A TWO-DIMENSIONAL ACCOUNT OF EPISTEMIC MODALS

A Dissertation
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy
in Philosophy

By

Daniel E. Quattrone, M.A.

Washington, DC
May 18, 2011
A Two-Dimensional Account of Epistemic Modals

Daniel E. Quattrone, M.A.

Dissertation Advisor: Steven Kuhn

Abstract

Not everyone knows that water is H2O. Suppose Alice is one of those people. Alice says, "For all I know, water might not be H2O." Intuitively it seems like Alice has spoken truly. That is, it seems like it is epistemically possible (for Alice) that water is not H2O. However, conventional accounts of modality in linguistics and philosophy of language predict that any metaphysically impossible statement will also be epistemically impossible (for anyone). And there are plausible arguments, from Kripke and others, that purport to show that it is metaphysically impossible for water to be anything other than H2O. So according to the standard accounts of modality, Alice has in fact said something false. This is highly counterintuitive and suggests that the standard accounts of modality need to be reworked. I offer a new account of modality that is capable of representing what I call EPMIs: epistemically possible metaphysical impossibilities. Sentences like "water might not be H2O" and "Hesperus might not be Phosphorus" are examples of EPMIs, and others can be readily found (including many that do not rely on Kripkean considerations about metaphysical possibility). My new account explains the existence of EPMIs while retaining the versatility and explanatory power of the standard accounts.

Index words: modality, epistemology, epistemic modality, two-dimensionalism, semantics

iii
Table of Contents

Chapter

1 Background ................................................. 1
   1.1 The Semantics of Names and Natural Kind Terms: Kripke and Putnam ................................. 3
   1.2 Epistemic Modality ........................................ 8
   1.3 The Problem .............................................. 12
   1.4 The Two-Dimensional Framework ......................... 16
   1.5 Alternative Approaches ................................. 30

2 The Metapositional Approach ................................. 49
   2.1 A Motivational Example ................................. 51
   2.2 The Approach Itself ...................................... 52
   2.3 The Metapositional Account in Context ................. 68
   2.4 Basic Content ........................................... 76
   2.5 Objection and Reply ..................................... 78

3 Basic Content .............................................. 86
   3.1 Basic Content ........................................... 86
   3.2 The Intuitive Plausibility of Basic Content .......... 87
   3.3 Basic Content and Narrow Content ..................... 97

4 The Propositional Approach ................................ 100
   4.1 Chalmers’s Account ..................................... 100
   4.2 Criticisms ............................................... 117
   4.3 Conclusions ............................................. 128
Chapter 1

Background

Epistemic modals are statements like “for all I know, it might be the case that $p$.”
More precisely, epistemic modals are statements about what is possible or necessary
given some body of knowledge. So “for all we know, it might be the case that $p$”
and “for all S knows, it might be the case that $p$” are also epistemic modals, along
with many other examples. Often, the explicit epistemic restriction (expressed by
“for all I know”) is omitted, which means that statements like “it might be the case
that $p$” will often have at least two readings. On standard semantic treatments,
modal statements like “it might be the case that $p$” are said to be true if there
is some (accessible) possible world at which $p$ is the case. However, given certain
widely accepted claims about the nature of metaphysical necessity, there are some
epipistemic modals that are true in some contexts even if they are not true at any
possible world.\footnote{Any mention of worlds or possible worlds is intended to refer to
metaphysically possible worlds, unless I specify otherwise.}

More precisely, epistemic modals seem to be statements about what is possible rela-
tive to some knowledge base. Often, this knowledge base consists in what is known
by some agent (or group of agents). Given the semantic treatment sketched out
above, it is reasonable to say that $p$ is epistemically possible relative to a knowledge
base iff $p$ is, in some appropriate sense, consistent with that knowledge base. Similarly, $p$ is epistemically necessary relative to a knowledge base iff $\neg p$ is inconsistent with the knowledge base.

Consider the statement “water might be XYZ,” where XYZ is some non-$\text{H}_2\text{O}$ chemical compound. Kripkean considerations lead us to conclude that “water = $\text{H}_2\text{O}$” is (metaphysically) necessarily true. So any statement of the form “water might be XYZ” will turn out false. But on the epistemic interpretation, it seems like it should be true for those who are unaware of the chemical composition of water. So the challenge, then, is to find an account of epistemic modality that allows for claims which are metaphysically necessarily false to nonetheless be epistemically possible. Call claims like this EPMIs. I will outline an account of epistemic modality that allows for EPMIs using the two-dimensional semantic framework.

In this chapter I aim to lay out the background material required for this project. First, I will set up the problem by showing how necessity of identity, which is often taken to be a consequence of semantic externalism (à la Kripke and Putnam), conflicts with most accounts of epistemic possibility. Second, I will discuss the “standard” account of modality adopted by many philosophers and linguists. Third, I will lay out some of the historical background of the two-dimensional framework and show why it might seem like a promising approach to address the problem of EPMIs. Finally, I will briefly discuss some alternatives to the standard account of modality and show why they do not fare any better than the standard account in dealing with

---

2Abbreviating the phrase “epistemically possible metaphysical impossibilities.”
EPMIs. This, in turn, will prompt a deeper investigation of two-dimensionalism in later chapters.

1.1 The Semantics of Names and Natural Kind Terms: Kripke and Putnam

In this section I will summarize the main claims of Saul Kripke and Hilary Putnam, advanced primarily in [17] and [26] regarding the semantics of names and natural kind terms and draw out an important consequence of the view.

1.1.1 Semantic Externalism

Semantic externalism is, as Putnam memorably put it, the claim that “meanings just ain’t in the head.” [26] Put another way, semantic externalism is the claim that the semantic content of (at least some) terms are determined solely by factors external to the speaker and not by the speaker’s psychological state or by any collective psychological property of the speaker’s linguistic community. For Kripke and Putnam, these external factors are causal and historical. Kripke focuses on names and natural kind terms. He argues that the semantic content of these terms is fixed by the causal-historical chain leading from the speaker to the referent. Here is Kripke’s rough statement of the view (which accords well with Putnam’s version).

\[3\]

\[4\]

Since the meaning of a term is at least partially conventional, and these conventions are external to any particular speaker, there is a sense in which more-or-less everyone is an externalist. We might call this minimal externalism. The thesis which Kripke and Putnam put forward is stronger than minimal externalism and far more controversial.

\[4\]I will not detail the arguments advanced in favor of semantic externalism. The main reason for this is that I aim to present an account of epistemic possibility that is consistent with semantic externalism, and as such will not challenge the central claims of the position. For my purposes, then, the arguments can be assumed to stand. Note, though, that while
An initial ‘baptism’ takes place. Here the object may be named by ostension, or the reference of the name may be fixed by a description. When the name is ‘passed from link to link’ the receiver of the name must, I think, intend when he learns it to use it with the same reference as the man from whom he heard it.

On this view, then, there is an initial “baptism,” wherein some speaker assigns some object (or a sample, in the case of natural kinds) a name, either via ostension or via definite description. For instance, one might take up a sample of a yellow metal and say, “Let this metal be called ‘gold.’” This baptism serves to fix the reference of the name or natural kind term and does nothing else (even if the baptism involves a description). This gives the term its semantic content; when the baptizer uses the term, its semantic content just is its reference, as determined by the initial baptism. When others use the term, its semantic content is still just its reference, in virtue of their having the appropriate sort of historical connection to the baptism.

Putnam’s version of semantic externalism is similar to Kripke’s in the relevant respects. He presents the famous “Twin Earth” thought experiment. Here is one version:

Twin Earth is a world very much like Earth, except that it contains no H₂O. Wherever one would find H₂O on Earth, on Twin Earth one instead finds XYZ, a different liquid which is indistinguishable from water. Twin Earth’s residents use the word ‘water’ to refer to XYZ. Now, consider a

---

I do not aim to refute semantic externalism, I am not committed to its being true. As I explain in §3, the problem I address arises whether or not semantic externalism is true.
resident of Earth, who does not know that water is $\text{H}_2\text{O}$, and his counterpart on Twin Earth, who does not know that the stuff he calls ‘water’ is XYZ. We might imagine that they are both in the same psychological state with regard to the word ‘water.’ But the Earthling’s utterances are about $\text{H}_2\text{O}$, whereas the Twin-Earthling’s utterances are about XYZ. So their utterances of ‘water’ have different referents. This, in turn, means that the word ‘water’ in English means something different from the word ‘water’ in Twin-English. And this difference is not due to any internal psychological state or property of the speakers of English and Twin-English; rather, it is due to the external facts about their respective worlds.

Putnam and Kripke have much in common. There are a few differences between their positions, though; Putnam, for instance, is primarily concerned with natural kind terms. My focus will be on Kripke’s version of externalism. In places where Putnam’s view is relevantly different from Kripke’s, I will handle Putnam’s version of externalism separately.

1.1.2 Kripke’s Anti-Descriptivism

One important feature of Kripke’s view is his repudiation of descriptivism. Descriptivism, at its simplest, is the claim that the semantic content of a name (and perhaps other terms, such as natural kind terms) is given by a description which speakers associate with the name. Kripke rejects descriptivism on the grounds that it gets the modal properties of names wrong. For example, suppose that the semantic content of the name “Saul Kripke” is ‘the author of Naming and Necessity.” If this
were the case, it would be necessarily true that Kripke is the author of *Naming and Necessity*. But surely this isn’t necessary; that is, it seems like we want to say that Kripke might not have written *Naming and Necessity*.

This criticism of descriptivism holds even if we adopt a version of descriptivism in which there need not be any particular description associated with any given name. For example, John Searle holds that the semantic content of names is given by a cluster or family of descriptions. On this view, the subject of the name need not satisfy any single description from the family of descriptions, but it must satisfy some (perhaps many or most) of them. But this view is also susceptible to Kripke’s criticism. If the semantic content of “Saul Kripke” is “the man who wrote *Naming and Necessity* or the man who invented semantic externalism or . . . ,” then it is necessarily true that Kripke is the man who wrote *Naming and Necessity* or the man who invented semantic externalism or . . . . But presumably it is contingent whether or not Kripke satisfies any of the descriptions given at all. The upshot of this attack (and other related attacks), for Kripke, is that descriptivism must be false. Kripke does not offer a developed theory to replace descriptivism, but he sketches out the account given in §1.1 as a starting point.

1.1.3 **NECESSITY OF IDENTITY**

Kripke claims that one significant consequence of his position on the semantics of names and natural kind terms is the necessity of identity. Consider Cary Grant, born Archibald Leach. “Cary Grant” and “Archibald Leach” both name the same person. Proper names are rigid designators, and “Cary Grant” and “Archibald Leach” are both proper names. This means that “Cary Grant” designates the same thing in
each possible world, as does “Archibald Leach.” Since these names designate the same object in this world, and since they designate the same object in every other world as they do in this one, Kripke argues that they must designate the same object in every possible world. Hence Cary Grant is necessarily Archibald Leach. This means that there are no worlds in which Cary Grant is a different person than Archibald Leach. Thus the statement that Cary Grant is Archibald Leach is necessarily true.

The same goes for natural kinds. If water is H₂O, then it is necessarily true that water is H₂O. There are no worlds in which water is XYZ. Of course, there might be worlds like Twin Earth—worlds in which some non-water substance is called “water.” Similarly, there might be a world in which “Cary Grant” names a famous actor and “Archibald Leach” names a different person altogether. But these worlds are not worlds in which water is XYZ or in which Cary Grant is Archibald Leach.

Kripke states the thesis clearly:

We use ‘Hesperus’ as the name of a certain body and ‘Phosphorus’ as the name of a certain body. We use them as names of those bodies in all possible worlds. If, in fact, they are the same body, then in any other possible world we have to use them as a name of that object. An so in any other possible world it will be true that Hesperus is Phosphorus.

The thesis of necessity of identity helps motivate the problem described below in §3. However, before we can see the problem we must first discuss the standard account of epistemic modality.
1.2 Epistemic Modality

In this section I will sketch out Angelika Kratzer’s account of epistemic modals from Kratzer’s account has some features which are shared by most extant accounts of epistemic modals and is relatively straightforward, which makes it useful as a case study. I will also discuss some of the roles played by epistemic modals in our language and psychology.

1.2.1 Kratzer on Epistemic Modals

Kratzer begins with a fairly standard possible worlds semantics for modals, but adds the notion of a relational modal “must in view of.” The relational modal appears in statements like “In view of what is known, water must be H2O.” “What is known” in this statement functions as a “modal restriction,” in that it serves to restrict the space of possibility—in this case, to those worlds consistent with what is known. The statement “In view of what is known, water must be H2O” is true iff water is H2O in all worlds consistent with what is known.

The worlds that fall under the modal restriction comprise the “modal base,” in Kratzer’s terminology. For instance, if we are operating in view of what is known, then the modal base consists in those worlds consistent with what is known.

One salient feature of Kratzer’s analysis that is shared by most other accounts of epistemic modality is that it treats epistemic necessity as a restricted sort of

\footnote{In particular, Kratzer’s account is equivalent to Lewis’s account of epistemic modals, as discussed in [20].}
metaphysical necessity. For example, there is an ongoing dispute regarding contextualist accounts of epistemic modality (advocated by, among others, Andy Egan, John Hawthorne, and Brian Weatherson). The dispute here is over cases in which two (or more) parties in a conversation have different knowledge bases. For instance, consider a case in which we have a speaker, Alice, who has not been outside in some time. The speaker claims that it might be raining. Bob, her audience, just came in from outside and knows that it is not raining. It is consistent with Alice’s knowledge that it is raining outside, but not with Bob’s knowledge. From Bob’s point of view, the following inconsistent claims all seem plausible: first, that it is not true that it might be raining; second, that when Alice says it might be raining, she says something true iff it is consistent with what she knows that it is raining; and third, that it is consistent with what Alice knows that it is raining. Cases like this one are problematic, and the dispute is over how to explain these plausible but inconsistent claims.

The dispute over contextualist accounts of epistemic modality is significant, and I will return to cases like the one discussed above in chapter 5. For the moment, though, the important thing to note is that the various parties in this debate treat epistemic modality in terms of possible worlds and implicitly accept something like Kratzer’s account, at least insofar as they are committed to treating epistemic necessity as a restricted sort of metaphysical necessity. But as we will see in the next section, this causes a problem if we take seriously the necessity of identity.

Another significant feature of Kratzer’s work is that it is generally taken to demonstrate that the various sorts of modals in natural languages have a common “semantic
core." That is, Kratzer shows that any treatment of modals which does not explain their common semantic properties is suspect. This, in turn, suggests that any account of epistemic modals we develop should be part of an account of modals in general.

1.2.2 Uses for Epistemic Modals

I have already noted that statements like "it might be the case that p" will often have at least two readings—one epistemic and one metaphysical. But I have not said anything about how we tell which reading is most appropriate in a given context. One might wonder, then, whether or not the epistemic reading is particularly common. Consider the following claim:

Epistemic modals are, in fact, not used very frequently. Suppose Alice says "water might not be H$_2$O." Suppose also that Alice does not know that water is H$_2$O. If we give this statement the epistemic reading, then what was said is true. But we’re inclined to say that Alice has uttered something false, which would indicate that we generally take utterances like Alice’s to be expressing a claim about metaphysical modality. This suggests that ordinary, unembedded modal claims, ought generally to be given the metaphysical reading.\(^6\)

This claim purportedly shows that epistemic modals may not be particularly common in everyday conversation. However, this does not actually follow. We can explain our inclination to say that Alice has said something false even if we give her utterance the epistemic reading. Perhaps when we hear Alice say "water might not be H$_2$O," we evaluate it not in terms of what she knows, but instead in terms

\(^6\)This claim was made to me by Nate Olsen in a personal communication.
of what we know. Since we know that water is H₂O, we judge the statement “water might not be H₂O” to be false. Furthermore, it seems unlikely that non-philosophers are normally concerned with metaphysical possibility, as we can navigate the world outside the philosophy classroom fairly effectively while ignoring those possibilities that are merely metaphysically possible. This suggests that an epistemic reading of statements like “water might not be H₂O,” when uttered in non-philosophical contexts, is more plausible.

Epistemic modals also have other roles in language. For instance, epistemic modals play an important role in explaining why people say things like “water is not H₂O.” Suppose Alice said “water is not H₂O.” If we were asked why Alice said what she said, a natural explanation would go something like this: she said water is not H₂O because she believes water is not H₂O; on the assumption that she has consistent beliefs, this will entail that “water is not H₂O” is epistemically possible for Alice.

In general, epistemic modals are often introduced to provide third-person explanations for why a speaker produced a certain utterance or acted in a certain way. Even if they do not appear frequently in other contexts, they are still widespread and important enough to warrant our attention.

---

7I have another reason to suspect that non-philosophers are not normally concerned with metaphysical possibility, but it is anecdotal. In my experience, students need to be introduced to the concept of metaphysical possibility and (in many cases) need to be shown why it’s relevant to philosophical investigations.

8I want to stress that I make no claims about the frequency of epistemic modals as compared to other sorts of modals. My point is simply that epistemic modals have an important role to play in our linguistic conduct, and that our semantic theories should therefore take them seriously.
1.3 The Problem

As we have seen, one consequence of Kripkean semantic externalism is that there are more metaphysical impossibilities than we might have expected. If Kripke is right, it turns out that statements like "water is H\(_2\)O" are metaphysically necessary. This, in turn, means that there are no possible worlds in which water is not H\(_2\)O. But this causes a problem for epistemic modals. If we adopt an account like Kratzer's, then we are committed to saying that anything which is epistemically possible is metaphysically possible (that is, that there are no EPMIs).

The problem which arises here is fairly clear: if epistemic possibility is a restricted sort of metaphysical possibility, then anything which is metaphysically impossible will be epistemically impossible as well, relative to any knowledge base. Similarly, anything which is metaphysically necessary will be epistemically necessary, again relative to any knowledge base. When we consider this in light of Kripke's necessity of identity thesis, we see that this means that statements like "water might not be H\(_2\)O" are inevitably false on the epistemic reading. Similarly, statements like "water is H\(_2\)O" invariably turn out to be epistemically necessary. Neither of these results seem appropriate. Not everyone knows that water is H\(_2\)O, and for some of those who don't it should be epistemically possible that water not be H\(_2\)O.

The problem also arises with names. On the Kripkean picture, Cary Grant is (necessarily) Archibald Leach, but some people don't know this; it should be epistemically possible that Cary Grant and Archibald Leach are different people for them. But on the going accounts of epistemic possibility, this is not the case, since there
are no worlds in which Cary Grant is one person and Archibald Leach another. So if Cary Grant’s birth name is unknown, the expression “Cary Grant is not Archibald Leach” is an EPMI.

Note that this problem arises even if Kripke and Putnam are wrong and identity is contingent. In general, there are likely to be sentences which are either necessary or impossible and whose truth value is independent of our knowledge. Let us call such sentences unknown non-contingent sentences. Some statements of metaphysical necessity and possibility might be unknown non-contingent sentences. To borrow an example from Quine, consider the sentence “it is metaphysically possible for something without extension to be colored.” If we do not know the truth of this sentence, then it is an unknown non-contingent sentence (assuming the truth of the principles of the modal logic S5). If it’s false, it’s an EPMI. One interesting class of statements here are claims about necessity of constitution, of the sort put forward in [11]. Johnston argues (contra Kripke) that statements like “water is H₂O” are not identity statements but rather claims about the constitution of water. Nonetheless, on Johnston’s view there are no worlds in which water is constituted by something other than H₂O. Some sentences of mathematics might also be unknown non-contingent sentences. Consider Goldbach’s conjecture: “Every even integer greater than two can be written as the sum of two primes.” The truth value of Goldbach’s conjecture is unknown, but if it is true it is presumably necessary (and if it is false it is presumably impossible). Goldbach’s conjecture, then, is an unknown non-contingent sentence.
Note that epistemic necessity cannot easily be identified with knowledge. While the truth value of Goldbach’s conjecture is not known, it might be a consequence of what is known. Epistemic necessity (at least on my account) is closed under consequence, and so if Goldbach’s conjecture is a consequence of what is known, then it is epistemically necessary. As a result, even though Goldbach’s conjecture is an unknown non-contingent sentence, it might turn out that Goldbach’s conjecture is nonetheless epistemically necessary; its negation, then, would not be an EPMI.

However, even if knowledge does not precisely map on to epistemic necessity, it does not follow that epistemic necessity is not an interesting and important object of study. It seems plausible, for instance, that anything which is epistemically necessary is either known or a consequence of what is known, and knowledge and epistemic necessity might coincide for certain kinds of ideal reasoners.

Sentences regarding de re necessary properties of objects can also be unknown non-contingent sentences, if we don’t know whether or not the object in question has the property under consideration. For example, suppose that Alice, a painter, is about to display her newest work. No one but Alice has seen the work, which is named “Figure 1." It is known, though, that Figure 1 depicts a plane figure—in fact, it either depicts a square or a circle. We do not know whether Figure 1 is a square or a circle. If Figure 1 is a square, it is necessarily four-sided. And so “Figure 1 is four-sided" is an unknown non-contingent sentence. Suppose there is

\footnote{This claim will be defended in §2.2.3 of chapter 2.}
\footnote{For more discussion, see §2.2.3 and 2.2.5 of chapter 2.}
\footnote{This claim must be understood as a de re modal claim. That is, to claim that Figure 1 is necessarily four-sided is to claim that that very object could not have had three sides. Alice could have created a different painting, and that painting could have been called Figure 1, but it would not have been the very same object.}
at least one unknown non-contingent sentence. If that sentence is true, then it is
metaphysically necessary but it might well be epistemically contingent. If it is false,
then it is metaphysically impossible but it might well be epistemically contingent.
In either case, the problem I have outlined in this section for most extant accounts
of epistemic modality arises.

A few different solutions to this problem are immediately apparent. First, we
might reject necessity of identity. This would require us to present a new account of
semantic content which accounts for the data supporting necessity of identity but
does not commit us to it. More importantly, rejecting necessity of identity would
not adequately address the problem; as I have shown, the problem is likely to arise
whether or not necessity of identity is true. As a result, I will not pursue this line of
inquiry. There are at least two further options. We might reject the possible worlds
analysis of epistemic possibility. We might alternatively introduce epistemically pos-
sible worlds as the counterparts of the metaphysically possible worlds in the possible
worlds analysis of metaphysical possibility. Both these approaches have issues which
might make them seem implausible, at least on the surface. For instance, rejecting
the possible worlds analysis would require us to reject an explanatory framework
which has proven very powerful in dealing with other sorts of modality. It would
also be both inelegant and semantically suspect to use a possible worlds analysis of
metaphysical modality and some entirely different framework for epistemic modality.
These sorts of considerations are unlikely to be decisive, but nevertheless I will not be
pursuing these alternative approaches because there are advantages to working with
the materials we have at hand (so to speak). If it is possible to solve the problem of
EPMIs using only metaphysically possible worlds, then we don’t need to introduce
epistemically possible worlds or jettison the possible worlds framework, and so the question becomes moot.

So here is the task: create an account of epistemic modals which respects necessity of identity and uses the same framework for epistemic and metaphysical modality. Ideally, this can be done with tools we already know and accept, namely, possible worlds (but see §5, below, for more discussion of alternatives to possible worlds). I aim to meet this challenge using the two-dimensional semantic framework.

1.4 The Two-Dimensional Framework

In this section I will present a brief historical overview of the two-dimensional semantic framework. I have two goals in doing this: first, to make clear what the framework is, and second, to motivate the claim that the framework will provide a useful means of solving the problem described in the previous section.

1.4.1 Two Notions of Necessity

I will use [3] as representative here. In this paper Martin Davies lays out a proposal for a formal account that characterizes the relations between several notions of necessity and between necessity and a priority, building on [4]. Davies’s account is worth investigating for my project insofar as it helps motivate the use of a two-dimensional framework for an account of epistemic modality and because it forms
a very significant part of the background for my own account.\footnote{Davies’s position is, of course, of great interest and importance in its own right, and there is much that could be said about it. I do not mean to suggest that Davies’s account is significant only because it informs my own work. The point of my discussion of Davies is not to provide a thorough analysis of his account, since the success or failure of his account is irrelevant to the success or failure of my account. Rather, the point is to see why two-dimensional accounts in general might be seen as promising with respect to epistemic modality.} It is important to note, though, that I am not endorsing Davies’s specific claims.

Davies begins with a quantified modal language that includes the familiar quantifiers and modal operators (\(\forall, \exists, \square,\) and \(\Diamond\)). However, some natural-language expressions (often involving the word “actually”) are difficult to formulate in this language. Consider Davies’s example: the expression “It is possible that everything that is actually red should have been shiny." This expression is true if there is a world which contains all the things which are red in the actual world and in which those things are shiny. We cannot render this as

\[\Diamond(\forall x)(x \text{ is red } \rightarrow x \text{ is shiny})\]

since this is true if there is a world in which every red thing is shiny, whether or not those objects are red in the actual world. Similarly, we cannot render “It is possible that everything that is actually red should have been shiny" as

\[(\forall x)(x \text{ is red } \rightarrow \Diamond(x \text{ is shiny}))\]

because this holds if for each actual red thing there is a world in which that object is shiny. We are looking for a single world in which all of the actually-red objects are shiny, not several worlds, in order to capture the meaning of the initial expression.
Davies’s suggestion is to introduce the “actually” operator, symbolized $A$. Each model has a special privileged world, which we will call $w^*$. Davies suggests that this privileged world is to be understood as the actual world. Intuitively, sentences of the form $A$s are true iff $s$ is true at $w^*$ (but see below for a clarification). According to Davies, this lets us render “It is possible that everything that is actually red should have been shiny” as

$$\Diamond (\forall x) (A(x \text{ is red}) \rightarrow x \text{ is shiny})$$

This sentence is true if there is a world $w$ such that every object in $w$ that is red in the actual world is shiny in $w$. Any world which contains all the things which are actually red and in which those things are shiny will meet this description and hence make the expression “It is possible that everything that is actually red should have been shiny” true.

It should be noted that this translation does not accurately capture the truth conditions for the most natural readings of the English expression that we started with. In particular, Davies’s translation is true if there is a world that contains shiny versions of some proper subset of the actual red things and lacks the other actual red things altogether. This suggests that Davies’s proposal may not accurately capture the truth conditions of sentences involving the English word “actually.” A more plausible representation could be obtained if we were to supplement Davies’s actuality operator with an “actuality quantifier” ($A x$). Sentences of the form $(A x) \phi x$ are true iff $\phi x$ is true of all objects $x$ in the privileged world $w^*$. We could then render “It is possible that everything that is actually red should have been shiny” as

$$\Diamond (A x) (A(x \text{ is red}) \rightarrow x \text{ is shiny})$$
Furthermore, Davies’s actuality operator does allow us to adequately represent the logical form of many English sentences containing the world “actually” that could not be represented without it. Consider the following sentence and its logical translation:

It might be the case that only things that are actually red are shiny.

\[ \Diamond (\forall x)(x \text{ is shiny} \to A(x \text{ is red})) \]

Davies also suggests that this framework gives us a way to think about \textit{a priori} truth. He claims that a sentence can be regarded as knowable \textit{a priori} if its truth does not depend on which possible world turns out to be actual, i.e. if it does not depend on which \( w \) is \( w^* \). This account of the \textit{a priori} makes it fairly straightforward to find examples to corroborate Kripke’s claim that the necessary and the \textit{a priori} are not coextensive classes. For example, note that \( A s \) is true iff \( s \) is true in this world, since this world is the actual world (i.e. this world is \( w^* \)). And there is a sense in which we can know that this world is \( w^* \) even if we don’t know which world is actual (i.e. we don’t know which \( w \) is \( w^* \)). Consider the following analogous case: suppose I wake up one morning to find myself in an unfamiliar location. There is a sense in which I don’t know where I am. However, I do know that wherever I am, I can truly say “I am here.” I may not know what that amounts to (i.e. I may not know which place “here” picks out, except in the sense given above), but that doesn’t mean I don’t know that I am here. Davies is, I think, relying on a similar intuition with respect to the space of possible worlds. I know that this world is actual world, but I may not know which world “the actual world” picks out.
Davies’s claim, then, is that we can show that $A \iff s$ is true, and we can know that it is true \textit{a priori}. But at least some sentences which are actually true are contingent, i.e. false in some worlds. Hence $A \iff s$ is contingent. So $A \iff s$ is an example of the contingent \textit{a priori}.

Here is an example. Let $s$ be the following sentence: George Washington was the first President of the United States. As it happens, $s$ expresses a metaphysically contingent truth. We know this truth \textit{a posteriori}. Since $s$ is actually true, $A \iff s$ is true. In fact, $A \iff s$ is metaphysically necessary. So $A \iff s$ expresses a contingent truth, since the left side is metaphysically necessary while the right side is not. Nonetheless, $A \iff s$ is knowable \textit{a priori}. On reflection, we can know that George Washington was the first President iff he was actually the first President even if we don’t know whether or not he was the first President.\textsuperscript{13}

Along similar lines, Davies argues that if $s$ is true, then $A \iff s$ is an example of the necessary \textit{a posteriori}. That is, $A \iff s$ is true in all worlds, but its truth depends on which world is actual. Any instance of a necessary \textit{a posteriori} truth will be problematic for most accounts of epistemic possibility in just the same way that “water might not be H$_2$O” is problematic.

So what we have in what we’ve seen so far of Davies’s system is a logic which can represent certain examples of necessary \textit{a posteriori} and contingent \textit{a priori} claims. Davies goes further, and gives us a way of modelling a second sort of necessity which, he suggests, can be plausibly understood as epistemic necessity. To accommodate

\textsuperscript{13}This argument does rest on two assumptions: first, that we inhabit the actual world, and second, that we know \textit{a priori} that we inhabit the actual world.
this second sort of necessity, we allow models in which the privileged world $w^*$ is permitted to vary and introduce an operator whose truth conditions take advantage of this expansion of the class of models.\textsuperscript{14} Davies calls this operator “fixedly” and symbolizes it $F$. The evaluation rule for this symbol requires us to introduce a second dimension of evaluation. Truth in the two-dimensional model is indexed to a pair of worlds rather than a single world. The pair of worlds consists in a world $w$ and a privileged world $w^*$. We allow both worlds in this pair to vary. Sentences of the form $F s$ are true at $\langle w, w^* \rangle$ iff $s$ is true at $\langle w, w^* \rangle$ regardless of which world is privileged (i.e. if we allow any other world to stand in for $w^*$ as the privileged world). $F$ thus functions analogously to the the familiar $\Box$ as a quantifier over worlds, except that in Davies’s framework $\Box$ quantifies over the first element of $\langle w, w^* \rangle$ and $F$ quantifies over the second. Given that truth is indexed to a pair of worlds, we must also revisit the account of the actuality operator $A$. Sentences of the form $A s$ are true at $\langle w, w^* \rangle$ iff $s$ is true at $\langle w^*, w^* \rangle$.

By combining the $F$ and $A$ operators, we can produce sentences of the form $F A s$, which (according to Davis) we can plausibly understand as saying that $s$ is \textit{a priori}. Sentences of the form $F A s$ are true iff for any world $w$, $s$ is true at $\langle w, w \rangle$. Given Davies’s account of the \textit{a priori}, it can easily be seen that for any sentence $s$, if $F A s$ is true, then $s$ is \textit{a priori}.

\textsuperscript{14}Note that allowing the privileged world $w^*$ to vary complicates the identification of $w^*$ with the actual world, since there is only one actual world. That said, we are talking about models. And it is perfectly acceptable to model situations in which different worlds are actual. In fact, while we know that only one world is actual we may not know which world it is, and hence we may not know which model corresponds to reality.
Many semantic theories index truth to worlds. Very roughly, the proposal I will investigate is to index truth to pairs of worlds, represented graphically by a two-dimensional array with worlds along both axes. Consider, for example, the sentence "grass is green." In ordinary English, this sentence expresses the proposition that grass is green. That proposition is true in some worlds and false in others. We might represent the worlds in which this proposition is true graphically like so:

<table>
<thead>
<tr>
<th></th>
<th>w</th>
<th>u</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>w</td>
<td>T</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

In this representation, \( w \) is the actual world; \( u \) and \( v \) are other worlds. As the table shows, in \( u \), grass is not green, but in \( v \) it is. We could in principle extend the table indefinitely to include all the possible worlds, and having done so we would be in a position to say for any world whether or not grass is green in that world. This would give us a one-dimensional picture of the truth conditions for "grass is green."

My account is a two-dimensional account. Very roughly, my proposal is to adopt a version of the two-dimensional semantic framework. In a two-dimensional theory, truth is indexed to pairs of worlds, represented graphically by a two-dimensional array with worlds along both axes. There are many ways of interpreting this array. I offer the following interpretation: expressions are associated with a metaproposition. Metapropositions determine how the proposition expressed by a sentence depends on the world in which that sentence was produced, in roughly the same way that the character of an indexical expression (in Kaplan's parlance) determines how the
A sentence like “I am running” expresses a different proposition when I say it than it does when you say it, because the indexical “I” picks out a different person in different contexts of use. In much the same way, a sentence like “water is H₂O” expresses one proposition when produced on Earth and another when produced on Twin Earth, and it is the metaproposition that captures this relationship between world of production and proposition expressed.

The information conveyed by a declarative sentence is normally determined by the proposition it expresses, using the actual world as the world of production. Propositions are the bearers of truth and the object of many semantic operators, including metaphysical modals. However, epistemic modals should, on my view, be seen as applying to metapropositions.

The metaproposition is represented by the two-dimensional array. This array encodes information about how the proposition expressed by the expression depends on the world in which the expression is produced. The worlds of production lie on the vertical axis of the array; the horizontal axis represents the worlds at which the expression is evaluated. Each row of the array thus represents the proposition expressed by the expression when it is produced in the corresponding world.

---

15 The analogy between the metapropositional approach and Kaplan’s account of indexicals is discussed in more detail in §3.1 of chapter 2.

16 If I am correct and epistemic modals should be understood in this way, then it is plausible that there are other operators that also apply to metapropositions. For instance, doxastic modals, which are concerned with what is possible or necessary relative to a set of beliefs, seem to behave much like epistemic modals, and intuitively we can construct examples of doxastically possible metaphysical impossibilities. Other candidates for metapropositional operators include “It is surprising that…” and “… wonders if it is true that…”
Consider the sentence "grass is green" once again. In our world, this sentence (when uttered in ordinary English) expresses the proposition that grass is green. But we can imagine worlds in which the sentence "grass is green" expresses a different proposition, perhaps because the meanings of the words "grass" and "green" are different, or because "grass" refers to some other substance. Let us stipulate that \( x \) and \( y \) are worlds in which "grass is green" expresses a different proposition than it does here. We can represent this using a two-dimensional array, like so:

<table>
<thead>
<tr>
<th></th>
<th>( w )</th>
<th>( x )</th>
<th>( y )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( w )</td>
<td>( T )</td>
<td>( F )</td>
<td>( T )</td>
</tr>
<tr>
<td>( x )</td>
<td>( F )</td>
<td>( F )</td>
<td>( T )</td>
</tr>
<tr>
<td>( y )</td>
<td>( F )</td>
<td>( F )</td>
<td>( F )</td>
</tr>
</tbody>
</table>

This array tells us that, when produced in \( x \), the sentence "grass is green" expresses a proposition which is true in \( y \) but false in \( w \) and \( x \). And when it is produced in \( y \) it expresses a necessary falsehood.

As a rough approximation, a sentence will be metaphysically possible with respect to a world of production \( w \) iff the proposition it expresses in \( w \) is true at least one corresponding world of evaluation (i.e. at some point along the same row as \( w \)) and metaphysically necessary if the proposition it expresses is true at all corresponding worlds of evaluation (i.e. at all points along the same row as \( w \)). And we can take a first pass at defining epistemic possibility and necessity as well: a statement will be epistemically possible with respect to a world of production \( w \) if there is some world \( v \) such that the proposition it expresses in \( v \) is true when evaluated at \( v \) (i.e. if it is true at some point along the diagonal of the array) and epistemically
necessary with respect to \( w \) if, for each world \( v \), the proposition it expresses in \( v \) is true when evaluated at \( v \) (i.e. it is true at all points along the diagonal of the array). This account of epistemic modality is only a rough approximation and admits of counterexamples; in particular, it omits two significant constraints that will affect how we evaluate epistemic modals. First, epistemic modals must be sensitive to what is known; this is the knowledge base constraint. Second, the sentence must have the same basic content in \( v \) as it does in the actual world.\(^{17}\) I will refine this formulation in chapters 2 and 3.

The two-dimensional framework has been employed to elucidate a number of different philosophical issues, and it has been interpreted in a variety of ways. My aim here is to use it to provide a treatment of epistemic necessity and possibility. On my view, if I say “Water might be XYZ,” I have said something true if there is a world \( w \) in which three conditions are met. First, the sentence “Water is XYZ” has the same basic content in \( w \) as it does in the actual world. Second, \( w \) is consistent with what we know. Finally, utterances of “Water is XYZ” express a true proposition when produced in \( w \).

David Chalmers has presented a somewhat different version of the two-dimensional framework, which I will call the propositional view. On his view utterances express two distinct propositions, each of which is encoded with the two-dimensional array. Chalmers suggests that the truth value of a sentence is indexed to worlds, which are maximal metaphysical possibilities, and scenarios, which are maximal epistemic possibilities. Chalmers offers an account of scenarios in which they are analyzable as

\(^{17}\)The notion of basic content will be introduced in chapter 2 and developed in chapter 3.
centered worlds—that is, possible worlds with a center consisting in a place, time, and agent (and possibly other pieces of information) marked out in order to evaluate indexical expressions. Chalmers suggests that we can consider a possibility as actual—that is, we can consider a possibility as representing a way the actual world could turn out to be. We can also consider a possibility as counterfactual, in which case we consider the possibility as representing a way the world might have been (but is not). These two ways of considering a possibility correspond to two different intensions. Chalmers’s intensions are functions from pairs of worlds to truth-values. What he calls “1-intensions” take as inputs (centered) worlds considered as actual, and what he calls “2-intensions” take as inputs worlds considered as counterfactual.

So, if the world is as we think it is, the expression “Water is H₂O” has a necessary 2-intension, which means that when we evaluate the sentence in different possible worlds considered as counterfactual when this world is considered as actual it comes out true every time. But it has a contingent 1-intension because there is some centered world (for instance, Twin Earth) such that if the actual world turns out to be like that world, the sentence would be false. Very roughly, we can say that on Chalmers’s view, 2-intensions behave like Kripke thinks intensions should behave, and 1-intensions behave like descriptivists think that intensions should behave. Chalmers’s view is discussed in more detail in chapter 4.

A third approach is (perhaps) implicit in Robert Stalnaker’s [30] and [31]. Stalnaker does not provide an account of epistemic modals, and he is not motivated by the problem of EPMIs. He is, however, concerned with a related problem. Stalnaker’s main concern involves expressions like “Water is H₂O” and “Two plus two
equals four," which (under normal conditions) express necessary truths but can nonetheless be informative. Stalnaker suggests that utterances are associated with a propositional concept (represented by a two-dimensional array). Each row in this array is called a horizontal intension and represents the proposition expressed by an utterance in a given world of production, as on my metapropositional view. Utterances like the ones given above typically express a necessary truth, given by the horizontal intension, and thus would appear to be uninformative. The fact that they can be informative is explained by pragmatic principles.

When someone says something like “Water is H$_2$O,” Stalnaker suggests we first evaluate the sentence normally and find that, taken literally, it expresses a necessary truth and would thus be uninformative. But it plainly can be informative. Stalnaker explains this apparent paradox by appealing to pragmatic principles. In particular, uttering an uninformative remark violates a number of Gricean maxims and thus implicates some other proposition. Stalnaker suggests that the utterance instead expresses the diagonal of the array, which he calls the diagonal intension. He calls the process whereby an utterance is reinterpreted to express the diagonal intension rather than the horizontal intension “diagonalization." This diagonal intension can be contingent even when the horizontal intension is necessary, and so it can be informative.

Stalnaker calls his view the metasemantic account. He identifies Chalmers’s account as an exemplar of the semantic account. The difference between the

\[\text{\footnotesize{\textsuperscript{[18]}}Stalnaker uses the Gricean picture of implicature in his example, but as far as I can see he has no special commitment to a Gricean theory. Any theory of pragmatics on which uninformative sentences can trigger pragmatic processes allowing for the reinterpretation of an expression as described above could suffice for his view.}}\]
semantic and metasemantic approaches involves what Stalnaker calls the “order of explanation,” i.e. the relation between the semantic properties of an utterance and the extra-semantic facts in virtue of which the utterance has the properties it has. On Stalnaker’s view, the standard semantic properties of an utterance are primary, and we introduce the diagonalization process in response to extra-semantic influences (such as the violation of a Gricean maxim). Here is Stalnaker:

In the case of our problematic statements [necessary truths that nonetheless seem to express informative propositions], the relevant maxim is that speakers presume that their addressees understand what they are saying. In terms of the two-dimensional apparatus, this presumption will be satisfied if and only if the propositional concept for the utterance is constant, relative to the possible worlds that are compatible with the context. Our problematic example, and all cases of necessary truths that would be informative (in the sense that the addressee does not already know that they are true) will be prima facie violations of this maxim, and so will require reinterpretation. Reinterpreting by taking the diagonal proposition to be the one the speaker intends to communicate brings the statement into conformity with the rule, and seems to give the intuitively correct result. [31]

So Stalnaker’s model is one in which we find that some expressions, when assigned their normal content, violate certain pragmatic or extra-semantic norms, and we are thereby driven to reinterpret the expression and assign it a different content which does not violate these norms. Diagonalization is the process by which we determine this secondary content. We use the diagonal of the array (as opposed to one of the
vertical columns, or a different row, or an arbitrary collection of points) because, according to Stalnaker, it seems to capture our intuitive judgments about the truth and falsity of what we take the speaker to be trying to get across. To use Chalmers’s terminology, Stalnaker maintains that 2-intensions are prior to 1-intensions, and that any explanation of 1-intensions will depend on an antecedent account of 2-intensions. Chalmers, in contrast, maintains that the opposite holds. The reason “Water is H\textsubscript{2}O” has a necessary 2-intension (assuming the world is as we think it is) is because it has a certain 1-intension.\textsuperscript{19} My view is also an example of the semantic account, since on my view the metaproposition is prior to the proposition expressed (although metapropositions are not the same as Chalmers’s 1-intensions); that is, the meta- or extra-semantic properties of an expression determine what semantic properties it has, rather than the other way around.

On my scheme for distinguishing styles of two-dimensionalism, Stalnaker and Chalmers are both examples of the propositional account, because on their views it is a proposition that is the bearer of epistemic modal properties. So we have three theories to contend with: my view, which is metapropositional and semantic, Stalnaker’s view, which is metasemantic and propositional, and Chalmers’s view, which is semantic and propositional. The propositional/metapropositional distinction hinges on what object acts as a bearer of epistemic modality. The semantic/metasemantic distinction hinges on the relationship between the semantic value of an utterance and the empirical facts of the world in which the utterance is produced. On my view,

\textsuperscript{19}The relation between 1- and 2-intensions on Chalmers’s view has some formal parallels with the relation between character and content in Kaplan’s account of indexicals. Stalnaker notes this and describes Chalmers’s view as an example of what Stalnaker calls the “generalized Kaplan paradigm.”
the bearer of epistemic modality is the metaproposition, whereas on Chalmers’s view it is the proposition. My view is thus an example of the metapropositional approach, while Chalmers’s view is an example of the propositional approach.

I will develop and discuss my view in more detail in the next two chapters. Chapter 4 will be dedicated to an examination of the propositional approach, using Chalmers as an exemplar. Ultimately I aim to show that Chalmers’s view has serious flaws that make it an inadequate account of epistemic modality, whereas my own metapositional view is more successful. Stalnaker’s view gives us an alternative way to understand the two-dimensional framework, but it is not an account of epistemic modality. Moreover, it is not obvious how to produce an account of epistemic modality using Stalnaker’s version of the two-dimensional framework.

1.5 Alternative Approaches

In this section I aim to argue that a two-dimensional approach to the problem of EPMIs is more promising than any other alternative that uses possible worlds. I will also argue that two-dimensionalism is more promising than dynamic semantics and situation semantics for addressing this problem.²⁰

²⁰It is worth noting that situation semantics and dynamic semantics are primarily motivated by other problems than the problem of EPMIs. I am not claiming that two-dimensionalism is more promising than these frameworks in dealing with the concerns that motivate them; rather, my claims are 1) that any account of epistemic modality should deal with the problem of EPMIs and 2) two-dimensionalism is the most promising way of addressing this issue. In fact, I think that the two-dimensional framework is largely compatible with situation semantics and dynamic semantics, and I will offer some suggestions for combining these disparate approaches in chapter 5.
1.5.1 CONVENTIONAL APPROACHES

The standard account of modals in contemporary semantics is given by Angelika Kratzer and described above in §2. The problem with Kratzer’s treatment is that it takes epistemic possibility as a restricted sort of metaphysical necessity. This means that Kratzer’s view cannot resolve the problem of EPMIs, as on Kratzer’s view there cannot be any such propositions, and any metaphysically necessary proposition will also be epistemically necessary. Hence for Kratzer (and others who adopt similar approaches), it will not be epistemically possible that water is not H$_2$O, and it will be epistemically necessary that water is H$_2$O, regardless of what is known.

This problem affects most of the conventional analyses of epistemic modals that use the possible-worlds framework. I now turn to consider some alternative approaches that might initially seem more promising.

1.5.2 EPISTEMIC COUNTERPARTS

In [26] Putnam offers this suggestion:

...it is epistemically possible that pencils could turn out to be organisms. It follows that pencils are artifacts is not epistemically necessary in the strongest sense and, a fortiori, not analytic.

Let us be careful, however. Have we shown that there is a possible world in which pencils are organisms? I think not. What we have shown is that there is a possible world in which certain organisms are the epistemic counterparts of pencils.
This suggestion is not developed more fully, but doing so is fairly straightforward. Putnam’s suggestion seems to be that the truthmakers for EPMIs (and for epistemic modals generally) are epistemic counterparts. The epistemic counterpart of an object or kind of object need not be identical to it, though it can be. In some worlds, the epistemic counterparts of pencils are pencils, and in others the epistemic counterparts of pencils are pencil-shaped organisms. Two objects (or object kinds) will be epistemic counterparts just in case they are indistinguishable from one another given what is known.

Here is a sketch of a counterpart theory of epistemic modality in action: consider the case in which we don't know about the chemical composition of water. If we're in this epistemic situation, then it is epistemically possible that water is not H$_2$O. One reason that this might be true is that there is a world $w$ which contains a substance—not H$_2$O—which is qualitatively indistinguishable from H$_2$O. Two substances are qualitatively indistinguishable if we are incapable of telling which is which using whatever methods we have at our disposal. H$_2$O could thus be qualitatively indistinguishable from XYZ for people who don’t have access to modern chemistry, since XYZ and H$_2$O have all the same phenomenal properties and react similarly to the sorts of tests that people without access to chemistry can perform. This would not be a world in which water is not H$_2$O, but it would be a world which is qualitatively indistinguishable from our own. We can think of this world as a world in which I am “qualitatively in the same epistemic situation,” to use Kripke’s phrase, or as a world in which water has an epistemic counterpart, as Putnam puts it, that possesses the same phenomenal properties but a different chemical composition. This counterpart is not water, as water is necessarily composed of H$_2$O. But
it nonetheless underwrites the claim that it is epistemically possible that water is not H$_2$O.

Counterpart theories have a number of problems, but it seems likely that they can be addressed. Here is one problem: suppose I am hit on the head and begin hallucinating. I see, or seem to see, an elephant. In my discombobulated state I do not know if what I saw was real or a hallucination. I say, “For all I know, there might be an elephant in the room." Suppose there is no elephant. Given what I know, it seems that what I have said is true. Since there is no elephant in the room (and nothing that I am mistaking for an elephant), there is no elephant-counterpart which can underwrite this claim on a counterpart theory.

The problem here arises when we try to identify a counterpart for something other than a discrete object or a kind. One way we might handle this difficulty is to expand our understanding of the counterpart relation to allow states of affairs to have counterparts in other states of affairs. Let us return to the problematic case of the elephant that might be in the room. The issue is that, since there is in fact no elephant in the room, there cannot be an elephant-counterpart, and hence we have to way to truly say that there might be an elephant in the room. But rather than look for elephant-counterparts, perhaps we should look for counterparts to the state of affairs I find myself in. This state of affairs has certain features: in it, I have a headache (thanks to being hit on the head), I seem to see an elephant in it, along with all the other objects in the room. If there is a world in which I (or my otherworldly counterpart) am in a state of affairs which is indistinguishable from the actual state of affairs, then we can label that otherworldly state of affairs a
counterpart to my own state, and allow it to underwrite the modal claim that there might be an elephant in the room.

So far, so good. But counterpart theory isn’t out of the woods yet. Let us again consider the case in which we don’t know the chemical composition of water. When I say “water might not be H₂O,” we go looking for a world w which contains an epistemic counterpart to water with a chemical composition other than H₂O. Here is something I know about water: I call it “water,” as do other English speakers in my linguistic community. So the epistemic counterpart of water should also be called “water” by my counterpart and his linguistic community. But this isn’t as simple as it sounds. Since “water,” in my language, rigidly picks out H₂O, we need to clarify; the epistemic counterpart of water (call it XYZ) should be called “water” by members of my counterpart’s linguistic community in the world containing XYZ. So the world we’re looking for isn’t just a world with XYZ in it; it’s a world in which “water” (as used by the relevant residents of w) refers in w XYZ and in which XYZ is qualitatively indistinguishable from H₂O. So we need to keep track of two things regarding “water”: what it refers to in our language and what it refers to in the language of the relevant residents of w.

At this point, one might become concerned that by introducing a language-specific component to epistemic modality, it becomes impossible to say that (for example) water might not be H₂O for a Spanish speaker. This concern would not be well-grounded. If I spoke Spanish but not English, I might not know that water is called “water” by English speakers, and I would not know that I call water “water” (because it would be false). However, I would know that I call water “agua” and that other
Spanish speakers in my community call it “agua.” The epistemic counterpart for water in this case would have to be called “agua” by my counterpart’s linguistic community, but the very same thing might be called “agua” by the Spanish-speaking community counterpart and “water” by the English-speaking community counterpart. So including linguistic information in the knowledge base in this way does not require us to make any troublesome judgments about epistemic possibility and necessity.

Once we see that linguistic information like this is part of identifying epistemic counterparts, we see that counterpart theory acquires a two-dimensional flavor. Of course, this isn’t a problem from my perspective, but if we can develop a two-dimensional account without counterparts we have no reason to introduce them. To clarify: the issue here is that we can’t identify epistemic counterparts for named objects and natural kinds unless we also introduce some mechanism whereby the reference for the names of said objects and kinds is allowed to vary. And it is just this sort of mechanism that I introduce in the next chapter as part of my two-dimensional account. Perhaps a counterpart theory could be developed that does not explicitly use the two-dimensional framework I have adopted. Such theories have been suggested, but never developed in detail, and I suspect that if one was developed it would still be a two-dimensional theory broadly construed. In any case, I have not shown that counterpart theory always turns into two-dimensionalism, and I do not aim to show this; rather, my claim is that 1) we have reason to think that any successful counterpart theory will incorporate a two-dimensional element, and 2) we don’t need counterparts at all.
1.5.3 Impossible Worlds

Another approach is to allow for (metaphysically) impossible worlds. Impossible worlds have been presented as a means of dealing with other problems in metaphysics and philosophy of language, but not generally as a means of dealing with EPMIs.\footnote{For example, Graham Priest, in his \cite{25}, sets out a model of impossible worlds. This task is also taken up by Edward Zalta in \cite{34}. Zalta suggests that impossible worlds can be used to resolve questions surrounding counterfactuals with impossible antecedents, which is related to the problem of EPMIs, but he does not develop this suggestion. Similarly, Terence Parsons uses impossible objects to deal with the problem of reference to nonexistent objects, but does not touch on EPMIs.} I will not undertake the task of setting out a complete theory of impossible worlds here and showing how it might solve the problem of EPMIs; rather, I will sketch out such a theory and suggest where difficulties for the view might lie. The introduction of impossible worlds comes with an ontological and intuitive cost and this fact, when considered alongside the worries I will raise, should be sufficient to show that impossible worlds are at least \textit{prima facie} unlikely candidates to resolve the problem of EPMIs.

I will begin by examining metaphysically impossible but logically possible worlds. There is conceptual space for such things (opened up by Kripke and others). Very roughly, an expression is logically necessary if its truth is guaranteed by the laws of logic. An expression is logically possible if the laws of logic do not guarantee its falsehood. The proposal, then, is to account for at least some EPMIs by having our epistemic modal operators quantify over those logically possible worlds that are consistent with our knowledge.
However, the logical form of an expression is not always obvious and cannot be “read” off the surface structure of the sentence. For instance, Kaplan claims that indexicals (like “I”) are directly referential, which means that when used in context, their contribution to the proposition expressed is just their referent. So when I say “I am a human,” it expresses the proposition that Dan Quattrone is a human. And if I were to say “I am Dan Quattrone,” it would express the proposition that Dan Quattrone is Dan Quattrone—which seems like a logical truth. Kaplan’s theory of indexicals is not universally accepted, so this does not definitively show that “I am Dan Quattrone” expresses a logical truth. But it does offer one way in which the logical form of an expression might differ from its surface structure.

A second example: it could turn out that “water is not H₂O” expresses a proposition which is logically impossible. If Kripke is right and the semantic content of “water” is exhausted by its reference, then “water is not H₂O” expresses a logical impossibility. On Kripke’s view, “water is not H₂O” expresses the proposition that water ≠ H₂O. But given that water = H₂O and H₂O = H₂O, we can easily drive a contradiction from this claim. Hence on Kripke’s view “water is not H₂O” entails a logical impossibility. Kripke’s view is perhaps implausible, but it again serves to illustrate the general claim that the logical form of an expression is not always obvious.

The view that semantic content is exhausted by reference is of course controversial. In order for us to make sense of such a notion, we will (at a minimum) need to reject a Millian account of the semantic content of names and natural kind terms. But we are then on the hook to offer up an alternative account and to show how to make
sense of metaphysically impossible worlds on this account. There are well-known alternatives to Millianism available, but I will not be investigating them further. The two-dimensional approach I favor gives us a promising way to handle EPMMIs and epistemic modality in general without getting entangled in a debate over which alternative to Millianism we should prefer. In general, if logical truth is understood as truth in virtue of logical form, then the space of logical possibility will depend on the logical form of our expressions. And there is no consensus on how to discern the logical form of an expression. Rejecting Millianism does not settle this question, and rather than attempt to delve into these waters I will explore a different path.

Furthermore, on an account of this sort, metaphysical modal operators quantify over the metaphysically possible worlds, while epistemic modal operators quantify over the logically possible worlds that are consistent with the relevant knowledge base. If this approach is to be successful, we must suppose that the metaphysically possible worlds would be a proper subset of the logically possible worlds. Formally speaking, we do not need to specify which subset it is; as stated above, there is conceptual space for this sort of account. But some such specification will be desirable (and perhaps necessary) if we want to show that this supposition is plausible. Ideally, we would want to have a way of delineating those worlds that are metaphysically possible and those worlds that are only logically possible, and this division should reflect our intuitions about logical and metaphysical possibility.

Here is one possible way of divvying up the space of possibility. Let us say that the logically possible worlds are those worlds in which the laws of logic hold. The metaphysically possible worlds are those worlds in which the laws of logic and the
laws of nature (and, if necessary, the laws of mathematics) hold. On this view, “water is H\textsubscript{2}O” would express a metaphysically necessary proposition (as long as the fact that water is H\textsubscript{2}O is a consequence of the laws of nature). But there are other worlds in which the laws of nature do not hold, including worlds in which water is something other than H\textsubscript{2}O.

This suggestion is problematic, because it does not reflect our intuitions about metaphysical possibility. On this view, metaphysical possibility collapses into physical possibility (i.e. what is possible in view of the laws of nature). Intuitively, though, these categories are not the same. For instance, as best we can tell, the laws of nature entail that humans cannot travel faster than light. So faster-than-light travel is not physically possible. The claim that it is metaphysically impossible, though, seems to me to be false. This does not show that there is no way of partitioning the space of possible worlds such that we get the desired results. This approach to the problem of EPMIs is a live alternative. However, we have to think that the two-dimensional approach I favor is a promising one, and we have reason to think that introducing impossible worlds is not a straightforward as it may seem. I opt to investigate the two-dimensional approach; if and when a promising alternative is presented, we can compare it to the two-dimensional approach then.

One final issue with impossible worlds as a solution to the problem of EPMIs is this: every treatment of impossible worlds that I am aware of takes impossible worlds to be logically impossible worlds. Rather than taking up the project of developing an

\footnote{This way of carving up logical space was suggested by Wayne Davis in a personal communication.} \footnote{Not everyone accepts this intuition. See, for instance, Shoemaker’s.}
account of metaphysically impossible worlds with the desired properties, I have opted to show how to handle EPMIs without introducing metaphysical impossibilia. It is, however, also worth looking into the extant accounts of metaphysically impossible worlds to see if we can explain EPMIs using one of them.

On the most common accounts of metaphysically impossible worlds, metaphysically impossible worlds are also logically impossible. That is, they are the worlds where, as Zalta puts it, “the logically impossible happens.” [34] We can introduce such worlds by allowing for worlds that verify \( p \) and \( \neg p \) for some proposition \( p \); these are the metaphysically impossible worlds. On this view, there are no metaphysically possible worlds in which water is XYZ, but there is a metaphysically impossible world in which water is XYZ. The accessibility relation for metaphysical modals excludes these metaphysically impossible worlds, but the accessibility relation for epistemic modals does not (or at least it does not always do so).

This view is an extension of the conventional possible worlds framework. By expanding the set of worlds, we allow for EPMIs. But this expansion comes at a cost. The conventional semantics for epistemic possibility in the possible worlds framework is this: \( p \) is epistemically possible given some knowledge base iff there is a world \( w \) such that \( p \) is true at \( w \) and \( w \) is consistent with the knowledge base. Similarly, \( p \) is epistemically necessary given some knowledge base iff for all worlds \( w \) such that \( w \) is consistent with the knowledge base, \( p \) is true at \( w \). These truth definitions are widely used because they reflect some of the connections between epistemic possibility, epistemic necessity, and knowledge.
However, once we allow logically impossible worlds, we are faced with a number of difficulties. First among them is the challenge of establishing which impossible worlds are consistent with any given knowledge base. Recall that impossible worlds are worlds where the logically impossible happens. We know the actual world is not an impossible world, but given that our knowledge of the world is incomplete we may not be able to tell if a given world is impossible.

Nonetheless, there is reason to think that all impossible worlds are inconsistent with our knowledge. If, as is customary, we represent our knowledge as a set of propositions, then we can represent a world $w$ being consistent with our knowledge by taking the set of propositions true at $w$ and seeing if the union of this set and the set of known propositions is consistent. For any impossible world $i$, the union of our the set of known propositions and the set of propositions true at $i$ will not be consistent, because $i$ is internally inconsistent. If all impossible worlds are inconsistent with our knowledge, the fact that there is an impossible world in which water is XYZ will turn out to be irrelevant when we evaluate the epistemic modal “water might be XYZ,” since that world is not consistent with what we know. That world is inconsistent with what we know because it verifies a contradiction and we know all contradictions are false. This suggests that impossible worlds cannot address the problem of EPMIs after all.

But perhaps we’ve moved too quickly. Let us suppose that we come up with some alternate understanding of consistency such that there are some impossible worlds that are consistent with what we know. Alternatively, we might suppose that we can come up with an alternate definition of epistemic possibility such that consistency
with the knowledge base is not required. The problem now becomes one of limiting which impossible worlds count for evaluating epistemic modals. Presumably, we will want to exclude some worlds, i.e. those worlds in which the laws of logic or mathematics do not hold, while allowing others, i.e. those worlds in which water is XYZ. But how are we to do this? The path here is far from clear. Since all the accounts on offer define impossible worlds as those which verify at least one contradiction, we cannot use that as a way of distinguishing “acceptable" impossible worlds. In the absence of a developed theory of impossible worlds as applied to epistemic modals, we need not delve too deeply into this problem. It is sufficient, I think, to recognize that it is far from obvious how impossible worlds can help resolve the problem of EPMIs.

To sum up: I will focus on the two-dimensional account sketched out above, which requires only the metaphysically possible worlds which are widely accepted. The extant accounts of impossible worlds take impossible worlds to be logically impossible worlds, and logically impossible worlds do not seem like a promising line of inquiry for solving the problem of EPMIs. I acknowledge the possibility of alternative accounts involving metaphysically impossible (but logically possible) worlds. No such proposals have been developed, though, which makes it hard to evaluate the plausibility of this line of investigation. In general, it would be advantageous if we could do without such worlds (if only because an account that allows for metaphysically impossible worlds would likely be less parsimonious), and that is what I will try to do.
1.5.4 **Situation Semantics**

Since the conventional approach to modality fares poorly in dealing with epistemic modals (even when supplemented with impossible worlds), we might opt to abandon the possible worlds framework to address this issue. One alternative is situation semantics, initially developed by Jon Barwise and John Perry. Situation semantics replaces worlds with situations, which can be represented as consistent (but not necessarily maximal) sets of propositions. Alternatively, we might adopt something like Frank Veltman’s formulation from [32] in which we begin with a set of atomic sentences \( A \). In this formulation, worlds are functions from \( A \) to \( \{0, 1\} \) and situations are partial functions with the same domain and range.\(^{24}\) Statements like “it might be raining outside” are true if there is a possible situation in which it is raining outside.

Situations are more fine-grained than possible worlds. For the conventional possible worlds framework, the problem of EPMIs arises due to a “shortage” of possible worlds, and so replacing worlds with something more fine-grained might seem like a promising approach. However, the problem of EPMIs arises for situation semantics as well.

By definition, each situation corresponds to a part of some possible world. But there are no worlds at which water is XYZ and hence there are no parts of worlds in which water is XYZ, and hence no situations in which water is XYZ. So even though

\(^{24}\)Veltman takes worlds and situations as basic and propositions to be sets of worlds. Barwise and Perry take situations as basic. We might also take propositions as basic and worlds and situations to be defined in terms of propositions, as I did above. These views have different metaphysical commitments but for our purposes can be treated identically.
situations are more fine-grained then worlds, situations do not provide us with the resources to solve the problem of EPMIs.

This is not to say that the problem of EPMIs is insoluble for situation semantics. But since there are no possible situations in which water is XYZ, any attempt to do so will need to take one of two approaches. First, we might allow impossible situations. But this will run afoul of the same sort of difficulties that afflict impossible worlds, discussed above. Second, we might allow possible situations to validate metaphysical and epistemic modals in different ways. That is, we might take the situation containing a qualitative duplicate of Earth in which all H2O is replaced by XYZ to validate the epistemic modal “water might be XYZ,” even though water is not XYZ in this situation. This approach might work, but it is effectively a hybrid of situation semantics and two-dimensionalism, and it is the two-dimensional component that is doing the work in solving the problem of EPMIs. So two-dimensionalism remains the most promising candidate for solving the problem of EPMIs.

It is worth noting that Kratzer seems to implicitly accept something like a two-dimensional situation semantics in [15]. Kratzer considers the following scenario: a bear is seen in Glacier National Park in August 2006. One observer utters the sentence “the bear might be a grizzly.” Kratzer suggests that the description “the bear” can be heard as an attributive rather than referential description. On the referential reading, “the bear” picks out the actual bear seen, and if that bear is not a grizzly then it will turn out that it is metaphysically impossible for that very bear to have been a grizzly. So in order for “the bear might be a grizzly” to be true, we need to be able to hear “the bear” attributively (i.e. “the bear, whoever he is,
might be a grizzly"). Kratzer’s suggestion is that the sentence “the bear might be a
gizzly" should be interpreted as quantifying over possible situations and is true if
there is a possible situation 1) which is accessible from the actual situation and 2)
which contains a grizzly bear. Accessibility is defined here as a relation between “two
situations $s$ and $s'$ in a context $c$ [that holds] just in case $s$ and $s'$ are equivalent
with respect to the information available in $c$, that is, whatever evidence about $s$ is
available in $c$ isn’t specific enough to distinguish between $s$ and $s'$." So what we’re
looking for is a situation $s'$ which is indistinguishable from the actual situation $s$
that contains a grizzly bear. But note that the bear in $s'$ need not be the very same
bear as the bear in $s$. Note, also, that the expression “that bear is a grizzly" could
be false when uttered in $s$ but true when uttered in $s'$. Kratzer’s account allows
the values of certain contextually-sensitive parameters to shift in a way which is
characteristic of the sort of two-dimensionalism I favor. This will be discussed in
more depth in chapter 2.

1.5.5 Dynamic Semantics

Dynamic semantics is intended to model the way in which sentences change the
context in which they are used. Rather than giving meanings in terms of truth
conditions, the dynamic view takes the meaning of a sentence to be the way in which
the sentence changes the conversational context. Here is a sketch of one formulation
of dynamic semantics (adapted from [9]). We begin with a collection of information
about the world. This is represented as a set of possible worlds, namely those worlds

\[25\] The basic formulation I present here is a simplified version of the one presented by
Groenendijk, Stokhof, and Veltman. Their formulation also involves tracking “discourse
information," which is information about past states of the discourse. Discourse information
is irrelevant for my purposes and so I have adjusted the formulation to do away with it.
which are consistent with the available information. Call this the information set. Declarative sentences function to add to the information about the world, and do so by changing the information set.

These changes come about when new information is accepted. If we accept a new piece of information, the information set will typically shrink (assuming the new information was not entailed by the information we previously had available to us). Similarly, if information is removed the set of worlds will typically grow. So on this view the meaning of each sentence can be viewed as a function from information sets to information sets, with the first set corresponding to the available information before the new information is accepted and the second set corresponding to the available information after the information is accepted.

A typical treatment of epistemic modals within this framework is as follows: if the information expressed by $s$ is consistent with the available information, then the sentence “Possibly $s$” results in no change to the information set. If it is not consistent with the available information, then the sentence “Possibly $s$,” if it were accepted, would return the null set (which prompts us to reject the sentence “Possibly $s$” rather than accept it).\footnote{This treatment can also be found in \cite{24}.}

Dynamic semantics might seem like a promising alternative to traditional truth-conditional semantics for dealing with the problem that concerns us here, but it ultimately fares no better. As described above, the dynamic framework starts with a set of worlds that are consistent with the information available to the various
speakers in the conversation. Utterances change this set of worlds; the way in which an utterance changes this set is thought of as the meaning of the utterance. The sentence “it might be that s" does the following: if there’s a world in the set in which s is true, then nothing changes; otherwise, the set becomes the null set. But there are no worlds in which water is XYZ, so there cannot be a world in the information set in which water is XYZ. So if I say “water might be XYZ," it will always return the null set, whatever information we have available. So we see that the problem arises again.

We might allow worlds to validate an epistemic modal differently than they validate a metaphysical modal, but this would be combining two-dimensionalism and dynamic semantics, and two-dimensionalism would be doing all the work.

1.5.6 Our Prospects

The upshot of this section is this: if you’re going to use worlds (or situations) to account for epistemic modality, two-dimensionalism is the most promising extant approach. If you’re not going to use worlds for epistemic modality, then you have two choices. Either you use worlds for other kinds of modality, and say that epistemic modality is special, in which case you have to account for the semantic data provided by Kratzer and others that suggests that epistemic modals work like other kinds of modals. Alternatively, you can try to use something other than worlds (or situations or sets of worlds) for all modals. Obviously, there might be some alternative to these proposals available, but any new approach will have to be evaluated when it is presented.
Since the possible worlds framework has been successful thus far in philosophy and linguistics, I will explore the two-dimensional alternative. In the next three chapters, I will examine two different two-dimensional accounts of epistemic modality. First, I will present my preferred version, on which epistemic modals function more-or-less like non-epistemic modals, but do so under a context-shifting operator that changes the extension of the problem-causing words (i.e. natural kind terms, names, and the like). I will then consider a competing proposal put forth primarily by David Chalmers. I will argue that my preferred account turns out to be the most promising account of epistemic modality for dealing with the problem that concerns us here.
Chapter 2

The Metapropositional Approach

In Lecture III of [17] Kripke writes:

What, then does the intuition that the table might have turned out to be made of ice or of anything else, that it might even have turned out not to be made of molecules, amount to? I think it means simply that there might have been a table looking and feeling just like this one in this very position in the room, which was in fact made of ice. In other words, I (or some conscious being) could have been qualitatively in the same epistemic situation that in fact obtains, I could have the same sensory evidence that I in fact have, about a table which was made of ice.

Putnam offers a similar suggestion in [26]:

...it is epistemically possible that pencils could turn out to be organisms. It follows that pencils are artifacts is not epistemically necessary in the strongest sense and, a fortiori, not analytic.

Let us be careful, however. Have we shown that there is a possible world in which pencils are organisms? I think not. What we have shown
is that there is a possible world in which certain organisms are the epistemic counterparts of pencils.

In this chapter I will detail a two-dimensional account of epistemic possibility that accommodates the intuitions of Kripke and Putnam, insofar as this is possible. The resulting theory will be counted a success if it adequately explains epistemic possibility. Ideally, this will be accomplished without committing us to any new kinds of entity and without requiring us to repudiate semantic externalism entirely. We will call this the metapropositional approach, because it holds that the bearer of the epistemic modal properties is a metaproposition (a function from worlds to propositions), rather than a proposition (a function from worlds to truth values).

Intuitively, a metaproposition is supposed to represent something like a Kaplanian character. Sentences with indexicals can express different propositions in different contexts, and the relation between the indexical and its contribution to the proposition is expressed is captured by the indexical’s character. In the same way, the metaproposition is intended to capture the relation between the world in which a sentence is produced and the proposition it expresses. The analogy between my view and Kaplan’s is discussed in more detail in §3 below. The approach sketched out in this section owes a debt to Kaplan’s account of indexicals from [12].

\[^1\]It is worth noting that the term “proposition” is used in many different ways. Here is a partial catalogue: propositions have been described as meanings, as sets of worlds (or as determiners of sets of worlds), as the objects of belief, desire, and other attitudes, and as the primary bearers of truth-value. I will use the term as follows: the proposition expressed by a sentence determines a function from worlds to truth values.
2.1 A Motivational Example

Alice and Bob are talking about water. They know that water is the clear liquid typically found in lakes and rivers. They know that (pure) water is clear, tasteless, and odorless. However, they do not know that water is H\textsubscript{2}O. Alice says, “For all we know, water might be XYZ.” Bob agrees. Given what Alice and Bob know, it seems like Alice has said something true. Traditionally, we would say that Alice’s sentence expresses the following proposition:

1) $\Diamond_E (\text{Water is XYZ})$

The question arises: how should we interpret the $\Diamond_E$ operator? Assuming Kripke and Putnam are correct about the semantics for “water” and other natural kind terms, if we say that $\Diamond_E (\text{Water is XYZ})$ is true if there is an accessible (metaphysically) possible world \( w \) such that

1. \( w \) is consistent with Alice and Bob’s knowledge
2. in \( w \), water is XYZ

we will find that there are no such worlds and hence that Alice said something false. But we want to say that Alice said something true. So we have an EPMI on our hands.

Given the discussion in the previous chapter, it seems clear that we have to say that Alice’s sentence doesn’t express 1) above. I want to suggest that Alice’s sentence is

\footnote{A note on notation: $\Diamond_E$ indicates epistemic possibility. $\Box_E$ indicates epistemic necessity. $\Diamond_M$ and $\Box_M$ indicate metaphysical possibility and necessity, respectively.}
instead associated with a metaproposition, and that it is this metaproposition which accounts for the truth of Alice’s claim.\footnote{One alternative to the account I am developing here is that Alice’s sentence expresses a different proposition instead of or in addition to 1. Chalmers’s account is the best example of this alternative. His account will be discussed in chapter 4.} On this account, Alice’s sentence expresses a truth if there is a world $w$ in which she could utter the same (disambiguated) sentence and have it express a proposition which would be true when evaluated in $w$.\footnote{This formulation is missing a few refinements which will be introduced later in this chapter. In particular, we will have to hold fixed the language in which the sentence is produced and the basic content of the expression. “Basic content” is technical notion which will be defined in §4 and developed in more detail in chapter 3.} It seems reasonable to suppose that there are such worlds. Twin Earth is one of them; as Putnam describes the scenario, “water” picks out something XYZ when used by a Twin Earth native. If this is correct, the sentence “water is XYZ” expresses a different proposition when produced on Twin Earth than it does when produced here. In fact, the proposition it expresses when produced on Twin Earth (roughly, that XYZ is XYZ) is a necessary truth. So if the sentence “water is XYZ” is produced on Twin Earth, it expresses a proposition which is true when evaluated in Twin Earth, and hence Alice’s original sentence (“water might be XYZ”) expresses a truth here on the actual world.

2.2 THE APPROACH ITSELF

In this section I will present a metapropositional account of epistemic modality using the two-dimensional semantic framework and show how it applies in a number of examples.
2.2.1 The Formal Machinery

On my theory, each expression is associated with a metaproposition, which is a function $f(w)$ from worlds in which the expression is produced (or worlds of production and abbreviated $W_P$) to propositions. Propositions are functions from worlds of evaluation (abbreviated $W_E$) to truth values. $W_P$ includes only those worlds in which expressions have the same basic content (in the sense given later) as they do in the actual world. $W_E$ is the set of all worlds. $W_P$ is thus a subset of $W_E$. We can represent the metaproposition using the familiar two-dimensional array with the worlds in $W_P$ along the vertical axis and the worlds in $W_E$ along the horizontal, like so:

<table>
<thead>
<tr>
<th></th>
<th>$w$</th>
<th>$x$</th>
<th>$y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When we adopt this approach, the entire array represents the metaproposition. Each row represents the proposition expressed by the expression when produced in the corresponding world in $W_P$.

All expressions are evaluated at a pair of worlds $(w_1, w_2)$, where $w_1$ represents the world of production and $w_2$ represents the world of evaluation. A disambiguated sentence $s$ is true at $(w_1, w_2)$ when associated with the metaproposition $f(w)$ iff $f(w_1)$ is true at $w_2$. We can initially define the modal operators as follows:
• $\Diamond_M s$ is true at $(w_1, w_2)$ when associated with the metaproposition $f(w)$ iff there is some world of evaluation $v$ such that $f(w_1)$ is true at $v$.

• $\Diamond_E s$ is true at $(w_1, w_2)$ when associated with the metaproposition $f(w)$ iff there is some world of evaluation $u$ such that $f(u)$ is true at $u$.

The necessity operators are defined in the traditional way. Given these truth definitions, we can see that a sentence $s$ is metaphysically possible when produced at $w$ iff there is some point along the relevant horizontal row at which $s$ is true. $s$ is epistemically possible iff there is some point along the diagonal at which $s$ is true. $s$ will be metaphysically necessary when produced at $w$ if $s$ is true at every point along the relevant horizontal row. Finally, $s$ will be epistemically necessary iff $s$ is true at every point along the diagonal (subject to a few restrictions, described below).

<table>
<thead>
<tr>
<th></th>
<th>$w$</th>
<th>$x$</th>
<th>$y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w$</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>$x$</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>$y$</td>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

For example, the diagram above represents an sentence $s$ with the following properties: $s$ is metaphysically necessary when produced at $w$, metaphysically impossible when produced at $x$, and metaphysically contingent when produced at $y$. $s$ is also epistemically contingent, because it is true at $(w, w)$ and false at $(x, x)$. If it were true at $(x, x)$, it would be epistemically necessary, since it would then be true at every point along the diagonal.
However, that last claim will not quite capture the behavior of actual epistemic modals. To do that we need to recognize two further constraints. First, we have the basic content constraint: $W_p$ includes only those worlds in which expressions retain the same basic content. I will discuss what is meant by “basic content" in more depth in the next section and the next chapter. For now, it suffices to say that the basic content of an expression in a given language determines the intension of the expression at every world of production, where intensions are understood in the traditional way as functions from worlds (of evaluation, in this case) to extensions.\footnote{The language itself also needs to be held constant. If the morpheme “gap" is used by an English speaker it denotes a gap; in Polish, this morpheme denotes an onlooker. Consider the English modal sentence “For all I know, the gap between the door and the platform might be six inches across." When we evaluate this modal, “gap" will retain its English basic content. Similarly, we require that the sentence be properly disambiguated. The sentence, “For all I know, that might be a ball" could be about a formal dance or about a round children’s toy (among other interpretations). We must disambiguate in order to determine which sense of “ball" is in play and thus which basic content is relevant.}

Holding basic content fixed means we only look at those worlds of production in which the function from world of production to intension is the same as it is here.

The second constraint is that the only worlds we need to examine when evaluating epistemic modals are those that are consistent with what is known. We introduce this constraint to capture the epistemic nature of epistemic modals. We restrict $W_p$ to those worlds consistent with what is known. How is this restriction to be understood? Often, this is done by taking what is known to be a set of propositions, each of which is or determines a set of worlds. The intersection of these sets of worlds is the knowledge base. We will adopt this approach.\footnote{For our purposes, all that matters is that the object of knowledge determines a set of worlds. We need not get embroiled in arguments regarding the object of knowledge as long as it has this property. If the objects of knowledge are propositions, then that satisfies our requirements. For simplicity I will continue to speak as though the objects of knowledge
propositional in this way is consistent with the claim that propositions are not the object of epistemic modal operators. Ultimately, the semantic machinery that I deploy to address epistemic modals does not depend on any particular characterization of knowledge beyond its functional aspect. As long as what is known (whatever it is) determines a set of worlds which can function as a knowledge base, my account should be fine.

With the notion of a knowledge base in hand, we can revise the truth definition for the epistemic possibility operator like so:

- \( \Diamond_E s \) is true at \((w_1, w_2)\) when associated with the metaproposition \( f(w) \) iff there is some world of evaluation \( u \) such that 1) \( f(u) \) is true at \( u \) without change in basic content and 2) \( u \) is in the knowledge base.

Again, the epistemic necessity operator is defined in the traditional way.

### 2.2.2 The Machinery In Action

To see how the proposal developed in the previous subsection functions, I will present three sample cases. In all three cases, the knowledge base will be shared are propositions and propositions are sets of worlds, but this is not strictly speaking a requirement of the theory.

Notice that while this restriction involves the epistemic status of some agent or collection of agents, it does not invoke any preexisting understanding of epistemic possibility. To determine what worlds go into \( W_P \), we attend to the epistemic status of the relevant agent or agents in a single world—that is, what the agent or agents know about the world they occupy and what further facts are consistent with that knowledge. But this is as it should be, since epistemic possibility and epistemic modality are relative to what is known by the relevant agent or agents.
by all members of the conversation for simplicity. More complicated cases involving unequal knowledge will be discussed in chapter 5.

CASE 1: WHERE'S CAROL?

Alice and Bob are in Washington, DC wondering where their friend Carol is. Carol is in fact in Zurich, but neither Alice nor Bob know this; in fact, all they know is that she is not in Washington. Alice says to Bob: "Carol might be in Cleveland." On my account, Alice has said something true iff there is a world $w$ such that 1) the English sentence "Carol is in Cleveland" has the same basic content as it does in the actual world and expresses a proposition which is true when produced in and evaluated at $w$ and 2) $w$ is in the knowledge base. The knowledge base in this case consists in those worlds in which Carol is not in Washington. There is indeed such a world: namely, the world in which Carol is in Cleveland. So Alice has said something true.

CASE 2: WHO'S CAROL?

Alice and Bob have noticed that their mutual friends Carol and Eve are never seen together. Alice has begun to suspect that Carol and Eve might in fact be the same person. She approaches Bob and voices her suspicion, saying "Carol might be Eve."

Here is what Alice and Bob know about Carol: Carol has red hair. Carol is short.

---

8The knowledge base will actually be smaller than indicated here, as Alice and Bob have all sorts of incidental knowledge which is not relevant to the evaluation of the modal sentence. For example, if Bob is wearing a red shirt (and is not red-green colorblind), then he will likely know he is wearing a red shirt. So the knowledge base will include all the worlds in which Carol is not in Washington and Bob is wearing a red shirt. But these other bits of knowledge are irrelevant and so have been glossed over.
Carol works at the bank. Here is what Alice and Bob know about Eve: Eve has red hair. Eve is short. Eve works for the federal government. Eve looks very much like Carol. Let us stipulate that Carol and Eve, despite their similar appearance, are different people. Then (given that identity and distinctness are necessary), it is necessarily false that Carol is Eve. Nonetheless it seems like Alice has said something true, so we have an EPMI.

On my account, in order for Alice to have said something true, there must be a world $w$ such that 1) the English sentence “Carol is Eve” expresses a proposition which is true when evaluated in $w$ and 2) $w$ is in the knowledge base. This time, the knowledge base will include those worlds in which Alice and Bob 1) have a friend who is a short person with red hair who works at a bank and answers to “Carol,” and 2) have a friend who is a short person with red hair who works for the federal government and answers to “Eve.” These people need not be distinct. So the world in which the person who answers to “Carol” is also the person who answers to “Eve” is the world in which “Carol is Eve” expresses a true proposition, and hence Alice has said something true.

It is important to note that this world is not a world in which Carol is Eve. Rather, it is a world in which at least one of the names “Carol” and “Eve” denote someone other than Carol and/or Eve and hence the expression “Carol is Eve” expresses a different proposition, though its basic content is unchanged. Proper names like “Carol” are still rigid designators, because once the world of production is set they always refer to the same thing (if it exists) in each world of evaluation.

---

9I will discuss how we determine the basic content of singular terms in chapter 3.
Case 3: Could Carol be a Cylon?

Alice and Bob are philosophers. They wonder if their mutual friend Carol has any necessary properties. In particular, they wonder if Carol must be made of matter. Alice says: “For all we know, it might be that Carol might have been an artificial biological life form.” Let us stipulate that the outermost modal in this case is epistemic, but the innermost modal is metaphysical. Let us further suppose that Carol must not be an articial being but Alice and Bob do not know this to be the case. On my theory, Alice has said something true if there is a world of production in the knowledge base in which “Carol might have been an artificial biological life form" expresses a true proposition. Since Carol is not an artificial biological life form, and Alice and Bob know this, we must look for a world \( w \) in which “Carol" denotes an object which is not an artificial biological life form in \( w \) but in which “Carol is not a artificial biological lifeform" does not express a necessary truth. Here, we will run into difficulties. We require “Carol" to have the same basic content in \( w \) as it does in the actual world. Since “Carol" is a rigid designator in the actual world, it will have to be a rigid designator in \( w \) as well (though it may not designate the same object in \( w \) as it does in the actual world). This means that we will not find a world in \( W_P \) that has the required properties and that Alice is mistaken in saying that it might be that Carol might have been an artificial being, despite the fact that Alice doesn’t know whether or not she might have been an artificial being. Alice’s claim is false; it is not epistemically possible for her that Carol is an artificial biological life form.

This may seem counterintuitive, but in fact this result is what we should expect. First, it is important to note that one traditional gloss for epistemic possibility is
problematic. Traditionally, the statement “s is epistemically possible for some agent A” is taken to mean something like “A doesn’t know that not-s.” This rendering entails that s is epistemically necessary iff s is known. However, epistemic necessity and knowledge have different properties and obtain in different circumstances. For instance, if the epistemic necessity operator is defined as a quantifier over points of evaluation (including worlds in a one-dimensional approach and pairs of worlds for the two-dimensional approach) that uses Kripke models, then epistemic necessity will be closed under entailment. Knowledge, on the other hand, is not.

Often, the problems involved with identifying knowledge with epistemic necessity are addressed by treating epistemic necessity as an idealization of knowledge—say, as knowledge for ideal reasoners or some such. I have a different strategy. Rather than treat epistemic necessity as an idealization of knowledge, I offer a different (and, I think, superior) intuitive rendering of epistemic necessity. My proposal is to intuitively treat epistemic necessity as follows: s is epistemically necessary for A iff s is known or is entailed by what is known by A. s is known by A if A knows the proposition expressed by s in the actual world. This entails that s is epistemically possible for A iff not-s is not logically entailed by what A knows, or equivalently that s is consistent with what A knows. One reason to prefer my approach is that it permits us to treat epistemic necessity as an independent object of study.

\[ ^{10} \text{I take no position on whether or not knowledge is closed under known entailment or under some other sort of operation. It is clear, though, that knowledge is not closed under entailment } \text{simply}\text{ because this would entail logical omniscience.} \]

\[ ^{11} \text{I prefer to define epistemic necessity and possibility relative to knowledge bases, not agents, but I am using agents here to bring out the issues that arise when we associate epistemic necessity with knowledge in the traditional way. A statement can be said to be epistemically necessary relative to an agent if it is epistemically necessary relative to the knowledge base of all worlds compatible with everything that the agent knows.} \]
with interesting connections to knowledge without requiring us to adopt a rarefied conception of knowledge itself.

Furthermore, my preferred gloss on epistemic possibility is itself a fairly common gloss on epistemic possibility, but the fact that it is not equivalent to the other traditional interpretation given above has not been widely discussed. So there is precedent for adopting my intuitive rendering of epistemic modals. It is important to note that this is just an intuitive gloss on epistemic modality, and any formal account may deviate from the intuitive notion at various points. This is to be expected and perhaps even desired. After all, if we had perfectly clear intuitions on epistemic modality we might not need an explicit analysis, and we likely wouldn’t find any hard cases where it is unclear whether or not a given expression is epistemically possible or necessary.

Moreover, in investigating our intuitions regarding epistemic modality we might find that they are confused and contradictory; in this case no consistent theory will adequately capture all our intuitions, but we should not want such a theory. My formal account of epistemic modality is intended to stick close to the second, superior intuitive gloss given above, but the theory may not capture the intuitive gloss completely. Cases where the theory and the intuitive notion come apart should be noted but in general, one should not assume that any individual case not covered by the theory should refute it.

When we understand epistemic necessity and possibility in this way, we can see that it is not consistent with Alice’s knowledge that Carol might have been an
artificial being. Alice knows that Carol is a natural being, and this entails that it is metaphysically necessary that Carol is natural (assuming that Kripkean views of metaphysical necessity are correct). Hence it is epistemically necessary for Alice that Carol is necessarily natural. It is instructive to contrast this with the case in which Alice doesn’t know whether or not water is H₂O and says “For all we know, water might be XYZ." “Water" is still a rigid designator in this scenario, and it is a rigid designator in every world in W₁. However, the crucial difference here is that Alice doesn’t know that water is H₂O, and so even though it is metaphysically necessary that water is H₂O, this is not logically entailed by her knowledge. We will see the same result if we consider the case in which Alice doesn’t know if Carol is natural or not, and says “For all we know, Carol might be artificial." It is the nested metaphysical modal that leads to the seemingly counterintuitive result.

In this case, “For all we know, it might be that Carol might have been an artificial biological life form" looks like an EPMI, but it is not. Alice and Bob know that Carol is a natural being, and hence “Carol is a natural being" is epistemically necessary for them given the interpretation of epistemic necessity given above. There is a peculiar feature of the sentence “Carol is a natural being" that deserves further attention. Given that Carol is a natural being, it follows from the basic content of the sentence that it is metaphysically necessary that Carol is a natural being. Since we hold basic content fixed, there is no way for the sentence to produce a contingent proposition in any world of production (though it could produce a necessarily true proposition or a necessarily false proposition).
2.2.3 Metaphysical Stability

Some sentences have an interesting property: when we hold their basic content fixed, a sentence that expresses a contingent proposition in the actual world will express a contingent proposition in any world of production, and a sentence that expresses a noncontingent proposition in the actual world will express a noncontingent proposition in any world of production. Let us call sentences with this property metaphysically stable sentences. Metaphysically stable sentences (like “Carol is a natural being”) obey certain logical principles when we combine epistemic and metaphysical necessity. In particular, we will see that for any metaphysically stable sentence $s$, $\Box_M s \rightarrow \Box_E \Box_M s$ and $\neg \Box_M s \rightarrow \Box_E \neg \Box_M s$.

This leads to an interesting conjecture: that all or almost all sentences are metaphysically stable. Cases like case 3 above give some intuitive support for this conjecture. Reflecting on the nature of basic content also gives us reason to think the conjecture is true. A term is a rigid designator in a given world of production if its reference is the same in every world of evaluation (in which its referent exists). This leaves open the possibility that a term might be a rigid designator in one world of production and a nonrigid designator in another. However, when we hold basic content fixed, this possibility seems to be closed off. Basic content determines the intension of an expression in every world of production. It is thus plausible that whether or not a given term is a rigid designator is a function of its basic content. If this is the case, when we hold basic content fixed, rigid designators will function as rigid designators in all worlds of production (although they may not rigidly designate the same thing in each world of production); similarly, nonrigid designators will be nonrigid. For instance, a sentence like “water is $\text{H}_2\text{O}$” which expresses a necessary
truth when produced in the actual world because “water” rigidly denotes H$_2$O on Earth. When produced on Twin Earth, the sentence produces a necessary falsehood because “water” rigidly denotes XYZ on Twin Earth. In both cases, though, “water” is a rigid designator. So the sentence “water is H$_2$O” need not always produce a necessary truth; however, it must express either a necessary truth or a necessary falsehood since “water” and “H$_2$O” are rigid designators in any world of production. Similarly, “The tallest man in the room has brown hair” expresses a contingent truth when produced in this world, and since “the tallest man in the room” is non-rigid regardless of the world of production we should expect the sentence to express contingent truths wherever it is uttered.

There may be sentences that are not metaphysically stable, but examples of such sentences are surprisingly hard to come by. Here is a possible example (inspired by Stalnaker, [31]): consider the sentence “Sherlock Holmes does not exist.” On the semantic treatment of existential statements of the sort advanced by Kripke and Stalnaker, this sentence expresses a necessary truth when uttered in the actual world. It expresses a necessary truth because “Sherlock Holmes” has no referent, and hence there is no object to which we can ascribe existence in some other possible world. However, it seems plausible that there are worlds of production in which “Sherlock Holmes” refers. In those worlds, “Sherlock Holmes does not exist” expresses a contingent proposition, since it is about a contingently existing being.

\[\text{[12]}\] This claim, made by Stalnaker, Kripke, and others, is supposed to be a consequence of their account of singular terms. If it is true, then it could give us a counterexample to my conjecture regarding metaphysical stability. If it is false, then it will not be possible to use negative existentials to construct a counterexample in this way.
The claim that there are worlds of production in which “Sherlock Holmes” refers is, perhaps, questionable. One might think that part of the basic content of “Sherlock Holmes” is that it names a fictional character. However, this seems unlikely. The fact that “Sherlock Holmes” happens to name a fictional character in the actual world is involved in determining what propositions sentences involving “Sherlock Holmes” express in other worlds of production. There does not seem to be anything special about “Sherlock Holmes” that distinguishes it from names that successfully refer. The names “Sherlock Holmes” and “Arthur Conan Doyle” have relevantly similar syntactic properties, for example, and we don’t typically make a practice of distinguishing between names of fictional characters and names of actual objects in our everyday discourse.

Even if the fact that “Sherlock Holmes” names a fictional character is part of the basic content of the name, it does not mean that we haven’t found counterexamples to the conjecture. We can introduce descriptive names without knowing whether or not they refer. For instance, consider the plot of the 1985 film *Clue*. In the film, the characters are invited to a dinner party by Mr. Boddy. After their arrival, Mr. Boddy is killed along with several members of the house staff and (seemingly) unrelated passers-by. The characters then attempt to determine the identity of the mysterious killer. In the film, they do not use a descriptive name for the killer, but they might have done so; let’s imagine they used the name “John” as a descriptive name for “the (unique) actual killer.” In the true ending of the film, it is revealed that no one person was responsible for the murders; each of them was committed by a different person. So the name “John” would not refer in this case. Nonetheless it is not part of the basic content of “John” that it does not refer. If there had been only one killer,
“John” would have referred to him or her, and the characters understand the name in this way.

All of these potential counterexamples rest on a semantic theory on which “Sherlock Holmes” fails to denote any object when produced in the actual world. One strange consequence of these kinds of theories is that true and false negative existential statements behave differently from one another. True negative existentials are necessarily true on this view, while false negative existentials are contingently false. We might take that as evidence against these semantic views. Indeed, I am not committed to such a theory, and it is possible to have a broadly Kripkean view of the semantics of names and natural kind terms while rejecting the claim that “Sherlock Holmes” fails to denote an object. One might alternately take “Sherlock Holmes” to (rigidly) denote a possible object that is merely possible in this world, but is actualized in other worlds. On this view cases like “Sherlock Holmes does not exist” will not provide counterexamples to the conjecture that all or almost all sentences are metaphysically stable. I am not committed to either semantic theory and so I take no position on whether or not sentences like “Sherlock Holmes does not exist” are metaphysically stable.

Despite the potential availability of counterexamples, the conjecture that all or almost all sentences are metaphysically stable may still hold. Apart from sentences involving non-referring names, it is surprisingly hard to find examples where the conjecture clearly fails. If non-referring names can be used to construct a counterexample, then presumably it would also fail for non-referring natural kind terms (e.g.
terms for fictional materials, such as the ice-nine of *Cat's Cradle*), for much the same reason that it fails for non-referring names. But it seems to hold for other sentences.

**Other Cases**

There are other counter-intuitive cases that warrant some analysis. Consider the following example: “For all Andrew Wiles knew in 1992, Fermat’s theorem might have been false.” This might strike us as true, but given the understanding of epistemic necessity and possibility given above (and given certain plausible assumptions about Wiles’ knowledge), we should expect this sentence to be false. Fermat’s theorem is presumably a necessary truth, and while Wiles did not know that Fermat’s theorem was true in 1992, it was likely a consequence of facts that he did know, as Wiles had been working on the proof for years. So we should expect that Fermat’s theorem was in fact epistemically necessary for Wiles in 1992.

It might be thought that this interpretation of epistemic necessity and possibility breaks the connection between these technical notions and the English words used to express them, such as “might” and “must.” This is not the case. Notice that we intuitively judge the sentence “Given what Wiles knew in 1992, Fermat’s theorem must have been true” to be true, as predicted by my preferred intuitive gloss of epistemic possibility. A natural reading of this sentence is that it is true if Fermat’s theorem is a consequence of Wiles’s knowledge in 1992. So our intuitions are inconsistent

---

13Andrew Wiles first announced that he had proved Fermat’s Last Theorem in 1993. While this version of the proof was flawed, his subsequent attempt in 1995 was successful.
here. Any systematic approach to epistemic modality will thus run into counterintuitive cases. There are no doubt other uses of “must” and “might;” nonetheless, the interpretation given here represents one significant use.\footnote{One other use of “must” is briefly discussed in chapter 5.}

2.3 The Metapositional Account in Context

My theory is in many ways analogous to Kaplan’s account of indexicals in \footnote{Note that this fits well with Putnam’s suggestion that natural kind terms have an indexical component. If Putnam is right, that would explain why natural kind terms behave as they do. I am inclined to think Putnam’s claim is correct, but nothing here rests on whether or not that is the case.}[12]. On Kaplan’s account, all utterances occur within a context of utterance and are evaluated with respect to a circumstance of evaluation. This context includes (at least) the speaker, the audience, and the time and place of utterance. Kaplan also maintains that all utterances have a character, which is a function from contexts of utterance to intensions. For many terms, this character returns the same intension regardless of context of utterance. Indexicals are unusual in that they have different intensions in different contexts of utterance; for example, the intension of the word “I” (in its normal use) depends on the identity of the speaker. Given this account, we can take the world in which the utterance was produced to be part of the context of utterance. Names and natural kind terms in my account function like indexicals do in Kaplan’s account\footnote{} except they are not (typically) sensitive to the common contextual variations like changes in the speaker’s identity, the time of utterance, and so on. Rather, they are sensitive to changes in the world of production. Since we cannot change which world we’re in, we can’t normally change the index for the
world of production, and hence the interpretation of names and natural kind terms does not vary due to variations in the world of production index.

However, when embedded under an epistemic modal, we evaluate the name or natural kind term in the embedded expression as though it was produced in another world. This makes the epistemic “might” and “must” into something like what Kaplan calls monsters (and Evans calls context-shifting operators). Monsters are operators that manipulate one of indices for the context of utterance rather than the circumstance of evaluation. Most operators are not monstrous; for example, “someday” is a non-monstrous operator. A sentence like “someday I will be rich” is true (very roughly) if there is a future time in which I am rich. So in this case, we go looking for a circumstance in which I, Dan Quattrone, am rich, and hence in which the sentence “I am rich” comes out true when uttered by me. However, note that the indexical “I” in “I am rich” still refers to the original speaker. If I say “someday, I will be rich,” the sentence will not be true if there is a time in which another person could truly utter the sentence “I am rich.” This is what distinguishes monsters from non-monsters. Monsters affect the index for the context of utterance and in so doing affect the value of indexicals.

Kaplan argues that there are no monsters in English and we could not add any to the language. His argument fails. Kaplan’s argument proceeds as follows: if there is any English expression that functions as a monster, it would be an expression like “In another context…” But consider the sentence “In another context, I am tall.” Intuitively, we understand this sentence to be true if there is a context in which I would be judged to be tall. But if “In another context…” was truly a monster,
then we should expect “In another context, I am tall” to come out true if there is a context containing a tall speaker, who need not be me. That is, most speakers would not judge “In another context, I am tall” to be true because there is a context in which “I am tall” is uttered by Robert Wadlow. We might think that, while “In another context...” does not shift the speaker index, it does shift some other feature of the context. I am in fact of roughly average height for an adult male, but I am tall compared to the average height of a child. If I were to say “In another context, I am tall,” we might judge it to be true because “tall” is itself context-sensitive (though it is a gradable adjective rather than an indexical), and because there are contexts in which the relevant height standard is different. However, the shift here is actually part of the circumstance of evaluation. We can see this by considering cases where there is a large gap between context of utterance and circumstance of evaluation. Suppose I record myself saying “I am tall.” We stipulate that the relevant standard for height at the time I initially make the recording is such that I count as tall. Years later, average heights have increased, and with it the standards for tallness. The recording is played in this situation. Intuitively, we would judge my utterance to be false when we hear it. This suggests that “tall” is (at least on some occasions; see below for more discussion) sensitive to features of the circumstance of evaluation rather than features of the context of utterance. So “In another context...” is not a monster. But if “In another context...” is not a monster, then surely nothing else could be.

I agree with Kaplan that “In another context...” and expressions like it are, at least initially, the most plausible candidates for monsterhood in English. I also

---

16 Robert Wadlow is listed in the Guinness Book of World Records as the world’s tallest person, measured at 8 feet 11.1 inches just before his death.
agree that such expressions are not monstrous. However, it does not follow that there are no monsters in English. First, note that Kaplan's argument is an inductive argument with a very small initial sample, and so needs further support. Second, and more importantly, there is empirical evidence suggesting that English does contain monsters. In [27] Philippe Schlenk has argued that expressions like “two days ago" and “in two days" display monstrous behavior, i.e. shift the context index. Consider the following example (a modification of one of Schlenker’s examples):

Bob met Alice exactly one year ago. In two days, she would be dead.

The indexical phrase “in two days" in the second sentence can be heard as referring to two days after Bob met Alice. However, the context of utterance is one year after Bob met Alice. If Kaplan were right and there were no monsters in English, it should be impossible to hear the phrase “in two days" as referring to anything but two days from the time of utterance. Similar phenomena occur in when we translate these expressions into French, which lends credence to the notion that these phenomena are not peculiar to English.

Schlenker also shows that the first-person pronoun can function monstrously in other languages. In particular, he shows that the first-person pronoun in Amharic can pick out persons other than the speaker in disquotational contexts, which suggests that the operators that signal disquotation (e.g. “he said") can act as monsters in Amharic. So the Amharic sentence that would be lexically translated as “Alice said I am hungry" would be better read as “Alice said she is hungry." This gives us evidence for monsters in other languages, which makes the hypothesis that English has monsters of its own more plausible.
Consider another example.\footnote{Adapted from examples by Wayne Davis.}

Scenario: There is a picture of Superman hanging on the wall. Lex Luthor comes in and places a placard next to the sign reading “I am Clark Kent.”

In this scenario, Luthor is the person producing the sentence “I am Clark Kent.” If Kaplan is right and there are no monsters in English, then we should expect that the “I” in “I am Clark Kent” to pick out Luthor. But that is not what happens. Instead, we have a case where Luthor can put words in Superman’s mouth (as it were), and we hear the “I” as referring to Superman and not to the person who actually produced the expression. This case is unusual in that there is no explicit operator which causes the context index to shift; that said, the indexical is still behaving as though it were governed by a monster, which suggests that such behavior can be observed when there are explicit operators as well.

Finally, let us return to the case of “tall.” We have seen that “tall” is sometimes sensitive to features of the circumstances of evaluation rather than the context of utterance. But that does not show that it is entirely insensitive to features of the context of utterance. Consider, for example, the following examples.\footnote{Adapted from examples by Steve Kuhn.}

1. Had I taken human growth hormone, I would have been tall.

2. Had I lived among the Kalahari Bushmen, I would have been tall.

Both of these sentences seem, intuitively, to be true, but for different reasons. In the first case, the sentence is true using the ordinary standards for tallness. But
in the second case, the standards for tallness seem to shift; rather than use the standards for early 21st-century Americans, we adopt the standards appropriate for the Kalahari bushmen. This seems to be monstrous behavior.

Taken together, what these examples show is that while there are English expressions which we would expect to be monsters but are non-monstrous, there are also cases in which English indexicals behave as though they are governed by a monstrous operator. Admittedly, this does not show that epistemic modals are monsters; however, it does show that the hypothesis is not to be dismissed. Given that the hypothesis is consistent with the available evidence and offers considerable explanatory power, we should at least tentatively accept it as (likely to be) true.

So here’s where things stand: names and natural kind terms behave like context-sensitive terms in Kaplan’s account of indexicals. However, they are only sensitive to changes in the world of production. Epistemic modals functional analogously to monsters; they index for change the world of production in the context of utterance but leave the other features of the context untouched.

This approach also accords well with Putnam’s proposal in [26]. Putnam asks us to consider two worlds \( w_1 \) and \( w_2 \). \( w_1 \) is the actual world. \( w_2 \) is Twin Earth. In both worlds, we have a speaker pointing to a glass containing a clear, odorless, tasteless liquid (\( \text{H}_2\text{O} \) in \( w_1 \), XYZ in \( w_2 \)). He writes:

\[ \ldots \text{there are two theories one might have concerning the meaning of } \]
\[ \text{“water”}: \]
1. One might hold that "water" was world-relative but constant in meaning (i.e. the word has a constant relative meaning). In this theory, "water" means the same in $w_1$ and $w_2$; it's just that water is $H_2O$ in $w_1$ and water is XYZ in $w_2$.

2. One might hold that water is $H_2O$ in all worlds (the stuff called "water" in $w_2$ isn't water), but "water" doesn't have the same meaning in $w_1$ and $w_2$.

Putnam endorses the second option. If we understand "meaning" as Putnam does (that is, as a sequence of components including the syntactic and semantic markers that apply to the word, the stereotype associated with the word, and the extension) the metapropositional approach articulated here would provide the same result. Putnam's definition of "meaning," however, is not the same as my notion of basic content, since his definition makes the extension of a term part of its meaning, and the extension of a term is not part of its basic content in my sense. Thus, on the metapropositional view the analogues of 1. and 2. for basic content do not hold. Instead, "water" has the same basic content in $w_1$ and $w_2$ as 1. holds, and water is $H_2O$ in both worlds as 2. holds (i.e., "water is $H_2O$" when uttered here is true with reference to both $w_1$ and $w_2$). The sentence "water is $H_2O$" is false when uttered in $w_2$ and evaluated at $w_2$ because "water," when uttered in $w_2$, refers to XYZ, because the basic content of "water" is such that its intension produced in $w_1$ takes us to $H_2O$ at every world of evaluation and its intension when produced in $w_2$ takes us to XYZ at every world of evaluation. Note, though, that its basic content does not change, and that its basic content is fixed by the way we actually use the term. We
can see, I think, that the metapositional view accounts for the intuitions that Putnam wants to trade on.

In certain cases, this theory looks like a counterpart theory of the sort suggested by Putnam and discussed in chapter 1. In such a theory, the truthmakers for epistemic modals are always epistemic counterparts; while my view can avail itself of counterparts when they are available, it can account for cases in which counterparts are not to be found. The metapositional theory is thus not a counterpart theory but it does account for the intuitions behind Putnam’s suggestion. As I showed in chapter 1, counterpart theories face difficulties. The main source of these problems is that there are cases in which counterparts are not obviously available. Epistemic modals involving states of affairs (such as “it might be raining outside”) or involving nonexistent objects (“there might be an elephant in the room”) pose a challenge for counterpart theories. The metapositional theory can account for cases like these, though. If there is a world \( w \) in \( W_P \) such that the sentence “There is an elephant in the room,” when I utter it at \( w \) expresses something true at \( w \), then the original sentence will be true.

So the metapositional view has advantages over the counterpart theory suggested by Putnam. It also can account for the intuitions behind Putnam’s suggestion. Consider “water might not be \( \text{H}_2\text{O} \).” On the metapositional view, this is true if there is a world \( w \) in \( W_P \) in which “water is not \( \text{H}_2\text{O} \)" is true. And this will be a world in which “water” refers to something other than \( \text{H}_2\text{O} \). Since the only worlds in \( W_P \) are the worlds in the knowledge base, everything which is known will have to be true at \( w \), including facts about the phenomenal properties of objects in the world. So \( w \)
will contain a substance which has all the phenomenal properties of water, but which is not \( \text{H}_2\text{O} \). This substance is what Putnam would call the epistemic counterpart of water. The upshot of this is that it can be useful to use the language of counterparts as a heuristic, and I will do so when appropriate. We can use counterpart talk as a heuristic by adopting a sort of fictionalism about the counterpart theory; that is, we can take the counterpart theory to be a fiction, and hold that an epistemic modal \( s \) is true iff \( s' \) is true in this fiction, where \( s' \) is the counterpart theory’s translation of \( s \).

2.4 Basic Content

One constraint on the theory developed above is that \( W_P \) includes only those worlds in which words have the same basic content as they have in the actual world. In this section I will briefly discuss this constraint and present some reasons for thinking it plausible that there is such a thing as basic content in the relevant sense.

Basic content determines an expression’s intension at every world of production. We can think of it as a function from worlds of production to intensions, and intensions as functions from worlds of evaluation to extensions. We make use of this concept because we do not want epistemic possibility claims for a metaproposition \( p \) to depend on what the sentence expressing \( p \) says in some other language or under unusual linguistic conventions. At the same time, the basic content of a term cannot include the term’s extension. On the metapropositional view, the intension of a term (and, *a fortiori* its extension) depends on the world in which it is produced. Since epistemic modals behave like monsters and (effectively) change the index for the
world of production, it follows that terms embedded in an epistemic modal can have different extensions than they normally have. Since the basic content of a term is always held fixed (even when embedded in an epistemic modal), and the extension can sometimes vary, basic contents cannot include or be identified with extensions.

This constitutes a break with some semantic externalists, notably Kripke, who argues that the semantic content of a name or natural kind term is exhausted by its reference. My claim is that names and natural kind terms (and, in fact, almost all terms) have some semantic content where "semantic content" is broadly construed beyond their reference. Putnam’s externalism (at least as described [26]) is _prima facie_ compatible with this notion of basic content, since Putnam allows for semantic contents other than reference.

The notion of basic content is fairly intuitive. In Putnam’s original Twin Earth thought experiment, we might notice that if Oscar and his twin switched places, such that Oscar was on Twin Earth and his duplicate was on Earth, neither Oscar nor his twin would have trouble navigating their new worlds. If a Twin Earth native were to say to Oscar, “Please bring me a glass of water," Oscar would be able to satisfy this request, even though he might think he’s providing a glass of H$_2$O. Twin Earth natives interacting with Oscar would judge him a competent user of the language.

However, if Oscar were transported to a world in which “water" denoted gasoline, he might not be so fortunate. Suppose that the world Oscar finds himself on is chemically identical to Earth. However, in this world, “water" denotes gasoline and another word denotes H$_2$O (and XYZ is nowhere to be found). Call this world Gas
Earth. If a Gas Earth native were to say to Oscar “Please bring me a glass of water,”
Oscar would not be able to satisfy the request (at least not until he became assimil-
ated to the peculiar linguistic practices of Gas Earth). Oscar’s linguistic behaviors
will be judged by Gas Earth natives to be incorrect, and Oscar may have the same
to say about the linguistic practices on Gas Earth.

The upshot of this is that there are behavioral similarities shared by the linguistic
communities on Earth and on Twin Earth, but which neither of them share with the
linguistic community on Gas Earth. Basic content is supposed to be the explanation
for these behavioral patterns. My suggestion is that the word “water” has the same
basic content on Earth as it does on Twin Earth, but not as it does on Gas Earth.
Very roughly, I take the basic content of an expression to be those facts that are
such that anyone who didn’t know them would be judged not to understand the
expression. For example, if someone did not know that water is a liquid under
normal conditions, we might very well think that they don’t know what “water”
means, and as such lacks the basic content associated with “water.” I will expand
on this in the next chapter.

2.5 Objection and Reply

2.5.1 Objection: Schmark Matter

Here is what I know about dark matter: it does not emit or reflect much elec-
tromagnetic radiation. It has noticeable gravitational effects on ordinary matter.
There is a lot more dark matter than regular matter. If this is all we know, then it
is epistemically possible for us that dark matter is baryonic. Now suppose that it turns out that dark matter is not baryonic. That means the expression “dark matter" refers to something that is necessarily non-baryonic. So no possible world contains baryonic dark matter. There is a possible world which contains baryonic matter that does not emit or reflect EM radiation, has gravitational effects on ordinary matter, etc. But the stuff in that world is not dark matter; call it schmark matter.

The objection is this: on my view, the epistemic possibility that dark matter is baryonic is accounted for by the presence of schmark matter in some other world. My view is that “dark matter might be baryonic" is true for me because there is a world in which, were the sentence produced there, it would express a true proposition. Given our supposition, the proposition expressed here 1) is metaphysically necessarily false and 2) is about dark matter. In the schmark matter world, the proposition expressed is that schmark matter is baryonic. This proposition is (metaphysically necessarily) true, but it is hard to see what relevance propositions about schmark matter have for my claims about dark matter.

2.5.2 Response

Metaphysical and epistemic modals are similar in many ways, but there are some crucial differences. One important difference is this: where metaphysical modality comes in de re and de dicto varieties, epistemic modality does not admit of this kind of distinction. Roughly, de re modal claims are predicated of an object and involve

---

19Baryons are particles composed of three quarks apiece, such as protons and neutrons; baryonic matter is matter composed of baryonic particles.
attributions of modal properties that are independent of language. In contrast, de dicto modal claims are predicated of a sentence and are language-dependent.\footnote{This definition is drawn from \cite{22}.}

This objection rests on the assumption that epistemic modals can be given a \textit{de re} reading, but on my theory this is impossible. This is because epistemic modals operate on meta-propositions rather than propositions, and so the sentence “dark matter might be baryonic” is not a \textit{de re} claim about dark matter. The surface grammar of a \textit{de re} metaphysical modal can often resemble the surface grammar of a epistemic modal, which explains the misunderstanding, but the underlying logical structure is different, according to the metapropositional theory. Metaphysical modals operate on propositions, while epistemic modals operate on metapropositions.

It is worth noting that the fact that no \textit{de re} reading is available for epistemic modals should not be taken to mean that they are all \textit{de dicto}. The \textit{de re}/\textit{de dicto} distinction has been represented in several different ways; epistemic modals are \textit{de dicto} on some of these interpretations, but not all. If, for instance, one might understand \textit{de dicto} modal claims as being predicated on propositions, rather than sentences, in which case epistemic modals would not be \textit{de dicto}, as they operate on metapropositions.

Here are three sample definitions of the \textit{de re}/\textit{de dicto} distinction, all from \cite{22}.

\begin{itemize}
  \item Syntactically \textit{de re}/\textit{de dicto}: A sentence is syntactically \textit{de re} just in case it contains a pronoun or free variable within the scope of an opacity verb
\end{itemize}
that is anaphoric on or bound by a singular term or quantifier outside the scope of that verb. Otherwise, it is syntactically \textit{de dicto}.

Semantically \textit{de re}/\textit{de dicto}: A sentence is semantically \textit{de re} just in case it permits substitution \textit{salva veritate}. Otherwise, it is semantically \textit{de dicto}.

Metaphysically \textit{de re}/\textit{de dicto}: An attribution is metaphysically \textit{de re} with respect to an object \textit{o} just in case it directly attributes a property to \textit{o}. Otherwise, it is metaphysically \textit{de dicto} with respect to \textit{o}, if it indirectly involves \textit{o}, and independent of \textit{o} if it doesn't.

As we can see, these definitions are not equivalent. Epistemic modals do not admit of a metaphysical or a semantic \textit{de re} reading; they can be syntactically \textit{de re}, since they have much the same surface grammar as metaphysical modals. For example, consider "There is someone such that, given what Alice knows, they must be a spy." This is syntactically \textit{de re} using the definition given above, but even so it is not semantically or metaphysically \textit{de re}. Since \textit{de dicto} statements are characterized just as those that are not \textit{de re} for the semantic distinction, that will mean that epistemic modals are semantically \textit{de dicto}, but they might be metaphysically \textit{de dicto} or independent.

It might be hard to see why there is no semantic \textit{de re} interpretation available here. Suppose Alice sees a man in a trenchcoat creeping around a parking garage, speaking in code, and otherwise behaving like a spy. Suppose further that the man in question is Bob, but Alice does not know that Bob is the man she's seen creeping around. In this case it might be epistemically necessary for Alice that the man in
the trenchcoat is a spy, but not epistemically necessary that Bob is a spy. So we do not have substitution *salva veritate* for epistemic modals.

I do not mean to endorse any particular version of the *de re/de dicto* distinction here or to characterize epistemic modals as admitting of *de dicto* readings and nothing else. Rather, the point is that epistemic modals do not seem to admit of anything like a semantic or metaphysical *de re* reading. At least, I have not been able to find any cases of epistemic modals with semantic or metaphysical *de re* readings, whereas metaphysical modals with two or more readings available are fairly easy to come by.

This phenomenon is not limited to natural kind terms and proper names. Consider this example (adapted from [15]):

We are out hiking at night and we encounter a black bear. I am not well-informed about bears, and given that and the poor lighting I am unsure about what type of bear I see. I point and say “For all I know, that bear might be a grizzly.”

Since black bears are not grizzlies, the bear in question is not a grizzly. As such, it is metaphysically necessary that it is not a grizzly. On Kaplan’s semantics for indexicals, the phrase “that bear” rigidly designates a particular bear. The rest of the story is familiar from the examples above; since it is not a grizzly, there are no worlds in which it is a non-grizzly, but nonetheless it seems true to say that for all I know, it might be a grizzly. So we have a (prima facie) example of an EPMI. My

---

21Black bears and grizzlies are separate species.
account will handle the expression in the same way that it handles cases like "dark matter might be baryonic" and "water might not be H\textsubscript{2}O." The important thing to note, though, is that there's only one way to hear this modal.

Compare the prior case to this one:

We are out hiking at night and we are discussing metaphysical possibility. While lost in conversation, we encounter a black bear. I am not well-informed about bears, and given that and the poor lighting I am unsure about what type of bear I see. I point and say "That bear might be a grizzly." From the context, it is clear I am using a metaphysical modal.

In this case, I've uttered a falsehood, since the phrase "that bear" rigidly designates a particular bear, and that bear could not be a grizzly (assuming species membership is necessary, as Kripke suggests). This is the \textit{de re} reading of this modal. No \textit{de dicto} reading is available, given the standard semantics for demonstratives like "that."

We might also wish to consider a third case:

We are out hiking at night and we are discussing metaphysical possibility. While lost in conversation, we encounter a bear. I am not well-informed about bears, and given that and the poor lighting I am unsure about what type of bear we saw. At the end of the evening, I say "The bear we saw might have been a grizzly." Suppose the bear we saw was a black bear.

First, let us stipulate that the context makes it clear that my utterance is supposed to be interpreted as a metaphysical modal. In this case, we might be inclined to just
claim that I’ve uttered a falsehood, which would be the *de re* reading. But note that there is a second reading available in this case; we can hear “The bear we saw might be a grizzly” as a *de dicto* claim, in which case my utterance will be true if there is a world where the sentence “The bear we saw is a grizzly” expresses a truth. If we eliminate the stipulation that the modal is metaphysical, then we can see that there are not two comparable readings available for the corresponding epistemic modal; that is, there is no way to predicate epistemic modal properties of the object itself given that we are unsure of the object’s identity.

That said, it is not true that epistemic modals always have only one reading. Let us modify the case above slightly:

We are out hiking at night. While lost in conversation, we encounter a number of bears. I am not well-informed about bears, and given that and the poor lighting I am unsure about what type of bears we saw. At the end of the evening, I say “The first bear we saw might have been a grizzly.” Suppose the first bear we saw was a black bear.

Let us stipulate that this is an epistemic modal. There are at least two ways of making this epistemic modal true. When I say “The first bear we saw might have been a grizzly,” it might be true because I do not recall which bear we saw first. It might also be true because I don’t know what species of bear I saw on the first sighting. This might sound superficially like a version of the *de re/de dicto* distinction, but it is not. The second possibility (that I do not know the bear’s species) does involve my knowledge about a particular bear, but it does not involve the ascription of modal properties to that bear.
What we have is rather a sentence whose truth conditions involve a disjunction. Many sentences, modal or otherwise, have this property; for instance, consider the sentence “For all I know, it might be raining outside.” This is true if I know where I am but don’t know what the weather is like there. It’s also true if I know what the weather is like everywhere on the planet but don’t know where I am.

I also want to suggest that any account of epistemic modality that takes Kripke’s semantic externalism seriously will suffer from this problem or something much like it. If Kripke is right, then we cannot try to underwrite the epistemic possibility that water might not be H$_2$O by looking for a world in which water is not H$_2$O, as there are no such worlds. This suggests that we will not get far trying to have water underwrite the epistemic possibility that water might not be H$_2$O. My suggestion, then, is to accept this consequence of Kripke’s position and look instead for objects to which the word ‘water’ might refer in other worlds of production.
First, a potential problem for the metapropositional approach: suppose Alice is just learning English. Bob, her instructor, has taken her to the beach to teach her some common words. Bob gestures toward the ocean and says “That stuff is water.” Alice, however, is confused; Bob’s gesture encompassed the ocean, the sand, and even some of the sky. Alice mistakenly forms the belief that “water” picks out sand. Here is where the potential problem arises. Is it epistemically possible for Alice that water is sand? One might think that the metapropositional approach would say that it is, since Alice’s linguistic ignorance seems to suggest that there is a world consistent with her knowledge base in which the English word “water” picks out sand. But Alice knows that water is not sand (though she’d express the thought that water is not sand in a different language). To deal with problematic cases like this, we introduce the basic content constraint. In this chapter I will offer some intuitive arguments for basic content and briefly discuss the similarities between basic content and narrow content.

3.1 Basic Content

Basic content determines the intension of an expression at every world of production. It can be represented as a function from worlds of production to intensions; inten-
sions, in turn, are functions from worlds of evaluation to extensions. The metapropositional account has a basic content constraint: when we evaluate epistemic modals, we hold basic content fixed. Basic content is determined by the way in which an expression is actually used.

Basic content is supposed to be a kind of semantic content; it is not identical with the meaning of a term, though the two notions are related. Roughly, basic content is the collection of facts that we take to be most significant for knowing the meaning of the expression, stated without using any synonym of the expression. That is, if someone lacks the basic content associated with an expression then we would take them to be ignorant of or confused about (some aspect of) meaning of the expression in question. For example, we might think that anyone who doesn’t know that water is drinkable under normal conditions doesn’t fully grasp the meaning of “water.” In contrast, we would not think that someone who doesn’t know that water is present on Enceladus (a moon of Saturn) is in any way lacking in linguistic competence. So the fact that water is drinkable under normal conditions is plausibly construed as part of the basic content associated with “water,” while the fact that water is present on Enceladus is not.

3.2 The Intuitive Plausibility of Basic Content

One interesting feature of Putnam’s original Twin Earth thought experiment is that a person could be transported from Earth to Twin Earth without noticing any differences. In particular, the Earth native could, for the most part, navigate the local linguistic environment with no significant difficulty. If I were transplanted to
Twin Earth and someone were to say to me, “Please bring me a glass of water," I could do so. If I didn’t know I was on Twin Earth, I would believe (falsely) that I was producing a glass of H\textsubscript{2}O, but whether or not I knew where I was I would be able to satisfy the request. I would seem to be a competent user of the language as a result. If the conversation turned to the topic of the chemical composition of certain clear liquids, I would be judged by a native of Twin Earth to be mistaken about the facts but not ignorant about the language.

Now imagine I was transplanted to a world in which “water" denotes gold, but which is otherwise chemically identical to Earth. Call this world Gold Earth. In this world, “water" denotes gold and another word denotes H\textsubscript{2}O (and XYZ is nowhere to be found). On this world I would seem to Gold Earth natives to be confused or ignorant about the language. The contrast here between the Earth/Twin Earth case and the Earth/Gold Earth case is notable. The linguistic communities on Earth and Twin Earth display many relevantly similar behaviors, and moreover they would have little trouble interacting with one another; those confusions that would arise in conversation would involve one party being mistaken about the facts, but not about the language.\[1\]

We can explain these behavioral similarities by appealing to basic content. Speakers on Earth and on Twin Earth share the same basic content, which accounts for their behavioral patterns. We can also explain the different behaviors exhibited

\[1\]This judgment presupposes a fairly strong notion of the analytic/synthetic distinction. In order for us to categorize someone as mistaken about the facts as opposed to mistaken about the language, we have to have a good idea as to what constitutes factual correctness as opposed to linguistic correctness.
by members of the linguistic communities on Earth and Gold Earth. These communities do not share the same basic content, and as a result a speaker from one community would judge a speaker from the other to be ignorant or confused about the language.

We can find further evidence for basic content by considering the differences between “water” and “H₂O.” Intuitively, these terms are not synonymous. It is possible to know what “water” means even if you don’t know that water is H₂O or the meaning of “H₂O.” And it seems implausible to suggest that “water” meant something different before its chemical structure was discovered. To say that “water” is not synonymous with “H₂O” is to say that their meanings differ. If their meanings differ, then that opens up the possibility that their basic content may differ as well.

It is possible to understand “water” without understanding “H₂O.” Young (English-speaking) children, for instance, generally know the meaning of “water” but not the meaning of “H₂O.” Before modern chemistry was developed, no one would have known what H₂O was or the meaning of “H₂O,” but we don’t want to say they didn’t know the meaning of “water” (though we might want to say they didn’t grasp the nature of water, or perhaps that they didn’t fully understand what “water” meant). Again, we can explain these facts by appealing to basic content. If

---

2 A die-hard Kripkean might claim that, since “water” and “H₂O” are co-referential and reference exhausts the semantic content of natural kind terms, “water” is synonymous with “H₂O.” I’m not going to provide an independent argument against this claim. It is worth, noting, though, that even Kripke noted the counterintuitive nature of some of his claims. My view is simply incompatible with the full-blooded Kripkean account, though it is compatible with semantic externalism broadly construed. A more in-depth discussion of how my view relates to Kripke’s can be found in chapters 1 and 2. I suspect that the full-blooded Kripkean account will not be able to offer a solution to the problem of EPMIs, which is one reason I reject the view.
“water” and “H₂O” have different basic content, then the facts that we take to be essential for knowing the meaning of “water” will not be the same as the facts we take to be essential for knowing the meaning of “H₂O.” And one might know one set of facts, but not the other.

Admittedly, this argument does not show that there are basic contents (in the relevant sense). After all, we might think that my ability to navigate the linguistic environment of Twin Earth is the result of a serendipitous coincidence. On this line of reasoning, when the Twin Earth native says, “Please bring me a glass of water,” I make two mistakes. First, I assume that the speaker is referring to H₂O. Second, I assume that the watery stuff on Twin Earth is H₂O. Luckily, these mistakes do not prevent me from identifying the watery stuff in the area, and luckily the watery stuff in the area is what the speaker wanted. So I am able to fulfill the speaker’s request.

What this line of reasoning overlooks, however, is the contrast between the Twin Earth and Gold Earth cases. If I am transplanted to Twin Earth, I am able to navigate the linguistic environment. I am not able to do the same of Gold Earth. Why should this be? After all, I make the same assumptions on Twin Earth as I do on Gold Earth. That is, in both cases, I am likely to assume that the speaker is referring to H₂O and that the watery stuff in the area is H₂O. On Gold Earth, the latter assumption is even correct (recall that Gold Earth is chemically identical to Earth). But if I were transplanted to Gold Earth, and a Gold Earth native said, “Please bring me some water,” I would not be able to fulfill the request (at least, not on the first attempt). It seems like there should be a reason for this—why is it that I
perform worse on a world that is (chemically, at least) more like my own world? We might opt to claim that the Earth/Twin Earth case is simply a fortuitous accident, while the transplant from Earth to Gold Earth is not so lucky. But we have an available explanation for some tantalizing behavioral patterns, and it is unsatisfying to appeal to coincidence in such a case.

Moreover, note that the Twin Earth native’s request can be fulfilled by filling a glass with the local watery stuff (which turns out to be XYZ) and bringing it to her. The same is true of my own requests for water; if I say, “Please bring me a glass of water,” my request can be fulfilled by bringing me a glass of the local watery stuff (which turns out to be H₂O). This is not the case for the Gold Earth native. Her request is fulfilled by bringing her a sample of the local gold-like stuff (which turns out to be Au). We should seek to explain the similarities between Earth and Twin Earth, as well as the contrasting dissimilarity of the Gold Earth case. In particular, we might seek to explain why “water” picks out watery stuff on Earth and Twin Earth, but not on Gold Earth. Basic content offers us a way to think about the search for an explanation. Basic content determines the intension of an expression in every world of production. On Earth and on Twin Earth, the intension of “water” is such that “water” rigidly designates the local watery stuff, whatever it happens to be.

We might think, then, that the fact that “water” rigidly designates the local watery stuff on Earth and on Twin Earth has something to do with the basic content associated with “water”—for instance, that the reason “water” picks out the watery stuff on both worlds is because the term has the same basic content on those worlds,
and the reason “water" picks out gold-like stuff (as opposed to watery stuff) on Gold Earth is because it has a different basic content there than it does here. That is, there is a rule on Earth that connects “water" up with watery stuff (which happens to be H$_2$O), and the same rule is in force on Twin Earth (although the watery stuff there is XYZ), whereas a different rule is in force on Gold Earth. And we can say why this might be: someone on Earth who doesn’t know that water is the local watery stuff doesn’t have the basic content associated with “water." The same goes for Twin Earth residents. But on Gold Earth, things are different, there, the person who thinks water is the local watery stuff (in our sense—i.e. the clear, odorless liquid found in lakes and rivers, whatever it happens to be) would be judged by Gold Earth natives to be linguistically ignorant or confused.

At this point, we may ask why we need to introduce basic content at all. Basic content already constitutes a break with the Kripkean position, since that position rejects any notion of semantic content beyond reference for names and natural kind terms. Why not just reject the Kripkean position entirely and adopt a Fregean view, or a descriptivist view? We could do just that. But we don’t have to. We might be able to preserve some Kripkean positions and still have a story that accounts for the similarities and differences between Earth, Twin Earth, and Gold Earth—and insofar as Kripke’s view is independently plausible, we should try to do just that. But to do that, we need basic content.

---

3 The basic content of “water" cannot be “the local watery stuff," on pain of circularity. I’m using “the local watery stuff" as a shorthand for a more careful statement of the basic content associated with “water," for the sake of readability.

4 As stated in chapter 1, a defense of Kripke’s positions would take us too far afield. The project is to show how to account for EPMIs while preserving as much of the Kripkean view as possible.
One area where things get a little tricky is proper names. We could maintain that proper names have meaning (or at least some kind of semantic content) apart from their reference, in which case the basic content of a proper name would be those features of the named object that we take to be most significant for knowing the meaning (or other semantic content) of the name. But this leaves us on the hook for an explanation of the semantic content of names. I will offer up one possible explanation that comports well with the metapropositional approach to epistemic modals.

In \[23\] Pelczar and Rainsbury argue that names are rigidly designating indexicals.\footnote{Pelczar and Rainsbury do not argue that natural kind terms are indexicals. Insofar as names and natural kind terms function analogously, though, we should expect that the same sort of considerations that would tend to show that names are indexical should apply (mutatis mutandis) to natural kind terms as well.} This means, as they put it, that names have variable character, but once the context of utterance is fixed, names have constant content. That is, the content of a name varies based on the context of utterance, but the extension of a name is the same in all circumstances of evaluation (given a fixed context of utterance).

Pelczar and Rainsbury start with something resembling the Kripkean causal-historical theory, in which the content of a name is determined by an initial baptism or dubbing. They depart from the Kripkean view, though, and claim that names are sensitive to an element of context they call “dubbings-in-force.” A name can be used to denote many different objects; consider the name “Dan,” which can be used to denote Dan Quattrone in some contexts and Dan Quayle in others.
Pelczar and Rainsbury, when someone says a name the intended referent is determined by the dubbings that are in force in the context of utterance. A dubbing is “a speech-act whereby a name acquires a referent," and if a dubbing is in force then the object dubbed bears the name it received in that dubbing.

For instance, consider the following scenario: Alice is talking to Bob. A third person, also named Bob, is walking by. Alice looks at Bob and says, “Bob, I need to know what you’re doing tomorrow." She then turns slightly, and in a louder voice says “Bob, it would be good if you told me what you’re doing tomorrow as well." It is clear that in this scenario, Alice picks out one person with her first use of the name “Bob’ and the other with her second use. According to Pelczar and Rainsbury, this is because the dubbings in force when “Bob" is first uttered differ from the dubbings in force when “Bob" is subsequently uttered.

Pelczar and Rainsbury offer up a two-part argument for their picture of names. First, they argue that their view accounts for long-standing puzzles associated with the Kripkean story about names. For instance, it is possible to believe that Hesperus is bright but Phosphorus is dim, without being irrational or lacking linguistic competence (assuming the believer does not know or believe that Hesperus is Phosphorus). This is hard to explain if names are rigid designators with no semantic content beyond their reference; since “Hesperus" and “Phosphorus" are co-referring terms, they should be synonymous on this view, and hence the sentence “Hesperus is bright but Phosphorus is dim" should be synonymous with “Hesperus is bright but Hesperus is dim," which is transparently inconsistent.

6The fact that many people can answer to the same name is typically handled by claiming that names are lexically ambiguous; Pelczar and Rainsbury reject this notion.
On Pelczar and Rainsbury’s account, the sentence “Hesperus is bright but Phosphorus is dim” is indeed inconsistent. However, we can nonetheless explain why we do not think that anyone who believes that Hesperus is bright but Phosphorus is dim is irrational or linguistically incompetent. Such a belief stems from what Pelczar and Rainsbury call a “contextual error.” Contextual errors are mistakes “as to which facts are determined by the characters of the sentences one uses.” For instance, if I am talking to Alice and Bob and I say “You are smart,” intending to indicate that Alice is smart, it might be that Alice and Bob each think I’m talking about themselves when in fact I am only talking about one of them. In this case, Bob has made a contextual error. Someone who believes that Hesperus is bright but Phosphorus is dim is making the same kind of error; they believe that the dubbings in force are such that “Hesperus” and “Phosphorus” denote different objects, when in fact the terms denote the same thing.

The second line of argument that Pelczar and Rainsbury offer is empirical. Pelczar and Rainsbury describe a class of terms they call “quasi-names.” Quasi-names are words like “Mom,” “Dad,” or “Sarge.” According to Pelczar and Rainsbury, quasi-names function like proper names and like indexicals. Quasi-names are singular referring expressions, and grammatically they function like proper names. Moreover, quasi-names, like indexicals, cannot be understood unless we know relevant contextual factors—if Alice says “Mom is angry,” we won’t know who is angry unless we know who Alice’s mother is. Quasi-names also typically take wide scope over other operators when embedded in attitude reports. Consider, for instance, suppose Carol says “Alice said that Bob believes that Mom is angry with him.”

\[^{7}\text{Pelczar and Rainsbury’s example.}\]
this sentence, “Mom" picks out Carol’s mother, rather than Alice’s mother or Bob’s mother (assuming Alice, Bob, and Carol are not siblings). Pelczar and Rainsbury then claim that, since proper names are like quasi-names, and quasi-names behave like indexicals, we have reason to think that proper names behave like indexicals as well.

Taken together, Pelczar and Rainsbury offer theoretical and empirical evidence for the claim that names are indexical. And if names are indexicals, then they have character (in Kaplan’s sense). The character of a proper name is semantic content, and it gives us a way to say what it is to understand the meaning (in a broad sense) of a name. That, in turn, gives us a way to understand the basic content of a name. Kaplan suggests in [12] that the meaning of an indexical is (roughly) its character. Character is a function from contexts of utterance to intensions. Consider a non-name indexical—for instance, “you." No one knows the entire character function for “you." That is, no one knows what the intension of “you" is in every possible context.

What we do know, though, is the rule that we use to figure out what “you" picks out on each occasion of use (roughly, the audience member or members being addressed by the speaker). Knowing this rule does not ensure that we will always be able to figure out what “you" picks out, because we don’t always have all the relevant information. For instance, suppose I am listening to speaker who is (unbeknownst to me) talking on the phone and unaware of my presence. The speaker says “You should go home." I might mistakenly think that I am being addressed, when in fact the speaker to addressing the person on the other end of the phone call. Nonetheless, we would not say that I am ignorant or mistaken about the meaning of “you," even
though I am mistaken about its application in this particular case. In general, I think we are inclined to say that anyone who knows the rule for "you" knows the meaning of the term, and conversely that anyone who does not know the rule does not know the meaning. My suggestion, then, is that this rule is the basic content of an indexical. By analogy, then, the same should apply to names. Consider a common proper name, such as "John." I may not know who "John" picks out on every occasion of use, but that does not mean I do not grasp the rule that allows us to figure out the referent of "John." And it is that rule that constitutes the basic content of the name.

3.3 Basic Content and Narrow Content

Basic content is in many ways similar to narrow content. Mental representations have content; that is, they represent the world as being some way or other. We can make a distinction between those representations whose content depends only on the intrinsic, internal properties of the representation’s owner and those whose content depends (at least in part) on external properties. Content that depends only on intrinsic properties is narrow content. That means that any exact duplicate of me (which must, by definition, share all my intrinsic properties) would have all the same narrow content that I do. Content that depends on external properties is wide content. My duplicate might have different wide content, if it were in a different environment. Putnam introduced the notion of narrow content in \[26\] specifically to cover the psychological properties that I have in common with my Twin Earth doppelgänger. The analogy to basic content should be clear: while narrow content
is the psychological content that I share with my doppelgänger, basic content is the semantic content that I have in common with my doppelgänger.

Not everyone agrees that there is narrow content, and the case for narrow content is not universally seen as compelling. Even among semantic externalists, views differ; Burge, for instance, thinks that almost all content is wide [1], while Evans thinks there is some narrow content [5]. Putnam has at times argued for the existence of narrow content [26] and at other times rejected it. In any case, it would take us too far afield to argue one way or the other. We will have to be satisfied in noting that there are powerful arguments for the existence of narrow content [8] and that if we find them persuasive that gives us reason to accept basic content as well. I introduce the notion of narrow content because it resembles basic content in some significant ways. In particular, if there is narrow content, then my duplicate on Twin Earth would by definition have the same narrow content that I do, even with regard to concepts like water (although their wide content would differ). We would have the same narrow content because our internal psychology would be the same. My duplicate on Gold Earth, in contrast, would not have the same narrow content, because his internal psychology would be different from my own. Since we have powerful arguments for the existence of narrow content, we also have (by analogy) reason to think that there is basic content—although we do not have decisive evidence.

To sum up: we have multiple independent lines of argumentation for basic content. First, there is the intuitive plausibility of such a notion, which we can see by

---

8See, for example, [6], [7], [21], and [16].
looking for an explanation of behavioral patterns in linguistic communities on different worlds of production. Second, we have the fact that coreferring terms (such as “water" and “H₂O") seem to differ with respect to some of their semantic properties, even if they are both rigid designators. For instance, we can explain the intuitive claim that we know the meaning of “water" without knowing the meaning of “H₂O" by appealing to basic content. Third, in the case of names, we have empirical and theoretical evidence from Pelczar and Rainsbury that suggests names are indexicals. If names are indexicals, then they have a Kaplanian character, which means that they have semantic content apart from their referent, which in turn gives us a way to define their basic content. Finally, we have the analogy between basic content and narrow content, which suggests that our reasons for thinking there is narrow content are also reasons for thinking there is broad content.
There is another approach to the issue of epistemic possibility that uses the two-dimensional framework—the propositional approach. The most prominent defender of this approach is David Chalmers, and so this chapter will focus on his account, developed primarily in [2]. As will be shown, the problems facing Chalmers are significant.

4.1 Chalmers's Account

At the outset, it is important to note that while Chalmers's project does involve epistemic modality, his primary interest is in validating what he calls the "Core Thesis." On Chalmers's view, a sentence $S$ has two intensions, called 1-intensions and 2-intensions, respectively. The nature of these intensions will be explained shortly. Chalmers seeks to show that the following is true:

**Core Thesis:** For any sentence $S$, $S$ is *a priori* iff $S$ has a necessary 1-intension.

To accomplish this task, he introduces what he calls the epistemic understanding of the two-dimensional semantic framework.
4.1.1 Introducing the Epistemic Understanding: Intuitive Support

The intent of this section is to establish the intuitions that lie behind the epistemically understanding of the two-dimensional semantic framework. Before getting into specifics, Chalmers first offers the following general characterization of the two-dimensional framework:

The core idea of two-dimensional semantics is that there are two different ways in which the extension of an expression depends on possible states of the world. First, the actual extension of an expression depends on the character of the actual world in which an expression is uttered. Second, the counterfactual extension depends on the character of the counterfactual world in which the expression is evaluated. Corresponding to these two sorts of dependence, expressions correspondingly have two sorts of intensions, associating possible states of the world with extensions in two different ways...

These two intensions correspond to two different ways of thinking of possibilities. In the first case, one thinks of a possibility as representing a way the actual world might turn out to be; or as it is sometimes put, one considers a possibility as actual. In the second case, one acknowledges that the actual world is fixed, and thinks of a possibility as a way the world might have been but is not: or as it is sometimes put, one considers a possibility as counterfactual.

So we have two intensions. Chalmers calls them 1-intensions and 2-intensions. 1-intensions correspond to possibilities considered as actual, and 2-intensions to
possibilities considered as counterfactual. These notions were introduced in §4 of chapter 1 and serve as the intuitive basis for Chalmers’s particular version of the two-dimensional framework. Possibilities are traditionally understood as possible worlds, and in his initial gloss Chalmers keeps with this tradition.

For instance, the sentence “water is H\textsubscript{2}O” is true in a world considered as counterfactual when H\textsubscript{2}O is H\textsubscript{2}O. In a world considered as actual, though, “water is H\textsubscript{2}O” is true (roughly) when the colorless, tasteless, potable liquid in that world is H\textsubscript{2}O. So the 1-intension of “water is H\textsubscript{2}O” assigns the value “true” to all those worlds meeting the latter condition, and the 2-intension assigns the value “true” to all those worlds meeting the former condition (i.e. all of them). Similarly, the 1-intension of “water” maps “water” onto the colorless, etc. liquid in each world, without regard to its chemical properties, while the 2-intension of “water” behaves like we would expect if “water” were a Kripkean rigid designator.

The question arises: why should we think that terms like “water” and sentences like “water is H\textsubscript{2}O” behave differently when evaluated with respect to possibilities

1Chalmers takes the analysis of the two-dimensional framework in terms of possibilities considered as actual and as counterfactual to be characteristic of two-dimensional theories in general, but this is not obvious. For instance, the metapositional approach developed in chapter 2 is best understood in terms of worlds of production and worlds of evaluation. That said, nothing important rests on whether or not all two-dimensional theories can be understood in terms of these two ways of looking at possibilities. Chalmers is free to use this as his starting point even if he’s wrong about it being a common starting point among two-dimensional theories. This does suggest two metrics for evaluating Chalmers’s approach. First, we can evaluate it with respect to its success at validating the Core Thesis. Second (and more important for our purposes) we can evaluate it with respect to its ability to handle EPMIs.

2As we will see, Chalmers eventually opts to replace worlds (at least in part) with scenarios, which are the epistemic analogs to worlds. The notion of a scenario will be presented in more detail in the next section.
considered as actual than they do when evaluated with respect to possibilities considered as counterfactual? The underlying intuition looks something like this. We know that water is $\text{H}_2\text{O}$. But we can imagine that things were different. Dalton, for instance, thought water was $\text{HO}$, in part because his tests could not distinguish between $\text{H}_2\text{O}$ and $\text{HO}$. Perhaps we’re actually in a similar position, and the actual world is not as we believe it to be. In fact, water is $\text{XYZ}$, and our tests are incapable of distinguishing between $\text{H}_2\text{O}$ and $\text{XYZ}$. If things are actually this way, we do not know that water is $\text{H}_2\text{O}$ after all, because water isn’t $\text{H}_2\text{O}$; it’s $\text{XYZ}$.

That is, given that water is actually $\text{H}_2\text{O}$, it turns out that there is no way things might have been (i.e. no possibility considered as counterfactual) in which water is $\text{XYZ}$. But if the actual world is different than the way we think it is, then perhaps water is not $\text{H}_2\text{O}$. And we don’t always know which world is the actual world—or at least we don’t always know how to distinguish the actual world from similar worlds by description. And even when we can identify a given world as non-actual, we can still consider how things appear to someone who believes that world is actual.

The upshot of all this is that intuitively, it seems like possibilities considered as actual are closely tied to our epistemic position. That is, whether or not a world can be considered as actual is dependent on what we know about the actual world. Insofar as our knowledge of the actual world is incomplete, it leaves open that the actual world might turn out to be different than we believe it to be.\(^3\) This, in turn, gives us reason to think that associating 1-intensions with possibilities considered

\(^3\)Notice that this intuition seems to rely on some prior understanding of what is possible given what we know—i.e. some understanding of epistemic modality. See §4.2.2 below for more on this concern.
as actual might let us construct an account of epistemic modality that validates the Core Thesis (which is Chalmers’s main goal) and, at the same time, resolves the problem of EPMIs. In order to do this, we will need to move beyond intuition-mongering and try to develop a more rigorous account. And to do that we will need to introduce scenarios.

4.1.2 Scenarios

The fundamental building block of Chalmers’s account of epistemic modality is the scenario. Scenarios are maximal epistemic possibilities. Possible worlds are maximal metaphysical possibilities, and hence scenarios are the epistemic analogs of possible worlds. In this section, our task is to see how scenarios are to be understood. As we will see, Chalmers offers two distinct ways of understanding scenarios and moves between the interpretations fairly freely. I will present both accounts in turn.

Deep Epistemic Possibility

In order to understand scenarios, Chalmers begins with the notion of epistemic space. Epistemic space is the space of epistemic possibilities; epistemic possibilities are ways the world might turn out to be, for all we know. Chalmers says:

There are many ways the world might be, for all we know. And there are even more ways the world might be, for all we know \textit{a priori}. The oceans might contain H$_2$O or they might contain XYZ; the evening star might be identical to the morning star or it might not. These ways the

---

4What it means to be a maximal epistemic possibility depends somewhat on what scenarios turn out to be; the topic will be addressed in §4.2.4 and 4.2.5.
worlds might be correspond to epistemically possible hypotheses, in a broad sense. Let us say that a claim is epistemically possible (in the broad sense) when it is not ruled out a priori. Then it is epistemically possible that water is H$_2$O, and it is epistemically possible that water is XYZ. It is epistemically possible that Hesperus is Phosphorus, and epistemically possible that Hesperus is not Phosphorus.

And later:

In the ordinary sense, we say that S is epistemically possible roughly when S may be the case for all we know, and that S is epistemically necessary roughly when we are in a position to know that S is the case. A notion of deep epistemic necessity goes beyond this sort of dependence on the shifting state of an individual’s knowledge, to 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 a natural candidate. We can say that S is deeply epistemically necessary when it is a priori: that is, when the thought expressed by S expresses actual or potential a priori knowledge. Then S is deeply epistemically possible when the negation of S is not epistemically necessary: that is, when the thought that S expresses cannot be ruled out a priori.

Chalmers does not further characterize the notion of deep epistemic modality and so, for all practical purposes, we can regard him as taking this (or equivalently the
notion of *a priori* knowability) as primitive. Deep epistemic modality is central to his account of epistemic modality. The collection of all the maximally specific deep epistemic possibilities is epistemic space; scenarios are intuitively understood as particular points within this space. This continues the analogy between scenarios and worlds; we also have a space of metaphysical possibilities, and worlds are particular possibilities within that space.

The focus on deep epistemic modals rather than ordinary modals is an unusual move. As Chalmers recognizes, most accounts of epistemic modality are relativized to either agents or knowledge bases. One concern here is that by focusing on deep epistemic modals rather than ordinary epistemic modals, Chalmers is introducing an irrelevant issue. After all, much (perhaps most) of what we know we know through *a posteriori* means. Putting aside that knowledge when evaluating epistemic modals makes epistemic modals fairly distant from our ordinary concerns. Chalmers’s formulation presupposes that there are no deeply epistemically necessary empirical truths.

Chalmers’s decision to take deep epistemic modality as primitive need not lead to this problem, though. Deep epistemic modals quantify over the whole of epistemic space. But nothing prevents us from introducing operators that quantify over a portion of this space—for instance, over that portion which is consistent with our *a posteriori* knowledge. These operators can plausibly represent our ordinary notions of epistemic modality.  

---

5Chalmers’s view also has the consequence that any *a priori* truth will be epistemically necessary. This might seem unproblematic, but see §2.1 below for more discussion.
Scrutability and Verification

Scenarios also have a property that Chalmers calls “scrutability.” Scrutability is intended to capture the way in which the extension of an expression depends on which world is actual. That is, if we learn that some aspect of the world was objectively thus-and-such, then (according to Chalmers) this tells us something about the extensions of our expressions. For instance, when we learned that the clear, potable liquid found in lakes and rivers is H\textsubscript{2}O, we also learned that the extension of “water” is H\textsubscript{2}O. But suppose we discovered that we’re mistaken about the constitution of the clear, potable liquid found in lakes and rivers. In fact, the tests we performed on this substance cannot distinguish between H\textsubscript{2}O and XYZ (i.e. our tests give a positive result when applied to H\textsubscript{2}O or to XYZ). Suppose further that we develop some improved test that can distinguish between H\textsubscript{2}O and XYZ and after applying this test, we discover that the stuff in question is XYZ and not H\textsubscript{2}O. According to Chalmers, if this happened, we’d conclude that “water” actually refers to XYZ.

The scrutability of scenarios is what allows us to make judgments like this one. Chalmers says:

If we come to know that the world has a certain character, we are in a position to conclude that the expression has a certain extension. And if we were to learn that the world has a different character, we would be in a position to conclude the expression has a different extension. That is: we are in a position to come to know the extension of an expression, depending on which epistemic possibility turns out to be actual.
Because scenarios are scrutable, we are (in principle) in a position to determine the extensions of our expressions using only \textit{a priori} reasoning once we know which scenario is actual. When we consider a scenario as actual, then, what we are doing is entertaining the hypothesis that world is as described in the scenario in question. We can entertain the hypothesis that the clear, potable liquid found in lakes and rivers is actually XYZ, and when we do so we are considering the epistemic possibility that the XYZ-world is actual.

Chalmers thinks that intuitively, if the XYZ world turns out to be actual, we should conclude that “water” refers to XYZ and hence that “water is XYZ” is true. He calls this relation between the XYZ-scenario and the sentence “water is XYZ” verification. In general, he says that a scenario $t$ verifies a sentence $s$ when $t$ makes $s$ true when $t$ is considered as actual.

\textbf{Epistemic Intensions}

So here’s where things stand: each scenario is a maximal (deep) epistemic possibility located in epistemic space. This means that, for each scenario $s$, it is deeply epistemically possible that $s$ is actual. That is, it is consistent with everything we can know \textit{a priori} that $s$ represents the actual world.\footnote{Scenarios are scrutable; if we grasp a scenario, we are in a position to determine the reference of our expressions if that scenario turns out to be actual.} It is not entirely clear how scenarios are supposed to be related to worlds; Chalmers offers two different possibilities. I am using “representing” as the relation here in order to avoid these issues for the moment.
We can now define epistemic intensions. The epistemic intension of a sentence $s$ is a function from scenarios to truth-values. If a scenario $t$ verifies $s$, then $s$ is true at $t$. If $t$ verifies $\neg s$, $s$ is false at $t$. If $t$ does not verify $s$ or $\neg s$, then $s$’s epistemic intension is indeterminate at $t$.

Whatever scenarios turn out to be, if they are to be to epistemic possibility what possible worlds are to metaphysical possibility, they will need to meet the following requirement:

**Plenitude Principle:** For all [sentences] $S$, $S$ is [deeply] epistemically possible if and only if there is a scenario that verifies $S$.

Alternatively, we could formulate the requirement thusly: $S$ is deeply epistemically necessary if and only if $S$ is verified by all scenarios. Given that $S$ is deeply epistemically necessary if and only if it is *a priori*, this means that the Plenitude Principle is equivalent to the Core Thesis (as Chalmers points out).

What we have is an account of what scenarios are supposed to do, but we do not yet have an account of what scenarios are. Chalmers offers two analyses of scenarios which provide possible answers to the latter question: first, that scenarios can be understood as centered worlds, and second, that scenarios can be understood as sentences in a constructed language. Chalmers offers these alternatives because he wants to avoid concerns about his choice of primitives. The latter alternative is, I think, Chalmers’s preferred understanding of scenarios, but both accounts warrant our attention. We will now take them in turn. Once we have these alternatives in
hand we can see if either of them make the Plenitude Principle true and hence make the Core Thesis true.

Scenarios as Centered Worlds

First, Chalmers suggests that scenarios can be understood as centered (metaphysically) possible worlds. A centered world is an ordered pair consisting in a possible world and a privileged point in that world (called a center). This point consists in a place, time, and agent and is used to evaluate indexical expressions. For example, the expression "I am a person" is true at a given centered world if the agent at the center is a person.

On this view, the world coordinate of the world-center pair tells us how things objectively are in that world, and the center allows us to evaluate the indexical component. We need to be able to evaluate indexicals because there are hypotheses like the following: "water is H\textsubscript{2}O and I am a person." We want this hypothesis to turn out to be deeply epistemically possible, and moreover we want it to be distinct from the hypothesis "water is H\textsubscript{2}O and I am a cat." The former hypothesis is verified by a centered world in which water is H\textsubscript{2}O and the agent at the center is a person; the latter is verified by a centered world in which water is H\textsubscript{2}O and the agent at the center is a cat. If we do not have a center picked out, then we have no way of evaluating the indexical component of the expressions. As a result, we will have no way of saying how the truth conditions for these expressions differ from one another.

\footnote{The center might need to include more than just a place, time, and agent in order to evaluate some indexical expressions. Nothing important rests on just what information is required; all that matters is that centers include enough information to evaluate indexicals, whatever that might be.}
If we are to analyze scenarios as centered worlds, then each scenario corresponds to one and only one centered world. We still have to explain why it is plausible that the things that satisfy the Plenitude Principle are in one-to-one correspondence with the centered worlds. Chalmers divides this task into three parts. First, we must show that each centered world corresponds to a scenario. Second, we must show that there are no centered worlds that verify deeply epistemically impossible hypotheses. Finally, we must show that there are no scenarios left over once every centered world is accounted for. Chalmers suggests that the first two issues are fairly trivial, and for the moment I will accept that. He thus focuses his attention on the third part of the task. In order to show that there are not more scenarios than centered worlds, he must show the truth of the following principle:

**Metaphysical Plenitude:** For all S, if S is [deeply] epistemically possible, there is a centered metaphysically possible world that verifies S.

There are two sources (broadly construed) of potential counterexamples to this principle. The first source is familiar; the Kripkean examples of EPMIs provide potential counterexamples. It seems that it is epistemically possible that water is XYZ, but if there is no world in which water is XYZ then it seems that there could not be a world in which water is XYZ.

Chalmers claims that cases like these are not, in fact, counterexamples. It is true that there is no world in which water is XYZ—but only if we are considering worlds as counterfactual. Things change once we start to examine possibilities as actual.
Consider Twin Earth, the world where all the H$_2$O has been replaced by XYZ. The sentence “water is XYZ” is false at Twin Earth considered as counterfactual, but it is true at Twin Earth considered as actual and therefore it is “verified” by Twin Earth.

That is, if we learned that the clear, potable liquid found in rivers and lakes is actually XYZ (and not H$_2$O), then “water is XYZ” will turn out to be true. We can truly say that there are no metaphysically possible worlds in which water is XYZ (and that Twin Earth is not a world in which water is XYZ) because when we talk about metaphysical possibility we are considering worlds as counterfactual, and when we consider worlds as counterfactual we acknowledge that the character of the actual world is already fixed. When we consider worlds as actual, we are considering the possibility that the character of the actual world is as specified in the scenario in question. We can do this even if that scenario is not compatible with our beliefs about the actual world, since we’re engaging in hypothetical reasoning. It is this distinction between worlds considered as actual and worlds considered as counterfactual that allows this account of scenarios to address the problem of EPMIs. A sentence is epistemically possible if it is true at some centered world considered as actual; it is metaphysically impossible if it is false at every centered world considered as counterfactual.

The second source of potential counterexamples is a collection of different and mostly unrelated controversial philosophical claims. Chalmers gives a brief catalogue

---

8Strictly speaking we should be considering the pair consisting in Twin Earth and a center point. But the center point is largely irrelevant for our present purposes so I am setting it aside.
of potential counterexamples. For instance, the Thomist tradition holds that God is a necessary being whose existence cannot be proved \emph{a priori}. If this is true, it is a counterexample to the Metaphysical Plenitude principle. If God’s existence can only be known using a posterori methods, then there should be a scenario in which God does not exist—but if God exists necessarily and scenarios are centered worlds then it would seem that there could be no such scenario.

This is not the only potential counterexample, but it has some important features which are, according to Chalmers, shared by all potential counterexamples. The main feature is that it is controversial; its truth is not at all obvious. If Chalmers is right on this point, then while there may be counterexamples to the Metaphysical Plenitude principle there will not be any obvious counterexamples. For the moment, at least, I will grant Chalmers this point; discussing the potential counterexamples would require lengthy digressions into largely irrelevant philosophical issues, and as we will see there are problems with Chalmers’s view that do not require us to introduce any digressions. That said, it would be better if there were a way to understand scenarios that isn’t susceptible to counterexamples (even controversial ones). Chalmers offers us one possibility.

\textbf{Scenarios as Linguistic Constructs}

One alternative to taking scenarios to be centered worlds is to construct them out of other objects. In particular, Chalmers attempts to provide a way to construct scenarios as sentences in a special constructed language. Scenarios, on this approach,

\footnote{Chalmers does have more to say about these potential counterexamples, so it is not as if this is the entirety of his response to these potential counterexamples. But his responses involve irrelevant philosophical issues as well, so I won’t be examining them here.}
will be identified with equivalence classes of “epistemically complete” sentences in an infinitary language L.

Before we proceed, it is necessary to define the notion of an epistemically complete sentence and say a few words about the properties of L. First, a sentence S in L is epistemically complete if it possesses the following properties:

1. S is epistemically possible.\[10\]

2. There is no sentence T in L such that S & T and S & ¬T are both epistemically possible.

The intuitive notion here is that S is an epistemically complete sentence of L if it fully specifies some epistemic possibility. That is, it describes an epistemic possibility completely, such that if we knew that the possibility in question obtained (and we understood S) we would know or be in a position to know all the truths in that possibility without any further empirical information. With this notion of epistemically complete sentences in hand, we can define equivalence as mutual \textit{a priori} implication and then identify scenarios with equivalence classes of epistemically complete sentences. Having done this, we see that scenarios (understood in this way) easily satisfy the Plenitude Principle as long as each epistemically possible sentence is implied by some epistemically complete sentence (Chalmers calls this Epistemic Plenitude). As Chalmers puts it, “Defined in this way, scenarios are tailor-made to satisfy the Plenitude Principle." And if the Plenitude Principle is upheld, so too is the Core Thesis.

\[10\]Note that here, as before, Chalmers takes epistemic possibility as a primitive.
Now we turn our attention to the language in which these sentences are to be expressed. It is unlikely that there are any existing natural languages that are capable of expressing epistemically complete sentences since epistemically complete sentences are infinitary. As a result, our language L will need to be capable of expressing infinitary sentences. Furthermore, it would be good if we had a relatively limited vocabulary V with the following property:

Scrubability of Truth: There is a relatively limited vocabulary V such that for any truth S, there is a V-truth D such that D implies S. [11]

V-truths are truths that can be expressed using only terms in V. Chalmers want a vocabulary with this property because he wants epistemic intensions to have interesting structure to them. Here is why such a vocabulary is necessary to achieve this result. Suppose we do not have a vocabulary with the Scrubability of Truth property. Then it might turn out that the only way to construct an epistemically complete sentence S in our language L is by directly specifying the truth value of all of the possible sentences in L. While this may give us a coherent description of epistemic space, it will not be an informative one.

Chalmers thus argues that it is plausible that we do have a vocabulary with the Scrubability of Truth property. In broad strokes, his argument has two stages. First, he points out that if we have reason to think that we can give an epistemically complete description of the actual world with a limited vocabulary, then (because the actual world is not relevantly special) we should be able to give epistemically complete descriptions of all epistemic possibilities. Second, he claims that we can

---

[11] Chalmers calls this Scrubability of Truth II; I’m omitting Scrubability of Truth I and leaving off the number to avoid confusion.
in fact give an epistemically complete description of the actual world using a relatively limited vocabulary. He suggests that the vocabulary needs to include only those terms which are necessary to fully specify what he calls “PQTI”—the physical, qualitative, and indexical truths, as well as a “that’s-all truth.” Chalmers does not defend this claim in FTDS; for the purposes of that paper, he says it’s sufficient to show that we have reason to think there is some limited vocabulary that can do the trick. It doesn’t need to be the PQTI-vocabulary in particular. For the moment, I will accept Chalmers’s arguments here as unproblematic and assume that we can, in fact, produce an epistemically complete description of the actual world using some limited vocabulary.\footnote{That said, I do have issues with his claim that having a vocabulary that covers the actual world entails having a vocabulary that handles all epistemic possibilities. See \S 2.6, below.}

So to summarize: Chalmers’s claim is that we can analyze scenarios in terms of sentences in a specially-formulated language L. L has a particular vocabulary, some subset of which (V) possesses the following property: any truth expressible in L is implied by some sentence D consisting entirely in terms in V. Furthermore, some sentences in L are epistemically complete. Epistemically complete sentences fully specify epistemic possibilities. Scenarios, then, turn out to be equivalence classes of epistemically complete sentences, where equivalence is understood as mutual \textit{a priori} implication.\footnote{This is Chalmers’s definition of equivalence; see, for instance, \cite{Chalmers} p. 84.}

On this analysis of scenarios, epistemic possibility is underwritten by scenarios, while metaphysical possibility depends on worlds. A sentence is epistemically possible if it is verified by some scenario; it is metaphysically possible if it is true at
some world. This allows the analysis to handle EPMIs. On this analysis of scenarios, an EPMI is a sentence which is verified by at least one scenario but which is false in every possible world.

4.2 Criticisms

In this section I will outline a number of problems that Chalmers’s view faces.

4.2.1 Deep Epistemic Necessity and the A Priori

Chalmers characterizes deep epistemic necessity in terms of a priori knowledge. His characterization also includes a modal element, and this modal element raises questions about the nature of deep epistemic necessity. Deep epistemic necessity, according to Chalmers, corresponds to “actual or potential a priori knowledge.” What does it mean, though, for some piece of knowledge to be potentially a priori? This is a modal notion. The most natural rendering of potential a priori knowledge is something like this: p is potentially a priori if p is knowable using a priori reasoning alone.

This leaves open a number of interesting possibilities. First, it allows for some a posteriori knowledge to count for the purposes of determining deep epistemic necessity and possibility. For example, consider the proof of the four color theorem. This theorem states it is possible to color any map using no more than four colors in such a way that no adjacent regions of the map will share a color. The proof of this theorem (proved in 1976 by Kenneth Appel and Wolfgang Haken) was done with computer assistance; Appel and Haken used special programs to show that certain
special maps could be colored using four or fewer colors. The fact that this proof was done with computer assistance means that our claims to know that the four color theorem is true are based on \textit{a posteriori} reasoning and observations—in this case, observations about the computer, Appel and Haken’s program, and so on. But it seems likely that we could come to know that the four color theorem is true using only \textit{a priori} reasoning, if we were sufficiently clever or if we had improved memory and processing speed, comparable to that of a computer. So the four color theorem is plausibly knowable \textit{a priori}, and hence plausibly deeply epistemically necessary even though we actually have only \textit{a posteriori} knowledge that it is knowable \textit{a priori}.

Second, it is likely that there are deeply epistemically impossible statements which we intuitively would judge to be epistemically possible using our ordinary notion of epistemic possibility. For example, consider Goldbach’s conjecture. Goldbach’s conjecture is that every even integer greater than 2 can be expressed as the sum of two prime numbers. The truth of this conjecture is (as of this writing) unknown. Intuitively, then, we want to say that it is epistemically possible that Goldbach’s conjecture is true, and that it’s epistemically possible that it is false. Suppose, though, that it is false. If it’s false, and we can know that it’s false, then it’s likely that we can know that it’s false using \textit{a priori} reasoning (owing to its status as a mathematical claim). So if it’s false, it’s plausible that it’s deeply epistemically impossible. But Chalmers’s modal operators for deep epistemic modality quantify over the entirety of epistemic space; we get the ordinary modal operators by quantifying over a por-
tion of epistemic space. So on Chalmers’s view, if a statement is deeply epistemically impossible, it is epistemically impossible for any knowledge base whatsoever.

Finally, if Chalmers wants to introduce potential *a priori* knowledge, we might wonder how best to understand this potentiality. It seems like it has to be some kind of physical or metaphysical possibility that’s in play here, as opposed to epistemic possibility. That is, in order to avoid circularity, a proposition that is potentially known *a priori* has to be one such that it is metaphysically or physically possible to know *a priori*. The notion of possibility that’s in play here might instead be logical possibility, if this is conceived of as something distinct from metaphysical possibility. This option was discussed in §1.5.3; while there is conceptual space for it, it’s not clear how to proceed in developing an account of logical possibility. Moreover, if we have an account of possibility on which logical and metaphysical possibility are distinct and the metaphysical possible worlds are a proper subset of the logical possible worlds, then we might be able to give an account of epistemic modality that solves the problem of EPMIs without introducing the two-dimensional machinery that Chalmers and I favor by allowing epistemic modals to quantify over logically possible worlds while metaphysical modals quantify over only the metaphysically possible worlds.

There are other issues that arise here: as a being with limited cognitive abilities, there are presumably limits to what I can know *a priori*. There are people whose

\[14\]This problem parallels the problem of EPMIs introduced in chapter 1. My view has a similar problem, but I adopt a different and independently plausible intuitive understanding of epistemic modality which accounts for the problem; see the discussion of metaphysical stability in §2.2 for more. It is not clear that Chalmers’s intuitive understanding of epistemic modality can do the same.
cognitive abilities outstrip my own in relevant ways. For instance, I may not be able to grasp an *a priori* proof of the four color theorem, even if one were presented to me. A trained mathematician, on the other hand, has a better chance of being able to perform this feat. So it’s likely that what’s knowable *a priori* for a trained mathematician is not the same as what’s knowable *a priori* for me. It also seems likely that there could be beings whose cognitive abilities and *a priori* reasoning outstrip those of any human. If something is deeply epistemically necessary for such a being, does that mean it’s deeply epistemically necessary for me as well?

The reason these issues matter is this: Chalmers introduces deep epistemic necessity in order to go beyond a “dependence on the shifting state of an individual’s knowledge, to capture some sort of rational *must*.” But these questions show that it’s not clear that deep epistemic necessity accomplishes this goal. At best, deep epistemic necessity seems to provide a kind of upper bound on *a priori* knowledge. While we may not be capable of learning all deeply epistemically necessary truths (due to our cognitive limitations), we can say that anything which is deeply epistemically impossible is not knowable by any being. This is all well and good, but insofar as deep epistemic necessary is supposed to be connected with what rationality demands *of us* it isn’t particularly informative. And if deep epistemic necessity is not connected with what rationality demands of us, then we lose sight of the point of introducing deep epistemic necessity in the first place. There are a number of distinct (and independently interesting) notions lurking in the background here; there is no problem with giving an account of what is knowable in principle, for example, or with giving an account of what is knowable by humans, either collectively in individually. The problem for deep epistemic necessity is that it is supposed to capture
a rational "must" that is constitutively tied to our rational abilities, but what it actually does is capture something closer to knowability in principle.

4.2.2 The Analysis of Epistemic Modality

In order for a theory to count as a successful analysis of epistemic modality, the theory cannot take epistemic possibility or epistemic necessity as primitives. Using this criterion, Chalmers’s view fails as an analysis of epistemic modality. Chalmers is fairly explicit in taking deep epistemic necessity as a primitive (See, for example, [2] p.79), and he defines epistemic possibility in terms of epistemic necessity in the standard way. So it’s clear that Chalmers does not meet my criterion for what counts as a successful analysis of epistemic modality; my task is to show why the criterion is a good one.

I take it that part of what it is to offer an analysis of some phenomenon is to offer an explanation for that phenomenon. Ideally this explanation with allow us to understand the phenomenon to be explained in terms of some other phenomena which we already understand. But taking epistemic necessity as a primitive makes it impossible to offer this sort of explanation. Since Chalmers’s theory takes epistemic necessity as a primitive, it seems like no matter what other virtues it has, we will not see any improvement on whatever grasp we had on epistemic modality prior to encountering his account. That is, if we’re confused about the nature of epistemic necessity (and/or epistemic possibility) before encountering Chalmers’s theory, we’ll be confused after encountering the view as well.
But perhaps that’s not the goal. It would be good to have a theory which allows us to determine the truth conditions for an epistemic modal, for instance, even if that theory doesn’t explain why any given epistemic modal has its particular truth conditions. Even if Chalmers’s account fails as an analysis of epistemic modality (using the criterion given above), it might succeed at some other worthy tasks. It is worth noting that, whatever Chalmers’s goal may be in presenting his account, the fact that it fails as an analysis of epistemic modality makes it less attractive compared to the metapropositional approach. The metapropositional approach, like any theory, has its primitives, but neither epistemic necessity nor epistemic possibility are among them.

4.2.3 Conditionals with Impossible Antecedents

One problem that is not unique to Chalmers, but is found in many views that involve considering a world as actual, involves conditionals with impossible antecedents. I will use Chalmers as an example. Suppose scenarios can be identified with centered worlds, as Chalmers suggests. Given this identification, Chalmers makes the following claim: to say that it is epistemically possible that water is XYZ is to say that there is a centered world such that, if it is the actual world, then (it is metaphysically necessary that) water is XYZ.

Now consider the conditional “If Twin Earth is the actual world, then water is XYZ." There is only one actual world, and we know that it is not (and, it seems, could not be) Twin Earth. This conditional has an impossible antecedent. We do not have an account of conditionals with impossible antecedents on which only the conditionals that defenders of the propositional approach want to be true in fact turn
out to be true, and Chalmers at least does not offer one. It is probably not necessary for defenders of the propositional approach to offer a fully fleshed-out account of conditionals with impossible antecedents, but they do need to provide some way for us to make reliable judgments regarding the truth of such conditionals insofar as they are relevant to epistemic modality. Since the metapropositional approach does not require the evaluation of counterfactuals, it avoids this problem entirely.

4.2.4 Scrutability and Weird Scenarios

Chalmers claims that scenarios are scrutable. Recall that the scrutability of scenarios manifests itself in a certain kind of cognitive transparency. That is, if we have a good grasp on a scenario and we know that the world the scenario describes is objectively thus-and-such, the scrutability of the scenario assures us that we will be in a position to know the reference of various expressions (such as "water" or "Cary Grant").

Scrutability is a strong requirement. Chalmers claims that there are relatively limited vocabularies that allow us to derive (using a priori reasoning alone) all the facts about a world from a global description of that world, but there is reason to think this claim is false. While it may be that some worlds can be described in such a way as to make them scrutable, it is not clear that all worlds have this property. Consider the following examples (inspired by [28]):

Case 1: Mixed Earth   Mixed Earth is a world much like Earth. However, while Earth has H$_2$O, Mixed Earth is a world in which we find both H$_2$O and XYZ in roughly equal quantities. Lakes, rivers, and oceans are comprised of a mixture of H$_2$O
and XYZ, as are the bodies of Mixed Earth’s residents. Mixed Earth’s residents can
drink H$_2$O, XYZ, or the mixture of XYZ and H$_2$O and respond in the same way to
all three options.

Case 2: Partitioned Earth  Partitioned Earth is also a world in which we find both
H$_2$O and XYZ in roughly equal quantities. However, any given lake, river, or ocean
is either entirely H$_2$O or entirely XYZ. Roughly half of Partitioned Earth’s residents
have H$_2$O as part of their makeup; the other half have XYZ. H$_2$O-people can drink
XYZ and vice versa.

Case 3: Coke Earth  Coke Earth is a world like Earth. In particular, Coke Earth
contains H$_2$O in all the places that Earth does. However, due to a lucrative marketing
arrangement, the people of Coke Earth never imbibe H$_2$O without mixing it with
other chemicals. In fact, they exclusively drink Coke Classic, Cherry Coke, and other
soft drinks produced by the Coca-Cola Company.$^{15}$ The residents of Coke Earth have
in fact been genetically modified to be unable to drink H$_2$O except when it is found
in Coke products.

In all of these cases, we can ask the same question: what is the referent of the
term “water" if these worlds are considered as actual? Chalmers thinks we have fairly
clear intuitions about the answer to this question in the case of Twin Earth (and not
without reason). But what about Mixed Earth and its cohorts? I, at least, lack clear
intuitions about these worlds. First, note that all of these cases appear to be deeply
epistemically possible—that is, they all can be considered as actual given what we

$^{15}$ I have received no promotional considerations from Coca-Cola or any other soft drink manufacturers.
know *a priori* (though they are all inconsistent with our *a posteriori* knowledge of the actual world). The scrutability requirement applies to all scenarios, and since all deep epistemic possibilities are covered by some scenario of other, we should be able to say what “water” refers to in each case. Let’s start with Mixed Earth. When we consider Mixed Earth as actual, we don’t seem to have any reason to think “water” refers to H$_2$O rather than XYZ (assuming it can’t refer to both). Perhaps “water” on Mixed Earth functions like “jade” does (or did) in the actual world—but perhaps it doesn’t. I, at least, lack strong intuitions one way or the other. And now consider Partitioned Earth. Partitioned Earth has the same kinds of stuff as Mixed Earth, but in different arrangements. Here, too, I lack clear intuitions, excepting the following: it seems less likely to me that “water” functions like “jade” does in the actual world if Partitioned Earth is considered as actual than it does if Mixed Earth is considered as actual. And so it goes for Coke Earth as well; since the liquid found in lakes and rivers is not the drinkable liquid that comes from our taps on Coke Earth, it’s not clear which liquid should be the referent of “water.”

Here is why cases like these are problematic for Chalmers. Chalmers claims that scenarios are scrutable, which means we can infer (using only *a priori* means) the reference of various expressions from the description of a possibility. Cases like 1-3, though, suggest that there are at least some worlds where this inference is not possibility. Admittedly, cases 1-3 trade on our intuitions (or more precisely our lack of intuitions) about particular cases, and intuition is not the same as *a priori* reasoning. But Chalmers trades on our intuitions as well, using examples like Twin Earth to motivate his scrutability claim. Furthermore, the fact that we lack clear intuitions about cases 1-3 gives us reason to doubt that they are scrutable.
Once we have a few examples like cases 1-3 above, it is easy to produce more. What about the world which is like Coke Earth, except that the inhabitants cannot drink H₂O at all? Or the world in which its inhabitants are surrounded by H₂O but drink XYZ, or vice versa? Again, these cases are not ruled out by our a priori knowledge, and so there is a scenario in which these cases obtain. We ought, on Chalmers’s view, to be able to consider such possibilities as actual. The ease with which we can construct problematic cases like these suggests that Chalmers can’t beat this objection simply by addressing a few individual cases.

Chalmers might try to escape this criticism by claiming that the descriptions given above are insufficiently detailed. Scenarios are, after all, epistemically complete, while the descriptions in cases 1-3 certainly not. This is, I think, an unsatisfactory response. First of all, cases 1-3 are described in as much detail as Twin Earth, and while it may very well be true that some worlds require more description than others in order to be scrutable this shows that complete descriptions are not always necessary in order to discern the referent of at least some terms. Second, epistemically complete descriptions are inaccessible for actual humans; even if we are to leave aside the fact that they are not expressed or expressible in any natural language, the fact that they are infinitary sentences puts them outside the reach of our cognitive machinery. If scrutability is to mean anything it has to apply to incomplete descriptions of scenarios. Perhaps, with more information, it would be possible to resolve cases 1-3 above—but that doesn’t mean that all problematic cases can be handled in this way. Ultimately, I think these considerations show that we would have to accept a large promissory note in order to buy the scrutability claim.
The metapositional approach presented in chapter 2 does not have to worry about this objection. The basic content constraint and the knowledge base constraint serve to rule out many of the worlds like Mixed Earth and its cohorts. It may be that some cases like 1-3 remain, but as I have not made any claim comparable to Chalmers’s scrutability claim their existence is largely unproblematic for my preferred view.

4.2.5 Scrutability of Truth, Limited Vocabularies, and Alien Properties

In §1.2.5, we saw that Chalmers claims that the actual world is not special, in the following sense: if we can produce an epistemically complete description of the actual world with some limited vocabulary (perhaps the PQTI-vocabulary, perhaps some other vocabulary), then we should be able to do the same with any epistemic possibility. He offers no further argument in support of this claim. I think we have good reason to think it is false.

Consider first the case of alien properties. Alien properties are properties that are uninstantiated in the actual world, but are instantiated elsewhere. For the moment, let us stipulate that there are alien properties. Then there are worlds that instantiate properties that are not instantiated here. And we have no reason to think that such a world could be described using whatever limited vocabulary allows us to produce an epistemically complete description of the actual world.

The existence of alien properties is controversial, but as we will see nothing in my final argument depends on their existence. They are being used as a way of introducing some unusual epistemic possibilities, and that’s it.
That said, Chalmers does not say that we can offer an epistemically complete description of every world using some limited vocabulary. His claim, rather, is that every (deep) epistemic possibility (i.e. every scenario) can be so described. But remember that for Chalmers, a sentence is deeply epistemically possible if it’s not ruled out by a priori reasoning. It seems highly likely that we do not have a priori knowledge of precisely which properties are instantiated in the actual world. That is, our a priori knowledge does not allow us to specify which properties there are; given a list of all possible properties, we may not be in a position to say which ones are instantiated in the actual world. Consider, for instance, the positions of some ancient Greek philosophers, who thought that all material objects consisted in various amounts of fire, earth, water, and air; it seems odd to say that these philosophers could have realized their error as a result of a priori reasoning rather than as a result of empirical science.

The upshot of all this is that whether or not there are alien properties, it is plausible that there are deep epistemic possibilities (i.e. scenarios) that instantiate properties not found in the actual world. And we have no reason to think that these scenarios can be described using whatever limited vocabulary we find that allows us to produce an epistemically complete description of the actual world.

4.3 Conclusions

As we have seen, Chalmers’s account centers on his notion of a scenario. Scenarios can be understood as centered worlds or as linguistic constructions, and these for-\footnote{These properties, strictly speaking, wouldn’t be alien properties, since they might not be instantiated at any world.}
mulations are not equivalent. However we understand scenarios, though, we can see that the account of epistemic modality that Chalmers builds with them is capable of handling EPMIs. That is, whether we take scenarios to be centered worlds or linguistic constructs, Chalmers is capable of explaining how there can be true EPMIs.

That said, his view also suffers from some significant problems which, collectively, make it less attractive than the metapropositional approach that I favor. In particular, Chalmers’s account fails as an analysis of epistemic modality, since it takes epistemic possibility as a primitive, while the metapropositional approach does not. Chalmers’s demand for scrutability also creates problems for him; epistemic modals are, according to Chalmers, always supposed to be accessible using reasoning alone, but as we have seen this is implausible. Since the metapropositional approach has no scrutability requirement, it does not suffer this deficiency. These flaws probably stem from the fact that Chalmers is concerned primarily with validating the Core Thesis rather than solving the problem of EPMIs. Nonetheless, the fact that these problems exist makes his approach less than satisfactory for my purposes here.
Chapter 5

Further Developments

In this chapter, I briefly discuss two directions for possible further work. First, we have thus far been assuming that the knowledge base used in evaluating epistemic modals is fixed and shared by all participants in the conversation. However, on many cases one or both of these assumptions do not hold. What is known changes over time, and in some cases each participant in a conversation has their own individual knowledge base which contains knowledge not shared by all parties to the conversation. I will discuss some examples that suggest (in a broad way) how to generalize the metapropositional framework to cover these cases. Second, our discussion has mainly been informed by the semantic properties of epistemic modals in ordinary language. We might want to make this discussion more precise by introducing a formal language with operators for epistemic necessity and possibility (as well as operators for metaphysical modality). I will briefly explore some of the considerations that arise in formalizing my account and producing an epistemic logic.

5.1 Variable Knowledge Bases

In this section I will discuss the behavior of epistemic modals in cases where the knowledge base is allowed to vary. In particular, I will examine cases where the
knowledge base is updated to reflect new information and cases where different participants in a conversation have different knowledge bases.

5.1.1 Updating

The simplest cases to consider are cases where the knowledge base is updated with new information because of some external event. Here is one such case.¹

**Case 1: Simple Updating**

Alice, Bob and Carol are playing seven-card stud poker. Alice has a pair of kings in the hole and is showing an ace, a jack, and a four. Bob is showing an ace, an eight, and a nine. Carol is showing an ace, a two, and a seven. Carol has been betting aggressively. Alice says “Carol might have another ace in the hole.” Bob agrees; both judge the sentence to be true. New cards are dealt. Alice gets another four, Bob a 10, and Carol another ace. In light of this new information, Alice says “Carol must not have another ace in the hole.”

In this case, each player has their own private information about the state of the game, as each player knows their own hole cards but does not know any of the other players’ hole cards. For the moment, we will ignore this private information and

¹This case, and several subsequent cases, use a game of seven-card stud poker as a framing device. Seven card stud is played like so: first, the dealer deals out two face-down cards to each player. These are the “hole” cards. The dealer then deals a single face-up card to each player. There is a round of betting, followed by the dealer dealing out another face-up card to each player. The players alternate rounds of betting and cards being dealt until each player has six cards, at which time there is a round of betting followed by the seventh card being dealt face-down. There is then a final round of betting, after which all players remaining in the game show their hands. The best five-card poker hand wins.
focus instead on the public state of the game (i.e. the face-up cards that are visible to all players). Given this public information, Alice’s initial utterance expresses a truth. The public information in this case includes the following relevant facts: Alice has an ace, a jack, and a four. Bob has an ace, an eight, and a nine. Carol has an ace, a two, and a seven. So the knowledge base includes those worlds in which these facts obtain.

We evaluate Alice’s modal by looking for a world of production in the knowledge base in which 1) Alice’s utterance retains the same basic content it has in the actual world, and 2) Alice’s utterance expresses a truth. There is such a world; it’s the one in which Carol has an ace in the hole. When the next round of cards is dealt, Carol gets another ace. So now the public information includes the fact that Alice has an ace, Bob has an ace, and Carol has two aces. Since there are only four aces in a deck of cards, there will no longer be a world in the knowledge base in which Alice’s original utterance will express a truth. For the same reason, Alice’s second utterance will express a truth.

This case shows the need to index epistemic modals to a particular knowledge base. Alice’s first utterance is indexed to the public knowledge of the game state after the fifth card has been dealt; the second is indexed to the public knowledge of the game state after the sixth card is dealt. This is the first expansion of my semantic theory.

The question now arises: how do we tell which knowledge base is in play for any given epistemic modal? I have no systematic answer to this question, and I
am not sure that there is one available to give. There are, however, a few general points to be made. The relevant knowledge base will generally be determined by the conversational context. It can be set explicitly, as when we say things like “For all I know…” or “Given what cards are on the table…” In cases where the relevant knowledge base is not made explicit, I want to suggest that the default knowledge base will be the common knowledge base, where common knowledge is understood in something like Lewis’s sense, as given in [19]. That is, \( p \) is common knowledge for some conversation if everyone involved in the conversation knows that \( p \), everyone involved knows that everyone knows that \( p \), and so on. Alternately, we might require that everyone involved in the conversation knows that \( p \), and the fact that everyone involved knows that \( p \) is mutually manifest to all members of the conversation.

However we formulate the conditions required for common knowledge, it is important to see that common knowledge is distinct from mere mutual knowledge. Consider the following case:

**Case 2: The Cheaters**

Alice, Bob, and Carol are playing seven-card stud poker. Alice has a pair of kings in the hole, and is showing an ace, a jack, and a four. Bob is showing an ace, an eight, and a nine. Carol is showing an ace, a two, and a seven. Carol has been betting aggressively. Bob says, “Carol might have another ace in the hole.” Unbeknownst to the other players, Alice is a cheater; she has marked the aces, and so she knows that Carol does not have one in the hole. Bob, however, is also a cheater; he’s used a small mirror to see Carol’s cards as they were dealt and as such knows that she
does not have an ace in the hole. Carol, of course, knows what she has, and knows she doesn’t have the last ace.

In this case, everyone knows that Carol does not have the last ace, but the fact that Carol doesn’t have the last ace is not common knowledge. Since Alice doesn’t know Bob’s a cheater, she does not know that he knows what Carol has. When she evaluates Bob’s modal, she can do so using her own knowledge base, in which case she’ll judge that he has said something false or using her model of Bob’s knowledge base. In the latter case, she’d have to judge that he has said something true, since her model of his knowledge base does not include the knowledge he gained through perfidious means. This is ultimately a case of asymmetric information, discussed in case 5 below. For now, though, it is important to see that no member of the conversation can intentionally evaluate the modal using their shared knowledge as the knowledge base, because they do not know that the fact that Carol doesn’t have an ace in the hole is shared knowledge.

Let us now turn to a slightly more complicated case.

Case 3: Slightly More Complicated Updating

Alice, Bob, and Carol are playing seven-card stud poker. Alice has a pair of kings in the hole and is showing an ace, a jack, and a four. Bob is showing an ace, an eight, and a nine. Carol is showing an ace, a two, and a seven. Carol has been betting aggressively. Alice says “Carol might have another ace in the hole.” Bob agrees. Carol says “I don’t have another ace.” Alice and Bob now judge the modal claim to be false.
Formally, this case resembles the first case. The modal is first indexed to the public knowledge of the game state after the fifth card is dealt, but after Carol asserts that she does not have a fourth ace it is instead indexed to the public knowledge of the game state plus the fact that Carol does not have another ace. What makes this case more complicated is that it shows how certain speech acts (with assertion being the paradigmatic example) can influence the knowledge base. It also raises the specter of pragmatic considerations. Poker is a game in which players have reason to deceive others about the state of the game; in the example above, Alice and Bob accept Carol’s assertion as true, but they need not do so (in fact, they likely would be unwise to do so without further corroborating evidence). These considerations are not part of the semantic story of epistemic modals on my account. They do, however, introduce added complexities to the task of determining what knowledge base to use in evaluating an epistemic modal, and it is because of these complexities that I suspect that there is no fully general, systematic story to be told about how we make such determinations.

It is worth noting, though, that in both this case and the preceding one we can represent the changing knowledge base in more-or-less the same way. When the external state of affairs changes, or when an assertion is accepted, it adds propositions to the knowledge base. We then index our epistemic modals to the expanded knowledge base, which will occasionally give us different results when it comes to the truth or falsity of our modal claims. In the case of assertion, we don’t have to add the asserted proposition to our knowledge base; if we think the speaker is lying, mistaken, or otherwise not speaking the truth, we generally will not add the proposition asserted to the knowledge base. Similarly, when we seem to observe that
something is the case (for instance, that there are now four aces visible), we can add
the relevant proposition to our knowledge base, but we need not. If, for instance, I
have only glanced at the cards, I might think I was mistaken in seeing four aces and
so suspend judgment—in which case I should not add the proposition that all four
aces are accounted for to the knowledge base.

One pragmatic issue that deserves further attention is this: “must” occasionally
gets used to indicate something weaker than epistemic necessity (in the intuitive
sense given in chapter 2). That is, we sometimes say that something must be the
case even when it is not implied by the knowledge base. For example, consider the
following case:

**Case 4: Weak Necessity**

Alice, Bob, and Carol are playing seven-card stud poker. Alice has a pair of kings in
the hole and is showing an ace, a jack, and a four. Bob is showing an ace, an eight,
and a nine. Carol is showing an ace, a two, and a seven. Carol has been betting
aggressively. Alice says “Carol must have another ace in the hole.” Bob agrees.2

In this case, it seems at first that I am committed to saying that Alice has said
something false, and that Bob is mistaken when he agrees with her. But this is
perhaps counterintuitive; we do, in fact, sometimes use “must” in just this way
without causing problems. My intuition (as a poker player) is that Alice’s statement
has at least one reading on which she says something true.

---

2This case is related to the phenomenon of weak necessity, discussed in [24] and [14].
Roughly, $s$ is weakly epistemically necessary iff $s$ is more probable than any other alternative to $s$ given what is known.
There is more than one way for me to account for this intuition. First, I can appeal to the pragmatic phenomenon of loose use: strictly speaking, Alice has said something false, but in doing so she has implicated something true—for example, that Carol’s having the final ace is the best explanation for her betting behavior. Second, I want to suggest that we might be able to account for this example by indexing our epistemic modals to a knowledge base. In this case, we expand the knowledge base to include things that are only surmised. This suggestion requires more development, since surmises are not factive (unlike knowledge) and thus can be false or inconsistent with the rest of the knowledge base.

5.1.2 Asymmetric Information

Cases of asymmetric information are cases where one (or more than one) participant in a conversation has information not available to some or all of the other participants. Let us return to the poker game to find such a case.

Case 5: Asymmetric Information

Alice, Bob, and Carol are playing seven-card stud poker. Alice has a pair of aces in the hole and is showing an ace, a jack, and a four. Bob is showing a four, an eight, and a nine. Carol is showing an ace, a two, and a seven. Carol has been betting aggressively. Bob says “Carol might have another ace in the hole.” Alice, however, has information that Bob lacks. She judges Bob’s claim to be false while recognizing that, if she were to restrict herself to the public information about the game, Bob’s claim would be true.
What is interesting about this case is that there seems to be a conflict with regard to Alice’s judgments about Bob’s modal claim. On one hand, Alice reasons, Carol could not have another ace because Alice has the fourth and final ace. On the other hand, Bob doesn’t know that Alice has the last ace. He knows where three of the aces are, and Carol’s betting behavior suggests she has a good hand. It is consistent with what he knows that Carol has an ace in the hole. Bob, then, seems to have said something true in claiming that Carol might have the fourth ace.

We can account for the discrepancy here using the same technique discussed in the previous section. By indexing the modal operator to a knowledge base, we can account for the seeming conflict in Alice’s judgments. When indexed to Alice’s knowledge base, Bob’s claim is false; when indexed to his knowledge base, his claim is true.

That said, our ability to account for the apparent conflict in Alice’s judgments does not constitute a resolution of those conflicts. We might still wonder which knowledge base Alice ought to use in evaluating Bob’s modal claim and whether or not his claim is true or false simpliciter. This question is more difficult to handle, and any answer I could offer would be controversial. My intuition is that Alice would claim that Bob believes that Carol might have an ace; she would also claim that

---

3More precisely, Alice’s intuitive judgment that Bob’s claim is true is underwritten by indexing the modal operator to Alice’s model of Bob’s knowledge base, which can be represented as the set of propositions that Alice knows that Bob knows. In the poker case, Alice knows that Bob knows the public information about the state of the game, which is the relevant information. In some cases, though, we don’t know what our conversational partners know, and in those cases we might not be able to evaluate their epistemic modals with respect to our models of their knowledge bases.

4See, for example, [10] and [33], for discussion of the difficulties that arise here.
Bob’s belief is reasonable given what he knows. Nonetheless she would say he has a false belief. We should privilege Alice’s judgments in this case because she has the larger knowledge base. That said, case 5 is unusual in that the relevant portion of Bob’s knowledge base is a subset of Alice’s knowledge base. This will not always be the case; sometimes, the members of the conversation will have different knowledge bases without one being a subset of another. In those cases, it is not obvious what we should say about their truth or falsity simpliciter.

5.1.3 Asymmetric Updating

We can also combine cases of updating with cases of asymmetric information.

Case 6: Asymmetric Updating

The poker game has progressed, and now the final cards are about to be dealt. Alice, Bob, and Carol each have (among other things) an ace showing. None of them have an ace in the hole. Carol has continued her aggressive betting. Bob says, “Carol might have an ace in the hole.” Alice and Bob agree that this claim is true. Each of them is dealt a final card, which is dealt face-down. Alice gets an ace. Now she has an ace showing and another ace face down; moreover, since the other two aces are both visible, she knows that no one else has a pair of aces. Bob again says, “Carol still might have an ace in the hole.” Now Alice finds herself in a situation like the one discussed above.

Again, Alice has a conflict: she has information that Bob lacks, and that information bears on her judgments regarding his modal claim. Given the public information
that Bob has available, his claim seems to be true. In fact, since Alice knows that Bob does not have the fourth ace, Alice can even say that it is consistent with Bob’s private knowledge of his own hole cards that Carol has another ace. It is not, however, consistent with what Alice knows (after the final card is dealt) that Carol has another ace.

Unsurprisingly, I want to suggest that we can handle this case by indexing epistemic modal operators to knowledge bases. The fact that Alice and Bob agree with respect to Bob’s first claim is explained by the fact that it is indexed to the public knowledge of the game state. Her conflicting intuitions with respect to Bob’s second claim can be explained in the same way discussed in case 3, above.

5.2 Epistemic Logic

5.2.1 Accessibility

The knowledge base bears a close connection to the accessibility relation in standard modal and epistemic logics. The knowledge base restricts the scope of the modal operators in much the same way that accessibility relations do, by establishing which worlds are relevant for evaluating modal expressions. However, the knowledge base is a set of worlds. In a two-dimensional logic formulas are evaluated at pairs of worlds, and so it makes sense to think of the accessibility relation as a relation between pairs of worlds.
We can use the knowledge base to generate an accessibility relation. In this way, our truth conditions for epistemic modals can be made to coincide with the familiar truth conditions in terms of accessibility. We will say that, for worlds \( w, u, x, \) and \( y, \langle w, u \rangle R \langle x, y \rangle \) iff \( x = y \) and \( x \) is in the knowledge base. \( w, u, \) and \( x \) need not be distinct. This relation is transitive and Euclidean, but it is not symmetric or reflexive. This means the system will validate the 4 and 5 axioms of modal logic for epistemic modals, but not the T or B axioms. The properties of the accessibility relation determine the logical properties of epistemic necessity, which in turn bears on our understanding of the relation between epistemic necessity and knowledge.

Rejecting the T axiom might seem problematic, since the T axiom \((\Box_E p \supset p)\) tells us that anything which is epistemically necessary is true, but this is not as troublesome as it seems. Let \( a \) represent the actual world. Whenever we evaluate an epistemic modal, \( a \) will never be excluded as a potential world of production; \( a \) has to be consistent with the knowledge base, since knowledge is factive, and \( a \) could not be excluded by the basic content constraint, since the basic content of an expression is determined by its use in \( a \). So \( \langle a, a \rangle \) will always be relevant in evaluating epistemic modals. Since the T axiom is true at every point on the diagonal, it will be true at \( \langle a, a \rangle \), and thus if something is epistemically necessary for a resident of \( a \) then it will be true. The same holds for the B axiom.

Moreover, since the T and B axioms are true at every point on the diagonal, and every point of evaluation (whether on the diagonal or not) is related to all and only the points on the diagonal, it follows that the necessitation of T and B will be true

\(^5\)I am ignoring the basic content constraint for the moment; reintroducing the constraint does not substantially change the logical landscape.
at all points of evaluation. This means that the necessitations of the T and B axioms will be valid in my system. We end up with the following axioms: K, 4, 5, and the necessitations of T and B. We might call this system □S5.

The T and B axioms will hold at every pair of worlds for metaphysical modals. For metaphysical modals, a pair of worlds will be related to every other pair of worlds in the same row of the two-dimensional array and nothing else. This means that the collection of all pairs of worlds will be partitioned into equivalence classes; each pair of worlds in one partition will be related to all other pairs in the same partition and nothing else. There will not be any overlap between partitions. So for metaphysical modals, the T and B axioms will hold, along with all of the other axioms of S5.

We can get a different picture of the logical landscape by adopting a different definition of validity. Validity is typically defined as truth in all models. In this system, that amounts to saying a formula is valid if it is true at all pairs of worlds. With this understanding of validity, T will not be valid (although its necessitation will), as discussed above. But we could adopt a different definition of validity. For instance, we could adopt the following definition:

$$\models S \text{ iff } (x, x) \models S \text{ for all worlds } x \text{ in the knowledge base}$$

When we adopt this definition, the T axiom will hold for epistemic modals. By adopting this definition we are effectively understanding validity as truth in all pairs of worlds on the diagonal.
It is plausible that these two notions of validity capture two distinct phenomena of interest. In [18], for example, Steve Kuhn suggests that the notion of truth at all diagonal points in a two-dimensional framework like this corresponds to a notion of logical truth as truth in virtue of form, whereas the notion of truth at all pairs corresponds to a notion of logical truth as necessary truth in virtue of form. Both characterizations of logical truth are common. Alternatively we might, following Chalmers, understand the weaker form of validity as a prioricity. The project of characterizing the formulas valid on either understanding of validity would be of interest from a logical point of view.

5.2.2 Mixing Modals

The accessibility semantics given the previous section makes it possible to investigate logical connections between epistemic and metaphysical necessity. In §2.1 of chapter 2, we gave the following truth conditions for epistemic and metaphysical possibility (using ◊_M to represent metaphysical possibility and ◊_E to represent epistemic possibility):

- ◊_M is true at (w₁, w₂) when associated with the metaproposition f(w) the iff there is some world of evaluation v such that f(w₁) is true at v.

- ◊_E is true at (w₁, w₂) when associated with the metaproposition f(w) iff there is some world of evaluation u such that f(u) is true at u.

The corresponding necessity operators are defined in the usual way.
With these definitions in hand, we can explore the connections between epistemic and metaphysical modals. One such connection has already been discussed in §2.2.4 of chapter 2. In that section, I introduced the notion of metaphysical stability. A metaphysically stable sentence is one that expresses a contingent proposition in any world of production if it expresses a contingent proposition in the actual world and expresses a noncontingent proposition is any world of production if it expresses a noncontingent proposition in the actual world. In our formal language, a sentence \( s \) is metaphysically stable if

\[
\Box_M s \rightarrow \Box_E \Box_M s \quad \text{and} \quad \neg \Box_M s \rightarrow \neg \Box_E \Box_M s.
\]

I also conjectured that all sentences are metaphysically stable. While this conjecture might turn out to be false, as shown in §2.2.4 potential counterexamples seem to be fairly thin on the ground, and I have not found any clear counterexamples that do not rest on controversial claims.

In the formal language, it is easy to express all manner of connections between epistemic and metaphysical modalities. For example, we can find examples of sentences like “For all I know, it might be that squares must have four sides.” The outer modal (“might”) is clearly an epistemic modal and takes wide scope over the necessity operator. It is possible to hear the necessity operator as an epistemic modal or as a metaphysical modal. Formally, these readings would be represented as \( \Diamond_E \Box_E s \) and \( \Diamond_E \Box_M s \) (where \( s \) represents “squares have four sides.”) In the former case, the claim is equivalent to “For all I know, squares must have four sides.” In the latter case, the epistemic modal is true if there is a world \( w \) such that \( \Box_M s \) is true when produced in and evaluated at \( w \).
In this sentence, the possibility modal, which is given an epistemic interpretation, takes wide scope. When the explicit epistemic qualifier is left off, we are left with "It might be that squares must have four sides." In this case, the possibility modal always takes wide scope, and we can give the sentence three interpretations: first, as a pair of epistemic modals; second, as a pair of metaphysical modals; finally, as a mixed sentence with one epistemic modal and one metaphysical modal. What is interesting is that in this last case (a mixed modal), it is hard to hear the sentence with the metaphysical modal taking wide scope over an epistemic modal. In general, cases where metaphysical modals take wide scope over epistemic modals are hard to come by in natural language. It is not clear why this might be; this is an open question in semantics which has not heretofore received much attention. Whatever the reason for this asymmetry, the formal language given here might provide an opportunity to study the logical connections that hold among these different forms of modality.


