief, whether or not it qualifies as knowledge, corresponds to a division in the space of scenarios. When a subject believes that *p*, we might say that some scenarios (in particular, scenarios in which $\neg p$) are ruled out as *doxastically impossible*, while others are left open. A scenario is doxastically possible for a subject if and only if it is not doxastically ruled out by any of the subject's beliefs. When a belief qualifies as knowledge, the scenarios ruled out as doxastically impossible are also ruled out as epistemically impossible.

A picture of this sort is often present in philosophical discussions of knowledge and belief. Within epistemology, it is common to think of knowledge in terms of the "elimination of possibilities", with some sort of underlying space of possibilities presumed. In discussions of skepticism, for example, the fact that certain skeptical scenarios are not eliminated is used as evidence that certain knowledge claims are not true. In epistemic logic and the theory of belief revision, it is common to model epistemic possibility using epistemic relations to an underlying space of possible worlds. The same goes for the theory of subjective probability: a subject's credences are usually taken to be distributed over a space of epistemically possible worlds.

It is surprisingly difficult, however, to make the intuitive picture precise. What sort of possibilities we are dealing with here? In particular, what is a scenario? And what is the relationship between scenarios and items of knowledge and belief?

It is natural to think of scenarios as possible worlds, and to think of a scenario in which *p* as a world in which *p*. But it is immediately clear that this will not work, at least on the most common contemporary understanding of possible worlds. There are subjects for whom it is epistemically possible that Hesperus is not Phosphorus; but on the usual understanding, there is no possible world in which Hesperus is not Phosphorus. It is epistemically possible for me that my great-grandparents were cousins and it is epistemically possible that they were not; but on the usual understanding, my great-grandparents are cousins either in all worlds in which they exist or in none. In the Cartesian sense, it is epistemically possible for me that water is not H_2O, but on the usual understanding (assuming that water really is H_2O), there are no possible worlds in which water is not H_2O. So if we are to maintain that it is epistemically possible that *p* iff there is an epistemically possible scenario in which *p*, we cannot identify a scenario in which *p* with a possible world in which *p*, at least on the usual understanding.

Some might react to this by denying the intuitions about what is epistemically possible (e.g. holding that it is never epistemically possible that Hesperus is not Phosphorus), and some might react by denying the coherence of the picture connecting epistemic possibility to epistemically possible scenarios. Both reactions would be premature: the first loses touch with the phenomenon we are trying to analyze, and the second assumes that possible worlds as currently understood are the only available tool.

Instead, we should try to understand epistemic possibility on its own terms. We are not dealing here with counterfactual space: the space of ways things might have been. Here, we are dealing with epistemic space: the space of ways things might be. This epistemic space calls for its own epistemic tools of analysis. Where the analysis of counterfactual space invokes possible worlds as maximally specific ways things might have been, the analysis of epistemic space should invoke scenarios as maximally specific ways things might be. The two notions are quite distinct, although they have a deep underlying relationship.

In this paper, I will try to make sense of epistemic space. I will explore different ways of making sense of scenarios, and of their relationship to thought and language. I will discuss some issues that arise, and I will outline some applications to the analysis of the content of thought and the meaning of language.

2 Principles of Epistemic Space

On the picture suggested above, we might say that the notion of *strict epistemic possibility*—ways things might be, for all we know—is undergirded by a notion of *deep epistemic possibility*— ways things might be, prior to what anyone knows. Unlike strict epistemic possibility, deep epistemic possibility does not depend on a particular state of knowledge, and is not obviously relative to a subject. Whereas it is strictly epistemically possible (for a subject) that p when there is some epistemically possible scenario (for that subject) in which p, it is deeply epistemically possible that p when there is some deeply epistemically possible scenario in which p. Since all scenarios are deeply epistemically possible on this picture, we can put this more simply: it is deeply epistemically possible that p when there is some scenario in which p.

The notion of deep epistemic possibility can be understood in different ways for different purposes. One might adopt a conception on which every proposition is deeply epistemically possible. One might also adopt a conception on which every proposition that is not logically contradictory is deeply epistemically possible, or one which every proposition that is not ruled out a priori is deeply epistemically possible. In this paper, I will mainly work with the latter understanding, which I flesh out in the next section. But first I will lay out some background structure that is largely independent of the notion of deep epistemic possibility that we adopt.

What are the objects of epistemic possibility? So far I have spoken as if they are propositions. I think that this view is ultimately correct, but the contested nature of propositions raises difficulties. For reasons given above, we want to hold that it can be epistemically possible for a subject that Hesperus is not Phosphorus, even if it is not epistemically possible that Hesperus is not Hesperus. On the popular Russellian conception of propositions, however, the proposition that Hesperus is not Hesperus is identical to the proposition that Hesperus is not Phosphorus. If so, and if propositions are the objects of epistemic possibility, it will be hard to obtain the result above. We could simply assume a Fregean view of propositions according to which these propositions are distinct, but the viability of such a view is contested. Furthermore, one aim of the present treatment is to use epistemic space to help make sense of a Fregean conception of propositions. If so, one cannot simply presuppose such a conception.

The alternative is to adopt an approach on which the entities that are assessed for epistemic possibility are linguistic items, such as sentences (or utterances), or mental items, such as thoughts (or beliefs or items of knowledge). Here I will work with sentences, but the framework is naturally extendible to thoughts. The sentences in question are restricted to assertive sentences, at least initially. To accommodate contextual variability in the use of sentences between subjects and occasions, these sentences should be individuated as sentence tokens (or sentences in contexts) rather than sentence types. I will take it that every assertive sentence token expresses a thought, and that every thought is expressed by a possible sentence token, so there is a natural correspondence between sentence tokens and thoughts. Below, reference to sentences without further explanation should be taken as invoking reference to sentence tokens.

The key to the picture is a relation of *verification* between scenarios and sentences. Any sentence divides the space of scenarios into those scenarios that *verify* the sentence and those that *falsify* the sentence. More formally, we can say that there is a relation *ver* between scenarios and sentences, such that $ver(w, s)$ can take on the same range of truth-values (e.g. true, false, and indeterminate) that sentences can take on. We can say that w verifies s when $ver(w, s)$ is true, and that w falsifies s when $ver(w, s)$ is false. When w verifies s, we can say that s is true at w. When w falsifies s, we can say that s is false at w. In some cases, it may be that s is indeterminate at w, or perhaps that s has some other truth-value at w.

There is also a relation of *actualization* between scenarios and sentence tokens. This is needed to capture the idea that for any utterance, one scenario is singled out as the scenario of utterance. A scenario w is actualized at sentence token s when w is the scenario of utterance for s. This scenario corresponds intuitively to the way things really are (relative to the subject) when the expression is uttered.

As discussed above, we also have a basic notion of *deep epistemic possibility* (here abbreviated as simply "epistemic possibility") that applies to sentence tokens.

Scenarios, sentences, verification, actualization, and epistemic possibility should obey at least the following principles. In these principles, quantification over sentences is to be understood as quantification over possible (not just actual) assertive sentence tokens.

> *Plenitude*: For all sentences s, s is epistemically possible iff there exists a scenario w such that w verifies s.
>
> *Actualization*: For all sentences s and scenarios w, if w is actualized at s, then the truth-value of s is $ver(w, s)$.
>
> *Compositionality*: When a complex sentence s is composed from simpler sentences s_i and truth-functional connectives, $ver(w, s)$ is determined by $ver(w, s_i)$ in the corresponding truth-functional way. For example, $ver(w, \neg s) = \neg ver(w, s)$, and $ver(w, s\&t) = ver(w, s)\&ver(w, t)$.

Plenitude and Actualization are basic principles of the framework that it would be hard to deny. There might be versions of the framework that deny Compositionality (perhaps to model non-ideal reasoners who accept s and t while denying $s\&t$, for example), but I will largely presuppose this principle in what follows.

The following three principles also have some attractions, although they are not obviously compulsory. For the purposes of the second and third principles below, let us say that scenarios w_1 and w_2 are *equivalent* iff for all possible sentences s, $ver(w_1, s) = ver(w_2, s)$.

> *Uniqueness*: For any sentence s, if scenarios w_1 and w_2 are actualized at s, then $w_1 = w_2$.
>
> *Parsimony*: If scenarios w_1 and w_2 are equivalent, then $w_1 = w_2$.
>
> *Specification*: For every scenario w, there is some sentence d such that w verifies d and such that if any scenario w' verifies d, then w' is equivalent to w.

Uniqueness says that there is a unique scenario of utterance for every sentence. This makes sense given the picture of scenarios outlined earlier, although one can also imagine a very fine-grained conception of epistemic space on which more than one scenario could be actualized simultaneously. Parsimony says, in effect, that there are no more scenarios than there need to be to differentiate their application to possible sentences. Specification says, in effect, that for each scenario there exists a *specification* that singles it out up to equivalence (and which singles it out uniquely, if Parsimony is true). Specification will require infinite sentences, raising issues that I discuss later in the paper. The framework will still deliver acceptable results if some or all of these principles are false, but these principles make it better behaved in certain respects.

Given this framework, we can say that the *epistemic intension* of a sentence s is a function from scenarios to truth-values mapping a scenario w to $ver(w, s)$. The epistemic intension of a sentence corresponds to the way it divides epistemic space.

3 Epistemic necessity and apriority

Before proceeding, we need to say more about what deep epistemic possibility involves. There are various different ways that this notion can be understood, and these understandings may each be useful for different purposes.

On a maximally liberal conception, any sentence at all is deeply epistemically possible. One can motivate this by observing that there are subjects that do not know anything at all, and for such subjects any sentence is strictly epistemically possible. If strict epistemic possibility entails deep epistemic possibility, then every sentence is deeply epistemic possible.

On this picture, there will be scenarios verifying arbitrary sentences, including all sorts of logical contradictions. This leads naturally to a picture on which the space of scenarios is something akin to the power set of the set of sentences: to any set of possible sentences, there corresponds a scenario, and vice versa. This picture might be useful for some purposes—say, for making sense of the epistemic states of extremely non-ideal thinkers. However, for many purposes it is useful to have a space of scenarios with a less trivial structure. For example, the maximally liberal picture will be of little use when it comes to analyzing meaning and content. On this picture, the epistemic intension of any sentence *s* will be true in precisely those scenarios corresponding to sets of sentences that include *s*. It follows that nontrivial relations among the meanings of sentences will never be reflected in their epistemic intensions.

A more useful notion of deep epistemic possibility will involve some imposition of a rational idealization, for example to rule out scenarios on which logical contradictions are true. We might say that the corresponding notion of deep epistemic necessity should capture some sort of rational *must*: a statement is deeply epistemically necessary when in some sense, it rationally must be true. Such a notion can be understood in various ways, but for our purposes there is an obvious candidate.

We can say that *s* is deeply epistemically necessary when *s* is *a priori*: that is, when *s* expresses actual or potential a priori knowledge. More precisely, *s* is a priori when it expresses a thought that can be justified independently of experience, yielding a priori knowledge. I have discussed this conception of apriority at length elsewhere (Chalmers 2004), but I will recap the essentials here.

On this picture, a thought is a sort of token mental state: in particular, a thought is an occurrent propositional attitude with a mind-to-world direction of fit. So occurrent beliefs are thoughts, as are mere entertainings. Like beliefs, thoughts are assessible for truth. Thoughts can come to be *accepted*, yielding beliefs, and thoughts can come to be *justified*, often yielding knowledge. We can then say that a thought is a priori when it can be justified independently of experience, yielding a priori knowledge.

The relation of expression is such that every assertive utterance expresses a thought.[2] Typical assertive utterances express occurrent beliefs, and even when they do not, perhaps because the utterance is insincere or speculative, they plausibly express thoughts whose associated credence falls short of what is required for belief. The expression relation should be understood as one that preserves truth-value and truth-conditions: it is guaranteed that when an utterance expresses a thought, the utterance is true if and only if the thought is true.

The expression relation allows us to move back and forth between thoughts and sentence tokens. We can predicate apriority, deep epistemic necessity, and deep epistemic possibility of both thoughts and sentence tokens in the obvious way. For example, a sentence *s* is deeply epistemically possible when the thought that *s* expresses cannot be ruled out a priori.[3]

This idealized notion of apriority abstracts away from contingent cognitive limitations. If there is any possible mental life that starts from a thought and leads to an a priori justified acceptance of that thought, the thought is a priori.[4] So if a hypothesis can be known to be false only by a great amount of a priori reasoning, it is nevertheless deeply epistemically impossible. For example, 'There are integers $a, b, c, n > 2$ such that $a^n + b^n = c^n$' is deeply epistemically impossible. As a result, this idealization is best suited for modeling the knowledge and belief of idealized reasoners that may be empirically ignorant, but that can engage in arbitrary a priori reasoning.

There are other, less idealized ways to understand deep epistemic necessity. It may well be that there is a spectrum of notions ranging from this highly idealized notion to the maximally liberal notion discussed earlier. At the end of this article, I will discuss notions that relax the idealization. For present purposes, however, the idealized notion is the best-behaved and the easiest to work with.

---

[2]What of apparently assertive utterances that do not express thoughts, such as some utterances by actors, sleeptalkers, distracted individuals, and so on? One might reasonably say that these utterances are not assertions at all. In any case, the current framework does not directly apply to them, as one cannot use their association with thoughts to assess apriority and define epistemic intensions. However, if one develops the framework for paradigmatic assertive utterances, one might be able to apply it to these atypical utterances indirectly, perhaps in virtue of relations that hold between these utterances and (actual or possible) paradigmatic assertions.

[3]A small complication is required to handled cases of indeterminacy. If it is a priori that *s* is indeterminate, then the negation of *s* will also be indeterminate. If we assume that indeterminate sentences cannot be known a priori, it follows that this negation is not epistemically necessary, and that *s* will count as deeply epistemically possible by the definition above. But this seems the wrong result. To handle this we can say that *s* is deeply epistemically possible when $\neg det(s)$ is not epistemically necessary: that is, when the thought that *s* expresses cannot be ruled false or indeterminate a priori. (Or in the framework below, *s* is deeply epistemically possible when a negation of a determination of a thought that *s* expresses is not epistemically necessary, where determination is a mental analog of the determinacy operator.)

[4]An issue arises if one thinks there may be a posteriori necessities limiting the space of possible mental lives. For example, if it is necessary that no mental life can involve more than 10^{100} steps, then a mathematical statement whose proof requires this many steps will not be deemed a priori by the current definition. For reasons discussed in the next section, I do not think that there are a posteriori necessities of this sort. If one holds that there are such necessities, it is probably best not to define apriority in modal terms.

When apriority is understood as above, it is clear that typical tokens of sentences such as 'Hesperus is Phosphorus' are not a priori. The thoughts expressed by these tokens are such that there is no possible mental life that starts from that thought and leads to an a priori justified acceptance of that thought. It may be that there is some *other* a priori justifiable thought (say, one expressed by saying 'Hesperus is Hesperus') that involves a relation to the same Russellian proposition as the first thought, but because these two thoughts are not themselves connectable a priori, the apriority of one does not entail the apriority of the other.

Apriority here is associated with sentence tokens rather than sentence types, to accommodate possible differences in use among fully competent speakers on different occasions. For example, in one context 'If someone is bald, they have no hair' may express a priori knowledge, while in another context it may not. Likewise, one speaker might use the names 'Bill' and 'William' of a particular individual interchangeably, so that 'Bill is William' expresses a priori knowledge, while another speaker who has acquired the two names through different routes might not. (For more on these cases, see Chalmers 2002a).

When an expression (e.g., 'bald', 'Bill') supports potential differences in apriority among fully competent users in this way, I will say that it is epistemically variant (or just variant); if not, it is epistemically invariant (or just invariant). For the special case of a sentence composed of invariant expressions, we can associate apriority with a sentence type, not just with sentence tokens: such a sentence type is a priori if some possible token of the type is a priori.

Some further structure will be useful for the constructions that follow. I also assume that the thoughts of a given thinker can stand in a relations of negation, conjunction, and disjunction to each other: so one thought can be formed by another by an operation of negation, or from another two thoughts by operations of conjunction or disjunction. We can then say that one thought *implies* another when a disjunction of the latter with a negation of the former is a priori. When s and t are epistemically invariant sentence types, we can say that s implies t when $\neg s \vee t$ is a priori. In addition, when s is an epistemically invariant sentence type and t is a thought, we can say that s implies t when some possible thought expressed by a token of s implies t. When s is an epistemically invariant sentence type and t is a sentence token, s implies t when s implies the thought expressed by t. As defined here, implication is a sort of epistemic necessitation, akin to a priori entailment.

4 Scenarios as centered worlds

The most natural way to think of scenarios, at least initially, is as possible worlds. In a way this is trivial—scenarios are defined as possible (in some sense) ways things might be (in some sense). But the notion of possibility invoked here differs from the notion of possibility that is usually associated with possible worlds: it is a sort of epistemic possibility, whereas possible worlds are usually understood to be associated with a sort of "metaphysical" possibility. Still, the question arises as to whether possible worlds understood in the latter sense might serve to help us model the space of scenarios, at least indirectly. That is: can we use the space of metaphysically possible worlds to construct a space of scenarios, and to make the case for a verification relation between scenarios (so understood) and thoughts?

I think we might. The intuitive idea is simple: to every possible world w, there corresponds a very specific (deep) epistemic possibility: the epistemic possibility that w is actual. So we might start by suggesting that scenarios *are* worlds. We could then say that a world w verifies a sentence token s when d implies s, where d is a canonical specification of w.[5] We could likewise say that w is actualized at s when w is the world in which s is uttered.

I will say more about canonical specifications shortly. For now I will note as above that they are best taken as sentence types, rather than tokens, in an epistemically invariant language. It is desirable that canonical specifications be *epistemically complete*, in that they leave no matters epistemically open. More precisely, we can say that d is epistemically complete iff for all sentences s, if d is epistemically compatible with s, then d implies s.

This is an attractive picture, but it runs into immediate problems. These problems lead to various clarifications to and modifications of the picture above. There are four main sources of problems: indexicality, rigidity, strong necessities, and parsimony.

(1) Indexicality

The first problem arises from the indexical phenomena involving expressions such as 'I', 'here', and 'now'. Let d be a full non-indexical specification of an "objective" world w. Let s be an indexical claim, such as 'I am a philosopher' or 'It is raining here now' or 'Today is Friday'. Then in each case, it may be that utterances of both $d\&s$ and $d\&\neg s$ are epistemically possible. So both

[5] If we have only two-truth values, we can say that w falsifies s when d does not imply s. If there are more than two truth-values, these can be handled in a manner parallel to that discussed under the epistemic construction in the next section.

$d\&s$ and $d\&\neg s$ are verified by a scenario. These scenarios must be distinct, as no scenario verifies both s and $\neg s$. But there will plausibly be only one world (objectively understood) in which d is the case. And even if there is more than one objectively indistinguishable d-world, it is hard to make out a distinction between those that verify 'I am a philosopher' and those that do not.[6] So it appears that if scenarios are construed as objective possible worlds, they will not satisfy Plenitude.

The natural solution is to identify scenarios with centered worlds: ordered sequences of worlds along with (optionally) individuals and times. The move to centered worlds requires that an "objective" specification d of a possible world w be supplemented by certain indexical claims that characterize the location of the center. This can be done as follows. Let us say that such a predicate ϕ *identifies* an individual x in w when ϕ is true of x in w and is true of no other entity in w. A canonical specification of a centered world w' will then take the form d & 'I am ϕ_1' & 'now is ϕ_2', where d is a canonical specification of the uncentered world w, ϕ_1 identifies the individual at the center, and ϕ_2 identifies the time at the center.

It is useful to stipulate that the marking of centered elements in a centered world is optional. This way, we can accommodate the (arguable but plausible) aposteriority of claims such as 'Thinkers exist' and 'The universe is temporal'. If we allow centered worlds without marked subjects or times, then there will be subjectless scenarios and timeless scenarios to falsify these claims. There can even be an empty scenario to verify 'Nothing exists', which is arguably a deep epistemic possibility. (Here I assume that 'I exist' is a posteriori, being justified by experience. If someone holds that 'I exist' is a priori, then they can require that centered worlds contain marked subjects.)

It may also be that we sometimes need additional optional marked information at the center of a world. This need arises in Austin's (1990) case of a demonstrative thought t to the effect that *that spot is red*, in a subject with a symmetrical visual field involving experience as of two red spots. Here, a full objective-plus-indexical specification of the world and of the subject's location within it (including a description of the spots, the subject's experience, and the connections between them) may not settle the truth-value of t. For example, the specification may tell the subject that a red spot is causing one spot-experience, and that a blue spot is causing another, without telling the subject which spot-experience is *this* spot-experience. To handle cases of this sort, one needs to allow one or more marked experiences at the center of a world: in effect, there will be a marked

[6]Just possibly, one could retain Plenitude for uncentered worlds by allowing that individuals and times have very few essential properties, and that there are distinct but qualitatively indistinguishable worlds where I coincide with arbitrary individuals, where now coincides with arbitrary times, and so on.

experience corresponding to each perceptual demonstrative involved in a thought. A canonical specification of the world will then involve certain phenomenal demonstratives ('This experience is ϕ', where ϕ identifies the relevant marked experience), where this phenomenal demonstrative is linked to the perceptual demonstrative in such a way that the canonical specification allows one to determine the truth-value of the thought.

In a few cases involving completely symmetrical worlds, there may be no identifying predicates available: that is, there may be no predicate (or at least no neutral predicate, in the sense discussed below) that is true of only the individual (or the time, or one of the experiences) at the center. In that case, one can invoke a maximally specific predicate instead: a predicate ϕ_1 such that for all (neutral) ϕ_2 true of the individual, ϕ entails 'everything that is ϕ_1 is ϕ_2'. Here, two centered worlds that differ only in symmetrical placement of the center will yield the same canonical specification. This phenomenon will be discussed more under the heading of parsimony, below.

(2) Rigidity

The second problem has also already been discussed. 'Hesperus is not Phosphorus' is epistemically possible for some subjects. But on the usual understanding of possible worlds (following Kripke 1980), 'Hesperus' and 'Phosphorus' are *rigid designators*, picking out the same object (Venus) in all possible worlds. If so, then there is no possible world *satisfying* 'Hesperus is not Phosphorus', where satisfaction is the standard sort of post-Kripkean evaluation of sentences in worlds. Something similar applies to 'water is not H_2O', 'my greatgrandparents were cousins', and so on. Adding centers to the possible worlds does not help with this. So if scenarios are centered worlds, and if verification is the same as satisfaction, then Plenitude is false.

To avoid this problem, we must deny that a world verifies a statement when it satisfies that statement. It may be that no centered world satisfies 'Hesperus is not Phosphorus', but some centered world may still verify 'Hesperus is not Phosphorus'. This conclusion is already forced on us by considering a claim such as 'I am a philosopher'. If *w* is a centered world in which David Chalmers is a mathematician and George Bush is a philosopher, centered on Bush, then this world satisfies 'I am not a philosopher'. But according to the definition above, this world will verify 'I am a philosopher'.

The key difference is that satisfaction is tied to metaphysical necessitation, where verification is tied to epistemic necessitation. To a first approximation, *w* satisfies *s* if a canonical specification of *w* metaphysically necessitates *s*, while *w* verifies *s* if a canonical specification of *w* epistemically necessitates *s*. The definition of verification above already appeals to epistemic necessitation, so the problem is automatically avoided.

An important residual issue, however, concerns the nature of the expressions used in a canonical specification of scenarios. To specify a scenario, we choose sentences that are true of it. But scenarios are centered worlds, should these be sentences that the world *verifies*, or sentences that the world *satisfies*? If we choose the first, there is a danger of circularity: verification of a sentence by a world will be defined in terms of canonical specifications, which will be defined in terms of verification. If we choose the second option, there is a danger of incoherence. The framework requires that not all centered worlds verify 'Hesperus is Phosphorus', even though all worlds satisfy 'Hesperus is Phosphorus'. But if a canonical specification can include any sentence that a world satisfies, including 'Hesperus is Phosphorus', then all worlds will verify 'Hesperus is Phosphorus', which is the wrong result.

The solution is to restrict canonical specifications to *neutral* expressions (plus indexicals to specify the location of the center). Intuitively, a neutral expression is one that behaves the same with respect to both verification and satisfaction. We cannot simply *define* a neutral expression in this way, for fear of circularity, but nevertheless we have a good grasp on the notion. For example, 'water' and 'Hesperus' are not neutral; but 'and', 'philosopher', 'friend', 'consciousness', and 'cause' plausibly are. To a first approximation, an expression is neutral if it is not "Twin-Earthable": that is, if one cannot devise a Twin-Earth case where a twin of a fully competent user of the expression uses their counterpart expression with a different meaning. There is more to say about the notion of neutrality than this (see Chalmers 2004), but this gloss will suffice for present purposes.

For this framework to yield fully adequate canonical specifications of worlds, it is required that there be epistemically complete specifications of arbitrary worlds involving only neutral terms and indexicals. If we assume that there is no problem with epistemically complete specifications that allow non-neutral terms, we can derive this claim from the thesis that every non-neutral sentence that is epistemically possible is implied by some epistemically possible sentence involving only neutral expressions and indexicals.

(3) Strong necessities

We have seen that the existence of a posteriori necessities such as 'water is H_2O' poses no deep problem for the picture of scenarios as centered worlds, as long as we distinguish verification from satisfaction. When n is a standard a posteriori necessity, it is plausible that although all worlds satisfy n, some centered world verifies $\neg n$. When this is the case, we can say that n is a *weak a posteriori necessity*. In these cases, we have a centered world verifying the relevant deep epistemic possibility, as Plenitude requires.

By contrast, a *strong a posteriori necessity* (or just a *strong necessity*) is an a posteriori necessity that is verified by all centered worlds. Strong necessities provide a more serious threat to Plenitude. Let us say that Metaphysical Plenitude is the thesis that for all sentences s, s is epistemically possible iff there exists a centered world that verifies s. If n is a strong necessity, then $\neg n$ is a counterexample to Metaphysical Plenitude. But on the assumption that scenarios are centered worlds, then Metaphysical Plenitude is equivalent to Plenitude. So if there are strong necessities, and if scenarios are centered worlds, then Plenitude is false.

For an example, consider a theist view on which 'An omniscient being exists' is necessary, but is not a priori. On such a view, this sentence (s) is plausibly a strong necessity. This follows from the claims that (i) s is a posteriori, (ii) every world satisfies s (as it is necessary) and (iii) a centered world verifies s iff the corresponding world satisfies the sentence (as there are no relevant two-dimensional phenomena here). On this view, although it is deeply epistemically possible that there are no omniscient beings, there are no centered worlds that correspond to this epistemic possibility. In effect, there are not enough centered worlds to go round.

Some other potential strong necessities are provided by the following philosophical views:

(i) A particularly strong "strong laws" view on which the fundamental laws and properties instantiated in our world are the fundamental laws and properties of every possible world. Let us say the view also holds (plausibly) that fundamental laws are a posteriori. On this view, a denial of the law of gravity (say) will be deeply epistemically possible, but there will be no possible world satisfying this denial, and there will also be no possible world verifying the denial.

(ii) A materialist view on which truths Q about consciousness are necessitated by the conjunction P of physical truths, but on which Q is not a priori derivable from P. Here, a psychophysical conditional $P\&\neg Q$ will be epistemically possible. It is not hard to show that if there is even a possible world verifying this conditional (as in the Kripkean cases), problems for materialism ensue. So some materialists deny that even a verifying world exists. If so, the conditional 'If P, then Q' is a strong necessity.

(iii) A view on which there are mathematical claims m—perhaps the Continuum Hypothesis?—that are true and are necessary, but are not knowable a priori by any possible being. On such a view, it seems that m will be a strong necessity: $\neg m$ will be epistemically possible, but verified by no possible world.

Other such views could be developed: e.g. one on which moral claims can be true and necessitated by natural truths, without being a priori derivable from natural truths; or a similar view about vague claims. In each of these cases, the distinction between verification and satisfaction does not seem to help. If the views in question are correct, there are simply not enough possible worlds to verify all epistemically possible thoughts and statements.

The simplest response to this problem, and the response that I think is correct, is to deny that there are any strong necessities. Each of the views listed above is at least controversial. In some cases, proponents claim support from the Kripkean cases, but these cases give no reason to believe in this much stronger phenomenon. In fact, one can argue in reverse: the fact that the link between epistemic possibility and verification by possible worlds is so strong elsewhere gives reason to believe that these claims are incorrect. One can also argue that there are deeper problems with these views. I have argued for these claims elsewhere (e.g. Chalmers 2002c), and will not repeat those arguments here.

It is at least clear that these views provide no *clear* reason to reject the model of scenarios as centered worlds, since in no case is the view in question clearly true. Still, the existence of these views entails that the claim that scenarios can be modeled by centered worlds will be at least as controversial as the denial of the views. And it would be desirable to give an account of scenarios that even holders of these views could accept. If so, that provides at least some reason to look at other models of scenarios.

(4) Parsimony

So far, we have examined reasons for thinking that there are *not enough* possible worlds to act as scenarios. But there are also reasons for thinking that there are *too many* possible worlds to act as scenarios. That is, while the problems above are mostly problems for Plenitude, one can also raise problems for Parsimony. In particular, it seems that there exist groups of centered worlds such that any possible sentence is equally verified or falsified by any world in the group. If so, it seems that each world in the group corresponds to the same scenario.

One way this can happen is with symmetrical worlds. Say that a world is mirror-symmetrical, and consider centered worlds w_1 and w_2 centered on corresponding subjects on each side, at the same time. Then as defined above, a canonical specifications of w_1 and w_2 will be exactly the same. Furthermore, this seems to mirror intuitions about the case. Intuitively, there is no sentence

s such that s is verified by w_1 but not by w_2. The main candidates for such a sentence are of the form 'I am ϕ', but centering works in such a way that both worlds will verify these claims equally.

The same goes for a world with a cyclic Nietzschean eternal recurrence of indistinguishable cycles, extending indefinitely into the past and the future. If we take a group of centered worlds w_i centered on corresponding subjects and times in different cycles, then it seems that for any t, if one world w_i verifies t, then all worlds w_i verify t. In these cases, it seems that the different centered worlds all correspond to the same epistemic possibility, violating Parsimony.

Parsimony might also be violated if possible worlds can contain inconceivable features. Say that there are two possible worlds w_1 and w_2 that are otherwise indistinguishable, except that at a certain point they contain different features ϕ_1 and ϕ_2. And say that ϕ_1 and ϕ_2 are inconceivable, in the sense that there is no possible neutral concept picking out ϕ_1 or ϕ_2. Then it may be that any neutral claim true of w_1 will also be true of w_2, so that canonical specifications of these worlds will be identical. If so, there is no sentence s that is verified by w_1 but not w_2.

Finally, suppose that (as some believe) there are qualitatively indistinguishable possible worlds. Take two identical twins Bill and Bob in the actual world. Some argue that there can be qualitatively indistinguishable worlds w_1 and w_2 such that only Bill exists in w_1 and only Bob exists in w_2. If so, it will plausibly still be the case that w_1 and w_2 verify all the same sentences.

The last two cases arise from possible ways in which the space of metaphysical possibilities may be more fine-grained than the space of epistemic possibilities. These two rest on controversial presuppositions that might be denied. But the first two, which arise from ways in which the space of *centered* metaphysical possibilities is more fine-grained than the space of epistemic possibilities, are relatively uncontroversial. So it seems that the space of centered worlds and the verification relation, as understood above, do not satisfy Parsimony.

One could respond in different ways. One might simply jettison Parsimony, holding that it is an inessential principle. Certainly, it seems less essential than Plenitude. One might also modify the picture slightly, by identifying scenarios with equivalence classes of centered worlds, where the worlds in groups such as the above will all fall into the same equivalence class. Either response will still allow a serviceable construction. Still, both responses suggest that there is at least a mild mismatch between scenarios and centered possible worlds.

What is the upshot of the four obstacles to identifying scenarios with possible worlds that we have discussed? The obstacles due to indexicality and rigidity can be overcome relatively easily, by invoking centered worlds and distinguishing verification from satisfaction. The obstacle due to strong necessities can be denied, and the obstacle due to parsimony can be dealt with as above.

Still, the last two obstacles suggest that while centered worlds may do a good job of modeling scenarios, the match is not perfect. The existence of philosophical views on which there are strong necessities suggests that even if these views are misguided, an analysis of scenarios as centered worlds will be at least mildly controversial. Because it makes a substantive (if plausible) claim about the relationship between possible worlds and epistemic possibility, this analysis goes beyond a surface analysis of epistemic possibility itself. The problems with parsimony also suggest a slight conceptual mismatch between the notions. So while centered worlds may provide a very useful way of thinking about scenarios, it is also useful to look at other ways.

5 The epistemic construction of scenarios

The obstacles in the previous section all have a common source. They arise because we are taking a class of entities—the possible worlds – developed in the service of a *different* notion of possibility (metaphysical possibility, or what might have been the case), and adapting it to help analyze the notion of epistemic possibility (what might be the case). It is inevitable that this adaptation will lead to certain complications. An alternative strategy suggests itself. Instead of adapting a different modal space, we might construct the space of scenarios directly, by a construction grounded in epistemic notions. In particular, we might take (deep) epistemic possibility as basic, and proceed from there. In this way, we can give an account of epistemic space in its own right.

A further motivation for this sort of construction is that it might generalize to the case of non-idealized epistemic possibilities. There is little hope that a construction in terms of centered worlds will generalize in this way. For example, a complex mathematical truth M is true in all centered worlds, so centered worlds cannot model the (non-ideal) epistemic possibilities in which the sentence in question is false. But if we adopt a non-idealized notion of deep epistemic possibility as primitive, then it is at least reasonable to hope that a version of the construction below might model the non-ideal epistemic possibilities in question. I will proceed by assuming an idealized notion here, but later I will discuss the generalization to the non-ideal case.

The natural way to proceed is to identify scenarios with constructions out of sentences. We already have a notion of epistemic possibility that applies to these entities, and this notion can be exploited to construct scenarios directly. These sentences will need to be sentence types of an ideal language, since it is unlikely that any existing language will have sufficient expressive power to specify all scenarios.

The ideal language must have certain properties. First, it must allow infinite sentences, in order to specify scenarios with infinite extent. I will discuss the precise nature of these infinite sentences later on, in section 9. Second, the ideal language should be restricted to epistemically invariant expressions. This ensures that we can associate epistemic properties with sentence types, not just with sentence tokens: when s is epistemically invariant, then if some possible competent utterance of s is epistemically necessary, all possible competent utterances of s are epistemically necessary. It also ensures that we can appeal to implication relations between sentences in the ideal language and sentence tokens in a nonideal language, as defined earlier.

It is arguable that most terms of a natural language such as English are not invariant. It is plausible that most ordinary proper names are not invariant, so they should be excluded from the ideal language, though arguably some descriptive names can be allowed. Something similar applies to most natural kind terms, but here there will often be an invariant term in the vicinity. In the case of theoretical terms, for example, these might be used by different speakers with somewhat different theoretical reference-fixers, but we can stipulate an invariant term in the vicinity with a fixed theoretical reference-fixer. Something similar applies to most context-dependent terms. For most context-dependent terms as used in a context, there will be a possible term that is not context-dependent in this way. For example, if 'know' is context-dependent because of variation in standards, there will be possible terms such as 'know(high)' and 'know(low)' that are not context-dependent in this way.

Applying this process to a natural language such as English will plausibly leave a residue of many invariant terms. Certainly 'I' and 'now' are invariant (at least if precisified somewhat to remove certain sources of variation with speakers' intentions), as will be cleaned-up versions of many mental and physical terms, causal and dispositional terms, as well as logical and mathematical terms and so on. So there does not seem to be a problem with the idea of an ideal language consisting only of invariant expressions.

Finally, the ideal language must have a sufficiently broad lexicon. For now, we might as well stipulate that for *any* possible invariant simple expression e, the ideal language contains a synonym of that expression: that is, an expression e' such that any competent utterance of '$e \equiv e'$' is epistemically necessary.

We can say, much as before, that a sentence d of our ideal language L is epistemically complete when (i) d is epistemically possible, and (ii) there is no sentence s of L such that both $d\&s$ and $d\&\neg s$ are epistemically possible. When d is epistemically complete, it is in effect as specific as any epistemically possible sentence in the language can be. As before, let us say that d is *compatible* with s when $d\&s$ is epistemically possible, and *d implies s* when $d\&\neg s$ is epistemically impossible. Then if d is epistemically incomplete, it leaves questions open: there will be s such that d is compatible with s but d does not imply s. If d is epistemically complete, d leaves no questions open: if d is compatible with s, d implies s.

We can now identify scenarios with equivalence classes of epistemically complete sentences in L, where d_1 is equivalent to d_2 iff d_1 implies d_2 and d_2 implies d_1. It is plausible, though not completely trivial, that L contains epistemically complete sentences. For example, as long as there are maximal classes of mutually compatible finite sentences of L (classes such that the conjunction of every sentence in the class is epistemically possible, but the conjunction of these sentences with any sentence outside the class is not), then the conjunction of the sentences in such a class will be epistemically complete. It is not completely trivial that such maximal classes exist, but I sketch an argument for the existence of the needed epistemically complete sentences below.

If s is a sentence of an arbitrary language, we can say that a scenario w verifies a sentence s ($ver(w, s)$ is true) when d epistemically necessitates s, for some sentence d in the equivalence class of w. We can say that w falsifies s ($ver(w, s)$ is false) iff d epistemically necessitates the negation of s.[7] If we have an "indeterminate" truth-value, we can say that $ver(w, s)$ is indeterminate when d epistemically necessitates $indet(s)$. If there are any further truth-values v, something similar applies: $ver(w, s) = v$ when d epistemically necessitates $O(s)$, where O is an operator such that $O(s)$ is true iff s has truth-value v.[8]

The Plenitude thesis now requires the following:

> *Epistemic Plenitude*: For all sentence tokens s, if s is epistemically possible, then some epistemically complete sentence of L implies s.

This thesis is entailed by the conjunction of the following two theses:

> (E1) For all sentence tokens s, if s is epistemically possible, then some epistemically possible sentence of L implies s.
>
> (E2) For all sentences s of L, if s is epistemically possible, then some epistemically complete sentence of L implies s.

[8] In the case where s is a sentence token, this should be understood as the claim that d implies a negation of the thought expressed by s. Something similar applies to the other truth-values: for example, $ver(w, s)$ is indeterminate iff d implies an indetermination of the thought expressed by s, where indetermination is understood as a mental analog of the indeterminacy operator.

The first thesis requires, in effect, that every sentence token *s* is implied by some invariant sentence. We could rephrase the second thesis by saying that any epistemically incomplete sentence *s* of *L* is *completable*: this requires officially that *s* is implied by some epistemically complete sentence in *L*, which comes to the claim that *s* can be expanded into an epistemically complete sentence by adding further conjuncts. Neither claim is trivial, but both are plausible (subject to a complication regarding (E1) that I will discuss). I will not try to prove these principles here, but I will make a prima facie case for them.

A case for (E2) runs as follows.[9] First, we can note that if *s* is *true*, *s* is plausibly completable. The world itself is determinate, making all sentences of *L* true or false (setting aside borderline cases of vague sentences, and the like, whose impact on this sort of argument is discussed in section 9). Conjoining all true sentences of *L*, if it were possible, would yield an epistemically complete sentence that implies *s*. Such a conjunction is probably impossible (perhaps because this sentence would have to be one of its conjuncts), but it remains plausible that some conjunction of sufficiently many atomic sentences of *L* is epistemically complete and implies *s*. All this depends on the details of the language *L*, but assuming a suitable language, this reasoning is plausibly a priori. That is, for any epistemically possible *s*, it is a priori that if *s* is true, *s* is completable. It follows that if *s* is epistemically possible, it is not a priori that *s* is uncompletable. Furthermore, the uncompletability of *s* seems to be sort of thing that is knowable a priori if it is knowable at all. So unless the uncompletability of *s* is wholly unknowable (even given ideal reasoning), *s* is completable.

This is not a rigorous proof of (E2), but it gives (E2) some prima facie support. Under certain assumptions (discussed in section 9 of this paper), unknowability can be excluded entirely, strengthening the support. But even without these assumptions, the hypothesis that some *s* are uncompletable but not knowably uncompletable is not especially attractive. At least, if the rest of the reasoning is correct, we can know that we will never be able to discover a counterexample to (E2).

As for (E1): to a first approximation, (E1) is plausible because the ideal language (L) should be able to capture more fine-grained possibilities than any given sentence token in natural language. If *s* is a token of an invariant expression, (L) will contain a synonymous sentence, so there is no

[9]This argument is loosely inspired by an argument given by Cresswell 2006 for a modal principle analogous to (E2). Cresswell attributes this sort of argument to Aristotle. Cresswell also has a useful discussion of the conditions under which modal principles such as (E2) are true, focusing on other varieties of modality, but suggesting that such principles are especially plausible where epistemic possibility is concerned.

problem here. And plausibly, when *s* is a token of a variant sentence, there will be some invariant expression that matches its content on any given occasion of use. The most likely exceptions here are indexicals. 'I' and 'now' are no problem as they are invariant, but a complication arises because of the case of demonstratives discussed earlier.

The picture so far suggests a common space of scenarios for all speakers. This picture has to be qualified slightly to handle the case of demonstratives. We have seen already that to handle these cases, canonical specifications of scenarios sometimes need to include phenomenal demonstratives that are specific to subjects. One might regard these demonstratives as terms of the ideal language, albeit unusual terms in that any one of them can be used by a single speaker. Perhaps better, one can say that for a given subject at a given time, the language L^* for the specification of scenarios may involve one or more such demonstratives in addition to the common language L. Understood this way, then thesis (E1) will be false of L^*, due to cases where *s* contains a relevant demonstrative, but both (E1) and (E2) will be true of L.

This yields a small modification of the original picture, with a subject-and-time-relative space of scenarios (or alternatively, a common space such that some elements of the space can be related only to specific subjects). However, to analyze sentences that do not contain relevant demonstratives or expressions that depend on them, then the common space of scenarios characterizable in the common language L will suffice. And even for subjects using relevant demonstratives, one can still map scenarios from one subject to another, up to isomorphism.

(It is also worth noting that if principle (E1) above is more radically false, because many epistemically variant sentence tokens are not implied by invariant sentences, then one could still engage in a version of the current construction by allowing arbitrary possible epistemically variant sentence tokens into canonical specifications of scenarios. This would yield a construction that satisfies Plenitude, at cost of having the space of scenarios be entirely subject-relative, without a useful notion of isomorphism between scenarios of different subjects.)

It is easy to see that this construction will satisfy Compositionality, as the principle follows from the analogous principle about implication. The construction will also satisfy Parsimony: if two sentences of L imply the same sentence tokens, then they will imply each other (at least assuming that the sentences can be uttered), so they will be members of the same equivalence class. The language is also designed so that it satisfies Specification.

As for Actualization and Uniqueness: it is not obvious how to define the relation of actualization between scenarios and sentence tokens. To do this we need to define a corresponding relation of actualization between epistemically complete sentences *d* and sentence tokens. It is tempting to say: such a sentence *d* is actualized at *s* iff, were the subject uttering *s* to utter *d*, the utterance of *d* would be true. But this cannot work, for the obvious reason that uttering *d* would change the world in which *s* is uttered. An alternative definition appeals to the notion of a canonical specification of a centered world from the previous section. For a sentence token *s*, let *w* be a centered world centered on the speaker and the time of utterance (and any experiences associated with demonstratives, if necessary), and let d' be a canonical specification of *w*: that is, an epistemically complete sentence including neutral terms and indexicals that is true of *w*. Then *d* is actualized at *s* iff *d* implies d'. It would be nice to have a definition that does not appeal to the notion of neutrality, or to the thesis that there are epistemically complete neutral/indexical specifications, as this notion and the associated thesis are otherwise unnecessary for the epistemic construction. But the nature of such a definition is currently an open question. For now, I will take it that we have a reasonably good intuitive grip on the notion, and the definition just given is also available. So I will assume a relation of actualization between scenarios and sentence tokens henceforth.

An important residual issue concerns the question of how small the ideal language can be while still satisfying Plenitude. If the language needs a term for every invariant expression, then the resulting semantic values at least for invariant sentences will be fairly uninteresting: they may simply be implied by all sentences of the ideal language that contain the original sentence as a conjunct. However, if the language only needs a relatively limited class of invariant expressions, then the structure will be much more interesting. I have argued elsewhere that a relatively small vocabulary suffices at least for the purposes of specifying scenarios that correspond to the actual world: see, for example, Chalmers and Jackson (2001). An extension of this reasoning suggests that a reasonably limited (if larger) vocabulary suffices to specify any scenario. Such a vocabulary will serve as a sort of basis for epistemic space. I will not investigate the character of such a basis here, but I discuss the issue at length in forthcoming work.[10]

Although we have constructed scenarios out of sentences here, other constructions are quite possible. One can even take the linguistic construction and convert it into another sort of construction. For example, if there is an epistemically complete invariant language including just neutral terms and indexicals, then each neutral term will have some object, property, or relation as its extension. We can then convert the neutral part of any epistemically complete specification into an abstract object that is a complex of the relevant objects, properties, and relations. This abstract

[10]For some related discussion in published work, see the discussion of scrutability principles in Chalmers 2002c and 2004, as well as the discussion of *PQTI* (involving physical, phenomenal, and indexical vocabulary along with a "that's all" clause) as a specification of the actual world in Chalmers and Jackson 2001.

object can be seen as a sort of "quasi-world", akin to a possible world except that the relevant state of affairs may or may not be metaphysically possible. One could then see scenarios as centered quasi-worlds. This has the advantage of moving out epistemically constructed scenarios out of the realm of language and into the realm of being.

6 Epistemically constructed scenarios and metaphysically possible worlds

We might call the constructions of scenarios in the last two sections the metaphysical and the epistemic construction respectively.[11] How are these two constructions of scenarios related to each other? Assuming both are coherent and that the relevant assumptions (not including Metaphysical Plenitude) are satisfied, there will be epistemically complete canonical specifications for each centered world and each epistemically constructed scenario. We can then say that a centered world and an epistemically constructed scenario *correspond* if their specifications imply one another. It will now certainly be true that for every centered world, there is a corresponding epistemically constructed scenario. *If* Metaphysical Plenitude is true (as I think it is), then for every epistemically constructed scenario, there will be a corresponding centered world (possibly more than one, due to failures of parsimony). If Metaphysical Plenitude is false, on the other hand, there will be epistemically constructed scenarios with no corresponding centered world.

If Metaphysical Plenitude is false, this will pose a serious obstacle to the metaphysical construction, but not to the epistemic constuction. Even on the theist views discussed earlier that deny Metaphysical Plenitude, for example, there will be scenarios verifying 'There is no omniscient being'. Even on the relevant mathematical view, there will be scenarios verifying the negation of the Continuum Hypothesis. Even on the relevant views on laws, there will be scenarios verifying the negation of laws. Even on relevant views on the mind–body problem, there will be scenarios verifying the claim that there are zombies. It is just that on these views, there will be no metaphysically possible world corresponding to these scenarios.

Apart from questions involving Metaphysical Plenitude, the constructions differ mostly in re-

[11] The metaphysical and epistemic constructions correspond roughly to the "one-space" and "two-space" views of modality discussed by Jackson (this volume). Like Jackson, I think that the one-space model is adequate, but unlike Jackson, I think that the two-space model is coherent and useful for various purposes. Jackson's central arguments against the two-space model depend, in effect, on the assumption that individuals can be reidentified across scenarios. I argue in section 8 that this assumption should be rejected.

quiring somewhat different assumptions. The metaphysical construction requires notions of epistemic and metaphysical necessity and a notion of neutrality, along with the thesis that every sentence is implied by some invariant neutral/indexical sentence. The epistemic construction requires only the notion of epistemic necessity and of invariance (both of which are also required by the metaphysical construction), along with the thesis that every sentence is implied by some invariant sentence. These assumptions are significantly weaker, which is another reason for preferring the epistemic construction if one is aiming for maximal generality.

Even on the epistemic construction of scenarios, there are many interesting interactions between epistemically possible scenarios and metaphysically possible worlds. One such interaction concerns epistemic possibilities concerning what is metaphysically possible. For example, one might hold that it is epistemically possible that Metaphysical Plenitude is true, and epistemically possible that it is false. One might even hold that it is epistemically possible that there is only one metaphysically possible world. If these views are correct, then there will be scenarios at which Metaphysical Plenitude is true, scenarios at which Metaphysical Plenitude is false, and scenarios at which 'There is only one possible world' is true.

These cases are naturally modeled in a two-dimensional way, by supposing that every scenario is associated with a modal space of putatively metaphysically possible worlds. (These putative worlds might themselves be modeled linguistically, or in some other way.) On the view just described, some scenarios will be associated with a space involving just one putative world (one that presumably corresponds to the scenario itself), while others will be associated with a space that has a putative world for every scenario.

If one accepts (as I do) that Metaphysical Plenitude is both true and a priori, then the structure will be simpler than this. In particular, every scenario will be associated with a space of putative worlds such that there is a putative world for every scenario. This raises the possibility that we can use the same set of possible worlds to model the space of putative worlds associated with every scenario, as on certain versions of two-dimensional semantics.[12] But in any case, the epistemic construction of scenarios gives us the tools to model a wide range of views about metaphysical modality.

[12]This issue is discussed at more length in Chalmers 2004, section 3.10.

7 Subsentential epistemic intensions

So far I have defined the evaluation of expressions in scenarios only for sentences. For many purposes it is useful to define this sort of evaluation for arbitrary expressions that have an extension, such as singular terms, general terms, kind terms, and predicates. I will take it that we have already decided on independent grounds what sort of extensions these expressions should have: e.g. individuals, classes, kinds, and properties. We then want to define an epistemic intension for any such expression, mapping scenarios to extensions within those scenarios.

Formally, we need a function *ext* from scenarios and these expressions to extensions, such that $ext(w, e)$ (the extension of e in w) is an entity of the appropriate sort. The epistemic intension of an expression e is a mapping from scenarios w to extensions $ext(w, e)$. We can stipulate that when e is a sentence, $ext(w, e) = ver(w, e)$. This function should obey a principle analogous to Compositionality: insofar as the extension of a complex expression e depends on the extension of its parts, the extension of e in a scenario w depends on the extension of its parts in w in the same way. And it should obey a principle analogous to Actualization: if w is actualized at e, the extension of e should correspond to $ext(w, e)$.

The details depend to some extent on whether we take the metaphysical or the epistemic approach to scenarios. The difference is that centered worlds already come populated with individuals and the like, or at least we are familiar with how to regard them as so populated. By contrast, epistemically constructed scenarios as outlined so far do not come populated with individuals, or at least we are less familiar with how to regard them as so populated.

If we take the metaphysical approach to scenarios: let w be a centered world with canonical specification d, and let t be a singular term. Let us say that ϕ is an identifying predicate relative to d iff d implies 'Exactly one individual has ϕ'. Then for most referring singular terms t, there will be some neutral identifying predicate ϕ such that d implies 't has ϕ'. In such cases, let us say that the extension of ϕ in w is the individual that satisfies ϕ in w. We can then say that $ext(t, e)$ is the extension of ϕ in w.

In some symmetrical worlds, for some terms t there may be no such neutral identifying predicate ϕ. In many such cases, there will be an identifying predicate ϕ' involving neutral terms and indexicals. In this case, one can replace the indexicals in ϕ' by singular terms (which need not be neutral) picking out the entities at the center of the world, yielding an expression ϕ''. We can then say that $ext(w, t)$ is the extension of ϕ'' in w. If there is no such neutral/indexical identifying predicate ϕ', then $ext(w, t)$ is null.

One can do the same for general terms (assuming these have extensions). If g is a general term, one can appeal to a neutral (and possibly indexical) predicate ϕ such that d implies '$\forall x$(x is a g iff $\phi(x)$)', holding that $ext(w, g)$ is the extension of ϕ in w (or of a de-indexicalized version thereof): that is, the class of individuals in w that satisfy ϕ. Kind terms and property terms are treated just as singular terms are (although here, of course, the denotation will be kinds and properties respectively). For predicates h, we appeal to neutral (and possibly indexical) predicates ϕ such that d implies '$\forall x(h(x)$ iff $\phi(x))$'. This method can be extended to arbitrary expressions (and different proposals for their extensions), delivering epistemic intensions for all such expressions.

If we take the epistemic view of scenarios, then we need to populate scenarios with individuals and the like. If we simply admit scenarios as a basic sort of abstract object with certain properties, one could simply stipulate that they contain individuals that can serve as the extensions of relevant expressions—much as many of those who introduce possible worlds simply stipulate something similar. But it is useful to go through an explicit construction.

Let w be a scenario with canonical specification d. Let us say that a denoting term is a singular term or a definite description. Then we can say that two denoting terms t_1 and t_2 are equivalent under w if d implies 't_1 is t_2'. Then we can identify every equivalence class of denoting terms under w with an individual in w, and hold that for a singular term t, $ext(w, t)$ is the individual corresponding to t's equivalence class in w. As for general terms: for a general term g, $ext(w, g)$ is that class of individuals whose corresponding equivalence class includes a denoting term t such that d implies 't is a g'. One can do something similar for predicates and kind terms: the details will depend on the precise view one takes of properties and kinds and their relation to individuals, so I will not go into them here.

There is one worry: what if a scenario requires that there are individuals that are not denoted by any denoting term? In particular, what if the truth of certain existentially quantified claims in a scenario requires individuals that are not the referent of any denoting term? For example, there may be a predicate ϕ such that d implies '$\exists x\phi(x)$', and d does not imply any claim of the form '$\phi(t)$', where t is a denoting term. Because d is epistemically complete, it will at least tell us exactly how many individuals have ϕ, whether some individuals with ϕ also have ψ and some do not, and so on. Of course if it tells us that some individuals with ϕ have ψ and some do not, then we can move to the conjunctive predicates $\phi\&\psi$ and $\phi\&\neg\psi$. Repeating this process, it is not hard to see that this sort of case requires predicates ϕ (perhaps an infinitely conjunctive predicate) such that d implies that there exists more than one individual with ϕ, and such that for all predicates ψ, d implies that these individuals are indistinguishable with respect to ψ. In this case, the individuals will be indistinguishable even in our idealized language, perhaps because of deep symmetries in the scenario.

In such a case, if d implies that there are n individuals with ϕ, one can arbitrarily construct n individuals, perhaps as ordered pairs $(\phi', 1)...(\phi', n)$, where ϕ' is the equivalence class containing ϕ. We can then stipulate that all these individuals fall under the extension of ϕ. Likewise, all these individuals fall under the extension of ψ for all predicates such that d implies 'everything that is ϕ is ψ'', and fall under the extension of general terms g such that d implies 'everything that is ϕ is a g', and so on.

One can populate a scenario with kinds by applying the same treatment as above to kind terms. One can populate it with properties, relations, and other entities in a similar manner. In this way, we can populate a scenario with entities that are needed to serve as the extensions of expressions, and we can specify the extensions of all relevant expressions at arbitrary scenarios.

As in the case of possible worlds, the entities we have used to construct individuals in scenarios are not themselves concrete objects, but they serve as proxies for concrete objects that exist if the scenarios are actualized (or that would exist if the worlds were actual). Where the objects in the actual world are concerned, one can treat the relevant abstract objects (classes of descriptions and the like) as proxies for the corresponding actual object, thereby yielding a version of the principle of Actualization. Of course once one has engaged in this sort of construction, one need not usually bother with the details again. Just as in the case of possible worlds, it is reasonable thereafter to speak of a scenario as containing individuals and the like, and to speak about terms as picking out various individuals in a scenario, quite independently of the details of the construction.

Compositionality is ensured by the details of the construction. For an identity statement (e.g. '$t_1{=}t_2$'), compositionality will be ensured by the equivalence class construction. For a predication (e.g. 't is a g', or $\phi(t)$) this will be ensured by the appropriate construction of extensions for general terms (as above) or predicates. The machinations two paragraphs above ensure that existential quantification will work straightforwardly, and universal quantification is guaranteed to work (if d implies $\forall x \phi(x)$, then every individual constructed above will have ϕ). Logical compositionality is guaranteed at the sentential level (if d implies both s and t, d will imply $s\&t$, and so on). Something similar applies to any construction involving compositionality of extensions. So for any such construction, the epistemic intension of a complex expression will be a compositional function of the epistemic intension of its parts.

8 Trans-scenario identity

One of the most hotly contested issues concerning possible worlds concerns whether there is transworld identity: can the same individual be identified across two different worlds? In the domain of epistemic space, an analogous issue arises: the question of *trans-scenario identity*. Can we say that an individual in one scenario is the same individual as that in another scenario?

In many cases, it seems that the answer is no. Consider the actualized scenario (for me now), in which 'Hesperus is Phosphorus' is true. Relative to this scenario, 'Hesperus' and 'Phosphorus' pick out an individual *x*. In another scenario *w*, 'Hesperus is Phosphorus' is false. Relative to this scenario, 'Hesperus' picks out one individual, and 'Phosphorus' picks out another. Can one say that both of these individuals are *x*, or that just one of them is? Neither answer seems attractive. So it seems that one cannot say that in any given scenario, *x* is identical to the referent of 'Hesperus', or that in any given scenario, *x* is identical to the referent of 'Phosphorus'.

Is there any other way to ground trans-scenario identity? Of course if scenarios are understood as possible worlds, we could appeal to transworld identity. To avoid entangling the epistemic and metaphysical modalities here, however, I will first work with the epistemic construction of scenarios, and will later consider the metaphysical construction.

A natural way to ground trans-scenario identity would be to isolate a *canonical designator n* for any individual *x* in a scenario, and say that in any scenario, *x* is the referent of *n* with respect to that scenario. The trouble is that at least for the objects designated by most ordinary singular terms, there does not seem to be any obvious choice of a canonical designator. For example, in the case of Venus, the designators 'Venus, 'Hesperus', 'Phosphorus', and many others will all give different results. The same goes for tables, people, countries, and so on.

In some cases involving abstract objects, there do seem to be canonical designators. For example, as Ackerman (1978) has discussed, numerals seem to function as canonical designators for numbers. These canonical designators can be used to ground claims of trans-scenario identity. Assuming that '2' designates the number two in the actual world, then it will designate an entity in many or all scenarios (depending on one's view of the apriority of the existence of numbers), and we can stipulate that these entities are identical with each other. One could do the latter either by modeling trans-scenario identity between individuals with a relevant relation, or, if it is important that individuals in scenarios literally be identical to each other, one can modify the previous construction of individuals. To to the latter, one could identify individuals with classes of individuals (as previously constructed) in different scenarios that are picked out by a canonical designator. Or in cases where the designator picks out an object in the actual world (as might be the case for '2'), one could identify individuals in a scenario with the actual object itself, invoking a "present in" relation between individuals and scenarios, and invoking claims about the predicates that an individual falls under relative to a scenario.

What is the relevant difference between '2' and 'Hesperus', and between two and Hesperus? Intuitively, the difference is that '2' is *epistemically rigid*—that is, it picks out the same object in all scenarios—while 'Hesperus' is not. Of course this intuitive characterization presupposes a notion of trans-scenario identity (just as Kripke's notion of rigid designation presupposes a notion of transworld identity), so it cannot be used to provide an independent grounding for trans-scenario identity, but it at least helps give a sense of what is going on. One might try to characterize epistemically rigid expressions in other terms. One useful suggestion is that an epistemically rigid expression is one such that one can know what it refers to a priori. This definition inherits the imprecision of the notion of knowing what an expression refers to, but there is at least an intuitive sense in which one can plausibly know a priori what '2' refers to (or better, what object two is), while one cannot know a priori what 'Hesperus' refers to (or better, what object Hesperus is).

When an expression is epistemically rigid, it will usually also be rigid in the Kripkean sense (subjunctively or metaphysically rigid). In such a case, and when the term is rigid *de jure* rather than merely *de facto*, we can say that the expression is *super-rigid* (a term due to Martine Nida-Rümelin). Any super-rigid term is neutral, but not every neutral term is super-rigid. For example, a general term such as 'philosopher' is arguably neutral without being epistemically rigid, subjunctively rigid, or super-rigid (it picks out different classes in different worlds). On the other hand, if 'philosopher' is neutral, the nearby property term 'the property of being a philosopher' will be both neutral and super-rigid. For any neutral term, one can find a super-rigid property term in the vicinity in this way.

Many properties have super-rigid canonical designators. For example, I have argued elsewhere that our central phenomenal concepts designate phenomenal properties super-rigidly. Something similar may apply to many mental properties, many causal and dispositional properties, and so on. However, when properties are constitutively tied to external objects (e.g., the property of being taller than Fred) or kinds (e.g., the property of containing water), then if there are no super-rigid designators for those objects and kinds, then there will plausibly be no super-rigid designators for the corresponding properties.

It seems plausible that while there are super-rigid designators for many abstract objects and many properties, there are no super-rigid designators for concrete objects. At least, such designators are extremely hard to find. One might suggest that if an object x has an essential identifying property ϕ—that is, a property ϕ such that necessarily something is x iff it has ϕ—then one can use a super-rigid designator for this property to construct a super-rigid designator for ϕ. But the most plausible candidates for such essential identifying properties (such as the property of being descended from a particular sperm and egg) will themselves be object- or kind-involving, so that there will be no obvious canonical designators for them, or at best there will be a regress of designators.

Likewise, the property terms discussed above that are candidates for super-rigid designation do not seem to obviously yield candidates for essential identifying properties. *Perhaps* one could argue that one's ontology should admit an object x such that necessarily, an object is x iff it is the biggest object in the universe, or iff it is the only individual with phenomenal property ϕ, or iff it is the first philosopher in the world. If so, then one could allow canonical designators and trans-scenario identity for objects of this sort. But these are at best objects of a very unusual sort.

One might think that one can at least refer super-rigidly to oneself. In this case, there is at least a canonical designator: the first-person pronoun. But this designator does not obviously support super-rigid reference. On the centered worlds model of scenarios, 'I' picks out many different individuals in different centered worlds. And on the epistemic construction, the individual at the center can have almost any range of properties (and need not even exist). Perhaps one could hold that at the center of any scenario there is always a common individual, EGO. But this would be a very odd sort of object—even odder than those discussed before, in that there seems to be no subject-independent fact of the matter about who is EGO in a world. Given the absence of a clear definition of epistemic rigidity, these facts do not conclusively establish that 'I' is not epistemically rigid, but they at least give good reason to doubt it. Something similar applies to 'now', and to demonstratives for token experiences (although terms for *properties* of experiences may be epistemically rigid, as discussed above).

If one wants to hold that reference to oneself is epistemically rigid, the best way to do so would be to hold that different individuals are related to their own subject-relative spaces of scenarios, such that each scenario in a subject's epistemic space has that subject at the center. This model might fit well with a Russell-style account that allows direct reference to the self as well as to properties and sense-data. But this model is at odds with our previous construction of scenarios, requiring significant modifications to both the epistemic and the metaphysical constructions. More importantly, this model makes cross-subject identification of scenarios impossible, and likewise makes it impossible for two subjects to share epistemic intensions. If extended to times and experiences, as parity would suggest, then the model would have even less generalizability across occasions. Finally, there is arguably an underlying epistemic difference between reference to numbers and properties, on the one hand, and reference to oneself on the other: merely possessing a concept of the former seems to put one in a position to know the nature of the referent a priori, whereas possessing a concept of the latter does not.

Still, it should be acknowledges that the choice between these models turns on delicate questions about the explanatory role one needs epistemic space to play, about just what is involved in epistemic rigidity, and about just what is involved in first-person reference. A pluralistic picture giving a role to both models is not out of the question. Nevertheless, I am tentatively inclined to favor a model on which epistemically rigid reference to oneself is impossible.

In the absence of canonical designators, is there any other way to pin down trans-scenario identity between ordinary objects? One might try to use *de re* claims such as: it is a priori of Venus that it is such-and-such. The trouble is that no such claims seem clearly to be true, except perhaps for trivial claims involving self-identity and the like. Perhaps there is a loose sense in which it is a priori of Venus that it is visible in the evening (if it exists) in virtue of the fact that it is a priori that Hesperus is visible in the evening (if it exists). But in this sense, all or almost all of Venus's properties will be a priori of it (for example, where ϕ is such a property, one can stipulate a partially descriptive name 'ϕ-Venus' such that it is a priori that if ϕ-Venus exists, ϕ-Venus is Venus and ϕ-Venus is ϕ). So one does not get to any interesting sort of trans-scenario identity this way.

One might try an analog to Kripke's method of asking, of an object such as Venus, whether if such-and-such a world obtained, then *it* would have been visible in the morning. Here, we could ask of Venus whether, if such-and-such a scenario obtains, then *it* is visible in the morning. But there seems to be no good way to answer this question. Consider a scenario verifying 'Hesperus is not Phosphorus', in which separate objects are visible in the morning and evening. If this scenario obtains, is Venus visible in the morning? There seems to be no way to say. Perhaps, following the analogy with Kripke, one could simply stipulate that the scenario in question is one in which *Venus* (that very object) is visible in the morning and not the evening. But such a stipulation will lead to serious problems, on the current model.

Consider the question: in a scenario stipulated to be such that the object has ϕ, can 'Hesperus has ϕ' be false? If no, then presumably by parity the scenario must also verify 'Phosphorus is ϕ', 'Venus is ϕ', and so on. So any such scenario will verify 'Hesperus is Phosphorus', 'Hesperus is Venus', and so on for any pair of names of the object. This entails that the object can exist only in a tiny fraction of scenarios, and arguably only in the actual scenario, since it is arguable that for any non-actual scenario, there is some pair of names *a* and *b* for Venus such that '*a* is *b*' is false in that scenario (appealing to names such as 'ϕ-Venus' for appropriate ϕ, for example). If the answer to the question is yes, so that 'Hesperus has ϕ' can be false of such a scenario, then presumably the same goes for 'Phosphorus is ϕ', 'Venus is ϕ', 'that object is ϕ' (for any demonstrative way of picking it out), and so on. But now, the behavior of the object across scenarios will float free of any of our ways of talking or thinking about it about it, so that there are aspects of scenarios that float free of their role in verifying sentences and beliefs. Perhaps such aspects are not incoherent, but they seem to have no explanatory role to play in the current framework.

This is not to say that talk of *de re* epistemic possibilities is incoherent. It seems intuitively reasonable to say of the cup on my desk that I know that it is brown, while I do not know when it was made. So it is intuitive to say that there are epistemic possibilities open to me in which that very cup was made on such-and-such a date, or on such-and-such a date. But if we are to model epistemic possibilities of this sort in such a way that they stand in a verification relation to our sentences and beliefs, then either we need to say that an epistemic possibility in which x is ϕ verifies 'n is ϕ' for any name n of the object, or we will be led to say that whether x is ϕ in a scenario can float free of whether the scenario verifies 'n is ϕ' for any name n of the object. Both models are coherent, but neither is useful for our current purpose.

When the first model is fleshed out, it will almost certainly be a model on which all true identities involving proper names (such as 'Hesperus is Phosphorus') are true in all scenarios, undermining one of the main explanatory aims of the current project. When the second model is fleshed out, it will naturally lead to a model on which the object-involving aspects of a scenario are largely independent of the role they play in verifying sentences and thoughts, which will render them largely useless in the explanatory structure of the current project. Still, there may be other projects for which these models are useful. The first model in particular may play a useful role in illuminating aspects of *de re* thought, and our epistemic relations to Russellian contents (see Soames 2004 for a treatment of epistemic possibility that resembles the first model here). One can reasonably be a pluralist about epistemic space.

Returning to the preferred model I have outlined: it seems clear that this model supports trans-scenario identity only for certain abstract objects, and not for ordinary concrete objects. We might think of this as a "qualitative" conception of epistemic space. There are objects in scenarios, and they have properties, but only the properties are re-identifiable across scenarios (and here only some of them), and not the objects (except for abstract objects). For the purposes for which we are using this model, a notion of trans-scenario identity for concrete objects has no role to play. Of course, for nearby scenarios in which familiar identities ('Hesperus is Phosphorus') and so on are true, there is not much harm in talking of these scenarios as scenarios in which the object in question has various properties. But strictly speaking, this *de re* talk should always be cashed out by *de dicto* locutions, speaking of scenarios in which Hesperus has various properties (or to be maximally explicit, scenarios verifying 'Hesperus has ϕ'), and so on.

The last point brings out a terminological nicety: if a scenario verifies 'Hesperus is ϕ', is it reasonable to call it a scenario in which Hesperus is ϕ? I do not see why not, as long as one is careful. In particular, in describing scenarios in this way, one cannot freely substitute terms that are coreferential in our world. So a scenario in which Hesperus is ϕ need not be a scenario in which Phosphorus is ϕ. That is to say that talk of "a scenario in which..." creates an opaque context. There may also be some subject-relativity: if we use the terms somewhat differently, it could be that what you and I count as "a scenario in which Hesperus is ϕ" may differ. But as long as one is alert to these phenomena, then there is no objection to using this convenient way of speaking.

The discussion above all presupposes the epistemic construction of scenarios. What about the metaphysical construction? Here, one might think that there will be a notion of trans-scenario identity that derives from the notion of transworld identity. It is arguable that object-involving metaphysical possibilities are *relatively* unproblematic: names function as canonical designators for objects in modal contexts, *de re* modal claims are reasonably well-behaved (at least if we allow that statues are distinct from the lumps that constitute them), and so on. If so, and if scenarios are constructed from such possibilities, then it may seem that transworld identity yields transscenario identity.

We have already seen that things are not as simple, however. The discussion of parsimony earlier suggests that the distinctly object-involving aspects of centered worlds are largely irrelevant to the way they function as scenarios. For example, qualitatively identical centered worlds involving distinct objects will verify all the same sentences, so they can naturally be seen as corresponding to a single scenario. And where transworld identity between concrete objects is present, it need not correspond to anything interesting at the epistemic level (in effect, it yields only a version of the "second model" discussed above). For example, if we stipulate a centered world where Aristotle died in childbirth while someone else wrote the books that have come down to us under the name of 'Aristotle', then where *verification* of our sentences and beliefs is concerned, the latter is more relevant than the former.

All this suggests that even if one believes in transworld identity, it is best to set it aside in considering the role that centered worlds play when functioning as scenarios. Or perhaps even better, for this purpose one can invoke a purely qualitative construction of centered worlds out of properties, so that the worlds in question do not support a natural relation of transworld identity between objects. As with epistemic space, one can be a pluralist about the construction of modal space, depending on one's purposes. For the role that modal space is needed to play here, a qualitative construction seems best.

The discussion above tends toward a conclusion suggested by Burgess (1997): that insofar as Quine's critique of quantified modal logic was concerned with *epistemic* modalities, it was not far from the mark. Burgess argues plausibly that Quine is concerned with modalities such as analyticity and apriority. Where these modal notions are concerned, many of the points above mirrors Quine's: different designators for an object yield different results in these modal contexts, and there are no canonical designators, so there is no way to derive *de re* quantified modal claims from *de dicto* modal claims, and there is no clear way to make sense of *de re* modal claims of this sort independently. Kripke responds to Quine, in effect, by invoking a different sort of modality, the subjunctive modality, to which Quine's arguments do not apply. As with Kripke's response to Frege, there is room for a split verdict: Kripke is right about subjunctive modality, while Quine is right about epistemic modality.

9 Infinitary scenarios

Some tricky issues arise from the fact that scenarios can have infinite extent, and that we have used an infinitary language to characterize scenarios.[13] There are questions about the exact size of the space of scenarios, closely related to problems that Kaplan (1995) raises concerning the size of the space of possible worlds. More basically, there is the question of the choice of infinitary language. What sort of infinitary constructions should be allowed: infinite conjunctions, infinite disjunctions, infinite sequences of quantifiers? Furthermore: how infinite are infinitary conjunctions (and so on) allowed to be? A countable number of conjuncts? Uncountable? As many conjuncts as an arbitrary infinite cardinal from set theory?

There are reasons to believe that one should allow scenarios corresponding to arbitrarily large conjunctions. One way to see this is to note that for any cardinal κ, it seems to be epistemically

[13] I am grateful to Bruno Whittle for pressing Kaplan-style worries about the space of scenarios in the case of epistemic space, and to Kit Fine and Wolfgang Schwarz for very helpful discussion. Whittle (2009) presses these worries in depth, responding in part to an earlier version of this paper in which these issues were not discussed.

possible that there are at least κ independent atomic entities in the universe, such that each entity can have or fail to have a simple property ϕ. This suggests that there are at least 2^κ scenarios, such that each scenario can be described using a conjunction of κ conjuncts.[14] If κ is an infinite cardinal, and if we stipulate that each conjunct must have length less than κ, then this scenario will not be describable using a conjunction of fewer than κ statements. So our ideal our ideal language should allow infinitary conjunctions with size corresponding to arbitrary cardinals, and some scenarios will require arbitrarily large conjunctions for their specification.

These issues are closely related to Kaplan's paradox concerning possible worlds, which we can put as follows. The following three claims are all prima facie plausible but are inconsistent:

(i) There are at least as many propositions as sets of worlds.

(ii) There are at least as many worlds as propositions.

(iii) There are more sets of worlds than worlds.

Claim (i) can be understood as stipulative if we take propositions as sets of worlds (it is also plausible on many other understandings of propositions). Claim (ii) is intuitively justified by mapping any proposition to a world in which that proposition is uniquely asserted (or in which it is uniquely entertained). Claim (iii) seems to follow from Cantor's theorem, which suggests that the set of all worlds, like any set, has more subsets than members. Prima facie, this situation suggests that there is no good candidate to be the cardinality of the set of all worlds, and that there may be no such set.

Kaplan's paradox arises at least as strongly when worlds and propositions are replaced by scenarios and intensions. If anything, the situation is worse. Lewis (1986) responds to Kaplan's problem by holding that there are propositions that are not asserted or entertained in any possible world. One might likewise hold that there are intensions (sets or classes of scenarios) that are not uniquely asserted or entertained in any scenario. But it is far from clear that the unique assertion or entertaining of any given intension can be ruled out a priori. Prima facie, any scenario can be specified by an infinitary conjunction, and any set of scenarios can be specified using an infinite

[14]Strictly speaking, where epistemic possibility as opposed to metaphysical possibility is concerned, symmetries within the scenarios and the absence of trans-scenario identity might yield many fewer than 2^κ different scenarios: when κ is infinite, it might yield only $g(\kappa)$ different scenarios, where $g(\kappa)$ is the number of cardinals less than κ. But even this is enough to make the key point that for every κ there must be at least 2^κ scenarios, as for every κ there is some cardinal μ such that $g(\mu) > 2^\kappa$.

disjunction of such conjunctions. There is no obvious a priori obstacle to the entertaining of such a conjunction or disjunction by an infinite being.[15]

(Kaplan also gives a constructive version of the paradox which does not turn directly on considerations about cardinality. To simplify, he constructs a proposition p consisting of those worlds w in which the set of worlds determined by the unique proposition asserted at w does not include w. Then if v is a world in which p is uniquely asserted, the set of worlds determined by p cannot include or exclude v, leading to contradiction. Whittle (2009) develops a version of the constructive paradox for the framework of epistemically possible scenarios. I focus on the nonconstructive version of the paradox here because, like Anderson (2009), I take the constructive version of Kaplan's paradox to be a version of the liar paradox that does not have much especially to do with possible worlds.[16] To support this point, it is worth noting that Kripke (forthcoming) gives a version of the constructive paradox with times in place of worlds.)

Lewis's official reasons for denying that every proposition can be entertained rest on his functionalism, which he takes to be a priori, so one might think these reasons also apply to epistemic possibility. But his argument rests also on the unargued claim that there is some cardinal upper bound on the number of functional roles. And once we allow arbitrarily complex infinitary beings, it is easy to generate arbitrarily many functional roles. Given κ states (each corresponding to a thought, for example), one can straightforwardly define 2^κ functional roles in terms of those states (each corresponding to a conjunction of some of the original thoughts, for example). Likewise,

[16]To see this, note that the key proposition p will be expressed at v by a liar sentence such as "The unique proposition asserted at this world is false". As such, the status of p should be handled by whatever mechanism best handles the liar paradox. Whittle makes a case (by invoking a somewhat more complex construction) that where epistemically possible scenarios are concerned, the move of holding that the problem sentence does not express a proposition is more difficult than in the case of metaphysically possible worlds. This may be right, but I take it that this move is a highly problematic treatment of liar sentences in any case, for reasons tied to compositionality. Other more promising strategies for handling the liar paradox appear to apply to the current case as well as they apply in the original liar case. For example, if one holds that the liar sentence has a nonstandard truth-value, we can say that the problem sentence above will have an intension mapping v to this nonstandard truth-value. Of course there are many unresolved issues concerning the liar paradox (including especially problems arising from strengthened liar sentences), but these issues are problems for everyone.

The constructive and nonconstructive paradox are not unrelated: the former can be generated from arguments for the latter, by applying the standard diagonal proof of Cantor's theorem to the mapping that generates thesis (ii) of the nonconstructive paradox. Nevertheless, one should distinguish the issue generated by the truth of Cantor's theorem (which has no particular connection to the liar paradox) from the issue generated by its standard proof. Thanks to Bruno Whittle for discussion here.

a belief involving any real number can be functionally defined in terms of beliefs involving rational numbers (using comparisons to smaller and larger rationals); a belief involving any set of real numbers can be functionally defined in terms of beliefs about real numbers (using judgments about whether the real number in question is in the set in question); and so on. So there do not seem to be clear a priori limitations here. If one is prepared to accept strong necessities, one might accept brute limitations on the complexity of worlds and on the complexity of possible thinkers. But this strategy will not help where epistemic possibility is concerned.[17]

Another response to the paradox, suggested by Kaplan himself, is to ramify the space of propositions and the corresponding space of worlds. Level-0 propositions concern only extensional matters, and level-0 worlds are (or correspond to) maximal level-0 propositions. Level-1 propositions concern extensional matters and level-0 propositions, and level-1 worlds are (or correspond to) maximal level-1 propositions. And so on. Then there is a level-*n* proposition for every set of level-*n* worlds, and there is a level-$n + 1$ world (but not a level-*n* world) for every level-*n* proposition, so paradox is avoided (though there remains an issue concerning propositions and worlds "simpliciter", analogous to an issue I discuss below).

Kaplan's response might in principle be applied to epistemic space, But there are reasons for concern. One worry is that it is arguable that all truths about propositions are epistemically necessitated by level-0 truths, and likewise when truths are replaced by epistemically possible sentences: if so, the level-*n* scenarios will simply correspond to the level-0 scenarios for all *n*. Another worry is that Kaplan's treatment disallows assigning semantic values uniformly to all sentences of natural language. Some sentences, such as 'All propositions are true', cannot be assigned a semantic value at all, and for any *n*, one can generate sentences that can only be assigned semantic values for levels greater than *n*: 'All level-*n* propositions are true', for example.

Furthermore, both Kaplan's and Lewis's responses turn on considerations specific to Kaplan's paradox and to issues about entertaining or referring to propositions. But I am inclined to think that the source of the worry is not as specific as this. The case at the start involving κ atomic entities appears to have much the same moral as Kaplan's paradox: it suggests that for any κ, there are more than κ scenarios (or worlds), so that there are too many worlds to form a set. And this case has nothing especially to do with entertaining or referring to propositions. I draw the moral that the source of both worries is that the worlds are broadly analogous to the sets. Any space of scenarios (like the space of sets) is in some sense indefinitely extensible. Ever more complex

[17]In *Counterfactuals*, Lewis suggests that the cardinality of the space of worlds might be $beth_2$, for reasons tied to the character of spacetime. But it is hard to see why our spacetime should restrict the space of worlds.

spaces of scenarios, of larger and larger cardinalities, can be generated, so that the scenarios as a whole (like the sets as a whole) cannot be collected into a set. This suggests that we might use the same sort of tools used to understand the set-theoretic paradoxes to understand this situation.

At this point it is natural to respond to Kaplan's paradox by denying (iii): just as there are no more sets of sets, there are no more sets of worlds than worlds. (Cantor's theorem does not apply when the entities in question do not form a set.) I think that this is the correct response, although it raises important issues about how intensions (or propositions) are then to be understood, and about how epistemic space can work if its members do not form a set.

In what follows, I will develop two strategies for responding to the paradox. The first strategy involves a stratified picture of the scenarios, with different spaces of scenarios corresponding to different cardinalities. On this view, each space forms a set, and the corresponding intensions can be understood in set-theoretic terms. The second strategy involves understanding intensions in non-set-theoretic terms, for example in terms of defining formulae. I think the second strategy runs deeper than the first and is also less technical and more general, so one could in principle skip straight to the second strategy. Still, I think that the first strategy helps to illuminate the situation by fleshing out a stratified structure among scenarios that is at least somewhat analogous to the stratified structure among sets.

1. The stratified construction of scenarios. Let us assume an infinitary language *L*. I will assume that *L* has a countable lexicon, consisting at least of the sort of expressions that make up a basis for epistemic space as discussed earlier, and perhaps of expressions corresponding to arbitrary invariant expressions in possible natural languages spoken by finite speakers.[18] Various rules concerning infinitary constructions are possible, but I will assume that the language at least allows infinitary disjunctions and conjunctions of arbitrary length. We might also allow infinite sequences of quantifiers, as is familiar from infinitary logic. There is no obvious obstacle to the claim that thoughts corresponding to sentences of *L* could be entertained by sufficiently infinitary beings, so I will assume that sentences of *L* can be assessed for epistemic possibility and necessity as before.

[18] One might worry that the choice of language will make the space of scenarios language-relative. In the lexicon is restricted to an epistemic basis, one can argue that imposing certain further constraints (e.g. requiring that the members of the basis are conceptually primitive in a certain sense) will remove any language-relativity from the resulting space of scenarios (although there remains the problem discussed under (E1*) below). If we allow the lexicon to include expressions corresponding to arbitrary invariant natural language expressions (plus the relevant indexicals), then as long as such expressions form a basis for epistemic space, the problem is removed. If there is not a countably infinite basis here, then an alternative model will allow a larger basic lexicon, or perhaps a lexicon whose size varies with κ.

For any infinite cardinal κ, let us say that a κ-sentence is a sentence of length less than κ. Then there will be at most $f(\kappa)$ κ-sentences, where $f(\kappa)$ is the sum of ω^α for all cardinalities $\alpha < \kappa$. (If the Generalized Continuum Hypothesis is true, $f(\kappa) = \kappa$ for all κ.) We can then say a κ-conjunction is a conjunction of at most $f(\kappa)$ κ-sentences. A κ-complete sentence is an epistemically possible κ-conjunction d such that for all κ-sentences s, $d\&s$ and $d\&\neg s$ are not both epistemically possible. We can then identify a κ-scenario with an equivalence class of κ-complete sentences, each of which will then be a specification of that scenario.

An important special case is the class of ω-scenarios, where ω is the cardinality of the integers. An ω-sentence is a finite sentence. There will be ω (= $f(\omega)$) ω-sentences. An ω-conjunction will be a conjunction of at most a countably infinite number of finite sentences. An ω-scenario will be an equivalence class of ω-complete conjunctions of this sort.

We can then say that κ-Plenitude is the claim that all sentence tokens (in a human natural language such as English) are verified by some κ-scenario. As with Epistemic Plenitude earlier, κ-Plenitude will follow from versions of principles (E1) and (E2).

(E1*), the analog of (E1), holds that any epistemically possible sentence token in such a language is implied by some member of $L(\kappa)$, the class of κ-conjunctions. Given that these sentence tokens are all finite, the fact that $L(\kappa)$ is restricted to κ-conjunctions does not raise any obvious reasons for concern for the argument given earlier. The restriction to a countable lexicon raises a potential concern, given that the previous argument for (E1) turned on the language having a synonym for arbitrary invariant expressions in natural language. Using that argument here requires the thesis that there is only a countable number of synonymy classes of possible invariant lexical expressions in natural language. This thesis is highly plausible in light of the fact that natural language speakers are finite beings.[19] If the thesis is denied, though, then one will need to appeal to a more limited countable basis, or perhaps better, one will need to expand the lexicon accordingly.[20]

The argument for (E2*), the analog of (E2), requires more work.[21] (E2*) holds that any epistemically possible κ-conjunction is κ-completable: that is, is implied by some κ-complete sentence.

[19]More specifically: for each such synonymy class there will be at least one corresponding cognitive state (that of a speaker using an expression in the class), and for finite beings it is plausible that there are only a countable number of relevantly distinct cognitive states. The second claim is particularly clear if one holds that natural language speakers can be modeled computationally.

[20]One might think that one could simply invoke a lexicon consisting of a countable basis for epistemic space. But a residual issue is that although we know that any epistemically possible natural-language sentence S is implied by an epistemically possible sentence in this lexicon, we do not know that any such sentence S is implied by a κ-conjunction in this lexicon.

One can argue for this roughly as I argued for (E2) earlier. If s is true, then s is κ-completable: s is implied by the conjunction of all true κ-sentences, and this conjunction is κ-complete (at least given that all κ-sentences are true or false, or given weaker assumptions discussed below). This reasoning is a priori, so it is a priori that if s is true, s is κ-completable. It follows that if s is epistemically possible, it is epistemically possible that s is κ-completable. That is, one cannot establish a priori that s is κ-uncompletable. Given that uncompletability of s is knowable a priori if it is knowable at all, it follows that unless this uncompletability is wholly unknowable (even given ideal reasoning), s is κ-completable. Unknowability can be excluded given certain assumptions about apriority, and even without these assumptions it seems a much less plausible option. So depending on whether these assumptions are granted, we have either a demonstration or a prima facie case for (E2*). I will give a concrete illustration of this case for completeness in the mathematical case described below.

This argument makes assumptions at two points: to make the case for κ-completability and to make the case against unknowability. At the first point, it suffices to assume that the basic vocabulary in $L(\kappa)$ is not vague, so that every κ-sentence is either true or false. Certain weaker assumptions also suffice. One sufficient assumption is the claim that the vocabulary includes an "indeterminately" operator *indet* such that when s is neither true nor false, *indet*(s) is true. Then as long as the connection between determinacy and epistemic possibility behaves as described earlier, *indet*(s)&s and *indet*(s)&¬s will not be epistemically possible, and a conjunction of κ-sentences including *indet*(s) sentences will be epistemically complete. One could also assume a multiplicity of such operators for various intermediate truth-values in borderline cases. Or one could assume that there is a precise subset of the vocabulary such that sentences using this subset determine the truth-value of all vague sentences.

At the second point, making the case against unknowability, certain assumptions about apriority suffice. These assumptions are most easily formulated using an apriority operator A. The first assumption is an S5 principle for apriority: (A1) if $\neg As$, then $A\neg As$. The second is a conjunctive closure principle (for arbitrary infinite conjunctions): (A2) if As for all s in a set S, then At where t is a conjunction of all members of S. These assumptions are not plausible for an unidealized notion of apriority, but they are reasonably attractive for a sufficiently idealized notion. Given these assumptions (along with other very plausible assumptions), one can establish (A3): if s is κ-uncompletable, it is a priori that s is κ-uncompletable. Given (A3), the argument above establishes

[21]The next three paragraphs can be skipped by those not interested in the technicalities.

(E2*). The argument from (A1) and (A2) to (A3) is in a footnote.[22]

Given (E1*) and (E2*), κ-Plenitude follows. One can likewise make a case for κ-analogs of Actualization, Compositionality, Uniqueness, Parsimony, and Specification, where the sentences in question are restricted to natural language sentences or κ-sentences, and where the scenarios in question are restricted to κ-scenarios.

One might think that κ-scenarios could not satisfy κ-Plenitude for reasons given earlier: it is epistemically possible that there be μ atomic entities, where $\mu > f(\kappa)$, so there will be at least 2^μ μ-scenarios requiring specifications of length μ. That is, the space of μ-scenarios will be more plenitudinous than the space of κ-scenarios, so κ-Plenitude will be false of the latter. But here, the sentences that are not verified by any κ-scenarios are too large to be κ-conjunctions, so (E2*) is not violated, and are too large to be sentences of natural language, so κ-Plenitude is not violated.

Of course the space of κ-scenarios does not exhaust epistemic space. In $L(\kappa)$, most of the 2^μ epistemic possibilities above cannot be fully specified. For many κ-scenarios w, a specification of w will be verified by specifications of multiple μ-scenarios, so that specifications of κ-scenarios are not epistemically complete in an absolute sense. But they do not need to be epistemically complete in such a sense for our purposes. For these purposes, κ-completeness suffices.

We can illustrate the situation (and also illustrate the case for (E2*)) by considering the space $\mathbb{R}^\mathbb{R}$ of functions from real numbers to real numbers, and by considering ω-sentences and ω-conjunctions characterizing such functions in mathematical language.[23] For example "$\forall x(F(x) = x)$" specifies one such function. Many such functions can be specified uniquely with a ω-conjunction: for example, any continuous function can be specified by specifying its values on the rational

[22]Let us say that Kd holds when d is κ-complete, i.e. when for all κ-sentences s, $A\neg(d\&s)$ or $A\neg(d\&\neg s)$. Suppose that s is κ-uncompletable: that is, there is no d such that Kd and $A(d \rightarrow s)$. Principle (A1) then implies $A(\neg A(d_1 \rightarrow s)\&\neg A(d_2 \rightarrow s)\&\ldots)$, where d_1, d_2, and so on are all the κ-complete sentences. Now, *if* it is a priori that $d_1, d_2, \ldots$ include all the κ-conjunctions, it follows that it is a priori that s is κ-uncompletable, as required. The antecedent of this conditional follows from two theses: (A4) given a list of κ-conjunctions, it is a priori that these are all the κ-conjunctions, and (A5) given a κ-conjunction d, it is apriori whether d is κ-complete (that is, either AKd or $A\neg Kd$). (A4) is a very plausible assumption. To make the case for (A5), first suppose that Kd. Then there will be a subset T' of the set T of κ-sentences such that $A\neg(d\&s)$ for all $s \in T'$, and such that $A\neg(d\&\neg s)$ for all s in T'', the complement of T' in T. It follows from (A2) that the conjunction of all sentences $\neg(d\&s)$ (for s in T') and $\neg(d\&\neg s)$ (for all s in T'') is itself a priori. Combined with the very plausible assumption (A6) that it is a priori that these d are all the κ-sentences, it follows that Kd is a priori. So if Kd, then AKd. Similarly but more straightforwardly, by appealing to principle (A1), we can establish that if $\neg Kd$, then $A\neg Kd$. So (A5) is true. Putting all this together, assumptions (A1), (A2), and (A6) establish (A5), and these together with (A4) establish (A3).

[23]Thanks to Brian Weatherson for suggesting this case as an illustration.

numbers (using a series of successive approximations for each value) and by specifying that it is continuous. Not all such functions are uniquely specifiable by an ω-conjunction, though, as there are c^c members of $\mathbb{R}^{\mathbb{R}}$ (where c is the cardinality of the real numbers) but only c ω-conjunctions. Still, for any such function f, we can take $D(f)$ to be the conjunction of all ω-sentences satisfied by f. Then $D(f)$ will be ω-complete. The ω-scenarios here correspond to equivalence classes of functions, where two functions are equivalent if there is no finite sentence that is true of one but not the other. Some ω-scenarios will correspond to a single function, and some will correspond to many functions that cannot be distinguished using finite sentences or infinite conjunctions thereof, although of course they might be distinguished using other infinite sentences.[24] Any epistemically possible ω-sentence will be satisfied by some function and will be verified by the corresponding ω-scenario, as (E2*) requires.

In fact, for our main explanatory purposes, which involve the epistemic possibility of sentence tokens in English, all of which are finite, it will probably suffice to invoke the space of ω-scenarios, and ω-intensions defined over this space. Of course this space will collapse certain scenarios that would otherwise have been held distinct, and it will not make maximally fine-grained distinctions between possible thoughts, but the distinctions it does not capture will be distinctions that we cannot express or even entertain. Unlike a ramified type-theoretic construction, this construction will have no problem handling sentences about propositions, and no problem handling sentences describing arbitrarily large universes, as long as the sentences themselves are finite: if a sentence like this is epistemically possible, it will be true at some ω-scenario. So for the purpose of assigning semantic values to English sentences, and contents to the thoughts of finite thinkers, the space of ω-scenarios may well suffice.

It is perhaps desirable we have scenarios rich enough to fully specify the actual world. This requires the empirical claim that there is a true ω-conjunction fully specifying the actual world (one that is epistemically complete, not just ω-complete). This thesis appears to be reasonably plausible: at least, contemporary physics seems to invoke only separable spaces (spaces with a dense countable subset) and continuous functions between these spaces, and these entities can be specified with ω-conjunctions.[25] Of course if the actual world cannot be described in this way, we can move to a larger cardinality, but otherwise the space of ω-scenarios will be rich enough for

[24] For example, any two members of $\mathbb{R}^{\mathbb{R}}$ can be distinguished by specifying an ordered pair of real values that belongs to the first but not the second. This cannot be achieved in general by using countable conjunctions of finite sentences, but it could be achieved using other infinite sentences, such as quantified countable conjunctions.

[25] Thanks to Marcus Hutter for discussion here.

our purposes.[26]

That said, once we move beyond natural-language semantics, the stratified approach does not give us everything that we might ask for. For a start, it can happen that infinite sentences with nonvague vocabulary nevertheless have indeterminate truth-values at some κ-scenarios. More importantly, there will be many infinite sentences that are epistemically possible, but whose ω-intension is not true at any ω-scenario. (A sentence specifying a particular uncomputable function in R^R may have no truth-value at the corresponding ω-scenario, while being false at all the others.) So a version of ω-Plenitude that applies to all possible sentences will be false. The same goes for other κ-Plenitude theses. So it is worth examining whether we can recover something closer to the original picture with scenarios, intensions, and a plenitude thesis that applies to all possible sentences. I will approach this matter by first considering how the framework will deal with Kaplan's paradox.

Recall that Kaplan's paradox turns on the claims that (i) there is a proposition for every set of worlds, (ii) there is a world for every proposition, and (iii) there are more sets of worlds than worlds. If we replace worlds here by κ-scenarios, and propositions by κ-intensions (sets of κ-scenarios, or better, functions from κ-scenarios to truth-values), the prima facie case for (ii) is removed. For most κ-intensions, there will be no κ-scenario in which that intension is entertained, as specifying such a scenario would require more than a κ-conjunction.

Now, one could stipulate that a scenario (*simpliciter*) is an entity that is a κ-scenario for some cardinal κ. Or perhaps better, we can remove coarse-grained κ-scenarios by requiring that a scenario is a *complete* κ-scenario, one specified by a κ-conjunction s that is not just κ-complete but epistemically complete (there is no sentence T, even of length longer than κ, such that $s\&t$ and $s\&\neg t$ are both epistemically possible). To attempt to generate the paradox, we might also stipulate that an intension (*simpliciter*) is an entity that is a κ-intension for some cardinal κ. These stipulations have the slightly awkward consequences that not every κ-scenario is a scenario *simpliciter* and that intensions *simpliciter* are not defined over scenarios *simpliciter*, but we can at least assess

[26]One tricky issue is that specifying the actual world plausibly requires some sort of "that's-all" clause, saying roughly that the world does not contain anything beyond what is specified or implied by the rest of the description. As usually construed, this "that's-all" sentence will be infinite (at least if the world is infinite), and so will not be an ω-sentence. For the actual world, we can avoid the problem by making the case that there are finite sentences that can do the work of the that's-all clause. But an issue like this arises with other scenarios, and there may be some for which a scenario that can otherwise be specified using an ω-conjunction require a longer that's-all clause. One could simply deny that these scenarios are ω-scenarios, but alternatively one could modify the definition of an ω-scenario (and a κ-scenario more generally) to allow a single longer that's-all clause to be included.

the elements of Kaplan's paradox where they are concerned.

Under these stipulations, (ii) will plausibly be true: for every κ-intension, there is an epistemically possible μ-scenario (for some $\mu > \kappa$) that verifies a sentence saying that someone is entertaining that intension. It is tempting to say that (i) is false, on the grounds that there can be sets of scenarios with unbounded cardinality, which will not correspond to κ-intensions for any cardinality κ. But this is not quite right. Just as there are no sets of sets with unbounded cardinality (for any sets of sets, the cardinality of their union will serve as an upper bound), there are no sets of scenarios with unbounded cardinality. For any set of scenarios, there will be a corresponding set of cardinals (for each scenario, this will be the least cardinal κ such that the scenario is a κ-scenario), with upper bound μ. Then this set of scenarios will correspond to a μ-intension. So (i) will be true, also.

Instead, the right thing to say is that where scenarios are concerned, (iii) is false. Just as there are not more sets of sets than sets, there are not more sets of scenarios than scenarios. Given that there is no sets of all sets, the sets do not have a cardinality, and one cannot form a power set, so Cantor's theorem does not get off the ground. Likewise, if there is no set of all scenarios, then Cantor's theorem does not get off the ground. We can take the moral of Kaplan's paradox to be that there is no cardinal upper bound on the size of a scenario, so that there are too many scenarios to form a set.

It is tempting at this point to appeal to proper classes, holding that there is at least a *class* of all scenarios, and that there will be more classes of scenarios than scenarios. But as in other domains, the appeal to classes does not change anything fundamental. If we understand classes as analogous to sets (just larger), we could go on to define ultra-intensions (classes of scenarios, or functions from scenarios to truth-values) and ultra-scenarios (scenarios constructed using proper-class size conjunctions), and will find ourselves in the same situation. If the paradox is cast in terms of ultra-intensions, ordinary scenarios, and classes, (ii) will be false; if it is cast in terms of ultra-intensions, ultra-scenarios, and classes, (iii) will be false; if it is cast in terms of ultra-intensions, ultra-scenarios, and a more general notion of collection, (i) will be false. We could iterate further to metaclasses and so on, but nothing will change. Even if we try to cast the paradox in terms of a general notion of collection and a corresponding general notion of scenario (assuming these notions are coherent), we will encounter the same issues that we encountered with sets: just as there is no collection of all collections, there is no collection of all scenarios (in this putatively general sense), and so on. I think it is much more straightforward to set aside proper classes, holding (with Boolos 1998, Shapiro 1993, and others) that all collections are sets, and drawing the morals that one would have to draw at the level of collections in any case at the level of sets.

At this point in the set-theoretic case, various theorists hold either (a) that although we can quantify in an absolutely unrestricted way over sets (or collections), there is no set (or collection) corresponding to the domain of quantification here (e.g. Boolos 1998, Cartwright 1994), (b) that we cannot quantify in an absolutely unrestricted way over sets (or collections), and can instead quantify over them in various restricted but indefinitely extensible ways (e.g. Dummett 1993, Fine 2008), or (c) that we do not have a single notion of set (or collection), but instead have an indefinitely extensible hierarchy of notions of set-like entities. All of these options are available in the case of scenarios. The third option is perhaps the least popular in the set-theoretic case, and I will set it aside in the case of scenarios. But analogs of options (a) and (b) are both open here, and I will not try to choose between them.[27]

A consequence of there being no set (collection) of all scenarios is that some problems arise with the use of intensions. In particular, all intensions as understood so far will be κ-intensions for some κ. Any set of scenarios *simpliciter* is a set of κ-scenarios for some κ, so if intensions are understood as sets of scenarios *simpliciter*, any intension will be a set of κ-scenarios. On the standard understanding, if such an intension were to be evaluated at a μ-scenario for $\mu > \kappa$ (where this μ-scenario is not itself a κ-scenario), its value would be false there, as the μ-scenario is not included in the intension. This has the odd consequence that if we attempt to understand these intensions as intensions that can be evaluated at arbitrary scenarios, all intensions will be false at all sufficiently large scenarios, and there will be no sensible way of negating intensions. On the other hand, if we understand intensions as functions from scenarios to truth-values, where functions are understood in the usual way as sets of ordered pairs, then κ-intensions will not be defined at μ-scenarios, and no intension will have a truth-value at all scenarios. Nevertheless, there is something intuitive about the idea of an intension that is true at all scenarios, or one that is true at a scenario if it contains particles, and so on.

2. The non-set-theoretic understanding of intensions.

We have seen that there are too many scenarios to form a set, and that this raises problems with the understanding of intensions as sets of scenarios. It also raises problems with the understanding

[27] If we take option (b), the many sentences in this paper at which I appear to quantify over all scenarios need to be reinterpreted. Most such sentences in earlier sections are still coherent if understood as quantifying over κ-scenarios, for some κ. This will not work for some sentences in this section (e.g. those concerning the impossibility of quantifying over all scenarios), for reasons familiar from the set-theoretic case, but I think that these claims can be reinterpreted by more complex means, for example using the dialectical strategy suggesed by Fine (2008).

of intensions as functions from scenarios to truth-values, where functions are understood as sets of ordered pairs. At this point, an alternative strategy understands intensions as functions as mappings in a sense that does not require understanding them as sets of ordered pairs. This strategy is familiar from the set theoretic case, in which the axiom of replacement in effect understands functions in terms of formulae that define them. Given any set, such a formula returns another set. Even though there is no set of ordered pairs to yield a set-theoretic entity that counts as a function in the official sense, one can still see such formulae as determining mappings from sets to sets in an intuitive sense. For example, there is certainly a mapping that maps any set S to the set S^* that contains S as its only member.

Something similar applies in the case of intensions. On the current picture, one can certainly evaluate any sentence at any κ-scenario, yielding a truth-value. So one can evaluate any sentence at any scenario (simpliciter), yielding a truth-value. This is in effect to say that we have a coherent understanding of the notion of verification of a sentence s by an arbitrary scenario w, which we can represent as usual with the locution $ver(w, s)$. Of course ver does not correspond to a set of ordered pairs (or ordered triples), but it is still a mapping from ordered pairs to truth-values in the intuitive sense. Likewise, we can understand the mapping ver_s which maps an arbitrary scenario w to $ver(w, s)$. Again, this mapping cannot be represented as a set of ordered pairs, but it is still a well-defined mapping, just as the mappings from sets to sets defined by formulae of set theory are well-defined mappings. We can think of this mapping as the intension of s.

Of course functions in this sense cannot be identified with set-theoretic objects. But for most purposes, the absence of such an object does not matter. As long we have given sense to the notion of verification of a sentence by an arbitrary scenario, this is all we need for the core aspects of the current framework. In particular, the notion of verification (along with the notions of epistemic possibility, scenarios, sentences, and actualization) is all we need in order to satisfy the core principles of epistemic space. It is possible that functions so understood can be modeled using the tools of nonstandard versions of set theory, such as Fine's (2005) theory of classes and Linnebo's (2008) theory of properties, on which the relevant entities are individuated by defining formulae rather than by members. But all that matters for our purposes is that talk of functions in this sense is coherent.

Does this intuitive notion of a mapping give rise to Kaplan's paradox once again? One might think that there will be a mapping for every class of scenarios, whether or not that class forms a set: for example, the mapping that maps all scenarios to "true" will correspond to the class of all scenarios. And one might think that there will be a possible thinker for every such mapping.

But the first claim is false. The intuitive notion of function invoked in the axiom of replacement does not yield a function for every class of sets (even though there is a function in this sense, corresponding to a predicate, that maps every set to true), but just a function for every formula. Likewise, the notion of mapping used here does not yield a function for every class of scenarios, but just one for every sentence. It may be that there is a sentence for every set of scenarios, but there is no reason to think that there is a sentence for every class of scenarios, at least in any sense in which there are more classes than sets. As before, it is probably best to avoid talk of classes here, except in an intuitive sense in which classes can be glossed in terms of predicates, analogous to the intuitive sense in which functions can be glossed in terms of formulae. Once we do this, it is clear that there is no reason to think that there are more classes than sets, and the threat of paradox dissolves.

It is even possible to introduce a special sort of abstract object corresponding to these intensions. Of course these abstract objects cannot be sets of ordered pairs. But we might think of an intension formally as an abstract object which when combined with an arbitrary scenario yields a truth-value (or an extension). Then every sentence will yield an intension in this formal sense. One might leave the further nature of intensions unspecified, perhaps invoking an analog of Bealer's (1982) algebraic conception of propositions, according to which propositions are characterized by how they behave under various logical operations but are metaphysically simple. Or one might develop a theory of these non-set-theoretic objects, perhaps along the lines of Fine's theory of classes or Linnebo's theory of properties, on which these entities, individuated by defining formulae, are adjoined to the standard set-theoretic ontology. Importantly, on any of these approaches, there is no reason to believe that there are more intensions than scenarios. This understanding will satisfy all the central principles of epistemic space, and is consistent with elements (i) and (ii) of Kaplan's paradox, along with the falsity of (iii). This is perhaps the closest we can come to recapturing the original framework of sentences, scenarios, and intensions.

What about Kaplan's original paradox, applied not to scenarios but to metaphysically possible worlds? Of course if Metaphysical Plenitude is true, what I say here about scenarios will also apply to metaphysically possible worlds. Even if Metaphysical Plenitude is false, the same framework might apply. As long as there are no restrictions on how many atomic entities could exist, or on what propositions could be believed, then a standard space of possible worlds will generate Kaplan's paradox. One might initially adopt a stratified construction along the lines above. We will have κ-worlds for various κ, corresponding to equivalence classes of κ-conjunctions that are κ-complete (in a sense defined in terms of metaphysical possibility), using an appropriate lexicon (perhaps including expressions for fundamental properties along with logical expressions and the like). For many or most applications of possible-worlds semantics, taking worlds to be κ-worlds for an appropriate κ should suffice, and taking intensions to be κ-intensions, should suffice.[28]

For a broader framework of worlds and intensions simpliciter, we can understand a world as an entity that is a complete κ-world for some κ. Of course there will be no set of all worlds, and no functions in the set-theoretic sense mapping arbitrary worlds to truth-values. But we can still evaluate sentences at arbitrary worlds, yielding truth-values, and corresponding to this sort of evaluation, one will have intensions in the intuitive sense above. Worlds and intensions of this sort can arguably fill most of the explanatory purposes to which possible worlds have been put.

10 Non-ideal epistemic space[29]

The notion of (deep) epistemic possibility that we have been dealing with is an idealized one: if s is a priori, then $\neg s$ is not epistemically possible, even when s is far from obvious, and even when no-one in the world knows that s. But in the ordinary sense of epistemic possibility, $\neg s$ is often epistemically possible even when s is knowable a priori. It is natural to wonder if there is a less idealized notion of deep epistemic possibility that might be useful in modeling less idealized sorts of reasoners.

To develop such a conception, we must start with a *non-ideal* notion of deep epistemic possibility. Instead of saying that s is epistemically possible when $\neg s$ cannot be ruled out a priori, we might say that s is epistemically possible when $\neg s$ cannot be ruled out *through reasoning of a certain sort*. Equivalently, we can say that s is epistemically necessary when s can be established through reasoning of a certain sort. Here, there are various options.

For example, we might hold that it is (non-ideally) epistemically necessary that s when:

(i) it is obvious a priori that $\neg p$;

(ii) s can be known through such-and-such amount of a priori reasoning;

[28] It is arguable that κ-worlds are less apt for applications of possible worlds in metaphysics than for applications in semantics. I am inclined to think that possible worlds play a less essential role in metaphysical explanation than in semantic explanation: usually the relevant work can be done by the notions of possibility and necessity instead. But if necessary, the broader framework is available.

[29] Jens Christian Bjerring's ANU Ph.D. thesis (2010) is devoted to the analysis of non-ideal versions of epistemic possibility and epistemic space. I am indebted to Bjerring in what follows. For some related ideas, see Hintikka 1975, Rantala 1975 and Jago 2006.

(iii) *s* can be proved through logical reasoning alone;

(iv) *s* can be proved in *n* steps of logical reasoning

(v) *s* can be established through nonmoral a priori reasoning;

(vi) *s* is cognitively insignificant

Given a notion of non-ideal epistemic possibility, we can attempt to set up a corresponding *non-ideal* epistemic space, made up of *non-ideal* scenarios. The principles governing this space will be much as before. The key principle, once again, will be Plenitude: there is a scenario verifying *s* iff *s* is epistemically possible. Because many more sentences will be epistemically possible for non-ideal notions of epistemic possibility, it follows that there will be many more corresponding non-ideal scenarios.

It seems reasonable that the Actualization principle should hold on this model, but there is some question about whether Compositionality should be endorsed. For example, one may wish to allow non-ideal scenarios that verify *s* and *t*, without verifying *s&t*. More generally, if a strong version of Compositionality holds, it is likely that if a scenario verifies some statements, it will verify all logical consequences of those statements. This will be undesirable in modeling many forms of non-ideal reasoning. If so, we may wish to do without Compositionality, or restrict it in some fashion.

The process of constructing scenarios will be more complex where non-ideal epistemic possibility is concerned. It is clear that taking scenarios to be centered worlds will lead to a failure of Plenitude: for example, a priori falsehoods are likely to be verified by no centered world. On the epistemic construction, we may need to avoid appealing to epistemically complete sentences, as these sentences are so long that they may have no interesting non-ideal epistemic properties. Instead, it may be best to appeal to *classes* of sentences: perhaps classes such that no sentence in the class is epistemically impossible, or perhaps classes such that no conjunction of sentences in the class is epistemically impossible.[30]

One could define what it is for a class *c* to verify a sentence *s* in a variety of ways: perhaps if

[30]As Bjerring establishes, both options here have problems. On the first option, the absence of any joint consistency constraints causes the resulting scenarios to behave in an extremely unconstrained way. On the second option, the joint consistency constraints have the consequences that sentences that can be ruled out only through very long chains of the relevant sort of reasoning will be excluded from all scenarios, even if they are epistemically possible. Perhaps the biggest open problem in the study of non-ideal epistemic space is that of finding a construction of non-ideal scenarios that avoids the Scylla of "anything goes" and the Charybdis of logical omniscience.

c includes s, or perhaps if some conjunction of $\neg s$ with a subset of c is epistemically impossible, or perhaps if there is an reasoning process of the relevant sort that takes us from a subset of c to s. We can say that one class verifies another class if it verifies every sentence in that class. We can say that a class is maximal if it is verified by no class that it does not verify. There will be difficulties in setting up equivalence relations on maximal classes, due to failures of transitivity in implication, but this problem might be dealt with in a variety of ways. It seems that this sort of approach at least holds some promise.

If we can set up a non-ideal epistemic space corresponding to a non-ideal notion of epistemic possibility, we will then have a corresponding non-ideal epistemic intension. We can say that the non-ideal epistemic intension of a sentence is the sentence's intension over non-ideal scenarios, according to whether those scenarios verify the sentence. Then for any two sentences s_1 and s_2 such that it is epistemically possible that s_1 holds without s_2 and vice versa, s_1 and s_2 will have different non-ideal intensions.

When this way of thinking is applied to different notions of epistemic possibility, it will yield various different applications. For example, if we are concerned with Frege's notion of cognitive significance, we can say that t is epistemically possible when $\neg t$ is cognitively significant (perhaps this will be whenever $\neg t$ is nontrivial), and we can set up a corresponding non-ideal epistemic space. This will yield a variety of non-ideal intension that, although unstructured, is as fine-grained as a Fregean sense.

It is likely that there is no canonical notion of non-ideal epistemic possibility. If so, there will be no canonical notion of non-ideal content. Instead, we might have a spectrum of notions of deep epistemic possibility, from the ideal to the non-ideal, perhaps ending at the notion on which anything is epistemically possible and on which contents are trivial. There will be a corresponding spectrum of epistemic spaces. Every sentence might then be associated with a spectrum of epistemic intensions, each of which is an intensions across scenarios within a given epistemic space. For different purposes, different intensions from within this spectrum may be relevant. Between these intensions and these epistemic spaces, there will be enough material to do significant explanatory work in many different epistemic domains.

11 Applications

I will end by briefly spelling out some applications of the notion of epistemic space, as outlined here.[31]

First, there are applications to the analysis of meaning and content. In many ways, epistemic intensions behave like a broadly Fregean sort of meaning. For example, two singular terms a and b have the same epistemic intension iff '$a = b$' is epistemically necessary. This is quite reminiscent of the Fregean thesis that a and b have the same sense iff '$a = b$' is cognitively insignificant. For example, tokens of 'Hesperus is Phosphorus' are not epistemically necessary, and correspondingly the tokens of 'Hesperus' and 'Phosphorus' are associated with different epistemic intensions.

The main difference between epistemic intensions and Fregean senses arises from the fact that at least on the idealized version of epistemic necessity, epistemic necessity does not imply cognitive insignificance (though the reverse implication plausibly holds). For example, '7+3 = 10' is epistemically necessary but cognitively significant, so '7+3' and '10' will have the same epistemic intension where they have different Fregean senses. Still, epistemic intensions can serve here as at least a coarse-grained sort of Fregean sense.

The current framework can be extended in order to provide finer-grained senses. One way to extend the framework is to invoke structured epistemic intensions, so that complex expressions are associated with complexes constituted from the epistemic intensions of their parts. Then '7+3' is associated with a structured epistemic intension quite different from that of '10', and so on. Another way to proceed is to start with a nonidealized notion of deep epistemic necessity. For example, if one adopts understanding (v) above, on which deep epistemic necessity is understood as cognitive insignificance, then if we can make sense of a corresponding nonideal epistemic space, we can expect that a and b will have the same nonideal epistemic intension if $a = b$ is cognitively insignificant, just as the Fregean framework requires.

One can also apply this framework to the contents of thought, yielding a variety of content such that different modes of presentation of the same referent—*Hesperus* and *Phosphorus*, say—are associated with different contents. This sort of content will plausibly behave like a sort of cognitive content, not constitutively tied to reference. And under certain plausible assumptions, it will behave like a sort of narrow content, so that the contents of a subject's thoughts do not constitutively depend on the character of a subject's environment.

[31] For more on some of these applications, see my 2002a (Fregean sense), 2002b (narrow content), forthcoming a (Fregean sense and attitude ascriptions), and forthcoming b (probability).

The idealized sort of epistemic space can also be applied to the analysis of subjective probability, providing a candidate for the space of entities over which the subjective probabilities of an idealized agent are distributed. It is also not out of the question that one might be able to use a non-ideal epistemic space to model the entities over which the subjective probabilities of a non-ideal agent are distributed.

Nonideal epistemic spaces may also be useful in analyzing various specific domains, such as the moral domain. We may think that the connection between the nonmoral and the moral is ultimately a priori, or we may think that moral beliefs are ultimately not truth-evaluable, but as long as the connection and the non-truth-evaluability is not obvious, there will be an interesting hypothesis space to investigate. To do this, we can invoke a notion of deep epistemic possibility along the lines of notion (iv) above: it is epistemically possible that p when p cannot be ruled out through nonmoral a priori reasoning. This will plausible yield a space of "moral scenarios" which is much like the space of ideal scenarios, except that it may have an additional dimension of variation in the way that it associates moral claims with nonmoral claims. These moral scenarios (which are reminiscent of the "factual-normative" worlds of Gibbard 1990) may have some use in analyzing moral thought and discourse without presupposing substantive moral views.

Finally, epistemic space may be useful in giving an account of the semantics of various ordinary language constructions. Elsewhere, I have discussed how intensions of the sort discussed here may be useful in analyzing attitude ascriptions and indicative conditionals. Closer to home, they may have some use in understanding ascriptions of epistemic possibility in the ordinary sense, and in understanding about what might or might not be the case, for all one knows.

For the ordinary notion of strict epistemic possibility, it is plausible that p is epistemically possible when one could not *easily* come to know that $\neg p$ given what one already knows. The corresponding notion of deep epistemic possibility is something like the following: it is deeply epistemically possible that p when $\neg p$ is not easily knowable a priori. From this notion, we will be able to set up a corresponding non-ideal epistemic space. For this space, we can then say that p is strictly epistemically possible for a subject iff there is a p-scenario that is not excluded by any item of that subject's knowledge.

We can apply this framework to utterances involving epistemic modals, such as 'It might be the case that *s*'. According to a natural view, such an utterance is true iff there is a scenario that verifies *s* and that is epistemically possible for the speaker. This view involves a contextualist treatment of epistemic modals, where the standards of epistemic possibility are set by the context of utterance. If one instead adopts a relativist treatment of epistemic modals, where the standards of epistemic possibility are set by a context of assessment, one can instead say that 'It might be the case that *s*' is true (at a context of assessment) iff there is a scenario that verifies *s* and that is epistemically possible for the subject in the context of assessment. Other treatments of epistemic modals (such as those on which epistemic modals are not assessible for truth, but merely for acceptability), can also be combined with the present framework. In this way, we can use the framework of epistemic space to help shed light on the ordinary claims about epistemic possibility with which this paper began.

References

Ackermann 1978. De re propositional attitudes toward integers. *Southwestern Journal of Philosophy* 9:145-53.

Anderson, C.A. 2009. The lesson of Kaplan's paradox about possible world semantics. In (J. Almog and P. Leonardi, eds) *The Philosophy of David Kaplan*. Oxford University Press.

Austin, D.F. 1990. *What's the Meaning of "This"?* Ithaca, NY: Cornell University Press.

Bealer, G. 1982. *Quality and Concept*. Oxford University Press.

Bealer, G. 1996. A priori knowledge and the scope of philosophy. Philosophical Studies 81:121-42.

Bjerring, J.C. 2010. *Non-Ideal Epistemic Spaces*. Ph.D. thesis, Australian National University.

Boolos, G. 1998. *Logic, Logic, and Logic*. Harvard University Press.

Burgess, J. 1997. Quinus ab omni naevo vindicatus. *Canadian Journal of Philosophy*, *Supplement* 23:25-66.

Cartwright, R.L. 1994. Speaking of everything. *Nous* 28:1-20.

Chalmers, D.J. 2002a. On sense and intension. *Philosophical Perspectives* 16:135-82.

Chalmers, D.J. 2002b. The components of content (revised version). In *Philosophy of Mind: Classical and Contemporary Readings*. Oxford University Press.

Chalmers, D.J. 2002c. Does conceivability entail possibility? In (T. Gendler & J. Hawthorne, eds) *Conceivability and Possibility*. Oxford University Press.

Chalmers, D.J. 2004. Epistemic two-dimensional semantics. *Philosophical Studies* 118:153-226.

Chalmers, D.J. forthcoming a. Frege's puzzle and the objects of credence. http://consc.net/papers/credence.pdf.

Chalmers, D.J. forthcoming b. Propositions and attitude ascriptions: A Fregean account. *Nous*. http://consc.net/papers/propositions.pdf.

Chalmers, D.J. & Jackson, F. 2001. Conceptual analysis and reductive explanation. *PhilosophicalReview*.

Cresswell, M. 2006. From modal discourse to possible worlds. Studia Logica 82:307-327.

Dummett, M. 1993. What is mathematics about? In *The Seas of Language*. Oxford University Press.

Fine, K. 2005. Class and membership. *Journal of Philosophy*.

Fine, K. 2008. Relatively unrestricted quantification. In (G. Uzquiano and A. Rayo, eds.) *Absolute Generality*. Oxford University Press.

Gibbard, A. 1990. *Wise Choices, Apt Feelings. A Theory of Normative Judgement*. Oxford University Press.

Hintikka, J. 1975. Impossible possible worlds vindicated. *Journal of Philosophical Logic* 4:475-84.

Jago, M. 2006. Imagine the possibilities: Information without overload. *Logique et Analyse* 49:345-71.

Kaplan, D. 1995. A problem in possible-world semantics. In (W. Sinnott-Armstrong, D. Raffman, and N. Asher, eds.) *Modality, Morality and Belief: Essays in Honor of Ruth Barcan Marcus*. Cambridge University Press.

Kripke, S.A. 1980. *Naming and Necessity*. Harvard University Press.

Kripke, S.A. forthcoming. A puzzle about time and thought. In *Collected Papers, Volume 1*. Oxford University Press.

Lewis, D. 1986. *On the Plurality of Worlds*. Blackwell.

Linnebo, O. 2008. Sets, properties, and unrestricted quantification. In (G. Uzquiano and A. Rayo, eds.) *Absolute Generality*. Oxford University Press.

Rantala, V. 1975. Urn models. *Journal of Philosophical Logic* 4:455-474.

Shapiro, S. 2003. All sets great and small: And I do mean ALL. *Philosophical Perspectives* 17:467490.

Soames, S. 2004. *Reference and Description: The Case Against Two-Dimensionalism*. Princeton University Press.

Whittle, B. 2009. Epistemically possible worlds and propositions. *Nous* 43:265-85.