*Noûs* 43 (2009)

# Epistemically Possible Worlds and Propositions*

Bruno Whittle

For at least the past few decades, philosophers from a wide variety of areas have put the notion of a possible world to many uses. Possible worlds have been used not just to give a useful way of thinking about modality itself, they have also been used to give fruitful philosophical accounts of an enormous range of notions, from causation to conditionals to psychological or linguistic content in general.

However there is more than one philosophically interesting notion of possibility. In addition to metaphysical possibilities—ways the world might have been—there are also epistemic possibilities—ways the world might be, for all we know. For example, it is now widely accepted that however things might have been, water would still be $H_2O$: that is, it is metaphysically necessary that water is $H_2O$. However, for all people in the eighteenth century knew, water was not $H_2O$: that is, it was epistemically possible that water was not $H_2O$.

When philosophers talk about possible worlds they tend to mean metaphysically possible worlds: something like metaphysically possible ways the world might have been. However it is natural to think that some notion of an epistemically possible world might be capable of playing an analogous and similarly fruitful role. This might seem particularly likely in situations in which one is interested in cognitive rather than metaphysical issues: e.g. in giving philosophical accounts of cognitive rather than metaphysical notions. And various philosophers, most notably David Chalmers, have recently started putting the notion of an epistemically possible world to just this sort of work: Chalmers argues that the notion can provide fruitful accounts of, among other things, Fregean sense, narrow content and conditionals.[1]

---

* I am grateful to David Chalmers, Ian Rumfitt, Timothy Williamson and audiences at Oxford and St Andrews for comments and discussion.

[1] See Chalmers [2002a], [2002b], [2006a], [2006b] and [forthcoming].

Indeed, something like a notion of an epistemically possible world would appear already to be present in many discussions in epistemology: it is natural and common to think of knowledge in terms of eliminating possibilities of some sort, and the notion of a 'sceptical scenario' which, the sceptic contends, we cannot eliminate, is common in presentations of scepticism. However, these 'epistemic scenarios' that come up in epistemology cannot simply be thought of as possible worlds of the standard variety, i.e. as metaphysically possible worlds. For example, when we learnt that water is $H_2O$, we eliminated a possibility in the epistemic sense, but not in the metaphysical sense.[2]

So there are a number of reasons why philosophers should be interested in the notion of an epistemically possible world. In §1 below I explain why philosophers who are interested in propositions and propositional attitudes should be particularly interested in the notion.

Unfortunately, however, this apparent philosophers' paradise it not quite as good as it seems. And the main purpose of this paper is to argue that there is a serious threat to this paradise from something like the paradox that was the downfall of (so-called) naive set theory: Russell's paradox.

The structure of the paper is as follows. In §1 I do a bit more to explain why one should be interested in the notion of an epistemically possible world. In §2 I go through some preliminaries and set out the methodology of the paper. In §3 I give what seems to be in many ways the most obvious and natural account of epistemically possible worlds: however I argue that this account must be wrong. In §4 I then consider the natural way in which one might try to modify this account, and I argue that this natural modification does not succeed in rescuing the account. In §5 I compare the problems raised in §§3–4 for accounts of epistemically possible worlds to Kaplan's 'problem in [metaphysically-]possible-world semantics' (Kaplan [1995]). I argue that Kaplan's objection can be rebutted (essentially as in Lewis [1986]), but that no analogous response is available to the problems of §§3–4. In §6 I consider how one might give an account of epistemically possible worlds that avoids the problems of §§3–4: I argue that although there may be such accounts with fruitful

[2] These ways in which epistemologists already make use of something like epistemically possible worlds are pointed out in Chalmers [forthcoming].

applications, any such an account must fundamentally compromise the basic idea behind epistemic possibility.[3]

## 1. Background and Motivation

In this section I explain how an interest in propositional attitudes and propositions leads naturally into an interest in epistemically possible worlds. Most accounts of propositions think of them as being structured in a way that more or less mirrors the structures of the sentences that express them: this is true of, for example, the Fregean or Russellian accounts of propositions. For example, according to the Russellian account, the proposition that Italy borders France is made up of three constituents, the country Italy, the relation of bordering, and the country France. However, such accounts of propositions are susceptible to a version of Russell's paradox. This can be seen as follows: in discussing the paradox I think of propositions as Russellian, but similar paradoxes can be derived for alternative accounts of 'structured' propositions (i.e. accounts according to which propositions have sentence-like structures). First, let R be a property that something has just in case it is a proposition of the form F(a), where the property F does not itself apply to the proposition F(a). That is, suppose that the following holds.

(R1) For any x, R applies to x iff x is a proposition of the form F(a) such that F does not apply to F(a).

It is natural to think that there must be some such property R: after all, I have just spelled out exactly what it would take for something to have this property.

(R2) unpacks how R works for the case in which x is a proposition of the form F(a).

(R2) For any proposition of the form F(a), R applies to F(a) iff F does not apply to F(a).

---

[3] The accounts of epistemically possible worlds argued against in §§3 and 4 are similar to two of the main accounts proposed in Chalmers [forthcoming]. Arguments similar to those of §§3–4 apply against these accounts. The other main account considered in Chalmers [forthcoming] uses metaphysically possible worlds to, in a certain way, stand proxy for epistemically possible worlds: an argument similar to that of §4 also applies against this account. The main reason that I have refrained from explicitly discussing Chalmers's accounts is that, although [forthcoming] has been presented as a talk at a number of places, it is unpublished and is yet to reach its final version.

Substituting a proposition R(b), for some object b, into (R2), then gives the following contradiction.

(R3) R applies to R(b) iff R does not apply to R(b).

Thus, if one wants to hold onto the notion of a structured proposition, one has to place certain restrictions upon which structured propositions there are. The natural and common response is to stratify propositions into some sort of hierarchy along something like the following lines. One begins with some things that are not propositions or properties at level 0. Then, at level 1 one has properties and propositions about the things at level 0. At level 2 one then has a new lot of properties and propositions that are about the things at level 1, and so on for levels 3, 4 etc. At each level one has a completely new batch of properties and propositions. This blocks the paradox that I gave because it relied on a characterization of R ((R1)) according to which R can apply to propositions of the form F(a) regardless of where they come in such a hierarchy: but this is exactly the sort of thing that is ruled out by the sort of hierarchical account described.[4]

Thus stratifications of this sort do block the paradox. However they do so only at the cost of greatly restricting what one can say. According to such hierarchical accounts there can be no properties that can apply to propositions at any level of the hierarchy. But intuitively the properties of being known or of being expressed or of being true are just such properties. And so according to such a hierarchical account of propositions one cannot express general claims about propositions: such as, for example, 'Knowledge implies truth', i.e. the claim that for any proposition P, if P is known, then P is true. But such general claims seem both otherwise unproblematic and essential to philosophy, and so this might seem a price too high to be worth paying. One might thus look to try to give alternative accounts of propositions.

One alternative that it is natural to propose at this point is to identify propositions with sets or classes of possible worlds of some sort: such an account of propositions would identify a proposition with the class of possible worlds at which it is true. And accounts of propositions that identify them with classes of metaphysically possible worlds have been proposed (from now on I use MPW for metaphysically possible world).[5] Accounts of

---

[4] For a hierarchical account of propositions along something like these lines see Russell [1908].

[5] See for example Stalnaker [1978] or Lewis [1986: 53–55]. Stalnaker [1978: 79, note 2] attributes the identification of propositions with classes of MPWs to Saul Kripke in the early 1960s. Following Lewis [1979] it is also common to identify propositions with classes of centred MPWs: a centred MPW is a pair ⟨w,c⟩ of an

propositions as classes of MPWs block the paradox that I raised because that paradox relied upon propositions being made up in some unique way out of a certain specific lot of properties and objects. For the characterization of R, i.e. (R1), to be adequate, given any proposition P of the form F(a) for some property F, this proposition P presumably cannot also be of the form G(b) for some non-coextensive property G: otherwise it would be possible for F to apply to P, but for G not to apply to P, in which case the characterization would not have fixed whether or not R should itself apply to P. But this is in a sense what happens if one identifies propositions with classes of MPWs. If F is some property that necessarily applies to some necessary existent a, and G is some other property that necessarily applies to another necessary existent b, then the proposition F(a) is identical to the proposition G(b): they are both the proposition that is true at all MPWs, i.e. the class of all MPWs.[6]

However, exactly that aspect of the account of propositions as classes of MPWs that blocks the paradox is also the account's most problematic feature. For the fact that there is, on this account, only one necessary proposition means that, for example, the proposition that water is $H_2O$ is the same as the proposition that tigers are animals. This coarse-grainedness then leads to at the very least prima facie problems with these propositions playing the role of the objects of propositional attitudes. For it seems intuitively clear that someone can believe that water is $H_2O$ without believing that tigers are animals.

---

MPW w and a 'centre' c, where c is either a space-time point, or a pair of a subject and a time. The account of propositions as classes of centred MPWs was proposed in Lewis [1979] to deal with problems involving propositional attitudes about one's location. For simplicity I for the time being consider only accounts of propositions as classes of possible worlds; however I return to accounts of propositions as classes of centred worlds in §4.

Another popular theoretical move within a possible worlds framework is to associate with a sentence not merely a proposition (i.e. a class of possible worlds or a class of centred possible worlds) but rather a 'two-dimensional instension': a class of ordered pairs of the form ⟨x,w⟩ where x is a centred possible world (or sometimes just a possible world) and w is a possible world (see for example Stalnaker [1978]). Chalmers proposes such an approach employing both epistemically and metaphysically possible worlds (see for example [2006a]). For reasons of space I do not discuss two-dimensional approaches in this paper, however nothing essential would change if they were brought into the picture.

[6] In somewhat more detail, the following is what happens if one tries to reinstate the paradox, in the context o the account of propositions as classes of MPWs. One thinks of a class of MPWs X as being of the 'form' F(a) if, for some property F and object a, a has the property F at precisely the MPWs in X. One then considers a property R' that applies precisely to classes of MPWs X that are of the form F(a) for some property F that does not apply to X. (Alternative attempts at reinstating the paradox, in terms of variants on R', will fail in similar ways.) In this case R' will just unproblematically apply to any class of MPWs X: if X is true (i.e. if the actual world is in X), then let F be the property of being false, and let a be Y, the class of all worlds not in X (Y is false in precisely the worlds in X, so X is of the form F(a) for this F and this a; and since X is true, this F does not apply to X); if X is false, on the other hand, then let F be the property of being true and let a be X. Now consider a class of worlds Z of the form F(a) for this property R' and some object b. R' will unproblematically apply to Z. One will not be able to infer that R' thus does not apply to Z (as in (R3)), because the property F in virtue of which R' applies to Z is not going to be R' itself but rather some other property (e.g. truth or falsity).

Thus it would seem desirable to find some account of propositions that not only avoids the sorts of problems involving a version of Russell's paradox that structured accounts of propositions face, but that also solves at least some of the problems that accounts of propositions as classes of MPWs face as a result of their coarse-grainedness. And an extremely natural suggestion at this juncture is to try to stick to identifying propositions with classes of possible worlds, but to use a somewhat more fine-grained notion of possibility that can distinguish between, for example, the class of possible worlds at which water is $H_2O$ and the class of worlds at which tigers are animals. Thus a natural suggestion is to identify propositions with classes of epistemically possible worlds (I use EPW for epistemically possible world). For there are presumably EPWs where water is $H_2O$ but where tigers are mammals: since one can know the former without knowing the latter. Further any coarse-grainedness that results from this EPW account will stem from facts about a variety of possibility that is closely connected to cognitive notions and less to do merely with facts about the world, so one would expect it to be somewhat more palatable.

Thus, as promised, an interest in propositions leads into an interest in epistemically possible worlds.

## 2. Preliminaries and Methodology

In this section I clarify what I will be talking about, and also how I will go about considering accounts of epistemically possible worlds.

First some remarks about what exactly I mean by an epistemically possible world. I said in the introduction that whereas a metaphysically possible world corresponds to a way the world might have been, an epistemically possible world corresponds to a way the world might be, for all we know. As the discussion has already indicated, the idea is not that EPWs are a subclass of MPWs. Rather the idea is that some things will be epistemically possible that are not metaphysically possible. Exactly which things will depend on the particular account of epistemic possibility (as we will see below). However all of the accounts considered will be such that there are EPWs at which there is water but no $H_2O$, for example; similarly all of the accounts considered will be such that there are EPWs at which Hillary Clinton is George W. Bush's sister.

I initially characterized epistemically possible worlds as corresponding to ways the world might be, for all we know. Thus, in one sense, what is epistemically possible will vary from subject to subject. However in addition to this subject-relative notion of epistemic possibility, there is a subject-independent notion of epistemic possibility that does not vary from subject to subject: a notion according to which something is epistemically possible just in case it is a way the world might be, prior to any knowledge whatsoever. And it is this subject-independent notion that one will want to employ if one is to give an account of the objects of propositional attitudes according to which different people with very different background knowledge can nonetheless have, for example, particular beliefs in common. And it is this latter, subject-independent notion of epistemic possibility that I am exclusively concerned with here. Thus it is this notion that I will have in mind when I speak simply of epistemic possibility.

I am not going to discuss the issue of what EPWs actually *are*: i.e. the metaphysics of epistemically possible worlds. Rather I am interested here in the question of which things are true at EPWs. That is, I am in interested in general accounts of exactly which putative ways the world might be, prior to any knowledge whatsoever, really correspond to EPWs. Thus I am going to need some way of framing and discussing such general accounts of which EPWs there are. One might think that one could do this just by taking accounts of EPWs to identify what is true at a given world by saying simply which actual individuals have which properties at that world. However this is problematic for at least two reasons. Firstly there will presumably be EPWs at which there are far more people than there actually are; and secondly we might want there to be EPWs at which Hesperus is not Phosphorus, and at which Hesperus has some property Phosphorus does not. So simply classifying EPWs in terms of which actual individuals have which properties at them will not work. And in fact the most natural thing to do here is simply to use structured propositions as things in terms of which to give general accounts of which EPWs there are. That is, I will assume that general accounts of which EPWs there are take the form of saying which classes of structured propositions correspond to EPWs: i.e. of saying for exactly which classes of structured propositions there is some EPW at which exactly those propositions are true.

Now, especially in light of the previous section of the paper, a few remarks should be made about this methodological decision. Firstly, anything that I say should apply whether these structured propositions are Fregean or Russellian. (But of course if one wants EPWs at which (e.g.) Hesperus has some property that Phosphorus does not, then one will not want to use Russellian propositions.)

Secondly, I will assume that these structured propositions are stratified into some sort of hierarchy as sketched in §1; and I will only be concerned with the structured propositions at level 1 of the hierarchy (i.e. the lowest level to contain any propositions at all).

Thirdly, I talk about 'classes' rather than 'sets' of structured propositions because there may be problems in talking about sets of structured propositions if these sets are understood in anything like the sense of standard set theory (i.e. Zermelo-Fraenkel, or ZF, set theory). For example there may be as many structured propositions (at the first level of the hierarchy) as there are sets and hence no set that contains, for any proposition P, either P or ¬P (just as there is no set of all sets). I assume, however, that some way of talking about classes of structured propositions that avoids such problems can be found: one could use some sort of second-order logic, for example.

Fourthly, in making this methodological decision I am certainly not assuming that some account of propositions as structured is the best philosophical account of propositions available; if desired, one can think of these structured propositions as a kind of Tractarian ladder that one uses to indicate what the picture of EPWs and EPW propositions looks like, but which one kicks away once one has this new account of propositions. (It may of course be that the question of which account of propositions is the best overall account is neither interesting nor fruitful.)

**3. A First Account of EPWs and a Problem**

So, as I have said, the way in which I am going to consider accounts of EPWs is by considering which particular classes of structured propositions these accounts say correspond to EPWs: where by 'corresponds' I mean the relation that holds between a class of structured propositions and a world just in case those propositions are exactly the propositions that are true at that world.

A very natural thing to think at this point is that for any logically consistent class of structured propositions there should correspond an EPW. A note about my use of 'logically consistent': I assume that I have at my disposal a notion of logical consistency for structured propositions that basically parallels the notion for sentences. (If desired this notion of logical consistency for structured propositions could be formally defined in terms of a notion of logical consistency for sentences that express the propositions; however for reasons of space I do not do this here.) So, a natural thing to think is that for any logically consistent class of structured propositions there should correspond an epistemically possible world. For, if some class of propositions is logically consistent, then one would think that it must surely require genuine knowledge about the world to establish that the members of this class of propositions are not all true. Hence, it seems that there should be some EPW at which they are all true.

I think that as a first thought this seems more or less correct. Although one caveat must be added: the classes of structured propositions that correspond to EPWs should not just be consistent but maximally consistent. That is, they should not be included in any larger consistent classes. (To say that a class X is included in another class Y is to say that X is a subclass of Y, i.e. to say that everything in X is also in Y; it is not to say that X is a member of Y; to say that some object z is contained in Y is, on the other hand, to say that z is a member of Y.) The reason one needs this caveat about maximal consistency is that otherwise one would be committed to the claim that, for example, there is some EPW w such that the one and only proposition that is true at w is the structured proposition that Edinburgh is north of Paris. And I take it that that seems unnatural. If these things are in any sense supposed to be 'worlds' then they should presumably be 'total' in the sense that the classes of structured propositions that are true at them should be not just consistent but maximally so. However, although this seems intuitively right to me, nothing in my arguments will turn on the requirement that only maximally consistent classes of propositions correspond to EPWs.

So the following seems to be a very natural account of EPWs.

(EPW) If X is a maximally consistent class of structured propositions, then X corresponds to an EPW.

That is, if X is a maximally consistent class of structured propositions, then there is some EPW at which precisely the propositions in X are true.

Unfortunately, however, it would appear that this natural and intuitive account (EPW) cannot be correct: as I now argue. The first premise of this argument is as follows.

(E1) There is some logically simple property F.

I continue to use ‘property’ simply as a place-holder for whatever it is in structured propositions that correspond to predicates in sentences; so if one’s favoured account of structured propositions is the Fregean account, then where I say ‘property’ one should read ‘concept’ or ‘concept-sense’. More generally I continue to talk about structured propositions as if they are Russellian: but everything that I say applies regardless of which account of structured propositions one uses. Why should there be some logically simple property F? By a ‘logically simple’ property I mean a property whose logical behaviour is analogous to that of atomic predicates in language. Thus, a logically simple property is a property with no internal logical structure. A plausible example of a logically simple property is the property of being green; a plausible example of a property that is not logically simple is the property of being either green or tall. I take it that it seems intuitively very likely that given any property, e.g. the property of being green, either it is logically simple or it is a logical construction out of other logically simpler properties, and if the latter, then these properties are either logically simple themselves or they are in turn logical constructions out of logically simpler properties, and so on. And I take it that this process must stop somewhere: that is, I take that it is not the case that every property is a logical construction of logically simpler properties and that in every case this process goes infinitely far down, as it were. Hence (E1).

The next premise is as follows.

(E2) If F is a logically simple property, and X is any class of EPWs, then the following is a consistent class of structured propositions: $\{F(w): w \in X\} \cup \{\neg F(w): w \notin X\}$.

$\{F(w): w \in X\}$ is the class of all structured propositions that ascribe the property F to some world w in the class X; $\{\neg F(w): w \notin X\}$ is the class of all structured propositions that ascribe the negation of F to some w not in X. (E2) is about the union of these two classes of structured propositions, $\{F(w): w \in X\}$ and $\{\neg F(w): w \notin X\}$: i.e. (E2) is about the class that contains every structured proposition that is in one or other of these two classes. (E2) says that if F is a logically simple property, then this union class is consistent. This is a consequence of the fact that F is a logically simple property. Taking the property of being green as an example of a logically simple property for a moment, it would seem to be the case that given any class of objects, it should be consistent for just those objects and no others to be green: that is, even if metaphysics might in some cases rule out this possibility, logic should not. (There are some subtleties here concerning the case in which one uses Fregean rather than Russellian propositions that I ignore.)

Next one has the following.

(E3) If F is a logically simple property, and X is any class of EPWs, then there is some maximally consistent class of structured propositions Y that includes $\{F(w): w \in X\} \cup \{\neg F(w): w \notin X\}$ and some EPW w that corresponds to Y.

(E3) says that given a logically simple property F and a class of EPWs X, then the class of structured propositions mentioned in (E2) is not only consistent but also included in some maximally consistent class Y, which hence (by (EPW)) corresponds to some EPW w. I take it that given any such an F and X, then given the consistency of the class of propositions mentioned in (E2), there will be a maximally consistent class of propositions Y that includes the former class.

Paradox is now just a few short steps away. First, let F be a logically simple property as required by (E1). Then, given any class of EPWs X, let $Y_X$ and $w_X$ be as required by (E3): that is, let $Y_X$ be a maximally consistent class of structured propositions, including the union class of (E2), i.e. $\{F(w): w \in X\} \cup \{\neg F(w): w \notin X\}$; and let $w_X$ be an EPW that corresponds to this class of structured propositions $Y_X$, i.e. let $w_X$ be such that the propositions true at $w_X$ are precisely those that belong to $Y_X$. One then has the following.

(E4) If X and X' are distinct classes of EPWs, then $w_X \neq w_{X'}$.

That is, each class of EPWs is associated with its own distinct EPW. To see why (E4) is true suppose that v is a member of X but not of X'. Then F(v) is going to be a member of $Y_X$ (because F(v) is a member of $\{F(w): w \in X\}$), and so F(v) will be true at $w_X$. However, $\neg F(v)$ is going to be a member of $Y_{X'}$ (because $\neg F(v)$ is a member of $\{\neg F(w): w \in X'\}$), and so, since $Y_{X'}$ is consistent, F(v) cannot also be a member of $Y_{X'}$, and so cannot be true at $w_{X'}$. Therefore, since something is true at one but not the other, $w_X$ and $w_{X'}$ must be distinct.

Next let Z be as follows:

$$Z = \{w: \text{for some class of EPWs } X,\ w = w_X \text{ and } w \notin X\}.$$

That is, Z is a class of EPWs: the class of precisely those EPWs w, that are of the form $w_X$, for some class of EPWs X that w is not a member of.

Now consider $w_Z$: there must be some such world by (E3). By the definition of Z and (E4) one has:

(E5) $w_Z \in Z$ iff $w_Z \notin Z$.

For suppose firstly $w_Z \in Z$; then by the definition of Z, for some class of EPWs X, $w_Z = w_X$ and $w_Z \notin X$; but by (E4) if $w_Z = w_X$ then Z = X; so $w_Z \notin Z$. Conversely suppose $w_Z \notin Z$; then this means that $w_Z$ satisfies the membership condition of Z and thus that $w_Z \in Z$.

Since (E5) is a contradiction the apparently natural account of EPWs (EPW) cannot be right. In §4 I consider how one might try to modify this account so as to avoid this problem.

**4. A Second Account of EPWs and a Problem**

Put loosely, the problem with the natural account of EPWs that I considered in §3 is that, according to it, there is an EPW for each maximally consistent class of structured propositions: but the argument (E1–E5) shows that this is impossible; there are simply going to be too many such classes of structured propositions for there to be enough EPWs to go round. Thus, the obvious way to approach the problem of trying to find a better account is simply to look for some natural way of disqualifying certain maximally consistent classes of structured propositions from corresponding to EPWs. And the natural way in which to go about disqualifying maximally consistent classes of structured propositions is by ruling out those classes that can be rejected on the basis of a certain special sort of knowledge. And the natural sort of knowledge to first fix upon is presumably a priori knowledge. So in this section I consider an account of epistemically possible worlds based on the following principle in place of (EPW).

(EPW*) If X is a class of structured propositions, then X corresponds to an EPW if X is maximally consistent and X cannot be rejected a priori.

In some ways this does look like at least something of a departure from the original aim of having an EPW for any way the world might be, prior to any knowledge whatsoever. For, if P is, for example, some mathematical structured proposition that we can know a priori, then ¬P will not be true in any EPW, according to (EPW*), even though, intuitively, before we had established P one would have thought that for all we knew at the time, it might have been that ¬P. On the other hand, if one wants the totality of these worlds to have an interesting structure, then one does not want any old class of propositions to correspond to an EPW: for example it is natural to require that the class is at least logically consistent. And once one puts such requirements in place, one is going to get cases of things being epistemically impossible despite in a sense being things that one can initially be ignorant of and then come to know: e.g. negations of complicated logical truths. One can motivate the additional requirements—and additional epistemic impossibilities—brought about by (EPW*) in a similar way. In any case, the argument of §3 shows that some additional requirement beyond logical consistency is called for, and that of (EPW*) appears the most natural one to try.

In order to be able to properly assess (EPW*) I need at least a rough account of what it is for a class of structured propositions to be rejectable a priori. Firstly, at least for the purposes of this paper I will take it that this is simply to be understood in terms of rejecting some member of the class a priori. Secondly, how do I understand what it is for a structured proposition to be rejectable a priori? I have in mind something like the following. P is rejectable a priori just in case ¬P is knowable a priori. And P is knowable a priori just in case it is metaphysically possible for some subject S to give a valid argument with P as conclusion and premises that are (something like) self-evident for S. I do not try to spell out what exactly the relevant notion of self-evidence is, or what alternative property one should require the premises of the argument to possess. (However, I have in mind a notion of a priori according to which, for example, knowledge of one's own particular mental states, e.g. that one is in pain at some specific time, is not the sort of thing that is knowable a priori.)

By a centred MPW I mean a pair ⟨w,c⟩ where w is an MPW and c is a 'centre', a pair of an object and a time (similarly for centred EPW). I assume the following.

(KAP) If a structured proposition P is knowable a priori, then there is some centred MPW ⟨w,c⟩ at which P is the unique structured proposition that is established a priori at ⟨w,c⟩.

I now move onto considering whether the account (EPW*) really succeeds in escaping from the sort of problems raised in §3.

In considering this issue I am going to talk about one infinite class being ‘bigger’ or ‘smaller’ than another infinite class. This is to be understood in the usual way in terms of one-to-one correspondences. Explicitly: a class X is the same size as a class Y iff there is a one-to-one correspondence between X and Y (i.e. a function from X to Y that sends distinct members of X to distinct members of Y, and such that every member of Y has some member of X sent to it). X is smaller than Y iff X is the same size as some subclass of Y but not the same size as Y. Y is bigger than X iff X is smaller than Y; X is at least as big as Y iff X is not smaller than Y; etc. For any X and Y: either X is smaller than Y, or X is the same size as Y, or X is bigger than Y.

The problem with the account of EPWs given in §3 ((EPW)) is essentially as follows. (EPW) says that there is a unique EPW for each maximally consistent class of structured propositions: that is, (EPW) says that there are at least as many EPWs as there are maximally consistent classes of structured propositions. But in fact however many EPWs there are, there are always going to be more maximally consistent classes of structured propositions: that is what is demonstrated by the argument (E1–E5), which is the derivation of a contradiction from (EPW). In fact, the specific problem was that, given some logically simply property F, there are going to be too many classes of structured propositions containing just propositions of the forms F(w) or ¬F(w) for EPWs w. Now the question is: might not (EPW*) rule out enough of these classes for there not to be too many left? And to answer this question one needs to know, loosely, how many structured propositions of the forms F(w) or ¬F(w) are knowable a priori. And, in turn, this question depends, by (KAP), on how many centred MPWs there are: since by (KAP) there are no more a priori knowable structured propositions than there are centred MPWs.

When considering centred worlds it is natural and convenient to think of structured propositions as being true or false at centred worlds, as opposed to being true or false at worlds simpliciter. The following principle is plausible.

(CME) For each centred MPW ⟨w,c⟩, there is a centred EPW ⟨w',c'⟩ such that for any structured proposition P, P is true at ⟨w,c⟩ iff P is true at ⟨w',c'⟩.

In fact for convenience I assume that every centred MPW ⟨w,c⟩ is itself a centred EPW as in (CME). (CME) is intuitively plausible if one assumes (EPW*): if something is metaphysically possible then it should not be rejectable a priori, but it would have to be rejectable a priori to be a metaphysical possibility that did not correspond to an epistemic possibility (by (EPW*)). (There are some complexities that I ignore here.)

The upshot of (CME) is that there are no more a priori knowable structured propositions than there are centred EPWs. But, on the other hand, for all that I have said so far, there may be as many a priori knowable structured propositions as there are EPWs. But in that case it might be that, given some particular logically simple property F, every proposition of the form F(w) for an EPW w is either knowable a priori or rejectable a priori. And in this case one certainly would not have to worry about there being too many classes of structured propositions containing just propositions of the form F(w) or ¬F(w), since all but one of these would be rejectable a priori: i.e. all but the class containing exactly the true such propositions. And in this situation the sort of argument that I gave in §3 against (EPW) would be thoroughly blocked, since it relied on there being one such class of structured propositions for every class of EPWs (i.e. one such class of structured propositions corresponding to an EPW).

At this point one could I think continue to focus on structured propositions of the form F(w) for EPWs w, and give a somewhat less straightforward version of the argument of §3. However philosophical applications of EPWs will employ not just EPWs themselves, but also classes of EPWs, or classes of centred EPWs. For example accounts of propositions in terms of EPWs will identify them either with classes of EPWs, or with classes of centred EPWs. And one is going to want one's EPWs to represent epistemic possibilities about which 'thoughts'—i.e. EPW propositions—are believed, asserted etc. Indeed, if one's interest in EPW propositions stems from a desire to avoid a hierarchically stratified account of propositions (§1), then one is going to be particularly concerned to ensure that one's EPWs represent epistemic possibilities about such EPW propositions: otherwise EPW propositions cannot themselves be about EPW propositions; one will thus need new propositions to express thoughts about EPW propositions (or the propositional attitudes modelled with them), and so one is back to a hierarchical account of propositions. Thus, rather than try to give a less straightforward version of the argument of §3, still focussing on structured propositions of the form F(w) for EPWs w, I will instead give an argument focussing on structured propositions of the form F(P) for P an EPW proposition, or P a centred EPW proposition (i.e. P a class of EPWs, or P a class of centred EPWs).

I initially give the argument in terms of centred EPW propositions (I describe how one can give a similar argument in terms of 'non-centred' EPW propositions later in this section). I thus consider rather than (EPW*) the following analogous principle for centred EPWs: in the following a centred EPW $\langle w,c \rangle$ 'corresponds' to a class of structured propositions X if X consists of precisely those structured propositions that are true at $\langle w,c \rangle$.

(CEPW*) If X is a class of structured propositions, then X corresponds to a centred EPW if X is maximally consistent and X cannot be rejected a priori.

For each centred EPW proposition P, let F(P) be the structured proposition that says that P is the unique centred EPW proposition that is questioned, and that nothing is established a priori: that is, F(P) is true at a centred EPW $\langle w,c \rangle$ iff P is the unique centred EPW proposition that is questioned at $\langle w,c \rangle$, and nothing is established a priori at $\langle w,c \rangle$.

For some centred EPW propositions P, there may not be a centred EPW $\langle w,c \rangle$ at which F(P) is true: for example F(P) might be rejectable a priori. In this case, by (KAP) and (CME), there will instead be a centred EPW $\langle w,c \rangle$ at which $\neg$F(P) is the unique structured proposition that is established a priori at $\langle w,c \rangle$.

To simplify things I assume that for any centred EPW proposition P, there is a centred EPW at which F(P) is true iff $\neg$F(P) is not knowable a priori: this is a simplification because it might conceivably be that $\forall x(x = P \rightarrow \neg F(x))$ is knowable a priori but $\neg$F(P) is not (the former being knowable a priori would be sufficient for (CEPW*) not to assert the existence of a centred EPW at which F(P) is true, because $\forall x(x = P \rightarrow \neg F(x))$ and F(P) are inconsistent with one another); one could if necessary do without this simplifying assumption, however.

The problem with (CEPW*) is then as follows. For each centred EPW proposition P, let $\langle w_P,c_P \rangle$ be a centred EPW at which either F(P) holds, or at which $\neg$F(P) is the unique structured proposition established a priori. One has the following.

(E1*) If P and Q are centred EPW propositions with P $\neq$ Q, then $\langle w_P,c_P \rangle \neq \langle w_Q,c_Q \rangle$.

For example suppose F(P) holds at $\langle w_P,c_P\rangle$ and F(Q) holds at $\langle w_Q,c_Q\rangle$: then the unique centred EPW proposition questioned at $\langle w_P,c_P\rangle$ is distinct from the unique centred EPW proposition questioned at $\langle w_Q,c_Q\rangle$, so $\langle w_P,c_P\rangle \neq \langle w_Q,c_Q\rangle$; similarly for the other cases.

Next consider the following centred EPW proposition.

$$S = \{\langle w,c\rangle\text{: for some centred EPW proposition P, } \langle w,c\rangle = \langle w_P,c_P\rangle \text{ and } \langle w,c\rangle \notin P\}.$$

S is the class of all centred EPWs of the form $\langle w_P,c_P\rangle$ for some centred EPW proposition P with $\langle w_P,c_P\rangle$ not in P (S is analogous to Z of §3).

One then has the contradiction:

(E2*) $\langle w_S,c_S\rangle \in S$ iff $\langle w_S,c_S\rangle \notin S$.

For suppose first $\langle w_S,c_S\rangle \in S$: then by the definition of S, for some centred EPW proposition P, $\langle w_S,c_S\rangle = \langle w_P,c_P\rangle$ and $\langle w_S,c_S\rangle \notin P$; but by (E1*) if $\langle w_S,c_S\rangle = \langle w_P,c_P\rangle$ then S = P; so $\langle w_S,c_S\rangle \notin S$. Conversely suppose $\langle w_S,c_S\rangle \notin S$; then this means that $\langle w_S,c_S\rangle$ satisfies the membership condition of S and thus that $\langle w_S,c_S\rangle \in S$.

So it appears that the account of EPWs under consideration, (CEPW*), cannot be right.

One can give a similar argument in terms of non-centred EPW propositions as follows. In place of structured propositions containing the property F, one considers those containing the property G, where for any EPW proposition P, G(P) says that P is the unique EPW proposition that is questioned, and that nothing is established a priori. So G(P) is true at an EPW w iff P is the unique EPW proposition that is questioned at w, and nothing is established a priori at w. For any EPW proposition P, either there is an EPW $v_P$ at which G(P) is true, or there is a centred MPW $\langle u_P,c_P\rangle$ at which $\neg$G(P) is the unique structured proposition that is established a priori (by (EPW*), (KAP) and a simplifying assumption similar to that used above in the argument in terms of centred EPW propositions). The slight complication is that, although $\neg$G(P) is the unique structured proposition established a priori at $\langle u_P,c_P\rangle$, it may not be the unique structured proposition established a priori at $u_P$ (different structured propositions may be established a priori at different ‘centres’ in $u_P$). This means that one does not yet have a way of associating each EPW proposition with its own distinct EPW. One gets around this by considering an MPW $u'_P$ that is just like $u_P$, except that the event at the centre of $\langle u_P,c_P\rangle$ is the last event in $u'_P$ (so $u'_P$ is just like $u_P$, except that everything that happens either simultaneously with, or after, the event of the centre, is removed). This allows one to associate each EPW proposition with its own distinct EPW. One then lets $x_P$ be either an EPW at which G(P) is true, or $u'_P$. One can now run an argument similar to that above, with EPWs of the form $x_P$ in place of centred EPWs of the form $\langle w_P,c_P\rangle$. It follows from this argument that (EPW*) cannot be correct.

## 5. Kaplan's Problem for Metaphysically Possible World Semantics

In §§3 and 4 I raised problems for accounts of epistemically possible worlds. These problems, especially that of §4, are similar in form to the 'problem for [metaphysically-]possible-world semantics' that Kaplan raises in [1995]: they are all similar in form to Russell's paradox for Frege's *Grundgesetze* system, and thus in turn to the proof of Cantor's theorem (from which Russell got the idea).[7] It is useful to compare and contrast the problems of §§3–4 with Kaplan's: especially since I think that while the former cannot be straightforwardly rebutted, the latter can be.

Kaplan's problem is essentially as follows: I put the problem in terms of centred MPWs to make clear the parallel with the problem of §4; Kaplan put the problem simply in terms of MPWs. For each centred MPW proposition P it is prima facie plausible that there is some centred MPW at which P is the unique centred MPW proposition asserted: for any P, let $\langle w_P,c_P\rangle$ be such a centred MPW. If $P \neq Q$ then $\langle w_P,c_P\rangle \neq \langle w_Q,c_Q\rangle$: because the unique centred MPW proposition asserted at $\langle w_P,c_P\rangle$ is distinct from the unique centred MPW asserted at $\langle w_Q,c_Q\rangle$. Now consider $T = \{\langle w_P,c_P\rangle: \langle w_P,c_P\rangle \notin P\}$. One then has $\langle w_T,c_T\rangle \in T$ iff $\langle w_T,c_T\rangle \notin T$ (the detailed argument for this contradiction is precisely analogous to that given in §4 for (E2*)). That is Kaplan's problem for MPW semantics.

The proponent of MPW semantics can simply respond as follows. Kaplan's argument shows that there are centred MPW propositions P such that there is no centred MPW at which P is the unique centred MPW proposition asserted. Indeed Kaplan even constructs such a

[7] Kaplan's problem was first discussed in print in Davies [1981: 262], where debts to Kaplan and Christopher Peacocke are recorded. Cantor's theorem states that for any set A, the set of A's subsets is larger than A: see for example Machover [1996: 52].

centred MPW proposition: T (that there is some centred MPW $\langle w_T, c_T \rangle$, at which T is the unique centred MPW proposition asserted, is really all one needs to get the contradiction; so one must specifically deny that there is a centred MPW at which T is the unique centred MPW proposition asserted). One must thus deny the prima facie plausible principle that for every MPW proposition P there is some centred MPW at which P is the unique centred MPW proposition asserted. However such a denial is entirely compatible with MPW semantics.[8]

One might at this point object as follows. Surely it is implausible and untenable to claim that one cannot uniquely assert T at some centred MPW; for surely it is straightforward to construct an English sentence that asserts T: e.g. 'There is a unique centred MPW proposition asserted at this centred world, and this centred world does not belong to it'; thus the suggested response to Kaplan will not do.

This is not the place for a full discussion of these matters. However it is simply a widespread fact of life that if one's language can talk about its own semantics (e.g. if it contain its own truth predicate), then there will be sentences that in some sense intuitively seem to say something, but that are in fact neither true nor false: and so do not express centred MPW propositions, for example. An example of such a sentence is of course: This very sentence is not true. For another example suppose Nixon says, 'Everything Dean says about Watergate is true'; suppose Dean says, 'Everything Nixon says about Watergate is untrue'; and suppose that neither Nixon nor Dean says anything else about Watergate.[9] Kaplan's argument simply provides another example of such a sentence.

I now explain why one cannot respond to the problem of §4 in a way similar to the above suggested response to Kaplan's problem (one also cannot respond to the problem of §3 in a way similar to the suggested response to Kaplan's problem, however I do not discuss that here). In the case of each sort of centred possible world—i.e. metaphysically or epistemically possible—there are more centred world propositions than there are centred worlds. So in each case there is going to be a problem if one is committed to a principle that requires there to be a distinct centred world $\langle w_P, c_P \rangle$ for each centred proposition P (such a principle amounts to

[8] This is essentially the response of Lewis [1986: 104–8]. However the original presentation of Kaplan's problem in Davies [1981] did not show how to construct a proposition along the lines of T; and Lewis does not appear to be aware that one can straightforwardly construct such an T. (One cannot straightforwardly assert this T, of course: see below in the text.)

[9] This last example is from Kripke [1975: 691]. Kripke [1975] gives an account of how languages can contain their own truth predicates, and of how this leads to sentences that are neither true nor false: despite perhaps in some sense seeming to say something.

saying that there are at least as many centred worlds as there centred propositions; any such principle will allow one to run an argument like Kaplan's, or like those of §§3–4). In the MPW case, there is a plausible principle that says that there is a way of assigning distinct centred worlds $\langle w_P, c_P \rangle$ to centred propositions P: the principle that says that for each centred MPW proposition P there is some $\langle w_P, c_P \rangle$ at which P is the unique centred MPW proposition asserted. However we saw above that this principle can in fact be naturally and straightforwardly rejected.

Now consider the account of centred EPWs of §4, (CEPW*). And consider the principle that for each centred EPW proposition there is some distinct centred EPW $\langle w_P, c_P \rangle$ where P is the unique centred EPW proposition that is questioned, and where nothing is established a priori (something like this principle plays a role in the problem of §4). One does not of course have to accept this principle assigning a distinct centred EPW to each centred EPW proposition. But the problem in this case is that for every centred EPW proposition P for which this principle fails, there must, by (CEPW*), (KAP) and (CME), be some centred EPW $\langle w'_P, c'_P \rangle$ where the unique structured proposition established at $\langle w'_P, c'_P \rangle$ is of the form $\neg$F(P) (recall from §4 that F(P) is true at a centred EPW $\langle w, c \rangle$ iff P is the unique centred EPW proposition that is questioned at $\langle w, c \rangle$, and nothing is established a priori at $\langle w, c \rangle$): for suppose that P is a centred EPW proposition for which the principle fails; that is, suppose that there is no centred EPW at which F(P) is true; then $\neg$F(P) must be knowable a priori (by (CEPW*) together with the simplifying assumption of §4 that there is a centred EPW at which F(P) is true iff $\neg$F(P) is not knowable a priori); so by (KAP) there is a centred MPW $\langle w, c \rangle$ at which $\neg$F(P) is the unique structured proposition that is established a priori; and then by (CME) there is a centred EPW $\langle w', c' \rangle$ at which $\neg$F(P) is the unique structured proposition established a priori; $\langle w', c' \rangle$ is $\langle w'_P, c'_P \rangle$ as desired, i.e. $\langle w', c' \rangle$ is a centred EPW where the unique structured proposition established is of the form F(P), for our given centred EPW proposition P. Thus in the case of the account of centred EPWs of §4, even if one denies the original principle assigning distinct centred EPWs to each centred EPW proposition, in doing so one simply commits oneself to a new principle assigning distinct centred EPWs to each centred EPW proposition (i.e. a centred EPW proposition P is assigned either a world $\langle w_P, c_P \rangle$ at which F(P) is true; or, if there is no such $\langle w_P, c_P \rangle$, P is assigned a centred EPW $\langle w'_P, c'_P \rangle$ at which $\neg$F(P) is the unique structured proposition that is established a priori; as I explained in

§4, this method of assigning centred EPWs to centred EPW propositions assigns distinct worlds to distinct propositions). That is the essence of the difference between the MPW and EPW cases: in the MPW case, denying that certain prima facie possibilities are genuine possibilities does not commit one to the existence of centred MPWs where this fact is somehow established. That is why the proponent of MPW semantics has a straightforward response to Kaplan's problem, but there is no such response to the problem of §4 available to the proponent of EPW semantics.

**6. Alternative Accounts of EPWs?**

The upshot of §§3–4 is that there cannot be an EPW corresponding to each maximally consistent class of structured propositions that cannot be rejected a priori. To give a coherent account of EPWs one is going to need some further way of ruling out maximally consistent classes of structured propositions (i.e. ruling them out from corresponding to EPWs). Might there be some natural way of doing this, delivering an account of EPWs immune to the problems of §§3–4? Any apparent epistemic possibility thus disqualified from corresponding to an EPW would be consistent and such that, for all we can possibly know a priori, might actually be the way the world is. Further, as can be clearly seen from looking at the problem of §4, it will not help to move from a priori knowledge to some more inclusive category of knowledge: the problem has nothing to do with the type of knowledge, but rather with the requirement that the knowledge be (metaphysically or epistemically) possible. Thus any coherent account of EPWs will have to rule epistemically impossible, apparent epistemic possibilities that are, for all we can possibly know, how the world actually is. This would appear to be pretty strongly in tension with the basic idea behind epistemic possibility: the idea that something is epistemically possible if it is a way the world might be, for all we know. Thus it would seem that any coherent account of EPWs must fundamentally compromise the basic idea behind epistemic possibility.

This is not to say there cannot be coherent and useful notions of EPWs. Any such notion will have to be as follows. There will be apparent possibilities concerning which 'thoughts' people have—i.e. which attitudes to centred EPW propositions they have—that we cannot (metaphysically or epistemically) possibly rule out, but that nevertheless fail to be epistemically possible. Such a notion of an EPW may have many fruitful applications.

However where it will not give the desired results—i.e. where it will give results that do not fit with the basic idea behind epistemic possibility—will be when it comes to epistemic possibilities about people's thoughts, or when it comes to thoughts about thoughts (i.e. centred EPW propositions about centred EPW propositions, as it were). Such shortcomings are particularly disappointing given that one of the reasons for being interested in EPWs was so as to avoid a hierarchical account of propositions, and to give a good account of propositions about propositions (e.g. propositions expressing generalizations about all propositions, that are unsatisfactorily prohibited by hierarchical accounts; §1).

There is also the possibility of giving an account of EPWs that is hierarchical in essentially the way that accounts of structured propositions tend to be (§1). This does not seem to be a very attractive option: especially given that one of the reasons for being interested in EPWs is as a way of avoiding hierarchical accounts of propositions (§1). In any case, the notion of a hierarchical account of worlds would again appear to fundamentally compromise the idea behind EPWs: a world would no longer represent everything that is the case in some possible situation; rather for any possible situation, different 'worlds' would represent what is the case in this situation at different levels of some hierarchy.[10]

We have seen that the natural accounts of epistemically possible worlds cannot be correct. There may be coherent accounts of epistemically possible worlds that have fruitful applications. However it would appear that any such an account must fundamentally compromise the basic idea behind epistemic possibility.

---

[10] I briefly mention two things one might be tempted to try, in an effort to give accounts of EPWs that evade the sorts of problems raised in §§3–4. The problems I raised in §§3–4 turn on there being too many classes of structured propositions about EPWs and EPW propositions (i.e. too many for each to correspond to its own EPW). Thus one might think that one could evade the problems by focussing on only those classes that are 'small' (that are below a certain size). The idea would then be to say only that every 'small' class corresponds to an EPW. This will not work, however. The point is essentially that the problems of §4 can be raised by focussing only on 'small' classes of structured propositions: e.g. singletons containing either a proposition of the form ¬F(P), or the proposition that ¬F(P) is the unique structured proposition established a priori.

Alternatively, one might try to restrict attention to only those EPW propositions that can be expressed, or defined, in some relatively simple language. One might hope that such a restriction would mean that there will no longer be too many classes of structured propositions about EPWs and EPW propositions. However, even with such a restriction in place, problems along the lines of those in §§3–4 will emerge. The point is essentially that the problematic EPW propositions (such as Z of §3 or S of §4) can be expressed and defined in very simple terms (e.g. the terms in which I define them in §§3–4). (This shows that problems similar to those of §§3–4 can emerge even in settings in which there are not too many classes of structured propositions about EPWs and EPW propositions; it is sufficient for one's account of EPWs to be committed to there being an EPW corresponding to each class of structured propositions that can be defined in a certain way.) Thus this attempt to evade the problems of §§3–4 will not be successful.

## References

Chalmers, D. J. 2002a. On Sense and Intention. In J. Tomberlin (ed.), *Philosophical Perspectives 16: Language and Mind*: 135–82. Oxford: Blackwell.

——2002b. The Components of Content. In his (ed.) *Philosophy of Mind: Classical and Contemporary Readings*: 608–33. Oxford: Oxford University Press.

——2006a. The Foundations of Two-Dimensional Semantics. In M. García-Carpintero and J. Maciá (eds), *Two-Dimensional Semantics*: 55–140. Oxford: Oxford University Press.

——2006b. Two-Dimensional Semantics. In E. Lepore and B. C. Smith, eds, *The Oxford Handbook of Philosophy of Language*: 574–606. Oxford: Oxford University Press.

——Forthcoming. The Nature of Epistemic Space. URL = <http://consc.net/papers/espace.html>.

Davies, M. 1981. *Meaning, Quantification, Necessity: Themes in Philosophical Logic*. London: Routledge and Kegan Paul.

Frege, G. 1893. *Grundgesetze der Arithmetik, Volume I*. Jena: Pohle.

Kaplan, D. 1995. A Problem in Possible-World Semantics. In W. Sinnott-Armstrong, ed., *Modality, Morality and Belief: Essays in Honour of Ruth Barcan Marcus*: 41–52. Cambridge: Cambridge University Press.

Kripke, S. 1975. Outline of a Theory of Truth. *Journal of Philosophy* 72: 690–716.

Lewis, D. 1979. Attitudes *De Dicto* and *De Se*. *Philosophical Review* 88: 514–43. Reprinted in his *Philosophical Papers: Volume I*: 133–59. Oxford: Oxford University Press.

——1986. *On the Plurality of Worlds*. Oxford: Blackwell.

Machover, M. 1996. *Set Theory, Logic and their Limitations*. Cambridge: Cambridge University Press.

Russell, B. 1908. Mathematical Logic as Based on the Theory of Types. In his *Logic and Knowledge: Essays 1901–1950*, R. C. Marsh (ed.), 1956: 59–102. London: Routledge.

Stalnaker, R. C. 1978. Assertion. In his *Context and Content: Essays on Intentionality in Speech and Thought*, 1999: 78–95. Oxford: Oxford University Press.