Two-Dimensional Semantics

David J. Chalmers
Philosophy Program
Research School of Social Sciences
Australian National University

Two-dimensional approaches to semantics, broadly understood, recognize two "dimensions" of the meaning or content of linguistic items. On these approaches, expressions and their utterances are associated with two different sorts of semantic values, which play different explanatory roles. Typically, one semantic value is associated with reference and ordinary truth-conditions, while the other is associated with the way that reference and truth-conditions depend on the external world. The second sort of semantic value is often held to play a distinctive role in analyzing matters of cognitive significance and/or context-dependence.

In this broad sense, even Frege's theory of sense and reference might qualify as a sort of two-dimensional approach. More commonly, two-dimensional approaches are understood more narrowly to be a species of possible-worlds semantics, on which each dimension is understood in terms of possible worlds and related modal notions.

In possible-world semantics, linguistic expressions and/or their utterances are first associated with an *extension*. The extension of a sentence is its truth-value: for example, the extension of 'Plato was a philosopher' is true. The extension of a singular term is its referent: for example, the extension of 'Don Bradman' is Bradman. The extension of a general term is the class of individuals that fall under the term: for example, the extension of 'cat' is the class of cats. Other expressions work similarly.

One can then associate expressions with an *intension*, which is a function from possible worlds to extensions. The intension of a sentence is a function that is true at a possible world if and only if the sentence is true there: the intension of 'Plato was a philosopher' is true at all worlds where Plato was a philosopher. The intension of a singular term maps a possible world to the referent of a term in that possible world:

In (E. Lepore & B. Smith, eds) *Oxford Handbook of the Philosophy of Language*. Oxford University Press, forthcoming.

the intension of ‘Don Bradman’ picks out whoever is Bradman in a world. The intension of a general term maps a possible world to the class of individuals that fall under the term in that world: the intension of ‘cat’ maps a possible world to the class of cats in that world.

It can easily happen that two expressions have the same extension but different intensions. For example, Quine’s terms ‘cordate’ (creature with a heart) and ‘renate’ (creature with a kidney) pick out the same class of individuals in the actual world, so they have the same extension. But there are many possible worlds where they pick out different classes (any possible world in which there are creatures with hearts but no kidneys, for example), so they have different intensions. When two expressions have the same extension and a different intension in this way, the difference in intension usually corresponds to an intuitive difference in meaning. So it is natural to suggest that an expression’s intension is at least an aspect of its meaning.

Carnap (1947) suggested that an intension behaves in many respects like a Fregean sense, the aspect of an expression’s meaning that corresponds to its cognitive significance. For example, it is cognitively significant that all renates are cordates and vice versa (this was a nontrivial empirical discovery about the world), so that ‘renate’ and ‘cordate’ should have different Fregean senses. One might naturally suggest that this difference in sense is captured more concretely by a difference in intension, and that this pattern generalizes. For example, one might suppose that when two singular terms are cognitively equivalent (so that ‘*a*=*b*’ is trivial or at least knowable a priori, for example), then their extension will coincide in all possible worlds, so that they will have the same intension. And one might suppose that when two such terms are cognitively distinct (so that ‘*a*=*b*’ is knowable only empirically, for example), then their extensions will differ in some possible world, so that they will have different intensions. If this were the case, the distinction between intension and extension could be seen as a sort of vindication of a Fregean distinction between sense and reference.

However, the work of Kripke (1980) is widely taken to show that no such vindication is possible. According to Kripke, there are many statements that are knowable only empirically, but which are true in all possible worlds. For example, it is an empirical discovery that Hesperus is Phosphorus, but there is no possible world in which Hesperus is not Phosphorus (or vice versa), as both Hesperus and Phosphorus are identical to the planet Venus in all possible worlds. If so, then ‘Hesperus’ and ‘Phosphorus’ have the same intension (one that picks out the planet

Venus in all possible worlds), even though the two terms are cognitively distinct. The same goes for pairs of terms such as 'water' and 'H_2O': it is an empirical discovery that water is H_2O, but according to Kripke, both 'water' and 'H_2O' have the same intension (picking out H_2O in all possible worlds). Something similar even applies to terms such as 'I' and 'David Chalmers', at least as used by me on a specific occasion: 'I am David Chalmers' expresses nontrivial empirical knowledge, but Kripke's analysis entails that I am David Chalmers in all worlds, so that my utterances of these expressions have the same intension. If this is correct, then intensions are strongly dissociated from cognitive significance.

Still, there is a strong intuition that the members of these pairs ('Hesperus' and 'Phosphorus', 'water' and 'H_2O', 'I' and 'David Chalmers') differ in some aspect of meaning. Further, there remains a strong intuition that there is *some* way the world could turn out so that these terms would refer to different things. For example, it seems to be at least *epistemically* possible (in some broad sense) that these terms might fail to corefer. On the face of it, cognitive differences between the terms is connected in some fashion to the existence of these possibilities. So it is natural to continue to use an analysis in terms of possibility and necessity to capture aspects of these cognitive differences. This is perhaps the guiding idea behind two-dimensional semantics.

Two-dimensional approaches to semantics start from the observation that the extension and even the intension of many of our expressions depend in some fashion on the external world. As things have turned out, my terms 'water' and 'H_2O' have the same extension, and have the same (Kripkean) intension. But there are ways things could have turned out so that the two terms could have had a different extension, and a different intension. So there is a sense in which for a term like 'water', the term's extension and its Kripkean intension depend on the character of our world. Given that *this* world is actual, it turns out that 'water' refers to H_2O, and its Kripkean intension picks out Venus in all possible worlds. But if another world had been actual (e.g. Putnam's Twin Earth world in which XYZ is the clear liquid in the oceans), 'Hesperus' might have referred to something quite different (e.g. XYZ), and it might have had an entirely different Kripkean intension (e.g. one that picks out XYZ in all worlds).

This suggests a natural formalization. If an expression's (Kripkean) intension itself depends on the character of the world, then we can represent this dependence by a function from worlds to intensions. As intensions are themselves functions from worlds to extensions, this naturally suggests a two-dimensional structure. We can represent this structure diagramatically as follows:

	H_2O-world	XYZ-world	…
H_2O -world	H_2O	H_2O	…
XYZ-world	XYZ	XYZ	…
…	…	…	…

This diagram expresses an aspect of the two-dimensional structure associated with the term 'water'. It is intended to express the intuitive idea that if the H_2O -world turns out to be actual (as it has), then 'water' will have a Kripkean intension that picks out H_2O in all worlds; but if the XYZ-world turns out to be actual (as it has not), then 'water' will have a Kripkean intension that picks out XYZ in all worlds. Intuitively, the worlds in the column on the left represent ways the actual world can turn out (these are sometimes thought of more precisely as possible contexts of utterances, and are sometimes thought of as epistemic possibilities), while the worlds across the top reflect counterfactual ways that a world could have been (these are sometimes thought of more precisely as possible circumstances of evaluation, and sometimes thought of as metaphysical possibilities). It is sometimes said that worlds on the left column (one world per row), making up the "first dimension" of the matrix, correspond to different worlds *considered as actual*; while the worlds in the top row (one world per column), making up the "second dimension" of the matrix, correspond to different worlds *considered as counterfactual*.

This two-dimensional matrix can be seen as a *two-dimensional intension*: a function from ordered pairs of worlds to extensions. Such a function is equivalent to a function from worlds to intensions, and seen this way can be regarded as capturing the intuitive idea that a term's intension depends on the character of the actual world. One can also recover the intuitive idea that a term's *extension* depends on the character of the actual world by examining the "diagonal" of this matrix, i.e. the cells that correspond to the same world considered as actual and as counterfactual. In the

example above: where the H_2O -world is considered as actual and as counterfactual, then 'water' picks out H_2O, while if the XYZ-world is considered as actual and as counterfactual, then 'water' picks out XYZ. We can say that an expression's "diagonal intension" is a function mapping a world *w* to the term's extension when *w* is taken as both actual and as counterfactual. So the diagonal intension of 'water' maps the H_2O -world to H_2O, the XYZ-world to XYZ, and so on.

We can then see how pairs of terms with the same extension and the same Kripkean intension might nevertheless have different two-dimensional intensions, and different diagonal intensions. For example, 'water' and 'H_2O' have the same Kripkean intension, but it is plausible that if the XYZ-world had turned out to be actual, they would have had different Kripkean intensions: 'water' would have had an intension that picked out XYZ in all worlds, while 'H_2O' still would have had an intension that picked out H_2O in all worlds. If so, then these terms have different two-dimensional intensions and different diagonal intensions.

One can make a case that something similar applies with 'Hesperus' and 'Phosphorus', and with 'I' and 'David Chalmers': the members of each pair have a different two-dimensional intension and a different diagonal intension. If so, then this begins to suggest that there is some sort of connection between an expression's two-dimensional intension (or perhaps its diagonal intension) and its cognitive significance. One might even speculate that an expression's diagonal intension behaves in some respects like a Fregean sense, in a way that might vindicate Carnap's project.

At this point it must be acknowledged that things are not so simple. A number of different two-dimensional approaches to semantics have been developed in the literature, by Kaplan (1979, 1989), Stalnaker (1978), Chalmers (1996, 2002a, 2004), and Jackson (1998), among others; and closely related two-dimensional analysis of modal notions have been put forward by Evans (1977) and by Davies and Humberstone (1981). These approaches differ greatly in the way that they make the intuitive ideas above precise. They differ, for example, in just what they take the "worlds" in the left column to be, and they differ in their analysis of how a term's intension and/or extension depends on the character of the actual world. As a result, different approaches associate these terms with quite different sorts of two-

dimensional semantic values, and these semantic values have quite different connections to cognitive significance.

In what follows, I will first go over the two-dimensional approaches pioneered in the 1970s by Kaplan, Stalnaker, Evans, and Davies and Humberstone.[1] Each of these approaches can be seen as sharing some of the formal structure described above, but with quite different conceptual underpinnings. Each of the approaches asserts some sort of connection between two-dimensional semantic values and apriority, but the connection is usually limited in scope, applying to indexicals (Kaplan) and to descriptive names (Evans), and 'actually'-involving expressions (Davies and Humberstone), while Stalnaker's later work rejects a connection to apriority altogether. I will then describe the more general two-dimensional approach to semantics developed in the 1990s by Chalmers, Jackson, and others. This approach associated two-dimensional semantic values with expressions of all kinds, and asserts a strong general connection between these semantic values and the domain of apriority and cognitive significance. I will close by briefly describing some applications of the framework, and by considering and responding to a number of objections.

2 Early two-dimensional approaches

2.1 Kaplan: Character and content

Perhaps the best-known broadly two-dimensional approach is Kaplan's analysis of the character and content of indexicals (Kaplan 1979, 1989). According to Kaplan, his work is partly grounded in work in tense logic by Kamp (1971) and Vlach (1973), which gives a sort of two-dimensional analysis of the behavior of 'now'. Kaplan applies his analysis to indexicals such as 'I', 'here', and 'now', as well as to demonstratives such as 'this' and 'that'. Kaplan's well-known analysis is described elsewhere in this volume, so I will describe it only briefly here.

---

[1] More detailed discussions of all of these two-dimensional frameworks and their interrelations can be found in two recent collections: the March 2004 special issue of *Philosophical Studies* on "The Two-Dimensional Framework and its Applications", and the book *Two-Dimensional Semantics: Foundations and Applications* (Garcia-Carpintero and Macia, 2006). See especially Chalmers 2006, Davies 2004, and Stalnaker 2004, and also the discussion in Soames 2005.

For Kaplan, the "worlds" involved in the first dimension are *contexts of utterance*: these can be seen as at least involving the specification of a speaker and a time and place of utterance, within a world. The "worlds" involved on the second dimension are *circumstances of evaluation*: these are ordinary possible worlds at which the truth of an utterance is to be evaluated.

Consider an expression such as:

(1) I am hungry now

According to Kaplan's analysis, when this expression is uttered by Joe at time t_1, it expresses a proposition that is true if and only if Joe is hungry at t_1. We can call this proposition expressed the *content* of the utterance. This content can naturally be represented as an intension that is true at all and only those worlds (those circumstances of evaluation) in which Joe is hungry. (Kaplan regards propositions as structured entities rather than intensions, but the difference does not matter much here.) In a different context — say, a context with Diana speaking at t_2 — an utterance of the same expression will have a different content. This content will be a proposition that is true at a world if and only if Diana is hungry at t_2 in that world.

The *character* of an expression is a function from contexts to contents, mapping a context of utterance to the content of that expression in that context. (If content is seen as an intension, then character is a sort of two-dimensional intension.) So the character of 'I am hungry' maps the first context above to the proposition that Joe is hungry at t_1, and the second context above to the proposition that Diana is hungry at t_2. Extending this idea to subsentential indexical terms, we can say that the character of 'I' maps the first context to Joe and the second context to Diana; more generally, it maps any context into the speaker in that context. Similarly, the character of 'now' maps any context into the time specified in that context.

The above definition of character is still somewhat imprecise, and many tricky issues come up in giving a precise definition. But to a rough first approximation, one can say that the character of an expression maps a context to the content that the expression would have if uttered in that context. There is more to say than this (especially as Kaplan intends his analysis to apply even to contexts in which there is no token of the relevant utterance), but this is enough for now. In general, character is associated with an expression type rather than with an expression token, although this

matter is complicated somewhat by the case of demonstratives such as ‘this’ and ‘that’, whose character may vary between different utterances.

On Kaplan’s analysis, the character of indexicals such as ‘I’, ‘now’, and ‘here’, as well as the character of demonstratives such as ‘this’ and ‘that’, reflects their cognitive significance. For example, ‘I am here now’ has a propositional content that is true in only some worlds, but its character yields a proposition that is true in all contexts of utterance. (Kaplan does not “diagonalize” character into an intension, but it would be easy enough to do so. If one did so, then ‘I am here now’ would be associated with a diagonal intension that is necessarily true.) So the character rather than the content seems to reflect the fact that the sentence can be known a priori (or near enough). Likewise, when a true utterance of ‘this is that’ is cognitively significant, the occurrences of ‘this’ and ‘that’ will refer to the same object, but their characters will differ. So at least in these domains, character behaves a little like a Fregean sense.

This behavior does not extend to other expressions, however. For example, Kaplan holds that names refer to the same individual in any context of utterance. On this view, co-extensive names such as ‘Mark Twain’ and ‘Samuel Clemens’ will have exactly the same character, and an identity such as ‘Mark Twain is Samuel Clemens’ will have a character that yields a true proposition in every context, even though the identity appears to be a posteriori and cognitively significant. Something similar applies to natural kind terms such as ‘water’. So on Kaplan’s analysis, names and natural kind terms have a “constant character” that is dissociated from their cognitive roles.

One can diagnose the situation by noting that character is most closely tied to the patterns of context-dependence associated with an expression, rather than to the expression’s cognitive significance. In the case of indexicals, the patterns of context-dependence of an expression are themselves closely associated with the expression’s cognitive significance. But for many other expressions, such as names and natural kind terms, cognitive significance is strongly dissociated from patterns of context-dependence. (The same goes for numerous ordinary context-dependent expressions, such as ‘tall’.) As a result, in the general case, Kaplan’s framework is better suited to the analysis of the context-dependence of expressions than to an analysis of their cognitive significance.

2.2 Stalnaker: Diagonal proposition and proposition expressed

Stalnaker's analysis starts with the idea that although sentences such as 'Hesperus is Phosphorus' express necessary truths, they are sometimes used to convey contingent information about the world. Stalnaker (1978) analyzes this contingent information as the *diagonal proposition* associated with an utterance.

On Stalnaker's analysis, the *proposition expressed* by an utterance is a standard intension, or a set of possible worlds. So the proposition expressed by an ordinary utterance of 'Hesperus is Phosphorus' is the set of worlds in which Hesperus is Phosphorus, which is the set of all worlds (leaving aside questions about existence). Stalnaker defines the *propositional concept* associated with an utterance as a function from possible worlds to propositions, mapping a world to the proposition that that utterance would express in that world. He then defines the *diagonal proposition* associated with an utterance as a function that maps a possible world to the truth value of that utterance when used in that possible world.

Stalnaker individuates utterances in such a way that a given utterance could have been used with an entirely different meaning. For example, an utterance of 'Hesperus is Phosphorus' could have been used to express the proposition that Mark Twain is George Bush, in a world *w* in which 'Hesperus' is used as a name for Twain and 'Phosphorus' is used as a name for Bush. It follows that while the propositional concept of my utterance maps the *actual* world to the proposition that Hesperus is Phosphorus, it maps world *w* to the proposition that Twain is Bush (which is itself presumably the empty set of worlds). The diagonal proposition of my utterance maps the actual world to the truth-value of the former proposition in the actual world (true), and maps world *w* to the truth-value of the latter proposition in *w* (false). So although my utterance of 'Hesperus is Phosphorus' expresses a necessary proposition in the ordinary sense, it is associated with a contingent diagonal proposition.

Stalnaker's propositional concept is a sort of two-dimensional intension, and his diagonal proposition is the associated diagonal intension. Like Kaplan, Stalnaker's framework can be seen as capturing a certain way in which the content of an utterance depends on the context in which it is uttered. But while Kaplan's analysis is in effect restricted to contexts in which the expression retains its original meaning, Stalnaker's analysis ranges over contexts in which the expression is used with entirely different meanings. As a result, Stalnaker characterizes his use of the two-dimensional framework as a "metasemantic" use: unlike Kaplan's character, diagonal propositions

are not really part of the meaning of an utterance, but rather capture something about how meaning is determined by the external world.

Stalnaker uses this framework mainly to analyze the information conveyed by assertions. In a context where the hearer knows the full meanings of the terms used in an utterance (e.g., if they know that 'Hesperus' and 'Phosphorus' both refer to Venus), and where this knowledge is common ground between speaker and hearer, then the utterance will convey its original propositional content. But if the hearer does not know the meanings of the terms, then the utterance will convey a different content. In particular, it will convey the diagonal proposition of the utterance: here, the proposition that 'Hesperus is Phosphorus' expresses a truth. If the common ground between speaker and hearer includes partial knowledge of meaning — say, the knowledge that 'Hesperus' is used to refer to the evening star and that 'Phosphorus' is used to refer to the morning star — then worlds outside this common ground are in effect excluded by presuppositions, and the diagonal proposition will in effect be equivalent to the proposition that the morning star is the evening star (at least across the relevant range of worlds). So in such a context, an assertion of 'Hesperus is Phosphorus' will convey the information that the morning star is the evening star.

In his 1978 paper, Stalnaker says that if one defines an operator '†' such that '†*P*' is true iff *P* has a necessary diagonal proposition, then '†' is equivalent to the "a priori truth" operator. In later work (e.g. Stalnaker 2004), however, he retracts that claim. It is easy to see why. Even paradigmatic a priori claims such as '1+1=2' do not have a necessary diagonal proposition: the diagonal proposition of '1+1=2' is false at a world where '1' refers to 3 and '2' refers to 7, for example. It is true that a statement such as 'Hesperus is the evening star', which is arguably an a priori truth, will have a diagonal proposition that is true in all worlds in a class that is restricted as in the previous paragraph (by imposing the restriction that 'Hesperus' is used to refer to the evening star). But in this case, it is the restriction that is doing all the work in connecting the diagonal proposition to a priori truth.

Because of this, there is no strong connection between diagonal propositions and a priori truth. There is sometimes a connection between an utterance's diagonal proposition and its cognitive significance, but this connection arises only in certain contexts where certain special restrictions due to limited knowledge of meaning are in force. Because of this, Stalnaker's diagonal propositions cannot be used to ground a two-dimensional approach to the cognitive significance of linguistic items in general.

Instead, they are most useful for analyzing what is conveyed by utterances when there is limited knowledge of meaning in place.

2.3 Evans: Deep necessity and superficial necessity

Evans' analysis (Evans 1977) is focused on *descriptive names*: names whose reference is fixed by a description. His main example is the name 'Julius', which is stipulated to be a name for whoever invented the zip, if anyone uniquely invented it (I will omit references to uniqueness in what follows, but they should be tacitly understood). He considers the following sentence:

(2) If anyone invented the zip, Julius invented the zip.

If one follows Kripke, then (2) expresses a contingent proposition. 'Julius' picks out the actual inventor (William C. Whitworth) in all worlds, so the proposition is false in all worlds where someone other than Whitworth invented the zip.

According to Evans, however, this sort of contingency is superficial. (2) is *superficially contingent*, in that the claim 'It might have been the case that someone other than Julius invented the zip' is true. Superficial necessity and contingency of a sentence turns on how it embeds within modal operators: *S* is superficially necessary iff 'It is necessary that *S'* is true. But Evans suggests that in a deeper sense, (2) is necessary. He holds that the sentence is necessary because it expresses a necessary *content*. On Evans' view, there is a semantic rule connecting 'Julius' with the invention of the zip, and this semantic rule makes it the case that the content of (2) is necessarily true.

Evans' framework has two modal operators, rather than two intensions. The framework does have two semantic values: the proposition expressed by an sentence, which is something like the familiar proposition that is true in all worlds where Whitworth invented the zip, and the content of the sentence, which behaves as characterized above. Neither propositions nor contents are characterized as intensions, but it is easy enough to define intensions in the vicinity. We can say that the *superficial intension* of *S* is the set of worlds in which the proposition expressed by *S* is true: roughly, the set of worlds *w* such that 'if *w* had obtained, *S* would have been the case' is true. We can say that the *deep intension* of *S* is the set of worlds in which the content of *S* is true. In these terms, (2) has a superficial intension that is false at some worlds, but a deep intension that is true at all worlds.

In the case of descriptive names such as 'Julius', deep necessity (as opposed to superficial necessity) seems closely connected to apriority, and deep intensions are closely connected to an expression's cognitive role. It is tempting to extend this connection beyond the case of descriptive names, but Evans does not discuss other expressions, and it is not entirely clear how an extension would go. A more precise analysis of Evans' notion of deep necessity would require a more precise understanding of his notion of "content", which serves as something of an unanalyzed primitive in his 1978 article.

From other work, it seems clear that Evans thinks in the case of ordinary proper names (as opposed to descriptive names), there is a semantic rule that ties a name to its referent, so that the referent is part of the content. Correspondingly, it seems that Evans held that identities involving ordinary proper names have a content that is necessary, so that an identity such as 'Mark Twain is Samuel Clemens' is not only superficially necessary but deeply necessary. If this is right, then the two names involved will have the same deep intension. So in these cases (and probably in analogous cases involving natural kind terms), deep necessity and deep intensions are not as strongly connected to cognitive significance or to apriority as in the case of descriptive names.

2.4 Davies and Humberstone: 'Fixedly Actually' and 'Necessarily'

The two-dimensional framework of Davies and Humberstone (1981) is based on an analysis of the operator 'actually' (***A***). ***A****P* is true in a world *w* iff *P* is true in the actual world. Davies and Humberstone note that '*P* iff ***A****P*' is contingent but knowable a priori. They suggest that although the sentence is contingent, there is an intuitive sense in which it is necessary: intuitively, no matter which world turns out to be the actual world, '*P* iff ***A****P*' will be true. Likewise, for a contingent empirical truth *P*, ***A****P* will be necessary, but there is an intuitive sense in which it is contingent: intuitively, there are some worlds such that if those worlds had been actual, then ***A****P* would have been false.

This intuition can be formalized by introducing a "floating" actual world into a possible-worlds model. Instead of simply designating a fixed world as the actual world, we take actuality to be a feature that can attach to different worlds. We can then evaluate sentences in a world *w*, where a world *w*' is taken to be actual ("considered as actual"). Or equivalently, we can evaluate sentences at pairs of worlds

(*w'*, *w*), where the first world represents the world that is designated as actual, and the second world represents the world in which the sentence is evaluated (relative to the designation of the first world as actual).[2]

Doubly-indexed evaluation behaves as follows. A sentence *P* without modal operators is true at (*w'*, *w*) iff *P* is true at *w* according to ordinary singly-indexed evaluation. □P is true at (*w'*, *w*) iff *P* is true at (*w'*, *v*) for all *v* (i.e. iff *P* is true at all worlds relative to *w'* considered as actual). ***A****P* is true at (*w'*, *w*) iff *P* is true at (*w'*, *w*) (i.e. iff *P* is true at *w'* when *w'* is considered as actual). In conjunction with the obvious semantics for truth-functional logical operators, this suffices to recursively define doubly-indexed evaluation of sentences in modal propositional logic (including '□' and '***A***') in terms of standard singly-indexed evaluation of atomic sentences.

Davies and Humberstone then introduce the further operator "fixedly" (***F***), which can be defined as follows: ***F****P* is true at (*w'*, *w*) iff *P* is true at (*v*, *w*) for all *v* (i.e. iff *P* is true at *w* relative to all worlds considered as actual). The "fixedly actually" operator ***FA*** is consequently such that ***FA****P* is true at (*w'*, *w*) iff ***A****P* is true at (*v*, *w*) for all *v*, i.e., iff *P* is true at (*v*, *v*) for all *v*. So ***FA****P* is true iff *P* is true at all worlds *w* when *w* itself is considered as actual.

The two crucial modal operators here are □ and ***FA***. We can say that *P* is necessary when □*P* is true (i.e. when *P* is true at all worlds when our world is considered as actual), and that *P* is *FA*-necessary when ***FA****P* is true (i.e. when *P* is true at all worlds *w* when *w* is considered as actual). Let us say that *P* is A-involving iff *P* contains an instance of ***A*** or of ***F***. It is easy to see that when *P* is not A-involving, *P* will be FA-necessary iff it is necessary. But when *P* is A-involving, the two may come apart. In particular, the sentence '*P* iff ***A****P*' is not necessary, but it is FA-necessary. Likewise, for a contingent atomic truth *P*, ***A****P* is necessary, but it is FA-contingent. So Davies and Humberstone suggest that FA-necessity captures the intuitive sense in which these two sentences are necessary and contingent respectively.

Davies and Humberstone also extend the discussion to Evans' case of descriptive names. They observe that descriptive names such as 'Julius' behave very much like A-involving descriptions of the form 'The actual inventor of the zip'. For example,

---

[2] For simplicity of presentation, I depart from Davies and Humberstone's own formalization, but the formalization here gives equivalent results.

just as 'Julius invented the zip' seems contingent and a priori, 'The actual inventor of the zip invented the zip' seems contingent and a priori. Furthermore, it is easy to see that when formalized in modal predicate logic, sentences of the latter form are not necessary but are FA-necessary. This mirrors Evans' claim that 'Julius (if he exists) invented the zip' is not superficially necessary but is deeply necessary.
Davies and Humberstone suggest the natural hypothesis that descriptive names are in fact abbreviated A-involving descriptions, and that Evans' deep necessity is just FA-necessity.

Davies and Humberstone speculate that all contingent a priori sentences may be (perhaps tacitly) A-involving sentences that are contingent and FA-necessary.[3] They also suggest that some necessary a posteriori sentences are A-involving sentences that are necessary and FA-contingent: for example, 'The actual *F* is *G*' (where 'the *F* is *G*' is contingent) and analogous claims involving descriptive names. They speculate tentatively that natural kind terms (such as 'water') might be seen as abbreviated A-involving descriptions (such as 'the actual waterish stuff around here'), in which case necessary a posteriori identities such as 'water is H_2O' may also be necessary and FA-contingent. However, they do not extend the claim to all necessary a posteriori sentences. In particular, they hold that ordinary proper names are not A-involving, so that identities involving ordinary proper names (such as 'John is Tom') are FA-necessary iff they are necessary. It follows from this that necessary a posteriori identities involving these names are FA-necessary, rather than FA-contingent.

Davies and Humberstone do not posit two semantic values to go along with their two modal operators, but one could naturally do so. We can say that the standard intension of *P* is true at *w* iff *P* is true at *w* when our world is considered as actual (i.e. iff *P* is true at (*a*, *w*), where *a* is the actual world), and that the FA-intension of *P* is true *w* iff *P* is true at *w* when *w* is considered as actual (i.e. iff *P* is true at (*w*, *w*)). We can also define the two-dimensional intension of *P* in the obvious way; then the FA-intension will be equivalent to the "diagonal" of the two-dimensional intension.

As defined here, FA-intensions are closely tied to apriority for some sentences: especially for A-involving sentences, and for tacitly A-involving sentences such as

---

[3] This claim will be true only if all contingent a priori sentences are A-involving. For some reasons for doubt about this (involving indexical contingent a priori sentences, for example), see Chalmers 2006.

those involving descriptive names and perhaps natural kind terms (if these are indeed tacitly A-involving). If the "actually" operator were the *only* source of the necessary a posteriori and the contingent a priori, then there would be a strong general tie between FA-intensions and apriority. But if there are other sources of the necessary a posteriori and the contingent a priori (such as ordinary proper names and indexicals), then in these cases, FA-intensions will not be closely tied to apriority at all.

3 Two-Dimensionalism

The two-dimensional approaches discussed above all introduce "first-dimensional" semantic values or modal notions that are more strongly connected to apriority and to cognitive significance than are the more familiar "second-dimensional" semantic values and modal notions. But in each of these approaches, the connection is somewhat attenuated. In the case of Kaplan's character, the connection only applies in the case of indexicals. In the case of Evans it is asserted only for descriptive names. In the case of Davies and Humberstone, it holds only for A-involving expressions and tacitly A-involving expressions such as descriptive names and perhaps some natural kind terms. In the case of Stalnaker, it applies only under certain strong restrictions on the domain of a diagonal proposition, or not at all.

In recent years, a number of philosophers (e.g. Chalmers 1996, 2002, 2004 and Jackson 1998, 2004; see also Braddon-Mitchell 2004, Lewis 1993, and Wong 1996) have advocated a two-dimensional approach on which first-dimensional semantic values are connected to apriority and cognitive significance in a much stronger and more general way. On this approach, the framework applies not just to indexicals and descriptive names, but to expressions of all sorts. Proponents hold that any expression (or at least, any expression token of the sort that is a candidate for having an extension) can be associated with an intension that is strongly tied to the role of the expression in reasoning and in thought. The term *two-dimensionalism* is usually used for views of this sort.

3.1 The core claims of two-dimensionalism

Five core claims of two-dimensionalism are as follows.

> (T1) Every expression token (of the sort that is a candidate to have an extension) is associated with a primary intension, a secondary intension, and a two-dimensional intension. A primary intension is a function from scenarios to extensions. A secondary intension is a function from possible worlds to extensions. A two-dimensional intension is a function from ordered pairs of scenarios and worlds to extensions.
>
> (T2) When the extension of a complex expression token depends compositionally on the extensions of its part, the value of each of its intensions at an index (world, scenario, or ordered pair) depends in the same way on the values of the corresponding intensions of its parts at that index.
>
> (T3) The extension of an expression token coincides with the value of its primary intension at the scenario of utterance and with the value of the secondary intension at the world of utterance.
>
> (T4) A sentence token *S* is metaphysically necessary iff the secondary intension of *S* is true at all worlds.
>
> (T5) A sentence token *S* is a priori (epistemically necessary) iff the primary intension of *S* is true at all scenarios.

In what follows I will first clarify and motivate these principles, without precisely defining all of the key notions or making a case for their truth. In later sections, I will discuss how the relevant notions (especially the notion of a primary intension) can be defined, in such a way that the principles might be true. These principles should not be taken to provide an exhaustive characterization of two-dimensionalism, but they lie at the core of the view.

Start with claim (T1). Here, a scenario is something akin to a possible world, but it need not be a possible world. In the most common two-dimensionalist treatments, a scenario is a *centered world*: an ordered triple of a possible world along with an individual and a time in that world. Other treatments of scenarios are possible (see Chalmers 2004), but I will use this understanding here.

An expression's secondary intension (or what Jackson calls its C-intension) is just its familiar post-Kripkean intension, picking out the extension of the expression in counterfactual worlds. For example, the secondary intension of a token of 'I' as used by speaker *S* picks out *S* in all worlds. The secondary intension of 'water' picks out H_2O in all worlds. The secondary intension of 'Julius' picks out William C. Whitworth in all worlds. And so on.

An expression's primary intension works quite differently. I will defer a full characterization, but some examples will give a rough idea. The primary intension of a token of 'I', evaluated at a centered world, picks out the designated individual at the "center" of that world. (So the primary intension of my use of 'I', evaluated at a world centered on Napoleon, picks out Napoleon, rather than David Chalmers.) The primary intension of a token of 'water', very roughly, picks out the clear, drinkable liquid with which the individual at the center is acquainted. (So the primary intension of my use of 'I', evaluated at a "Twin Earth" world centered on a subject surrounded by XYZ in the oceans and lakes, picks out XYZ, rather than H_2O.) The primary intension of a token of 'Julius' picks out whoever invented the zip in a given world. (So the primary intension of 'Julius', evaluated at a world where Tiny Tim invented the zip, picks out Tiny Tim, rather than William C. Whitworth.) And so on.

Thesis (T1) also holds that expression tokens can be associated with a *two-dimensional intension*: roughly, a function from (scenario, world) pairs to extensions. We can then say that at least on the centered worlds understanding, the primary intension coincides with the "diagonal" of the two-dimensional intension (i.e. the value of *S*'s primary intension at a centered world *w* coincides with the value of *S*'s two-dimensional intension at the pair (*w*, *w**), where *w**is the possible-world element of *w*). Likewise, the secondary intension coincides with the "row" of the two-dimensional intension determined by the scenario of an utterance (i.e. the value of *S*'s secondary intension at a world *w* coincides with the value of *S*'s two-dimensional intension at (*a*, *w*), where *a* is the scenario of utterance). However, for most purposes the two-dimensional intension of an expression is somewhat less important than its primary and secondary intension, and the two-dimensionalist need not hold that an expression's primary and secondary intension are derivative from its two-dimensional intension.

Thesis (T2) says that the primary and secondary intensions of a complex expression depend on the primary and secondary intensions of its parts according to the natural compositional semantics. For example, the primary intension of 'I am Julius' will be true at a scenario if the individual at the center of that scenario is the inventor of the zip in that scenario.

Thesis (T3) states a natural connection between the intensions and the extension of an expression token. This thesis requires that for every utterance, just as there is one world that is the world of the utterance, there is also one scenario that is the scenario of the utterance. If scenarios are understood as centered worlds, this will be a world centered on the speaker and the time of the utterance. When evaluated at the scenario and world of utterance, the primary and secondary intensions (respectively) of an expression token will coincide with the extension of the expression token. At other worlds and scenarios, however, the values of these intensions may diverge from the original extension, and from each other.

Turning to claims (T4) and (T5): Here, we can say that *S* is a priori when it expresses a thought that can be justified independently of experiences. *S* is metaphysically necessary when it is true with respect to all counterfactual worlds (under the standard Kripkean evaluation). Thesis (T4) is a consequence of the standard understanding of metaphysical necessity and the corresponding intensions. Thesis (T5) is intended to be an analog of thesis (T4) in the epistemic domain.

Thesis (T5) is the distinctive claim of two-dimensionalism. It asserts a very strong and general connection between primary intensions and apriority, one much stronger than obtains with the other two-dimensional frameworks discussed earlier. It is possible that a two-dimensionalist might grant some limited exceptions to thesis (T5) (say, for certain complex mathematical statements that are true but unknowable) while still remaining recognizably two-dimensionalist. But it is crucial to the two-dimensionalist position that typical a posteriori identities involving proper names or natural kind terms, such as 'Mark Twain is Samuel Clemens' or 'water is H_2O', have a primary intension that is false in some scenario.

Consequences of the previous theses include the following:

> (T6) A sentence token *S* is necessary a posteriori iff the secondary intension of *S* is true at all worlds but the primary intension of *S* is false at some scenario.

(T7) A sentence token *S* is contingent a priori iff the primary intension of *S* is true at all scenarios but the secondary intension of *S* is false at some world.

So two-dimensionalism proposes a unified analysis of the necessary a posteriori: all such sentences have a necessary secondary intension but a contingent primary intension. Likewise, it proposes a unified analysis of the contingent a priori: all such sentences have a contingent primary intension but a necessary secondary intension.

From the previous theses, one can also draw the following conclusions about the primary and secondary intensions of both sentential and subsentential expressions. Here '*A*' and '*B*' are arbitrary expressions of the same type, and '*A*≡*B*' is a sentence that is true iff '*A*' and '*B*' have the same extension. For example, if *A* and *B* are singular terms, '*A*≡*B*' is just the identity statement '*A*≡*B*', while if *A* and *B* are sentences, '*A*≡*B*' is the biconditional '*A* iff *B*'.

(T8) '*A*≡*B*' is metaphysically necessary iff *A* and *B* have the same secondary intension.

(T9) '*A*≡*B*' is a priori (epistemically necessary) iff *A* and *B* have the same primary intension.

It follows that for a posteriori necessary identities involving proper names, such as 'Mark Twain is Samuel Clemens', the two names involved will have the same secondary intensions, but different primary intensions. Something similar applies to kind identities such as 'water is H_2O'. If this is correct, then primary intensions behave in these cases in a manner somewhat reminiscent of a Fregean sense.

3.2 Epistemic two-dimensionalism

For these claims, especially claim (T5), to be grounded, we need to have a better idea of what primary intensions are. Clearly, they must differ from characters, diagonal propositions, deep intensions, and FA-intensions, at least as these are understood by their proponents. Here, I will outline one approach (the approach I favor) to understanding primary intensions. This approach, which we might call *epistemic two-dimensionalism*, is elaborated in much greater detail in other works (Chalmers 2002a, 2002b, 2004, 2006; Chalmers and Jackson 2001).

According to epistemic two-dimensionalism, the connection between primary intension and epistemic notions such as apriority requires that primary intensions should be characterized in epistemic terms from the start. On this approach, the scenarios that are in the domain of a primary intension do not represent contexts of utterance. Rather, they represent *epistemic possibilities*: highly specific hypotheses about the character of our world that are not ruled out a priori. The value of an expression's primary intension at a scenario reflects a speaker's rational judgments involving the expression, under the hypothesis that the epistemic possibility in question actually obtains.

For example, 'water is not H_2O' is epistemically possible, in the sense that its truth is not ruled out a priori. Correspondingly, it is epistemically possible that our world is the XYZ-world (or at least, that it is qualitatively just like the XYZ-world). And if we suppose that our world is the XYZ-world (that is, that the liquid in the oceans and lakes is XYZ, and so on), then we should rationally endorse the claim 'water is XYZ', and we should rationally reject the claim 'water is H_2O'. So the primary intension of 'water is H_2O' is false at the XYZ-world, and the primary intension of 'water is XYZ' is true there.

Likewise, 'Mark Twain is not Samuel Clemens' is epistemically possible, in the sense that it is not ruled out a priori. Correspondingly, it is epistemically possible that our world is a world *w* where one person wrote the books such as *Tom Sawyer* that we associate with the name 'Mark Twain', and a quite distinct person is causally connected to our use of the term 'Samuel Clemens'. If we suppose that *w* is our world, then we should rationally endorse the claim 'Mark Twain is not Samuel Clemens'. So the primary intension of 'Mark Twain is Samuel Clemens' is false at *w*.

According to two-dimensionalism, something similar applies to any Kripkean a posteriori necessity. For any such sentence *S*, the negation of *S* is epistemically possible. And it is plausible that for any such *S*, there is a world *w* such that if we suppose that our world is qualitatively like *w*, we should rationally reject *S*. If so, then the primary intension of *S* is false at *w*. If this pattern generalizes to all a posteriori necessary sentences, then any such sentence has a primary intension that is false at some scenario, as thesis (T6) above suggests.

Here, primary intensions are characterized in thoroughly epistemic terms. It should be noted that the claims above are in no tension with the Kripkean claims that

'water is H_2O' is metaphysically necessary, or that 'water' picks out H_2O in all worlds. Even Kripke allows that 'water is not H_2O' is *epistemically* possible. And it is a familiar Kripkean point that there can be an epistemic necessitation between two statements *A* and *B* even when there is no metaphysical necessitation between them (witness '*X* is the source of heat sensations' and '*X* is heat'). We simply have to strongly distinguish this sort of epistemic evaluation of sentences in worlds (which turns on epistemic necessitation) from the usual sort of counterfactual evaluation (which turns on metaphysical necessitation). Primary intensions are grounded in the former; secondary intensions are grounded in the latter.

3.3 Defining primary intensions

It remains to define primary intensions more precisely. To generalize from the above, we might suggest that the primary intension of a sentence *S* is true at a scenario *w* iff the hypothesis that *w* is actual should lead us to rationally endorse *S*. Somewhat more carefully, we can say that the primary intension of *S* is true at a scenario *w* iff *D* epistemically necessitates *S*, where *D* is a canonical specification of *w*. It remains to clarify the notion of a scenario, a canonical specification, and epistemic necessitation.

Scenarios are highly specific epistemic possibilities. On the centered-worlds version of epistemic two-dimensionalism, scenarios are identified with centered worlds. It is also possible to develop a version of epistemic two-dimensionalism where scenarios are more strongly dissociated from ordinary possible worlds (see Chalmers 2004; forthcoming a), and instead are characterized in more purely epistemic terms (for example, as maximal epistemically consistent sets of sentences in an idealized language). But I will focus on the centered-worlds understanding here.

For any possible world *w*, it is epistemically possible that *w* is actual; or at least, it is epistemically possible that a world qualitatively identical to *w* is actual. (More precisely: it is epistemically possible that *D* is the case, where *D* is a complete qualitative characterization of *w*. More on this notion shortly.) But epistemic possibilities are more fine-grained than possible worlds. For example, the information that the actual world is qualitatively like a possible world *w* is epistemically consistent with various different epistemic possible claims about one's self-location: for example, it is consistent with the claims 'It is now 2004' and 'It is now 2005'.

To handle these claims about self-location, we model epistemic possibilities using centered worlds. The individual and the time marked at the "center" of a centered world serve as a "you are here" marker, which serves to settle these claims about self-location. For a given thinker, the hypothesis that a given centered world *w* is actual can be seen as the hypothesis: '*D* is the case, I am *F*, and the current time is *G*', where *D* is a complete qualitative characterization of *w*, and *F* and *G* are qualitative descriptions that pick out the individual and the time at the center of *w*. We can think of this conjunctive claim as a *canonical specification* of the centered world in question.

In the foregoing, a qualitative vocabulary is, to a first approximation, a vocabulary that is free of terms (such as names and natural kind terms) that give rise to Kripkean a posteriori necessities and a priori contingencies. (Restricting world-descriptions to a vocabulary of this sort avoids obvious problems that would arise if we allowed, for example, 'water is H_2O' into the description of the XYZ-world. For more on the characterization of qualitative, or "semantically neutral", vocabulary, see Chalmers 2004.) A complete qualitative characterization of *w* is a qualitative statement *D* such that (i) *D* is true of *w*, and (ii) if *E* is a qualitative statement that is true of *w*, then *D* necessitates *E*.

We also need to define epistemic necessitation. To a first approximation, we can say that *D* epistemically necessitates *S* iff accepting *D* should lead one to rationally endorse *S* (without needing further empirical information, given idealized reflection). On a refined definition, we can say that *D* epistemically necessitates *S* iff a conditional of the form '$D \supset S$' is a priori. The refined definition is arguably better in some difficult cases, but for many purposes, the first approximation will suffice.

Because they are defined in epistemic terms, there is an inbuilt connection between primary intensions and the epistemic domain. In particular, there will be a strong connection to apriority. When a sentence token *S* is a priori, then it will be epistemically necessitated by any sentence whatsoever (this is especially clear for the second understanding of epistemic necessitation above), so its primary intension will be true in all scenarios. When a sentence token *S* is not a priori, then its negation will be epistemically possible, and *S* will be false relative to some highly specific epistemic possibility. As long as there is a scenario for every epistemic possibility, then the primary intension of *S* will be false in some scenario. (On the centered worlds

understanding of scenarios, the existence of a scenario for every epistemic possibility is a substantive but plausible claim; see Chalmers 2002c and 2004.) If so, then thesis (T5) will be correct.

One can define the secondary intension of a sentence in a similar, if more familiar, way. The secondary intension of *S* is true at a world *w* iff *D* metaphysically necessitates *S*, where *D* is a canonical specification of *w*. Here a canonical specificant can be characterized much as before as a complete specification, although here it is not necessary to impose the restriction to qualitative specifications. Metaphysical necessitation could be taken as basic, or perhaps better, we can define it in terms of subjunctive conditionals: *D* metaphysically necessitates *S* when a subjunctive conditional of the form 'if *D* had been the case, *S* would have been the case' is true.

One can likewise define the two-dimensional intension of a sentence. The two-dimensional intension of *S* is true at (*v*, *w*) iff *D* epistemically necessitates that *D'* metaphysically necessitates *S*, where *D* is a canonical specification of the scenario *v* and *D'* is a canonical specification of the world *w*. If we understand epistemic necessitation in terms of a priori material conditionals and metaphysical necessitation in terms of subjunctive conditionals, this will be the case iff '$D \supset (D' \Rightarrow S)$' is a priori, where the outer conditional is material and the inner conditional is subjunctive.

This discussion of the intensions of sentences can be extended to the intensions of subsentential expressions in a reasonably straightforward way. For details, see Chalmers 2004.

3.4 The roots of epistemic two-dimensionalism

The epistemic two-dimensional framework is grounded in a thesis about the *scrutability* of reference and truth: once a subject is given enough information about the character of the actual world, then they are in a position to make rational judgments about what their expressions refer to and whether their utterances are true. For example, once we are given enough information about the appearance, behavior, composition, and distribution of various substances in our environment, as well as about their relations to ourselves, then we are in a position to conclude (without needing further empirical information) that water is H_2O. And if instead we were given quite different information, characterizing our environment as a "Twin Earth" environment, then we would be in a position to conclude that water is XYZ.

Of course, if we allow the "enough information" to include arbitrary truths, such as 'water is H_2O', the scrutability claim will be trivial. But we can impose significant restrictions on the information without compromising the plausibility of the thesis. For example, one can argue that even if we restrict ourselves to truths that do not use the term 'water' or cognates, it remains the case that given enough truths of this kind, we are in a position to know the truth of 'water is H_2O' (see Chalmers and Jackson 2001). The same goes for many or most other terms, plausibly including most names or natural kind terms.

The upshot is that there is some reasonably restricted vocabulary *V*, such that for arbitrary statements *T*, then once we know enough *V*-truths we will be in a position to know (without needing further empirical information) the truth-value of *T*. Just how restricted such a vocabulary can be is an open question. Chalmers and Jackson (2001) argue that *PQTI*, a conjunction of microphysical, phenomenal, and indexical truths along with a "that's all" truth, can serve as a basis. But this claim is not required here. All that is required for present purposes is that some qualitative vocabulary, conjoined with indexical terms such as 'I' and 'now', is sufficient.[4]

This suggests that for any true sentence token *S*, there is a *V*-truth *D* such that *D* epistemically necessitates *S*, in that a subject given the information that *D* will be in a position to rationally endorse *S* (given ideal rational reflection). Furthermore, it appears that in principle, no further empirical information is needed to make this judgment; if such information were required, we could simply include it (or equivalent qualitative information) in *D* to start with. This strongly suggests that there is a non-empirical warrant for the transition from *D* to *S*. In particular, one can make the case that in these cases, the material conditional '$D \supset S$' will be a priori. (This case is made at length by Chalmers and Jackson 2001). If this is correct, then *D* epistemically necessitates *S* in the second, stronger sense given above.

---

[4] As before, a qualitative vocabulary is one that excludes terms, such as names and natural kind terms that give rise to Kripkean a posteriori necessities. A qualitative vocabulary may include all sorts of high-level expressions: 'friend', 'philosopher', 'action', 'believe', and 'square', for example. It will not designate individuals by using names: instead it will make existential claims of the form 'there exist such-and-such individuals with such-and-such qualitative properties'. Some theoretical terms (perhaps including microphysical terms) may be excluded, but information conveyed using these terms can instead be conveyed by the familiar Ramsey-sentence method, characterizing a network of entities and properties with appropriate causal/nomic connections to each other and to the observational and the phenomenal. For familiar reasons, no important information is lost by doing this.

The scrutability claim does not apply only to the actual world. It is plausible that for all sorts of scenarios, if we are given the information that the scenario is actual, then we are in a position to make a rational judgment about the truth-value of arbitrary sentences. For example, if we are given a complete qualitative characterization of the bodies visible in the sky at various times, with the feature that no body is visible both in the morning sky and the evening sky, then we should rationally reject the claim 'Hesperus is Phosphorus'. This sort of judgment is part of the *inferential role* associated with our use of the terms 'Hesperus' and 'Phosphorus'. The point is general: for any expression that we use, then given sufficient information about the actual world, certain judgments using the expression will be irrational, and certain other judgments using the expression will be rational. It is arguable that the expressions of any language user will have this sort of normative inferential role. This is just part of what being a language user involves.

It is this sort of inferential role that grounds the primary intension of an arbitrary expression (as used by an arbitrary speaker). A given sentence token will be associated with a raft of conditional rational judgments, across a wide variety of scenarios. This raft of conditional judgments corresponds to the sentence's primary intension. Something very similar applies to subsentential expressions: for a singular term, for example, there will be a raft of conditional rational judgments using the expression across a wide variety of scenarios, and these can be used to define the extension of the expression relative to those scenarios (see Chalmers 2004). So we will have substantial primary intensions for a wide range of sentential and subsential expression tokens.

It should be noted that nothing here requires that the expressions in question be definable in simpler terms (such as in qualitative terms), or that they be equivalent to descriptions (even to rigidified or "actually"-involving descriptions). The inferential roles in question will exist whether or not the term is definable and whether or not it is equivalent to a description (for more on this, see Chalmers and Jackson 2001 and Chalmers 2002a).

These claims are quite compatible with Kripke's epistemological argument that terms such as 'Gödel' are not equivalent to descriptions. In effect, Kripke describes a scenario *w* where someone called 'Schmidt' proved the incompleteness of arithmetic, and then it was stolen by someone called 'Gödel' who moved to Princeton, and so on. Kripke's argument might be put by saying that (i) *w* is not ruled out a priori, and (ii) if

we accept that *w* obtains, we should reject the claim 'Gödel proved the incompleteness of arithmetic', so (iii) 'Gödel proved the incompleteness of arithmetic' is not a priori. A two-dimensionalist will put this by saying that the primary intension of 'Gödel proved the incompleteness of arithmetic' is false at *w*, so that the primary intension of 'Gödel' differs from that of 'the prover of the incompleteness of arithmetic'. If Kripke's argument generalizes to other descriptions, it will follow that the primary intension of Gödel is not equivalent to the primary intension of any such description. But nothing here begins to suggest that 'Gödel' lacks a primary intension.

Although the primary intension of an expression may not be equivalent to that of a description, one can often at least approximately characterize an expression's primary intension using a description. For example, one might roughly characterize the primary intension of a typical use of 'water' by saying that in a centered world *w*, it picks out the dominant clear, drinkable liquid with which the individual at the center of *w* is acquainted. And one might roughly characterize the primary intension of 'Gödel' by saying that it picks out that individual who was called 'Gödel' by those from whom the individual at the center acquired the name. But these characterizations will usually be imperfect, and it will be possible to find Kripke-style counterexamples to them. Ultimately a primary intension is not grounded in any description, but rather is grounded in an expression's inferential role.

3.5 Two-dimensionalism and semantic pluralism

Two-dimensionalism is naturally combined with a *semantic pluralism*, according to which expressions and utterances can be associated with many different semantic (or quasi-semantic) values, by many different semantic (or quasi-semantic) relations. On this view, there should be no question about whether the primary intension of the secondary intension is *the* content of an utterance. Both can be systematically associated with utterances, and both can play some of the roles that we want contents to play. Furthermore, there will certainly be explanatory roles that neither of them play, so two-dimensionalism should not be seen as offering an exhaustive account of the content of an utterance. Rather, it is characterizing some aspects of utterance content, aspects that can play a useful role in the epistemic and modal domains.

Likewise, there should be no question about which of the two-dimensional frameworks described in this paper is the "correct" framework. Each framework

offers a different quasi-semantic relation that associates expressions with two-dimensional semantic values, and each of these may play an explanatory role in different domains. Each has different properties. Most obviously, primary intensions have a stronger connection to apriority and cognitive significance than the semantic values described earlier. Unlike characters, deep intensions, and FA-intensions, the primary intension associated with an a posteriori identity such as 'Mark Twain is Samuel Clemens' will be contingent. Unlike diagonal propositions, the primary intension of an a priori sentence such as '2+2=4' will be necessary.

These differences arise from the differences in the way the semantic relations are defined. Unlike characters and diagonal propositions, primary intensions are not defined in terms of context-dependence. Unlike deep intensions, they are not defined in terms of a prior notion of content. Unlike FA-intensions; they are not defined in terms of the behavior of an 'actually' operator. Rather, they are defined in epistemic terms.

Because they are defined in epistemic terms, primary intensions can often vary between tokens of an expression type. This will happen most obviously for context-dependent terms such as 'tall', for which tokens in different contexts will be associated with different inferential roles. Primary intensions may also vary between different tokens of the same name (especially by different speakers), for different tokens of the same demonstrative (e.g. 'this' or 'that'), and perhaps also for different tokens of the same natural kind term. It follows that in these cases, a primary intension does not constitute an expression's linguistic meaning, where this is understood as what is common to all tokens of an expression type, or as what is required for any competent use of the expression. Instead, a primary intension can be seen as a kind of utterance content.

Even if they are not always part of linguistic meaning, primary intensions are nevertheless a sort of truth-conditional content. The primary intension of an utterance yields a condition under which the utterance will be true. For example, the primary intension of 'there is water in the glass' will be true at some scenarios and false at others, and the utterance will be true iff the primary intension is true at the scenario of the utterance (roughly, if the glass picked out by the individual at the center of the scenario contains the dominant watery stuff in the environment around the center). This can be seen as an *epistemic* truth-condition for the utterance, specifying how the truth of the utterance depends (epistemically) on which epistemically possible

scenario turns out to be actual. This contrasts with the "metaphysical" truth-condition corresponding to the secondary intension, which might be seen as specifying how the truth of the utterance depends (metaphysically) on which metaphysically possible world is actual. Again, there is no need to decide the question of which of these is *the* truth-condition associated with an utterance.

Are primary intensions a sort of semantic content? This depends on how we understand the notion of semantic content. If we stipulate that the semantic content of an utterance is truth-conditional content, then primary intensions are a variety of semantic content. On the other hand, if we stipulate that semantic content is linguistic meaning in the sense above, or that semantic content is always associated with expression types and not tokens, then primary intensions are not in general part of semantic content (though they may be part of semantic content for some expressions, such as some indexicals and qualitative expressions). In any case, once we are clear on the various properties of these intensions, nothing important to the framework turns on the terminological question of whether they count as "semantic".

A semantic pluralist can allow that for some explanatory purposes, it may be useful to modify two-dimensionalist semantic values in some respects. For example, one might define the *structured* primary intension of a complex expression as a structured entity involving the primary intensions of the simple expressions involved in the expression's logical form. One might likewise define structured secondary and two-dimensional intensions. Given compositionality, a structured primary intension will determine an unstructured primary intension (and likewise for the other intensions), but the reverse need not be the case. This means that structured primary intensions are more fine-grained than unstructured primary intensions: for example, all a priori truths will have the same unstructured primary intension (one that is true at all scenarios), but they will have different structured primary intensions. The fine-grainedness of structured intensions makes a difference for certain purposes, described below.

What are *propositions*, according to two-dimensionalism? Some two-dimensionalists (e.g. Jackson 1998) hold that propositions are sets of possible worlds, in which case a given utterance expresses two propositions (a primary proposition and a secondary proposition). This view is naturally combined with the view that there are no necessary a posteriori propositions: necessary a posteriori sentences have a primary proposition that is contingent and knowable only a posteriori, and a

secondary proposition that is necessary and knowable a priori. Other two-dimensionalists may hold that propositions have more structure than this. For example, one could hold that propositions are structured entities involving both the primary and secondary intensions (and/or perhaps the two-dimensional intension) of the simple expressions involved. A two-dimensionalist of this sort may allow that there are necessary a posteriori propositions.

A semantic pluralist view tends to suggest that there are numerous entities which can play some of the explanatory roles that propositions are supposed to play, and that there is no need to settle which of these best deserves the label 'proposition'. My own view is that if one were forced to identify propositions with one sort of entity that can be modeled in the framework, there would be a good case for choosing structured two-dimensional entities of some sort (perhaps those discussed as candidates for Fregean senses, below). But one might also allow that at least for some purposes, propositions should be seen as entities more fine-grained than any two-dimensional objects, so that propositions can be associated with intensions without themselves being intensions. In any case, core two-dimensionalism as characterized above is compatible with a wide range of views here.

4 Applications of two-dimensionalism

I will briefly sketch some applications of the two-dimensionalism outlined in the previous section.

(i) *Fregean sense* (Chalmers 2002a): Thesis (T9) above says that two expressions *A* and *B* have the same primary intensions iff '$A \equiv B$' is epistemically necessary. This is reminiscent of the Fregean claim that two singular terms *A* and *B* have the same sense iff '$A=B$' is cognitively insignificant. It suggests that primary intensions can play at least some of the roles of a Fregean sense, individuating expressions by their epistemic role. Of course there are some differences. For example, primary intensions are not as fine-grained as Fregean senses: a priori equivalent expressions (such as '7+3' and '10') will have different Fregean senses, but they have the same primary intension (though they will usually have different structured primary intensions). Further, there are differences between primary intensions and Fregean senses in the case of indexicals: for example, uses of 'I' by different speakers have the same primary intension, whereas Frege held that they have different senses. Relatedly,

where Frege held that sense determines reference, primary intensions do not determine extensions in a strong sense (although they may still determine extension relative to context), as two expressions may have the same primary intensions and different extensions. Still, may nevertheless think of primary intensions as a broadly Fregean aspect of an expression's content.

One can also use the two-dimensional framework to define semantic values that behave even more like Fregean senses. (Here I go beyond the discussion in Chalmers 2002a.) We might stipulate that the sense of a simple expression token is an ordered pair of its primary intension and its extension, and that the sense of a complex expression token is a structured complex made up of the senses of its parts. Now, most pairs of a priori equivalent expressions, such as '7+3' and '10' will have different senses. (The only potential exceptions will arise if there are a priori equivalent but cognitively distinct simple expressions, which is not obvious.) Furthermore, uses of 'I' by different speakers will have different senses. And now, sense determines reference in the strong sense. So entities of this sort might be seen as very much akin to Fregean senses, and we might think of the structured entity associated with a sentence token as akin to a Fregean thought.

(ii) *Contents of thoughts* (Chalmers 2002b). One can extend the framework above so that primary and secondary intensions are not just associated with sentences but with thoughts, where these are understood as occurrent mental states. For example, my thought *water is H_2O* will have a contingent primary intension (false in the XYZ-scenario) but a necessary secondary intension. One can then argue that a thought's primary intension is a sort of *narrow content*: content that is shared between intrinsically identical thinkers. For example, when Oscar on Earth and Twin Oscar on Twin Earth say 'water is wet', the thoughts they express will have different secondary intensions (so secondary intensions are a sort of "wide content"), but they will have the same primary intension.

(iii) *Belief ascriptions* (Chalmers 2002b): One can use this framework to analyze ascriptions of belief and other propositional attitudes. As a first attempt, one might suggest that an ascription '*S* believes that *P*' is true iff the referent of *S* has a belief with the primary intension of '*P*' (in the mouth of the ascriber), or a belief with the secondary intension of '*P*' (in the mouth of the ascriber). Neither of these suggestions works: the first is falsified by cases such as 'John believes that I am hungry', while

the second is falsified by cases such as 'Lois believes that Clark Kent can fly'. However, more sophisticated analyses are possible. For example, Chalmers 2002b suggests

> An utterance of '*S* believes that *P*' is true iff the referent of *S* has a belief with the structured secondary intension of '*P*' (in the mouth of the ascriber) and with an appropriate structured primary intension.

Here, "appropriate" functions to pick out a range of primary intensions (allowing, for example, that the Pierre can satisfy a 'London'-involving ascription even if he uses the term with a different primary intension), where this range may depend on the context of utterance. (Structured intensions are needed in order that independent "appropriateness" constraints may be imposed separately on each element of a belief.) This analysis is closely related to "hidden-indexical" analyses of belief ascriptions, with primary intensions playing the role of "modes of presentation".

One can also use primary intensions to give an analysis of *de re* attitude ascriptions, in the style of Kaplan 1968.

> A *de re* attitude ascription '*S* believes of *X* that it is *F*' is true iff *S* has a belief with the secondary intension of '*X* is *F*', and which picks out the referent of *X* under a *de re*-appropriate primary intension.

Here, the conditions on a *de re*-appropriate primary intension may again be context-independent, but to a first approximation we can think of such an intension as one that is acquaintance-entailing: necessarily, if a subject *S* has a state with a *de re*-appropriate primary intension that picks out extension *E*, then *S* will be acquainted with *E*.

(iv) *Indicative conditionals* (Weatherson 2001): We can also use epistemic two-dimensionalism to give a possible-worlds-style analysis of the intuitive acceptability-conditions of indicative conditionals that is analogous to the familiar Lewis-Stalnaker analysis of subjunctive conditionals.

> A token of an indicative conditional 'If *P*, then *Q*' is acceptable iff the epistemically closest scenario satisfying the primary intension of '*P*' (in the mouth of the speaker) also satisfies the primary intension of '*Q*'.

Of course an elaboration of this account requires an elaboration of what epistemic closeness amounts to. But given that the familiar Ramsey Test for the

acceptability of an indicative conditionals is defined in epistemic terms (if one conditionally accepts *P*, should one rationally conclude *Q*?), and given that primary intensions are defined in very similar terms, it is not surprising that there is a close relation.

(v) *Conceivability and possibility* (Chalmers 2002): If thesis (T5) is correct, it licenses a certain sort of move from conceivability to possibility. Let us say that *S* is conceivable when it is epistemically possible: that is, when *S* is not ruled out a priori. If T5 is correct, then when *S* is conceivable, the primary intension of *S* will be true at some scenario. If scenarios are centered worlds, then there will be some centered (metaphysically possible) world *w* satisfying the primary intension of *S*. This does not entail that *S* is metaphysically possible, but it nevertheless allows us to draw conclusions about metaphysical possible worlds from premises about conceivability. Reasoning of this sort is central to some uses of conceivability arguments in the philosophy of mind (e.g. in Chalmers 1996).

5 Objections to two-dimensionalism

A number of objections to two-dimensionalism have been made in the literature. Some objections (the first nine considered here) rest on the attribution of views to which two-dimensionalism is not committed. They might be considered objections to certain versions of two-dimensionalism, but they do not apply to the epistemic two-dimensionalism that I have outlined. Other objections (the next two considered here) show that the claims of two-dimensionalism must be restricted in certain respects. Still others (the last three considered here) raise substantive issues whose adjudication is an ongoing project.

What is held constant? (Block and Stalnaker 1999): Evaluation of primary intensions turns on claims about what a term such as 'water' would have picked out in counterfactual circumstances. But this raises the question of what is held constant across worlds in counting an expression as a token of 'water'. If only orthography is held constant, then many tokens of 'water is watery' will be false; if reference is held constant, then no token of 'water is H_2O' will be false. So to yield the desired results, a two-dimensionalist must hold constant some intermediate sort of content, such as

Fregean or descriptive or narrow content. But it is question-begging for a two-dimensionalist to presuppose such a notion of content.

Response: Evaluation of primary intensions does not turn on metalinguistic claims about what a term would have picked out in counterfactual circumstances. One could define an expression's *contextual intension* as a mapping from worlds containing a token of the expression to the extension of that token in that world. The question of what is held constant would then become relevant: one would obtain different sorts of contextual intensions depending on just what one counts as a relevant token. But primary intensions are not like this. They simply turn on the epistemic properties of an expression in the actual world. For example, it is epistemically possible (not ruled out a priori) that there are no utterances, and so the primary intension of 'There are no utterances' will be true in an utterance-free world (whereas the contextual intension of 'There are no utterances' will not be defined there). Because properties of counterfactual tokens are irrelevant to the evaluation of primary intensions (except in some special cases), the problem of "what is held constant" does not arise.

Twin Earth intuitions are irrelevant (Soames 2005). Intuitions about the reference of 'water' as used on Twin Earth are irrelevant to the meaning of our term 'water', as the term 'water' on Twin Earth has a different meaning.

Response: Again, evaluation of primary intensions does not depend on the referents of homonymous terms in counterfactual worlds. Rather, it depends on certain epistemic properties associated with uses of 'water' in our world. For example, if we are given the information that the liquid in the oceans and lakes is and has always been XYZ, we should conclude that water is XYZ. This is a fact about the inferential role associated with uses of *our* term 'water'. Epistemic two-dimensionalism uses this inferential role to analyze an aspect of the content of these uses of the term.

Names and natural-kind terms are not indexicals (Nimtz and Beckermann forthcoming; Soames 2005): Two-dimensionalism entails that terms such as 'water' are really disguised indexicals that can pick out different referents in different contexts. But such terms are not indexicals. Any utterance of the English term 'water', in any context, picks out H_2O.

Response: Epistemic two-dimensionalism does not entail that names and natural kind terms are disguised indexicals, and it is consistent with the claim that any utterance of the English term 'water' refers to H_2O. If primary intensions were Kaplanian characters or contextual intensions, then the claim that 'water' refers to H_2O in any context would be inconsistent with the two-dimensionalist claim that the primary intension of 'water' picks out XYZ in the Twin Earth world. But primary intensions are not Kaplanian characters or contextual intensions. To ground the desired behavior of primary intensions, the two-dimensionalist simply requires the plausible claim that it is *epistemically* possible (i.e. not ruled out a priori) that water is XYZ. This claim is consistent with the claim that (given that 'water' actually refers to H_2O), all metaphysically possible tokens of the English term 'water' refer to H_2O.

Names are not rigidified descriptions (Soames 2005). Two-dimensionalism entails that names and natural kind terms are disguised rigidified descriptions (of the form 'the actual ϕ', for some ϕ). But Kripke's epistemic arguments show that names are not rigidified descriptions, as do considerations about the way that names and descriptions behave in belief ascriptions.

Response: Two-dimensionalism does not entail that names and natural kind terms are rigidified descriptions. We have noted already that Kripke's epistemic arguments are accommodated by the observation that primary intensions cannot always be encapsulated into a description. Furthermore, as noted above, it is consistent with two-dimensionalism to hold that names and natural kind terms, unlike rigidified descriptions, have the same referent in any context of utterance. It is also consistent with two-dimensionalism to hold that the primary intension of a name or natural kind term may vary between speakers. The account of belief ascriptions given above does not entail that names will behave like rigidified descriptions in belief contexts, and handles the relevant data straightforwardly.

Speakers lack identifying knowledge (Byrne and Pryor 2005; Schiffer 2003). Two-dimensionalism requires that every name *N* (at least as used by a speaker) be associated with a "uniqueness property" ϕ (such that at most one individual has ϕ), and requires that the speaker have a priori "identifying knowledge" of the form '*N* is ϕ'. But speakers in general lack this sort of knowledge.

Response: Two-dimensionalism does not require that speakers possess identifying knowledge. It is true that primary intensions can be associated with

uniqueness properties (or better, uniqueness relations, because of the role of centering). But speakers need not have beliefs about these uniqueness properties (expressible in the form '*N* is ϕ'). Epistemic two-dimensionalism simply requires that speakers have a *conditional ability* to determine the referent of *N* (or better, to determine the truth-value of claims using *N*), given relevant information about the character of the actual world and given idealized rational reflection. This conditional ability need not be grounded in the possession of identifying knowledge. Furthermore, the invocation of rational reflection makes this a normative claim that idealizes away from cognitive limitations of the speaker. For example, even if a child cannot actually identify a referent across all circumstances, there may still be idealized inferential norms on how they should update their relevant beliefs given relevant information about the world. These norms are all that is required.

Ordinary expressions are not ambiguous (Bealer 2002; Marconi 2005): Two-dimensionalism explains the difference in truth-value between

(3) It is metaphysically necessary that water is H_2O.

(4) It is epistemically necessary that water is H_2O.

by saying that 'water' expresses its primary intension in the first context and its secondary intension in the second context. But this entails implausibly that 'water' is ambiguous. Further, this view cannot handle combined contexts, such as 'It is metaphysically necessary but not epistemically necessary that water is H_2O'.

Response: Two-dimensionalism does not hold that ordinary expressions are ambiguous. 'Water' has exactly the same content in both (3) and (4) above: in both contexts (and in all contexts) it has both a primary intension and a secondary intension (or equivalently, it has a complex semantic value involving both a primary and a secondary intension). This does not entail that 'water' is ambiguous, any more than the distinction between character and content entails that indexicals are ambiguous. The distinction between (3) and (4) is handled instead by the difference between the modal operators. The semantics of these operators are such that 'It is metaphysically necessary that *S*' is true when *S* has a necessary secondary intension, while 'It is epistemically necessary that *S*' is true when *S* has a necessary primary intension. Combined contexts are handled in the obvious combined way.

Two-dimensionalism cannot handle belief ascriptions (Soames 2005): It is natural for two-dimensionalists to hold that '*x* believes that *S*' is true when the subject has a belief whose primary intension is the primary intension of *S*. But this view gives the wrong result in a number of cases, and no better two-dimensionalist treatment of belief ascriptions is available.

Response: The view of belief ascriptions mentioned above is considered and rejected in Chalmers (1995; 2002), and to the best of my knowledge no two-dimensionalist endorses the view. The account of belief ascriptions described in section 4, straightforwardly handles most of the puzzle cases developed by Soames (see Chalmers 2004 for details). Soames raises some further puzzle cases for this account involving the relationship between ordinary belief ascriptions and *de re* belief ascriptions, but the account of *de re* belief ascriptions sketched above (and also given in Chalmers 2002) handles these cases straightforwardly.

Two-dimensionalism requires global descriptivism (Stalnaker 2003, 2004): Two-dimensionalism holds that the primary intension of an utterance or a belief is determined by the internal state of the speaker or believer. This requires an internalist "metasemantic" theory, showing how intentional content is determined by internal state. The main candidate for such a theory is the "global descriptivism" of Lewis 1984, holding that the content of our utterances and beliefs is determined by whatever assignment of content yields the "best fit" between the beliefs and the world. But global descriptivism is false.

Response: Two-dimensionalism does not require global descriptivism. Of course there is not yet any satisfactory theory of the basis of intentionality, but there are many possible internalist alternatives. For example, one might hold that the primary intension of a mental state is determined in part by its internal functional role, and in part by associated phenomenal states (where the latter may be especially relevant for phenomenal and perceptual concepts).

The wrong sentences are a priori: Two-dimensionalism requires the claim that sentences such as 'Hesperus (if it exists) is Phosphorus' are not a priori, while sentences such as 'Julius (if he exists) invented the zip' are a priori. But these claims are incorrect: the former expresses a trivial singular proposition that can be justified a priori, while the latter expresses a nontrivial singular proposition that cannot be justified a priori.

Response: If one stipulated that apriority of a name-involving sentence is to be understood in terms of the apriori knowability of an associated singular proposition, these (controversial and counterintuitive) claims would be correct. But the two-dimensionalist takes this as good reason to reject the stipulation, or at least stipulates a different understanding of apriority for the purposes of the framework. For these purposes, an utterance can be said to be a priori when it expresses a belief (or at least an occurrent thought) that can be justified non-empirically, yielding a priori knowledge. There is an obvious epistemic difference between beliefs expressed by typical occurrences of 'Hesperus is Hesperus' and 'Hesperus is Phosphorus': no amount of non-empirical reasoning can convert the latter belief into a priori knowledge, but the former is easily justified a priori. (Note that on this definition of apriority, two different beliefs might be related to the same singular propositional content while differing in their epistemic status: the epistemic status attaches primarily to belief tokens, not to belief types or to propositional contents.) This epistemic difference at the level of thought can be used to ground the relevant claims about the apriority of utterances. More generally, the primary intensions of utterances are grounded in the (normative) cognitive role of associated thoughts.

Primary intensions are not linguistic meaning: Different speakers can use the same name ('Fred') or natural kind term ('water') with quite different cognitive roles, and with distinct patterns of epistemic evaluation. If so, the same expression will have different primary intensions for different speakers. So an expression's primary intension is not part of its linguistic meaning, where this is understood as meaning that is associated with an expression type simply by virtue of the conventions of a language.

Response: This point is correct: primary intensions are not always part of linguistic meaning. For example, it can happen that an identity statement (e.g. 'Bill Smith is William Smith') can be cognitively insignificant for one speaker (e.g. his wife, who uses the two names interchangeably) but not for another (e.g. a colleague who uses the names in quite different domains without knowing that they are coextensive). If so, then the primary intensions of the names will coincide for one speaker but not for another, so that the primary intension of at least one of them must vary across speakers. Primary intensions can also vary for context-dependent terms such as 'tall' and 'heavy'. The moral is that for maximal generality, primary

intensions should be associated with expression tokens (or with utterances of expression types) rather than with expression types.

Primary intensions are insufficiently fine-grained. Cognitively distinct expressions may have the same primary intensions. When expressions are equivalent a priori, their primary intensions will coincide. For example, logical and mathematical truths all have the same primary intension (true in all scenarios), and have the same secondary intension too. But these clearly differ in meaning and in cognitive significance. So two-dimensional semantic values do not exhaust meaning (or utterance content), and are not as fine-grained as Fregean senses.

Response: A two-dimensionalist can accommodate many of the relevant cases here by invoking structured intensions. This will distinguish between different logical and mathematical truths, for example. The only residual problem will arise if there are pairs of simple expressions that are equivalent a priori but that are cognitively distinct. It is not obvious that there are such pairs, but if there are, there is more to meaning than primary intensions. We might say that primary intensions individuate expressions by their *idealized* cognitive significance, and so do not capture differences in *nonidealized* cognitive significance. One might try to capture these differences by moving to intensions that are defined over a space of finer-grained epistemic possibilities. Or a two-dimensionalist might simply allow that in addition to intensions, expressions are associated with finer-grained semantic values that lie behind and determine these intensions. But in any case, this point is no threat to the two-dimensionalist who is a semantic pluralist. Primary and secondary intensions are not all there is to meaning, but nevertheless utterances can be associated with primary and secondary intensions, in a way that can play the various explanatory roles described above.

There are epistemic possibilities that correspond to no centered world (Yablo 1999, 2002). A key two-dimensionalist claim holds that when *S* is not ruled out a priori, then there is some centered world at which the primary intension of *S* is true. This may be so for typical Kripkean a posteriori necessities such as ‘water is not H_2O’, but there are other sentences for which the claim false. For example, it may be that the existence (or nonexistence) of a god is necessary without being a priori. If so, ‘There is no god’ (or ‘There is a god’) is not ruled out a priori, but it is necessarily false. There appears to be no relevant difference between primary and secondary

intensions here, so the primary intension is true in no possible world. Something similar applies if the laws of nature in our world are the laws of all possible worlds. If these views are correct, then the space of epistemic possibilities outstrips the space of metaphysical possibilities in a way that falsifies the two-dimensionalist claim.

Response: All of these purported counterexamples rest on controversial claims about modality or apriority, and I have argued (Chalmers 1999; 2004) that none of them succeed. Furthermore, there is good reason to believe that the concept of metaphysical modality itself has roots in the epistemic domain, so that there cannot be "strong necessities" that exhibit this sort of disconnect between epistemic and metaphysical modalities. Still, the existence or nonexistence of strong necessities is a delicate and controversial issue. An alternative version of two-dimensionalism remains neutral on this issue by understanding scenarios not as centered metaphysically possible worlds, but instead as maximal epistemic possibilities (corresponding roughly to maximal epistemically consistent sets of sentences). Then even if no metaphysically possible world verifies 'There is no god', some maximal epistemic possibility will verify 'There is no god', so there will be a scenario at which the primary intension of this sentence will be true. Understood in this neutral way, two-dimensionalism does not ground inferences from conceivability to metaphysical possibility (those inferences will turn on a further claim about the relationship between scenarios and metaphysically possible worlds), but it can still play much the same role as before in the epistemic and semantic domains.

Complete canonical descriptions are not available (Schroeter 2004): Epistemic two-dimensionalism requires that there be qualitative descriptions of a given scenario that are complete in that they epistemically determine the truth-value of arbitrary judgments. But there may be some features of the world, such as intrinsic physical features, which cannot be captured in a qualitative description.

Response: It is not clear whether there are intrinsic properties that cannot be captured in a qualitative description, but if there are, this will be irrelevant to epistemically determining the truth-value of any of our sentences. When information about these features is needed to epistemically determine the truth-value of a sentence in a scenario, a qualitative characterization of the features (e.g. an existential or a Ramsey-sentence characterization) will suffice. (Such a characterization may not suffice for metaphysical determination, and for evaluating truth-values of sentences in

counterfactual worlds according to their secondary intensions. But qualitative descriptions are only needed for primary intensions.) The minimal size of a vocabulary that can epistemically determine the truth of all sentences is an important open question, but there is good reason to believe that some qualitative (and indexical) vocabulary suffices. It should also be noted that if we take the purely epistemic approach to scenarios described in the previous response, a restriction to qualitative vocabulary is not needed, and so the issue here does not arise.

Objections to the role of apriority (Block and Stalnaker 1999, Yablo 2002). It is true that there is an epistemic relation between information about the world and claims about reference: for example, given the information that we are in the H_2O - world (appropriately characterized) we should conclude that water is H_2O, and given the information that we are in the XYZ-world, we should conclude that water is XYZ. And it is true that we can makes these conditional inferences from the armchair, without needing to perform further investigation of the environment. But nevertheless, these inferences are not justified a priori. The inferences are justified in part by background empirical knowledge of the world (Block and Stalnaker) or by "peeking" at our own judgments (Yablo). As a result, primary intensions are not connected to apriority as strongly as the two-dimensionalist supposes.

Response: Chalmers and Jackson (2001) argue that these connections are in fact a priori: although empirical facts about the world can play a causal role in determining the relevant patterns of inference, there is good reason to believe that they do not play a justifying role. (Chalmers (2002) responds to Yablo.) It is also worth noting that even a skeptic about apriority can use the epistemic two-dimensional framework. Even if the relevant inferential connections are not a priori, one can still use them to define primary intensions, and the resulting primary intensions will still behave much as they are supposed to (assigning a necessary intension to 'Hesperus is Hesperus' but not to 'Hesperus is Phosphorus', for example. The connection between primary intensions and apriority will be lost, but primary intensions will still be strongly connected to the epistemic domain.

Bibliography

Bealer, G. 2002. Modal epistemology and the rationalist renaissance. In (T. Gendler & J. Hawthorne, eds) *Conceivability and Possibility*. Oxford: Oxford University Press.

Block, N. & Stalnaker, R. 1999. Conceptual analysis, dualism, and the explanatory gap. *Philosophical Review* 108:1-46.

Braddon-Mitchell, D. 2004. Masters of our meaning. *Philosophical Studies* 188:133-52.

Byrne, A, & Pryor, J. 2006. Bad intensions. In (M. Garcia-Carpintero & J. Macia, eds) *Two-Dimensional Semantics: Foundations and Applications*. Oxford: Oxford University Press.

Carnap, R. 1947. *Meaning and Necessity*. Chicago: University of Chicago Press.

Chalmers, D.J. 1995. The components of content. Manuscript. [consc.net/papers/content95.html]

Chalmers, D.J. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.

Chalmers, D.J. 1999. Materialism and the metaphysics of modality. *Philosophy and Phenomenological Research*. 59:473-96 [consc.net/papers/modality.html]

Chalmers, D.J. 2002a. On sense and intension. [consc.net/papers/intension.html]

Chalmers, D.J. 2002b. The components of content (revised version). In *Philosophy of Mind: Classical and Contemporary Readings*. Oxford: Oxford University Press. [consc.net/papers/content.html]

Chalmers, D.J. 2002c. Does conceivability entail possibility? In (T. Gendler & J. Hawthorne, eds) *Conceivability, and Possibility*. Oxford: Oxford University Press. [consc.net/papers/conceivability.html]

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

Chalmers, D.J. 2006. The foundations of two-dimensional semantics. In (M. Garcia-Carpintero & J. Macia, eds) *Two-Dimensional Semantics: Foundations and Applications*. Oxford: Oxford University Press. [consc.net/papers/foundations.html]

Chalmers, D.J. forthcoming. The nature of epistemic space. [consc.net/papers/espace.html]

Chalmers, D.J. & Jackson, F. 2001. Conceptual analysis and reductive explanation. *Philosophical Review*. 110:315-61 [consc.net/papers/analysis.html]

Davies, M. & Humberstone, I.L. 1981. Two notions of necessity. *Philosophical Studies* 58:1-30.

Evans, G. 1977. Reference and contingency. *The Monist* 62:161-89.

Frege, G. 1892. Über Sinn und Bedeutung. Translated in (P. Geach & M. Black, eds.) *Translations from the Philosophical Writings of Gottlob Frege*. Oxford: Blackwell, 1952.

Garcia-Carpintero, M. & Macia, J. (eds) 2006. *Two-Dimensional Semantics: Foundations and Applications*. Oxford: Oxford University Press.

Jackson F. 1998. From Metaphysics to Ethics: A Defense of Conceptual Analysis. Oxford: Oxford University Press.

Kamp, H. 1971. Formal properties of 'now'. *Theoria* 37:227-273.

Kaplan, D. 1968. Quantifying in. *Synthese* 19:178-214.

Kaplan, D. 1979. Dthat. In (P. Cole, ed.) *Syntax and Semantics*. New York: Academic Press.

Kaplan, D. 1989. Demonstratives. In (J. Almog, J. Perry, and H. Wettstein, eds.) *Themes from Kaplan*. Oxford: Oxford University Press.

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

Lewis, D. 1984. Putnam's paradox. *Australasian Journal of Philosophy* 62:221-237.

Lewis, D. 1994. Reduction of mind. In (S. Guttenplan, ed.) *Companion to the Philosophy of Mind*. Oxford: Blackwell.

Marconi, D. 2005. Two-dimensional semantics and the articulation problem. *Synthese*. 143(3):321-349.

Nimtz, C. & Beckermann, A. forthcoming. Why 'water' is not an indexical.

Putnam, H. 1975. The meaning of `meaning'. In (K. Gunderson, ed.) *Language, Mind, and Knowledge*. Minneapolis: University of Minnesota Press.

Schiffer, S. 2003. Two-dimensional semantics and propositional attitude content. In *The Things We Mean*. Oxford: Oxford University Press,

Schroeter, L. 2004. The rationalist foundations of Chalmers' two-dimensional semantics. *Philosophical Studies* 18:227-55.

Soames, S. 2005. Reference and Description: The Case Against Two-Dimensionalism. Princeton, N.J.: Princeton University Press.

Stalnaker, R. 1978. Assertion. In (P. Cole, ed.) *Syntax and Semantics: Pragmatics, Vol. 9*. New York: Academic Press.

Stalnaker, R. 2003. Conceptual truth and metaphysical necessity. In *Ways a World Might Be: Metaphysical and Anti-Metaphysical Essays*. Oxford: Oxford University Press.

Stalnaker, R. 2004. Assertion revisited: On the interpretation of two-dimensional modal semantics. *Philosophical Studies* 118: 299-322.

Vlach, F. 1973. Now and Then: A formal study in the logic of tense anaphora. Ph.D. dissertation, UCLA.

Weatherson, B. 2001. Indicatives and subjunctives. *Philosophical Quarterly* 51:200-216. Wong, K.-Y. 1996. Sentence-relativity and the necessary a posteriori. *Philosophical Studies* 83:53-91.

Yablo, S. 1999. Concepts and consciousness. *Philosophy and Phenomenological Research*. 59:455-63.

Yablo, S. 2000. Textbook Kripkeanism and the open texture of concepts. *Pacific Philosophical Quarterly* 81:98-122.

Yablo, S. 2002. Coulda, woulda, shoulda. In (T. Gendler & J. Hawthorne, eds) *Conceivability and Possibility*. Oxford: Oxford University Press.