An Account of Possibility

Agustín Rayo
Massachusetts Institute of Technology

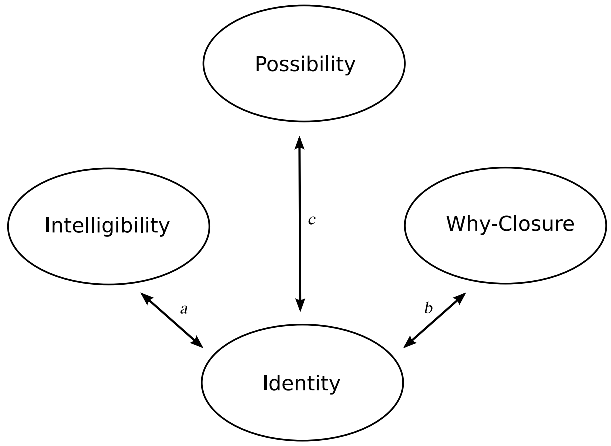
March 18, 2008

There are two different conceptions of possibility *simpliciter* (or 'metaphysical possibility', as it is sometimes referred to in the literature). According to the first conception, the possible scenarios are just the intelligible scenarios. According to the second conception, a scenario only counts as possible if it satisfies certain *metaphysical* constraints—call them 'the metaphysical laws'. A proponent of this latter conception might believe that it is in some sense intelligible that there be no abstract objects, or that I be an elephant, but claim that suitable metaphysical laws rule out the relevant scenarios as metaphysically impossible.

It seems to me that thinking of possibility in terms of metaphysical laws is a bad idea. By wheeling in metaphysical laws one drives a wedge between the notion of possibility and one of its most useful applications: the representation of content. For in setting forth a metaphysical law one is presumably meant to be saying something non-trivial about the way the world is. But if every possible scenario is a scenario which verifies the metaphysical laws, one cannot hope to represent the non-trivial requirement that the world must meet in order for a metaphysical law to hold by dividing up the possibilities into those that meet the requirement and those that do not.

The purpose of this paper is to develop an account of possibility according to which the limits of possibility are simply the limits of intelligibility. Because I will be foregoing any appeal to metaphysical laws, one might worry that I'm bound to end up with a highly revisionary list of modal truths. It is tempting to think, for example, that without metaphysical laws one would be left with the result that the statements of pure mathematics must be thought of as contingent, or that modal principles such as the essentiality of origin must be given up. That would be a mistake—a mistake which, I suspect, is responsible for the idea that the notion of metaphysical possibility is somehow tied to metaphysical laws. We shall see that when the relevant notions are properly understood, there is no tension between the view that the limits of possibility are the limits of intelligibility and a non-revisionary list of modal truths.

I would like to emphasize that in postulating a link between intelligibility and possibility I do not wish to claim that one of these notions is 'prior' to the other. In particular, I do not claim that a scenario is possible 'in virtue' of being intelligible, or that the way to acquire a grasp of the notion of possibility is to deploy an independent understanding of the



Arrow	Connection
a	Any scenario which violates an identity statement one accepts will be regarded as unintelligible.
b	If s is an identity statement one accepts, one will be unable to make sense of certain questions of the form 'Why is it the case that s?'.
c	To settle the truth-values of an arbitrary sentence involving boxes and diamonds it is sufficient to settle the truth-values of the identity statements, and arbitrary sentences involving no boxes or diamonds.
	Other connections—such as the link between intelligibility and possibility—will be derived on the basis of *a*–*c*.

Figure 1: Connections between possibility, intelligibility, identity and why-closure.

notion of intelligibility. The project is to shed light on both intelligibility and possibility by getting clear on how they interact with one another, not to use one of these notions to provide a foundation for the other.

The link between intelligibility and possibility is meant to be a consequence of a larger system of interconnections, involving not just the notions of intelligibility and possibility, but also the notion of identity and a notion I shall call 'why-closure'. (See figure 1.) As before, the suggestion is that one can shed light on all four notions by getting clear about how they interact with one another, not that some of the notions should be used as a foundation for the rest.

The ultimate goal, however, is to attain a better understanding the notion of possibility. I hope that reflection on these various interconnections will put us in a position to address the following two questions:

1. What are the mechanisms whereby one is entitled to conclude that a given scenario is possible?

   (How, for example, should one go about settling the question of whether I might have been an elephant?)

2. What is required of the world in order for the truth-conditions of modal sentences to be satisfied?

   (Does satisfaction of the truth-conditions of 'there is a mammal that might have been a human', for instance, require more than the existence of a human?)

Many of the ideas I will be discussing are not new.[1] But I hope to revisit the issues in a way that sheds new light on this well-travelled terrain.

1 Identity Generalized

The standard first-order identity-predicate '=' has been much discussed. But there is a higher-order identity-predicate which has received less attention. Consider the following sentences:

> Sibling
> To be a sibling *just is* to share a parent.
>
> [In symbols: '$\text{Sibling}(x) \equiv_x \exists y \exists z(\text{Parent}(z,x) \wedge \text{Parent}(z,y) \wedge x \neq y)$']
>
> Heat
> To be hot *just is* to have high mean kinetic energy.
>
> [In symbols: '$\text{Hot}(x) \equiv_x \text{High-Mean-Kinetic-Energy}(x)$']
>
> Water
> To be composed of water *just is* to be composed of H_2O.
>
> [In symbols: '$\text{Composed-of-water}(x) \equiv_x \text{Composed-of-}H_2O(x)$']

In these three sentences the expression 'just is' (or its formalization '$\equiv_x$') is functioning as an identity-predicate of sorts. To accept '$F(x) \equiv_x G(x)$' is not simply to accept that all and only the Fs are Gs. If you accept Sibling, for example, you believe that there is *no difference* between being a sibling and sharing a parent with someone; you believe that if someone is a sibling it is *thereby* the case that she shares a parent with someone. (Compare: if you accept 'Hesperus is Phosphorus', you believe that someone who travels to Hesperus has *thereby* traveled to Phosphorus.)

One might be tempted to describe Sibling, Heat and Water as expressing identities amongst *properties* (e.g. 'the property of being a sibling = the property of sharing a parent'.) I have no qualms with this description, as long as property-talk is understood in a suitably deflationary way. But I will mostly avoid property-talk here because it is potentially misleading. It might be taken to suggest that one should only assert Sibling

[1] See especially Yablo (1993), Fine (1994), Chalmers (1996), Jackson (1998), Block and Stalnaker (1999), Chalmers and Jackson (2001), the essays in Stalnaker (2003) and Kment (2006).

if one is prepared to countenance a naïve realism about properties—the view that even though it is intelligible that there be no properties, we are lucky enough to have them. The truth of SIBLING, as I understand it, is totally independent of such a view. If one wishes to characterize the difference between '$\equiv_x$' and the standard first-order identity predicate '$=$', the safe thing to say is that whereas '$=$' takes a singular-term in each of its argument-places, '$\equiv_x$' takes a first-order predicate in each of its argument places. One might therefore refer to sentences of the form $\ulcorner a = b \urcorner$ (where a and b are singular terms) as *first-order* identity statements, and sentences of the form $\ulcorner \phi(x) \equiv_x \psi(x) \urcorner$ (where $\phi(x)$ and $\psi(x)$ are first-order predicates) as *second-order* identity statements.

Sometimes one is in a position to endorse something in the vicinity of a second-order identity statement even though one has only partial information. Suppose you know that the chemical composition of water includes oxygen but don't know what else is involved. You can still say:

> Part of what it is to be composed of water is to contain oxygen,
>
> [In symbols: 'Composed-of-water$(x) \ll_x$ Contains-Oxygen(x).']

I shall call this as a *semi-identity statement*. Think of it as a more idiomatic a way of saying:

> To be composed of water *just is* (to contain oxygen and maybe something else).

or, more precisely:

$$\exists Z[\text{Composed-of-water}(x) \equiv_x (\text{Contains-Oxygen}(x) \wedge Z(x))]$$

Second-order identity statements can be dispensed with in the presence of semi-identity statements. WATER, for example, is equivalent to the conjunction of 'part of what it is to be water is H_2O' and 'part of what it is to be H_2O is to be water'. And, in general, '$F(x) \equiv_x G(x)$' is equivalent to the conjunction of '$F(x) \ll_x G(x)$' and '$G(x) \ll_x F(x)$'.

As in the case of second-order identity statements, it is tempting to think of semi-identity statements in terms of properties (e.g. 'the property of being water has the property of containing oxygen as a part'.) Again, I have no objection to this sort of description, as long as property-talk is taken in a suitably deflationary spirit. But I will avoid it here because of its potential to mislead.

As in the case of second-order identity statements, it seems to me that we have moderately robust pre-theoretic grasp of semi-identity statements. It is natural to say, for example, that part of what it is to be scarlet is to be red, that part of what it is to be an elephant is to be a mammal, and that part of what it is to be a star is to undergo nuclear fusion. I will rely on this pre-thoeretic understanding throughout the paper. But I also hope to do some regimenting as we go along, by developing interconnections between the notion of semi-identity and other notions.

Let me end this section by mentioning a few things that my discussion of semi-identity statements does *not* presuppose. I do not presuppose that semi-identity corresponds to a notion of essence. I do not presuppose that the semi-identity statements make distinctions too fine-grained to be captured using boxes and diamonds, as in Fine (1994). I do not presuppose that '$F(x) \ll_x G(x)$' entails that something is an F 'in virtue' of being a G. Finally, I do not presuppose that the semi-identity predicate '$\ll_x$' expresses a 'non-modal' notion, or that one can come to understand it independently of one's grasp of the notion of possibility.

2 Connections

In this section I will spell out a three-way connection between intelligibility, identity and why-closure. Mathematical sentences (and other sentences involving abstract-object-talk) will be addressed in section 7—please ignore them until then.

2.1 Intelligibility

Say that a story is *de re* if every name used by the story is used to say of the name's actual bearer how it is according to the story, and if every predicate used by the story is used to attribute the property actually expressed by the predicate to characters in the story. Accordingly, a *de re* story that says 'Hesperus is covered with water' is a story according to which Venus itself is covered with H_2O. Relatedly, no names that are actually empty (e.g. 'Sherlock Holmes'), and no predicates that are actually empty (e.g. '... is composed of phlogiston', '... is a unicorn'), are allowed to figure in *de re* stories.

De re stories supply a useful way of getting a grip on the notion of intelligibility that will play a role in this paper. Here are a couple of examples:

1. I can certainly imagine circumstances under which one might be inclined to conclude that Mark Twain is not Samuel Clemens—I can imagine, for example, discovering that *Huckleberry Finn* and other works written under the name 'Mark Twain' were authored by one of Clemens's literary rivals. And I could certainly make sense of a *de re* story describing such circumstances. But given that I take for granted that Mark Twain is, in fact, Samuel Clemens, I am unable to make sense of a situation in which it is true of Mark Twain—the man himself—that he is not Samuel Clemens—the man himself. For I take for granted that it would have to be a situation in which someone is not identical to himself, which is something I am unable to make sense of. Since a *de re* story that says 'Mark Twain is not Samuel Clemens' would depict such situation, I take the story to be unintelligible.

2. I can certainly imagine circumstances under which one might be inclined to conclude that water does not contain hydrogen—I can imagine, for example, discovering that water electrolysis results in the release of oxygen and nitrogen rather than oxygen and hydrogen. And I could certainly make sense of a *de re* story describing such circumstances. But given that I take for granted that part of what it is to be composed of water is to contain hydrogen, I am unable to make sense of a situation in which a

portion of water—the very substance that results from binding oxygen and hydrogen atoms in a certain kind of way—contains no hydrogen. For I take for granted that it would have to be a situation in which something containing hydrogen contains no hydrogen, which is something I am unable to make sense of. Since I take a *de re* story that says 'there is a portion of water containing no hydrogen' to depict such a situation, I take the story to be unintelligible.

An important lesson of these examples is that which propositions one takes to be intelligible depends on what one assumes is the case. If one takes for granted that Mark Twain is Samuel Clemens, one will regard a *de re* story stating 'Mark Twain is not Samuel Clemens' as unintelligible; if one takes for granted that Mark Twain is *not* Samuel clemens, one will regard that same story as intelligible. Notice, however, that depicting a scenario which is inconsistent with one's assumptions is not sufficient for a *de re* story to be regarded as unintelligible. (I take for granted that the Earth is the third planet from the Sun, for example; yet I have no trouble making sense of a *de re* story that says 'There are 14 planets between the Earth and the Sun'.) Identity and semi-identities are special in this respect. Once taken on board, they impose limits on what one finds intelligible. Specifically: if a subject accepts a family of identity and semi-identity statements, and if she is able to conclude on the basis of those statements that a *de re* story would depict an inconsistent situation, then she will be unable to make sense of the story.

This yields two sufficient condition for unintelligibility: (1) someone who accepts $\ulcorner a = b \urcorner$ will treat a *de re* story that says $\ulcorner a \neq b \urcorner$ (or says something that is assumed to entail $\ulcorner a \neq b \urcorner$) as unintelligible; and (2) someone who accepts $\ulcorner F(x) \ll_x G(x) \urcorner$ will treat a *de re* story that says $\ulcorner \exists x(F(x) \wedge \neg G(x)) \urcorner$ (or says something that is assumed to entail $\ulcorner \exists x(F(x) \wedge \neg G(x)) \urcorner$) as unintelligible.

A consequence of these observations is that intelligibility, as it pertains to *de re* stories, is *not* in general an *a priori* matter. For it turns on which identities and semi-identities one accepts, and the acceptance of identities and semi-identities is not in general an *a priori* matter.

2.2 Why-Closure

In the preceding section we considered a link between the acceptance of identity and semi-identity statements, on the one hand, and the intelligibility of *de re* stories, on the other. Now we will see that there is a three-way link between: (a) the acceptance of identity and semi-identity statements, (b) the intelligibility of *de re* stories, and (c) what I shall call 'why-closure'.

As before, I begin with an example. Suppose someone asks 'why are elephants mammals?'. Different sorts of answers might be appropriate, depending on the context. One way of answering the question is by explaining that elephants form part of a lineage leading back to common mammalian ancestors. In doing so one supplies an elucidation of what it is for elephants to be mammals: one explains what the fact that elephants are mammals consists in. A different way of answering the question is by justifying the claim that elephants are mammals. One might point to the distinctive mammalian traits of elephants, or note convergences between elephantine DNA and the DNA of other mammals. And, of course, one might use elements of these two types of answer in a single response: one might explain that elephants form part of a lineage leading back to common mammalian ancestors, and go on to cite evidence for this lineage. But notice the following: it would hard to know what to make of someone who said "I know full well what it is for elephants to be mammals, and I can see as clearly as can be that elephants are mammals—what I want to know is *why* elephants are mammals!" (Notice, in contrast, that there is no analogous problem with someone who says "I know full well what it is for the window to be broken, and I can see as clearly as can be that the window is broken—what I want to know is *why* the window is broken!". Here a perfectly good response might be "because it was hit by a soccer ball".)

Let me introduce some notation. Say that 'why is it the case that q?' is read as a *grounding question* when it is read either as requesting an elucidation of what it is for it to be the case that q or as requesting a justification for the claim that q (or both).[2] Say that a sentence ϕ is treated as *why-closed* if one is unable to make sense of ⌜why is it the case that ϕ?⌝ unless one reads it as a grounding question. (Thus, 'Hesperus is Phosophorus' and 'part of what it is to be scarlet is to be red' are treated as why-closed, but 'the Earth is the third planet from the Sun' is not.)

I would like to suggest that there is a straightforward connection between why-closure and the unintelligibility of *de re* stories. It is this: one will treat s as why-closed just in case one is unable to make sense of a *de re* story stating s's negation. Here are a couple of examples:

1. Suppose it is taken for granted that Mark Twain is Samuel Clemens. With some awkwardness, one may still be able to read 'why is Mark Twain Samuel Clemens?' as a grounding question. (One might answer, for example, 'because the works signed 'Mark Twain' were written by Samuel Clemens'.) But suppose someone says "I know full well what it is for Mark Twain to be Samuel Clemens, and I can see as clearly as can be that Mark Twain *is* Samuel Clemens—what I want to know is *why* Mark Twain is Samuel Clemens!" Then one is unable to make sense of one's interlocutor's request. This is correlated with the fact that one would regard a *de re* story stating 'Mark Twain is *not* Samuel Clemens' as unintelligible.

2. Suppose it is taken for granted that part of what it is to be water is to contain hydrogen. With some awkwardness, one may still be able to read 'why does every portion of water contain hydrogen?' as a grounding question. (One might answer,

[2]There is a use-mention infelicity in this passage. It could be avoided by treating the passage as a schema in which arbitrary sentences are allowed to take the place of 'q', but doing so would complicate the exposition, here and below. Since there is no real risk of confusion, it seems to me that a rise in expository complexity is too high a price to pay for logical hygiene. Throughout the paper I shall adopt the policy of indulging in use-mention infelicities whenever they cannot be easily avoided, as long as there is no real risk of confusion.

for example, 'because water is made up of H_2O molecules, and H_2O molecules are built from an oxygen atom and two hydrogen atoms'.) But suppose someone says "I know full well what it is for every portion of water to contain hydrogen; I can see as clearly as can be that part of what it is to be water is to contain hydrogen, and therefore that every portion of water contains hydrogen—what I want to know is *why* every portion of water contains hydrogen!" Then one is unable to make sense of one's interlocutor's request. This is correlated with the fact that one would regard a *de re story* stating 'there is a portion of water containing no hydrogen' as unintelligible.

It is important to distinguish between cases in which one is unable to make sense of a why-question and cases in which one is unable to see what sort of explanation would count as an answer to the why-question. Consider the following examples:

Why did the particle decay now rather than a second later?

Why are the laws as they are?

Why is there something rather than nothing?

One is able to make sense of these questions, even if one is unable to see what sort of explanation would count as an answer. This is evidenced by the fact the fact that one might reply by saying 'that's just the way the world turned out' or 'we were just lucky'. (Compare with 'why are scarlet things red?' or 'why does water contain hydrogen?'—when grounding readings have been excluded by context, one is tempted to respond, e.g. 'scarlet things just *are* red' or 'What do you mean?', but not 'that's just the way the world turned out'.) And, of course, there is a link with the intelligibility of *de re* stories. Whereas one is unable to make sense of a *de re* story that says 'there is something that is scarlet but not red', or 'there is a portion of water that contains no hydrogen', one has no problem making sense of a *de re* story that says 'the particle decayed a second later', or 'the cosmological constant is greater than it actually is', or 'there is nothing'.

In the preceding section I set forth two sufficient conditions for the unintelligibility of *de re* stories. If the link between why-closure and intelligibility that we have considered here is along the right lines, this yields corresponding sufficient conditions for why-closure: (1) when $\ulcorner a = b \urcorner$ is taken for granted, any sentence $\ulcorner a = b \urcorner$ is assumed to entail (including $\ulcorner a = b \urcorner$ itself) is treated as why-closed; (2) when $\ulcorner F(x) \ll_x G(x) \urcorner$ is taken for granted, any sentence $\ulcorner \forall x(F(x) \rightarrow G(x)) \urcorner$ is assumed to entail (including $\ulcorner \forall x(F(x) \rightarrow G(x)) \urcorner$ itself) is treated as why-closed.

3 Deciding between rival identity and semi-identity statements

The purpose of this section is to say something about the considerations that go into determining which identity and semi-identity statements to accept.

Part of the story is straightforward. One will only accept $\ulcorner F(x) \ll_x G(x) \urcorner$ if one also accepts $\ulcorner \forall x(F(x) \rightarrow G(x)) \urcorner$, and one will only accept $\ulcorner a = b \urcorner$ if one also accepts (*ab*-independent[3]) instances of $\ulcorner \phi(a) \leftrightarrow \phi(b) \urcorner$. But this cannot be the end of the story, since one might reject $\ulcorner F(x) \ll_x G(x) \urcorner$ even if one accepts $\ulcorner \forall x(F(x) \rightarrow G(x)) \urcorner$, and one might reject $\ulcorner a = b \urcorner$ even if one accepts every (*ab*-independent) instance of $\ulcorner \phi(a) \leftrightarrow \phi(b) \urcorner$. Our problem is to determine what else is required.

In rough outline, the story is this. In an effort to satisfy our goals, we develop strategies for interacting with the world. Fruitful strategies allow us to *control* what the world is like and *predict* how it will evolve under specified circumstances. They also allow us to *direct* our research in ways that lead to the development of further fruitful strategies. In order to articulate the strategies we adopt, we do three things at once: firstly, we develop a language within which to formulate theoretical questions; secondly we set forth theoretical claims addressing some of these questions; finally, we endorse a family of identity and semi-identity statements. The third task is connected to the other two because the identity and semi-identity statements we endorse will help determine which theoretical questions are treated as why-closed, and therefore which theoretical questions correspond to avenues of investigation that we take to be pointless.

It is useful to consider some examples.

Example 1: The chemistry crank

A chemistry crank believes that the chemical composition of various substances varies with temperature. Methanol, she thinks, is normally composed of hydrogen, oxygen and carbon; but at a temperature of precisely $\sqrt{2}$ degrees celsius, its chemical composition changes to hydrogen and platinum. Similarly, our crank expects the chemical composition of water to vary with temperature. Her model predicts that at a temperature of precisely π degrees celsius, water is composed of oxygen and gold. Our crank sets out to test her hypothesis by carrying out water-electrolysis at a range of temperatures. What she finds, of course, is that hydrogen and oxygen bubble up, regardless of how closely the temperature approximates π degrees celsius. She concludes that every portion of water on Earth is composed of hydrogen and oxygen. She takes this to be a remarkable fact, in need of explanation. "Perhaps Earth's gravitational field is getting in the way"—she thinks—"perhaps under low-gravity conditions water is composed of oxygen and gold at π degrees celsius."

What semi-identity statements will our crank accept? The decision will be based on her explanatory needs. She wishes to make sense of 'why is this portion of water composed

[3]The qualification of *ab*-independence is needed because the link between accepting $\ulcorner a = b \urcorner$ and accepting instances of $\ulcorner \phi(a) \leftrightarrow \phi(b) \urcorner$ is not very illuminating if allows, e.g. $\ulcorner a = a \leftrightarrow a = b \urcorner$ to count as an instance. To solve the problem one must do more than simply insist that $\phi(x)$ not contain a and b, since that would fail to rule out a case in which a is 'Socrates' and $\phi(x)$ is 'x Socratizes', or a case in which $\phi(x)$ is '$x = c$' and it is taken for granted that $\ulcorner a = c \urcorner$ is true. We shall therefore stipulate that the *ab*-independent instances of $\ulcorner \phi(a) \leftrightarrow \phi(b) \urcorner$ are to be built up from a certain set of basic vocabulary, and insist that this set be such that it is not taken for granted that one can only accept every instance of $\ulcorner \phi(a) \leftrightarrow \phi(b) \urcorner$ (for $\phi(x)$ built up from vocabulary in that set) if one also accepts $\ulcorner a = b \urcorner$. (This does not fully determine what *ab*-independence is, but it determines enough for present purposes.)

of oxygen and hydrogen?' (read as a non-grounding question). So she has reason *not* to accept 'part of what it is to be water is to be composed of oxygen and hydrogen'. Perhaps she will instead be inclined to accept 'part of what it is to be water is to be a colorless, odorless, liquid'. This would lead her to think that there is sense to be made of 'why is this portion of water composed of hydrogen and oxygen' (read as a non-grounding question), but not of, e.g. 'why is this portion of water a liquid?' (unless read as a grounding question).

Now consider a real chemist. She does not think it at all remarkable that every portion of water on Earth be composed of hydrogen and oxygen, and does not think it is in need of explanation. In fact, she believes that there is no sense to be made of 'why is every portion of water on Earth composed of hydrogen and oxygen?' (unless it is read as a grounding question). Accordingly, she is inclined to accept 'part of what it is to be water is to be composed of hydrogen and oxygen'. She also believes that 'why is this portion of water a liquid?' is a sensible question to ask (when read as a non-grounding question). So she had better not accept 'part of what it is to be water is to be a liquid'.

The crank and the chemist use different strategies for investigating chemical phenomena. Because of these differences, they articulate their methods of inquiry in different ways: they both accept the theoretical claim 'every portion of water on Earth is composed of hydrogen and oxygen', but only the chemist treats the claim as why-closed because only the chemist accepts 'part of what it is to be water is to be composed of hydrogen and oxygen'. It is important to be clear, however, that nothing in their methods of inquiry *mandates* a particular selection of semi-identity statements. The crank could, if she really wanted, accept 'part of what it is to be water is to be composed of oxygen and hydrogen'. By doing so she would be committed to thinking that there is no sense to be made of 'why is every portion of water on Earth composed of oxygen and hydrogen?' (unless read as a grounding question). But this needn't interfere with her ability to articulate her methods of inquiry. For rather than asking 'why is every portion of water on Earth composed of hydrogen and oxygen?', she could ask, e.g. 'why is every portion of colorless, odorless, liquid on Earth with such-and-such additional properties composed of hydrogen and oxygen?'.

The reason for preferring a particular set of semi-identity statements over its rivals is not that it is somehow entailed by one's explanatory methods. The reason is that it yields an articulation of one's explanatory methods that one finds especially congenial. Accepting 'part of what it is to be water is to be composed of oxygen and hydrogen' is uncongenial for the crank because it forces her to reformulate some of her chemical claims. Rather than saying, e.g. 'under ideal conditions, water is composed of oxygen and gold at π degrees celsius', she will now have to say, e.g. 'under ideal conditions, colorless, odorless liquids with such-and-such additional properties are composed of oxygen and gold at π degrees celsius'. Such reformulations are undesirable because of the extra work, but also because it may not be obvious to the crank how a given claim is best reformulated. There would be no problem if—before the change—the crank had an exhaustive characterization of what it is to be water. For then she would be in a position to reformulate her chemical claims by replacing every occurrence of the world 'water' with the relevant characterization. But suppose that—before the change—all the crank has is a *partial* characterization of what it is to be water: she believes that part of what it is to be water is to be a colorless, odorless liquid, but leaves open that there might be more to being water than that. Then she won't be sure how to cash out the 'such-and-such' in a claim like 'under ideal conditions, colorless, odorless liquids with such-and-such additional properties are composed of oxygen and gold at π degrees celsius.'

If our crank comes to accept 'part of what it is to be water is to contain hydrogen' she will change her views about the satisfaction-conditions of '. . . is a portion of water'. She might start out believing that part of what it takes to satisfy '. . . is a porition of water' is to be a colorless, odorless liquid. But after accepting 'part of what it is to be water is to contain oxygen and hydrogen', she will come to believe that part of what it takes to satisfy '. . . is a portion of water' is to be composed of oxygen and hydrogen. This is *not* to say that 'water', and used by the crank, undergoes a change in satisfaction-conditions. (It rigidly designates H_2O throughout.) The point is that if the crank accepts deviant semi-identity statements, she will also have deviant beliefs about the satisfaction-conditions of the expressions of her language. (A similar point can be made with respect to the crank's *concepts*: if she accepts deviant semi-identity statements, she will also have deviant beliefs about the satisfaction-conditions of her concepts.)

Example 2: The zoology crank

A zoology crank believes that snails placed in a 'lobster environment' will evolve into lobsters after a few generations, and that lobsters placed in an 'elephant environment' will evolve into elephants after a few generations. He also believes that 'original creatures' spontaneously come into existence from time to time, and their descendants evolve into members of various species as the generations go by. As a result, our crank believes that it is rare for all the members of a species to share a common ancestry. They are typically composed of the descendants of several different 'original creatures'. All the same, our crank decides to sample the DNA of extant elephants in order to determine whether they are, in fact, related. When the results come in, he concludes that elephants do, as a matter of fact, share a common ancestry. This conclusion strikes him as remarkable, and in need of explanation. "Perhaps it was an incredible coincidence"—he thinks—"perhaps it is only the descendants of a particular original creature that happened to end up in an elephant environment."

What semi-identity statements will our crank accept? As before, the decision will be based on his explanatory needs. He wishes to makes sense of 'why do elephants share a common ancestry?' (when read as a non-grounding question). So he has reason *not* to accept 'part of what it is to be an elephant is to have a certain lineage'. Perhaps he will instead be inclined to accept, e.g. 'part of what it is to be an elephant is to have such-and-such a phenotype'. This would lead him to think that there is sense to be made of 'why do elephants share a common ancestry' (when read as a non-grounding question), but not of, e.g. 'why do elephants have trunks?' (unless it is read as a grounding question).

Unfruitful research might lead our zoology crank to change his approach. He may alter his zoological beliefs and explanations. He might come to believe, for example, that changes in phenotype between an individual and its offspring are much slighter (and far more random) than he previously thought, and that the link between surviving under environmental conditions of a given type and having a particular phenotype is much less robust than he previously thought. Accordingly, he might come to think that the sharing of elephant-phenotypes—and the ability of elephants to interbreed—is best explained by elephants' common ancestry.

What semi-identity statements will our crank now accept? After the change of approach, the notion of having such-and-such a lineage can be expected to play a more central role in his zoological thinking than the notion of having such-and-such a phenotype. So by accepting 'part of what it is to be an elephant is to have a certain lineage' rather than 'part of what it is to be an elephant is to have such-and-such a phenotype' he can expect to articulate his methods of zoological inquiry in a relatively congenial way. He will be committed to thinking that there is no sense to be made of 'why do elephants have a common ancestry?' (unless read as a grounding question). But he can still hold that there are intelligible non-grounding questions in the vicinity (e.g. 'why are elephants the only extant creatures with an elephantine phenotype?').

One respect in which the zoology example is different from the chemistry example is that the considerations that recommend accepting 'part of what it is to be an elephant is to have a certain lineage' over some of its rivals are not particularly decisive. There is room for arguing, for instance, that the notion of being a member of a group of organisms that are able to interbreed and produce fertile offspring will play a more central role in the zoologist's thinking than the notion of having such-and-such a lineage, and therefore that his methods of zoological inquiry would be best articulated by accepting 'part of what it is to be an elephant is to be a member of a certain group of organisms that are able to interbreed and produce fertile offspring' (Two groups, actually: African and Indian elephants can't interbreed).

In general, the question of what identities and semi-identities to accept may turn on the purposes at hand. The same theorist might find it useful to accept 'what it is to be an elephant is to be a member of a certain lineage' for the purposes thinking about evolution by natural selection, and find it useful to accept 'what it is to be an elephant is to be a member of a certain group of organisms that are able to interbreed and produce fertile offspring' for the purposes of studying ancient patterns of elephant-migration. Moreover, one would expect that situations in which one's methods of inquiry can be articulated with similar success by way of rival semi-identity statements would be the rule rather than the exception.

Example 3: The vixen-conspiracy crank

A conspiracy crank believes that vixens are super-intelligent creatures with magical powers who are out to dominate the Earth. She believes they are extraterrestrials—no Earthly creature has ever been vixen-like. (Male foxes, on the other hand, are not alive at all; they are robots created by the vixens as a cover for their operations.) Vixens have the ability to assume any form they like, and have chosen to assume the form of mammalian females. This is a fact in need of explanation.

What semi-identity statements will our crank accept? She wishes to make sense of 'why are vixens female?' (when read as a non-grounding question). So she has reason to reject 'part of what it is to be a vixen is to be female'. Perhaps she will claim instead that what it is to be a vixen is to be a member of a certain extra-terrestrial lineage. She can then be expected to have unorthodox views about the satisfaction-conditions of 'vixen'. In particular, she might deny that part of what it takes to satisfy 'vixen' is to be female—she might claim instead that all it takes to satisfy 'vixen' is to be a member of the right extra-terrestrial lineage, whether or not one happens to assume a female form. This is *not* to say, however, that 'vixen', as used by the crank, undergoes a change in satisfaction-conditions. Even in her mouth, what it takes to satisfy 'vixen' is to be a female fox.

As in previous examples, nothing in the crank's methods of inquiry mandates a particular choice of semi-identity statements. (She could, if she really wanted, accept 'part of what it is to be a vixen is to be female'. By doing so she would be committed to thinking that there is no sense to be made of 'why are vixens female?', unless read as a grounding question. But this needn't interfere with her ability to articulate her methods of inquiry. For rather than asking 'why are vixens female?' she might ask 'why have the extraterrestrials chosen to assume female form?'.) But the present example can be used to illustrate an important point: even if nothing in the crank's methods of inquiry mandates a particular choice of semi-identity statements, there is room for thinking that certain choices are excluded by facts about the crank's *language*, together with her methods of inquiry. Suppose, for example, that one believes that 'vixens are female' is *analytic*. One might then think that—on pain of betraying the meanings of her words—the crank is barred from accepting, e.g. 'to be a vixen *just is* to be a member of a certain extraterrestrial lineage'.

More generally, suppose, one believes that our words have primary intensions.[4] One believes, in particular, that the meaning of 'vixen' determines more than just an intension for 'vixen' (i.e. a function assigning to each possible world the set of objects 'vixen' applies to at that world); it also determines, for each possible world, what the intension of 'vixen' should be taken to be on the assumption that that world is actual. One might then think that—on pain of betraying the meanings of her words—the crank is not free to choose between rival semi-identity statements. Suppose the meaning of 'vixen' determines that, on the assumption that the crank's beliefs are true, one should take the intension of 'vixen' to assign each world the set of female members of a certain terrestrail lineage. Then—on pain of betraying the meanings of her words—the crank is barred from accepting 'to be a vixen *just is* to be a member of a certain extra-terrestrial lineage'.

I myself am skeptical of the notion of analyticity as traditionally understood, and I can see no good reason for believing that the meaning of 'vixen' determines what the intension of 'vixen' should be taken to be on the assumption that that the crank's beliefs are true. But I shall remain neutral about such questions throughout the paper.

[4]See Chalmers (1996) and Jackson (1998). As far as I can tell, there is no reason to think that Chalmers and Jackson are committed to the strong version of the position I discuss here.

Example 4: First-order Identity

The examples we have considered so far have all concerned *second-order* identity statements, and they have all had the same form. We discussed scenarios in which a statement of the form $\ulcorner\forall x(F(x) \rightarrow G(x))\urcorner$ is taken for granted, and asked about the sorts of considerations that would lead one to take the additional step of accepting or rejecting $\ulcorner F(x) \ll_x G(x)\urcorner$. The first-order analogue of these examples would be a case in which every (*ab*-independent) instance of $\ulcorner\phi(a) \leftrightarrow \phi(b)\urcorner$ is taken for granted, and one asks about the sorts of considerations that would lead one to take the additional step of accepting or rejecting $\ulcorner a = b\urcorner$. It is not easy to find an example of this kind in which a and b are terms from science or everyday life. For in non-philosophical contexts the acceptance of a sentence of the form $\ulcorner a \neq b\urcorner$ will usually be linked to the acceptance of some (*ab*-independent) instance of $\ulcorner\phi(a) \wedge \neg\phi(b)\urcorner$. (Our acceptance of 'Saturn $\neq$ Jupiter', for instance, is linked to our acceptance of sentences like 'Saturn has a big ring, Jupiter does not' or 'Saturn is there, Jupiter is not'.) I shall therefore use an example in which the relevant terms come from science-fiction.

By the dawn of the twenty-seventh century, the old theories of quantum mechanics had been superseded by *superquantum-theory*. This new fundamental physics countenances four different kinds of fields (numbered 1 through 4). Type-1 and type-2 fields are generated, respectively, by particles of type-1 and type-2. Type-1 and type-2 particles usually have different locations, but on occasion they undergo a 'superquantum merger', in which they come to occupy the same location. ('Location' in superquantum-theory is a messy affair: particles occupy 'clouds' in space rather than exact regions of space, and particles 'merge' when they come to share a 'cloud'.)

Everyone agrees that type-3 fields and type-4 fields are different. (Type-3 fields can be blocked by a lead barrier, for example; but type-4 fields cannot.) Everyone agrees, moreover, that type-3 fields and type-4 fields are both generated by particles, and there is a stipulation in place to the effect that the source of a type-3 fields is to be referred to as a 'type-3' particle and the source of type-4 fields is to be referred to as a 'type-4' particle. There is, however, an important disagreement. As far as anyone has been able to tell, type-3 fields and type-4 fields are always generated from the same location. *Monists* suggest that this is because type-3 particles and type-4 particles are one and the same. *Dualists* suggest that type-3 particles and type-4 particles are distinct, but that they are always 'merged' with one another, in the same sort of way that type-1 particles and type-2 particles are sometimes 'merged'.

Perhaps you think that there must be an objective fact of the matter about whether monism or dualism is true. That is an issue that will be taken up in section 4, but it is not our present concern. Our present concern is with the sorts of considerations that would lead one to embrace one of these positions over its rival. Here the crucial observation is that monists and dualists face different explanatory burdens. From the dualist's perspective, the colocation of type-3 particles and type-4 particles is a fact that calls for explanation. (Perhaps she *has* an explanation: perhaps the theoretical model which she uses to explain the occasional merging of type-1 and type-2 particles predicts that the particles generating type-3 and type-4 fields will be systematically merged, or that they will be merged in all but extraordinary circumstances.) From the perspective of the monist, on the other hand, there is no need to explain the colocation of type-3 and type-4 particles. For when grounding questions have been excluded by context, there is no sense to be made of 'I can see that this type-3 particle is identical to this type-4 particle, but why are they colocated?'.

If she really wanted to, the dualist could try to articulate her methods of inquiry in monistic terms. Whether or not she would find a monistic articulation congenial depends on the details of her methods of inquiry. The change may prove to be a welcome one if her theorizing is unable to supply a satisfying answer to the question 'why are type-3 particles and type-4 particles systematically colocated?'. For in that case the shift to monism would relieve her of an unwelcome explanatory burden. But suppose instead that she has a satisfying answer to the question—perhaps her answer involves a bit of theory which is ripe with interesting predictions, some of which have been confirmed. Then whether the move is congenial will depend on whether she is able to find an attractive way of articulating the insight captured by the extra bit of theory in a monistic setting.

What this story suggests is that the sorts of considerations that would lead one to go from accepting (*ab*-independent) instances of $\ulcorner \phi(a) \leftrightarrow \phi(b) \urcorner$ (e.g. 'this type-3 particle is located here $\leftrightarrow$ this type-4 particle is located here') to accepting $\ulcorner a = b \urcorner$ (e.g. 'this type-3 particle is this type-4 particle') are similar to the sorts of considerations that would lead one to go from accepting $\ulcorner \forall x(F(x) \rightarrow G(x)) \urcorner$ to accepting $\ulcorner F(x) \ll_x G(x) \urcorner$. In both cases, one will be motivated to take the additional step to the extent that the resulting explanatory landscape offers good prospects for a congenial articulation of one's methods of inquiry.

4 Realism

Nothing in our discussion so far is in tension with a realist attitude towards identity and semi-identity statements:

> Realism
> There is good sense to be made of the question whether an identity or semi-identity statement is *true*, over and above the question of whether it yields a congenial articulation of our methods of inquiry.

This leaves open the question of whether there is a connection between the truth of an identity or semi-identity statement and its ability to deliver congenial articulations of our methods of inquiry. But a realist might be expected to endorse the following:

> Evidence
> Provided we have the right methods of inquiry, the ability of a particular set of identity and semi-identity statements to deliver congenial articulations of our methods of inquiry will supply *evidence* for their truth.

On the assumption that one accepts both Realism and Evidence, the discussion in preceding sections yields an answer to the question of how the truth of identity and semi-identity statements ought to be ascertained: pick the set of identity and semi-identity statements that supplies the most congenial articulation of your methods of inquiry. (If there is a tie, proceed as you would in other cases of underdetermination of theory by evidence.)

Throughout the next few sections I shall work on the assumption that Realism and Evidence are both true. But the issue will be revisited in section 8.

5 Varieties of Semi-identity

So far I have been focussing on the most basic case of semi-identity statements: monadic unconditional semi-identity statements. In this section we will consider semi-identity statements that go beyond the basic case.

Non-monadic semi-identity statements

The semi-identity operator '$\ll$' can bind more than one variable. Here is an example:

$$\text{Sisters}(x, y) \ll_{x,y} \exists z(\text{Parent}(z, x) \land \text{Parent}(z, y))$$

(*Read:* part of what it is for x and y to be sisters is for x and y to share a parent.)

As in the monadic case, there is a link between polyadic semi-identity and why-closure: when $\ulcorner F(x_1, \ldots, x_n) \ll_{x_1,\ldots,x_n} G(x_1, \ldots, x_n)\urcorner$ is taken for granted, any sentence $\ulcorner\forall x_1 \ldots \forall x_n(F(x_1, \ldots, x_n) \rightarrow G(x_1, \ldots, x_n))\urcorner$ is assumed to entail is treated as why-closed. (One is, for example, unable to make sense of the following request: "I know full well what it is for a pair of sisters to share a parent. And I can see as clearly as can be that part of what it is for two people to be sisters is for them to share a parent, and therefore that any pair of sisters share a parent—what I want to know is *why* every pair of sisters share a parent!") Because of this link with why-closure, the story I told in section 3 about how we determine which semi-identity statements to accept carries over to the polyadic case: we accept whichever semi-identity statements allow for the most congenial articulations of our methods of inquiry.

The semi-identity operator can also bind no variables at all, as in the following example:

$$\text{a wedding takes place} \ll \text{someone gets married}$$

(*Read:* part of what it is for a wedding to take place is for someone to get married.)

As before, there is a link with why-closure: when $\ulcorner\phi \ll \psi\urcorner$ is taken for granted, any sentence $\ulcorner\phi \rightarrow \psi\urcorner$ is assumed to entail is treated as why-closed. And, as before, the link can be used to tell a story about how we determine which semi-identities to accept.

Constitutive properties and conditional semi-identity statements

P is a *constitutive* property if the assumption that z has P is enough to license the conclusion that part of what it is to be z is to be P. Being human, for example, is a constitutive property. For the assumption that Socrates is human is enough to license:

$$x = \text{Socrates} \ll_x \text{Human}(x)$$

(*Read:* part of what it is to be Socrates is to be human.)

Being snub-nosed, on the other hand, is not a constitutive property. For the assumption that Socrates is snub-nosed does not license:

$$x = \text{Socrates} \ll_x \text{Snub-Nosed}(x)$$

(*Read:* part of what it is to be Socrates is to be snub-nosed.)

There is a link between constitutivity and why-closure: when it is taken for granted that P is a constitutive property and that z has P, one isn't able to make sense of someone who asks, of z, 'I can see that it has P, but *why* does it have P?' (unless the request is read as a grounding question). If, for instance, you take for granted that being human is a constitutive property, you should be unable to make sense of 'I can see that Socrates is human, but *why* is he human?' (unless the request is read as a grounding question).

Because of this link with why-closure, a version of the story I told in section 3 can be used to explain how we determine which properties to regard as constitutive. Consider a second zoology crank. This one believes that all living creatures start their lives as 'generic beings', but that different creatures look different later in life because of the particular environments they have been immersed in. After a few years living in a 'human environment' a generic creature will begin to acquire human characteristics, and after a few years living in an 'octopus environment' it will begin to acquire octopus characteristics. "I can see that Susan is human"—the crank might ask himself—"but *why* is she human?" He might add: "perhaps it is because she did not spend long enough in an octopus environment early in her life". Since the crank does not treat 'Susan is human' as why-closed, she has reason *not* to treat being human as constitutive property. But, of course, should unfruitful research lead the crank to adopt methods of inquiry more akin to those of professional zoologists, she may come to think of 'Susan is human' as why-closed, and thereby acquire reason to treat being human as a constitutive property. The lesson is this: as in the cases of identity and semi-identity statements discussed in section 3, the properties one will treat as constitutive are those that supply congenial articulations of one's methods of inquiry.

The claim that P is a constitutive property can be formulated as a *conditional* semi-identity statement:

$$\frac{P(z)}{x = z \ll_x P(x)}$$

(*Read:* assume z is P; then part of what it is to be z is to be P.)

Having w as a biological parent is a *parameterized* constitutive property. For the assumption that z has w as a biological parent is enough to warrant the conclusion that part of what it is to be z is to have w as a parent. Parameterized constitutive properties can also be captured by conditional semi-identity statements. One can say, for example:

$$\frac{\mathrm{B}(z, w)}{x = z \ll_x \mathrm{B}(x, w)}$$

(*Read:* assume z has w as a biological parent; then part of what it is to be z is to have w as a biological parent.)

Although conditional semi-identity statements can take other forms, here I will restrict my attention to the reflexive case: the case in which the antecedent is a first-order formula $\phi(z, \vec{w})$ and the consequent is $\ulcorner x = z \ll_x \phi(x, \vec{w})\urcorner$. Doing so has a crucial advantage. Since every reflexive conditional semi-identity statement expresses the thought that a certain property is constitutive, the considerations that go into determining which reflexive conditional semi-identity statements to accept are precisely the considerations that go into determining which properties to treat as constitutive. (Attention will also be restricted to *monadic* conditionals—the special case in which '$\ll$' binds a single variable—but this is only to keep things simple.)

Some notation: a predicate will be said to be *constitutive* just in case it expresses a constitutive property. Also, since $\ulcorner a = b\urcorner$ can be paraphrased as $\ulcorner x = a \ll_x x = b\urcorner$, and since $\ulcorner \phi(\vec{x}) \equiv_{\vec{x}} \psi(\vec{x})\urcorner$ can be paraphrased as the conjunction of $\ulcorner \phi(\vec{x}) \ll_{\vec{x}} \psi(\vec{x})\urcorner$ and $\ulcorner \psi(\vec{x}) \ll_{\vec{x}} \phi(\vec{x})\urcorner$, I will sometimes say 'semi-identity statements' instead of 'identity and semi-identity statements'. Finally, I shall use 'semi-identity statement' to cover both conditional and unconditional semi-identity statements.

6 Possibility

In this section I will suggest a link between possibility and semi-identity. As before, the notion of possibility I have in mind is what is sometimes called 'metaphysical possibility'.

6.1 The link

I would like to suggest that one possibility and semi-identity are related by the following two principles:

Principle of Vindication

Every true semi-identity statement should be *vindicated*, in the following sense:

1. if $\ulcorner \phi(\vec{x}) \ll_{\vec{x}} \psi(\vec{x})\urcorner$ is true, then so is $\ulcorner \Box(\forall \vec{x}(\phi(\vec{x}) \rightarrow \psi(\vec{x})))\urcorner$;
2. if $\ulcorner \dfrac{\phi(z, \vec{w})}{x = z \ll_x \phi(x, \vec{w})} \urcorner$ is true, then so is

$$\Box(\forall z \forall \vec{w}(\phi(z, \vec{w}) \rightarrow \Box(\exists y(y = z) \rightarrow \phi(z, \vec{w}))))$$

(*read:* necessarily, if something bears ϕ to $\vec{w}$, then it couldn't exist without bearing ϕ to $\vec{w}$).

PRINCIPLE OF MAXIMALITY
There are no limits on possibility beyond the constraint that every true semi-identity statement be vindicated.

The Principle of Maximality is in need of further elucidation.

6.2 The Principle of Maximality

There is a procedure for cashing out the Principle of Maximality that immediately suggests itself. Start with a space of logically consistent 'words'. (I find it useful to think of a world as a *de re* story containing every sentence of the language or its negation, but you may think of worlds differently if you like.) Go on to eliminate just enough worlds to get the result that every true semi-identity statement is vindicated. The worlds you're left with are the worlds that depict genuine possibilities.

When one restricts one's attention to the special case in which conditional semi-identities are ignored, this procedure is well-defined. For given a space of worlds W and an arbitrary set of *unconditional* semi-identity statements S, there is a unique maximal subset of W that vindicates every member of S. It is the result of eliminating from W all and only worlds that fail to satisfy $\ulcorner \forall \vec{x}(F(\vec{x}) \rightarrow G(\vec{x})) \urcorner$ for some $\ulcorner F(\vec{x}) \ll_{\vec{x}} G(\vec{x}) \urcorner$ in S.

Unfortunately, things get messier when conditional semi-identities are brought into the picture. To see this, suppose W fails to vindicate some conditional semi-identity statement—say it is a semi-identity statement to the effect that F-ness is categorical. Then W must contain a *counterexample* to the categoricity of F-ness: there must be worlds w and w' in W such that according to w there is something that is an F, and according to w' that very individual is not an F. The problem is that there is more than one way to get rid of the counterexample. One could do so either by eliminating w or by eliminating w'. So it is not immediately clear how one should go about eliminating 'just enough worlds to get the result that every true semi-identity statement is vindicated'.

Getting around the problem is not as straightforward as one might have hoped. It will not do, for example, to cash out the Principle of Maximality as the constraint that one's space of worlds be a *fixed point*: that it vindicate every true semi-identity statement and contain enough worlds for it to be impossible to change the truth-value of modal sentences by bringing in additional worlds without thereby falsifying a true semi-identy statement. The trouble is that the Principle of Maximality is meant to maximize the extent of *possibility*, and an increase in worlds can result in the *elimination* of possibilities—specifically, it can result in the elimination of possibilities concerning *de re* modal properties. Suppose, for example, one has a space of worlds according to which Lonely is true:

Lonely
$\Diamond(\exists x \Box(\exists y(y = x) \rightarrow \forall y(y = x)))$
(*Read:* it is possible that there be an essentially lonely object)

On an $S5$ semantics, one makes Lonely false by adding worlds in which its witnesses have companions.

Fortunately, there turns out to be a precise way of cashing out the Principle of Maximality. It is based on the following result:

Canonicity
Let T be the set of true sentences in some first-order language L, and let S be the set of true (conditional or unconditional) semi-identity statements in the language $L^{\lll}$ that results from enriching L with '$\lll$'. Then there is a canonical way of generating a space of worlds on the basis of T and S. (See appendix for details.)

With Canonicity in place, the Principle of Maximality can cashed out as follows:

The Main Thesis
Let $L^{\Diamond}$ be the result of enriching L with the modal operators '$\Box$' and '$\Diamond$'. Then a sentence of $L^{\Diamond}$ is true just in case it is counted as true by the Kripke-model based on the canonical space of worlds generated by S and T.[5]

In preceding sections I offered a story about how the truth of semi-identity statements ought to be ascertained: pick whichever set of semi-identity statements supplies the most congenial articulation of your methods of inquiry. If one accepts the Main Thesis, this can be extended to a story about how the truth of *modal* statements ought to be ascertained: pick whichever set of semi-identity statements supplies the most congenial articulation of your methods of inquiry; then use the Main Thesis to determine which sentences of $L^{\Diamond}$ are true.

6.3 The list of modal truths

Suppose the Main Thesis is true. Which modal sentences turn out to be true?

1. The use of a Kripke-semantics guarantees that one gets a normal modal system (and, in particular, that one gets every theorem of classical logic).

   Also, the canonical space of worlds is based on a trivial accessibility relation. So one gets every theorem of $S5$.

[5] The actual world of the Kripke-model is the world which counts every sentence in T as true and has a domain consisting of objects in the domain of L.

2. The canonical space of worlds vindicates the set of true semi-identity statements. So one gets $\ulcorner\Box(\forall\vec{x}(\phi(\vec{x}) \to \psi(\vec{x})))\urcorner$ whenever $\ulcorner\phi(\vec{x}) \ll_{\vec{x}} \psi(\vec{x})\urcorner$ is true, and one gets $\ulcorner\Box(\forall z\forall\vec{w}(\phi(z,\vec{w}) \to \Box(\exists y(y = z) \to \phi(z,\vec{w}))))\urcorner$ whenever $\phi(z,\vec{w})$ is constitutive.

   For instance, by assuming that 'Elephant$(x) \ll_x$ Mammal(x)' is true one gets the result that '$\Box(\forall x($Elephant$(x) \to$ Mammal$(x)))$' is true, and by assuming that 'Human(x)' is constitutive one gets the result that (an $L^{\Diamond}$ rendering of) 'necessarily, humans are essentially human' is true.

   As far as I can tell, any first-order modal sentence that constitutes a relatively uncontroversial example of a metaphysical necessity can be recovered from suitable semi-identity statements in this sort of way.[6] (Fgures 2–4 list some additional examples.)

3. One can show that any possibility statement in a certain syntactically demarcated class P is verified by the canonical space of worlds if and only if it is consistent with a set of 'core' modal truths. (Very roughly, the core modal truths are consequences of the set of true semi-identities, and P consists of sentences that begin with a diamond and are subject to the constraint that whenever one quantifies into a box, one is allowed no further quantifying-in within the scope of the box. Proofs and details are supplied in the appendix.)

   As far as I can tell, any reasonably simple sentence of $L^{\Diamond}$ that one would intuitively count as a possibility statement will be a member of P. For instance, $L^{\Diamond}$-renderings of 'I might have had a sister', 'I might have had a sister who was a cellist but might have been a philosopher' and 'there might have been an essentially lonely object' are all members of P. They will therefore be counted as true, provided they are consistent with the core modal truths. (Figure 5 lists some additional examples.)

There is no official catalogue of recognized modal truths. But I hope I have done enough to convince you that a proponent of the Main Thesis who accepts the right semi-identity statements will end up with list of modal truths which is roughly in line with the standard literature on metaphysical possibility.

6.4 Supervenience

The Main Thesis delivers a pleasing conclusion: once the truth-value of every semi-identity statement in $L^{\ll}$ is fixed, determining the truth-value of every sentence in L is enough to determine the truth-value of every sentence in $L^{\Diamond}$. Put in a different way: if one accepts the Main Thesis and assumes a fixed set of true semi-identities, Canonicity can be thought of as a *supervenience theorem* for modal languages.

Consider an example. What is required of the world in order for the truth conditions of the following sentence to be satisfied?

[6] For present purposes one might count a true first-order sentence as a 'metaphysical necessity' just in case it is equivalent to a sentence in normal form containing no diamonds. (See appendix for a characterization of normal form.)

Modal
$\exists x(\textsc{Mammal}(x) \wedge \diamond(\textsc{Human}(x)))$

(*Read:* something is a mammal and might have been a human.)

It is clear that something or other is required of the world, since whether or not there are mammals who might have been human depends on what the world is like. (As it happens, there *are* mammals who might have been humans, since there are humans, and every human is a mammal who might have been a human. But had there been no humans, nothing would have been a mammal who might have been a human—at least on the assumption that nothing could be a human without being essentially human.)

A perfectly accurate way of specifying truth-conditions for Modal is by stating that what is required of the world in order of Modal's truth-conditions to be satisfied is that there be something that is a mammal and might have been a human. But suppose one is aiming for more than mere accuracy. One wants one's specification to take the following form:

> What is required of the world in order for Modal's truth-conditions to be satisfied is that it be such that p.

where 'p' is replaced by a sentence containing no modal operators. Then one might reason as follows:

> Being a non-human is constitutive property: if you're a non-human, part of what it is to be you is to be a non-human. So the requirement that the world be such that there is a mammal that might have been a human boils down to the requirement that the world be such that there is a mammal that is also a human. But part of what it is to be a human is to be a mammal. So the requirement that the world be such that there is a mammal that is also a human boils down to the requirement that the world be such that there are humans.
>
> Accordingly, all that it is required in order for Modal's truth-conditions to be satisfied is for the world to be such that there are humans—and this gives us what we want, since 'there are humans' contains no modal operators.[7]

If the Main Thesis is true, it is a consequence of Canonicity that an analogous result holds for arbitrary modal sentences. More specifically, for ϕ an arbitrary sentence of $L^{\diamond}$, the truth-conditions of ϕ are correctly specified by some clause of the form:

> What is required of the world in order for ϕ's truth-conditions to be satisfied is that it be such that p.

[7]A corollary of this result is that the world is not required to contain non-actual possibilia in order for the truth-conditions of Modal to be satisfied—unless, of course, the world is required to contain non-actual possibilia in order for the truth-conditions of 'there are humans' to be satisfied.

where 'p' is replaced by a (possibly infinite) sentence built out of the vocabulary of L.[8]

It would be hasty to conclude from this that the modal can be reduced to the non-modal, since a predicate like 'human' is presumably not devoid of modal content. But one does get the result that the truth-conditions of sentence involving boxes and diamonds can be specified without using boxes or diamonds (and without using '$\ll$').

6.5 Intelligibility and Possibility

Suppose the Main Thesis is true. How is possibility related to intelligibility?

In section 2.1 I suggested a sufficient condition for the unintelligibility of *de re* stories: when '$F(x) \ll_x G(x)$' is taken for granted, a *de re* story that says something that is assumed to entail '$\exists x(F(x) \wedge \neg G(x))$' will be treated as unintelligible. For the special case of an *omniscient* agent—an agent who is omniscient about logical consequence and about the truth of semi-identity statements—this condition can be restated as follows: a *de re* story that says something incompatible with the set of true semi-identity statements will be treated as unintelligible.

Say that a *de re* story is *complete* (relative to a first-order language L) if it contains every sentence in L or its negation. Then it is a consequence of the Main Thesis that a complete *de re* story depicts a metaphysically impossible situation only if an omniscient agent would take it to be unintelligible.

Is the converse true? Whether one thinks so will depend on whether one thinks that an omniscient agent could treat a *de re* story as unintelligible even if isn't incompatible with the set of true semi-identity statements. I doubt that our pretheoretic notion of intelligibility is robust enough to allow for an adequate assessment of this question. But it won't hurt to introduce some terminology. Say that a sentence is *strongly unintelligible* if it can be ruled out as unintelligible on the basis of the sufficient condition on unintelligibility mentioned above. Then it is a consequence of the Main Thesis that a complete *de re* story will depict a metaphysically impossible situation just in case an an omniscient agent would take it to be strongly unintelligible.

In this qualified sense, the present proposal may be described as an account of possibility according to which the limits of possibility are simply the limits of intelligibility.

6.6 Why-Closure and Necessity

Suppose the Main Thesis is true. How is necessity related to why-closure?

In section 2.2 I suggested a sufficient condition for the why-closure of first-order sentences: when '$F(x) \ll_x G(x)$' is taken for granted, any sentence '$\forall x(F(x) \rightarrow G(x))$' is

[8] Let a complete L-theory be a maximally consistent set of sentences of L. Let S be the set of true semi-identity statements of $L^{\ll}$. Then a sentence guaranteed to do the job is the (possibly infinite) disjunction of (infinite) sentences, each of which consists of the (infinite) conjunction of sentences in a complete L-theory T such that that ϕ is true according to the Kripke-model based on the space of worlds generated by S and T.

assumed to entail is treated as why-closed. For the special case of an *omniscient* agent—an agent who is omniscient about logical consequence and about the truth of semi-identity statements—this condition can be restated as follows: any sentence entailed by a true semi-identity statement will be treated as why-closed.

It is consequence of the Main Thesis that a first-order sentence is metaphysically necessary only if it is why-closed.

Is the converse true? Whether one thinks so will depend on whether one thinks that a first-order sentence can be why-closed even if isn't entailed by the set of true semi-identity statements. As before, I doubt that that the issue can be settled by reflecting on intuitions. But, as before, it won't hurt to introduce some terminology. Say that a sentence is *weakly why-open* if it cannot be shown to be why-closed on the basis of the sufficient condition on why-closure mentioned above. Then it is a consequence of the Main Thesis that first-order sentence will be metaphysically possible just in case it is weakly why-open.

6.7 Limitations of the Main Thesis

The bulk of this section has been devoted to exploring the modal landscape that one gets if one accepts the Main Thesis. But is the Main Thesis true?

In section 6.3 I noted that the Main Thesis allows one to extract a fairly orthodox set of modal truths from a suitable set of semi-identity statements. In section 6.4 I pointed out that when the Main Thesis is in place, one gets a supervenience theorem for modal languages. In sections 6.5 and 6.6 I observed that the Main Thesis delivers robust connections between intelligibility, possibility and why-closure. All of these might be seen as points in its favor. But it is important to be clear about the Thesis's limitations. The following strike me as especially significant:

1. Is there any reason to think that we couldn't find the need to impose limits on metaphysical possibility too complex to be articulated by a set of semi-identity statements?

   Consider the following sentence as an example:

   > NEMESES
   > $\Diamond(\exists x \Diamond(\exists y \Box(x \neq x \vee y \neq y)))$
   > (*Read:* There might have been *incompatible* objects: objects each of which might have existed but such that they couldn't have existed together.)

   In the absence of specific reasons to the contrary, I'm inclined to count Nemeses as true. For a *de re* story that says '$\exists x \Diamond(\exists y \Box(x \neq x \vee y \neq y))$' strikes me as intelligible, and in the absence of specific reasons to the contrary I'm inclined to treat intelligible *de re* stories as describing metaphysically possible scenarios. But I could certainly be convinced otherwise. And treating Nemeses as false would lead to trouble. For—unless the world is finite—Nemeses cannot be ruled out on the basis of semi-identity statements alone: its truth is guaranteed by the Main Thesis.

2. Suppose one agrees that possibility is limited only by semi-identity. One might still worry that the some of the semi-identities that are needed to assign the right truth-values to sentences of $L^{\Diamond}$ are not expressible in L. This would spell trouble for the Main Thesis because the canonical space of worlds is built from the set of true semi-identity statements that are expressible in L. As far as I can tell, the problem doesn't arise in practice. But one would like one's assurances to be more solid than that.

3. So far, I have been focussing exclusively on first-order languages. But suppose one wanted to articulate a version of the Main Thesis in the context of a more expressive language. One may wish to introduce second-order variables, for instance, and allow semi-identity statements of the form '$\phi(X) \ll_X \psi(X)$'.[9] The formal framework I developed in this section won't automatically carry over to the new setting.

For these reasons, I don't want to put too much weight on the Main Thesis as it is currently formulated. What I do want to put weight on is the general methodology of the paper: the idea that one can place constraints on the notion of metaphysical possibility by acquiring clear grasp of the sorts of considerations that go into determining the truth values of a set of particularly tractable modal sentences, and using these sentences to settle the truth values of other modal sentences in a principled way. The Main Thesis can be thought of as supplying a 'proof of concept' for this general methodology.

7 Mathematics

At the beginning of section 2 I asked you to ignore mathematical sentences. Now is the time to address them.

7.1 Logical Vocabulary

I shall begin by considering a warm-up case: logical vocabulary. I will proceed by making three related observations. If it is not transparent why I have chosen to focus on these particular points, please bear with me. They are meant to supply the groundwork for the following section. My discussion will take for granted that one is prepared to use classical logic in one's metalanguage, and that one favors a classical semantics for one's object-language.

Observation 1: Emptiness

Say that a sentence's truth-conditions are *empty* just in case they can be shown to be satisfied by a world w without presupposing anything specific about w—anything beyond one's general framework for theorizing about worlds. (In particular, the sentence's truth

[9] A language including '$\ll_X$' could be used to capture Kit Fine's '$\Box_F A$'. (See Fine (1995a), Fine (2000); see also Fine (1994) and Fine (1995b).) For when second-order quantification is cashed out in terns of plural quantification, '$\Box_F A$' can be treated as an abbreviation of '$\forall z(Xz \leftrightarrow Fx) \ll_X A[X/F]$'.

conditions can be shown to be satisfied without presupposing anything about the domain of w or about the the extensions of predicates in w.)

The truth-conditions of '$\exists x\,\text{Elephant}(x)$' are non-empty. For in order to show that they are satisfied by w one needs to know whether the following is true:

$[\exists x\,\text{Elephant}(x)]_w$

(*Read:* at w, there are elephants)

And one cannot know whether this is true in the absence of information about the extension of 'Elephant' in w. In contrast, '$\exists x\,\text{Elephant}(x) \vee \neg\exists x\,\text{Elephant}(x)$' has empty truth-conditions. For on standard semantic assumptions, we have each of the following:

- the truth-conditions of $\ulcorner\phi \vee \psi\urcorner$ are satisfied by w just in case either the truth-condtions of ϕ are satisfied by w or the truth conditions of ψ are satisfied by w;
- the truth-conditions of $\ulcorner\neg\phi\urcorner$ are satisfied by w just in case the truth-conditions of ϕ fail to be satisfied by w.

So the truth-conditions of '$\exists x\,\text{Elephant}(x) \vee \neg\exists x\,\text{Elephant}(x)$' will be satisfied by w just in case the following holds:

$[\exists x\,\text{Elephant}(x)]_w \vee \neg[\exists x\,\text{Elephant}(x)]_w$

(*Read:* either, at w, there are elephants, or it is not the case that, at w, there are elephants)

And since this sentence is of the form '$p\vee\neg p$', one can establish its truth without presupposing anything specific about w. It is enough to deploy classical logic in one's metalanguage.

The point generalizes: any valid sentence of a negative free logic (hereafter 'logical truth') can be shown to have empty truth-conditions. For by presupposing a Kripke-style semantics for worlds and making standard assumptions about the meaning of the logical vocabulary, one can show that each of the following holds:[10]

- $[\phi \vee \psi]_w \leftrightarrow ([\phi]_w \vee [\psi]_w)$
- $[\neg\phi]_w \leftrightarrow \neg[\phi]_w$
- $[\exists x\,\phi(x)] \leftrightarrow \exists x([\exists y(y = \dot{x})]_w \wedge [\phi(\dot{x})]_w)$
- $[\dot{x} = \dot{y}]_w \leftrightarrow ([\exists y(y = \dot{x})]_w \wedge x = y)$
- $x = y \rightarrow ([\phi(\dot{x})]_w \rightarrow [\phi(\dot{y})]_w)$
- $[P(\dot{x}_1, \ldots, \dot{x}_n)]_w \rightarrow ([\exists y(y = \dot{x}_1)]_w \wedge \ldots \wedge [\exists y(y = \dot{x}_n)]_w)$ (for P atomic)

[10]The dotted variables are introduced to deal with the fact that the domain of w might contain objects that don't actually exist. (See my 'An Actualist's Guide to Quantifying-In' for details.)

And when ϕ is a logical truth, these principles can be used to prove $\ulcorner [\phi]_w \urcorner$ without presupposing anything specific about w. All one needs to do is deploy classical logic in the metalanguage.

The observation that any logical truth has empty truth-conditions is open to two potential sources of misunderstanding. Firstly, one might take it to entail the conclusion that logical truths all have the same meaning, or that they all express the same proposition. To see that there is no such entailment note that one might have a fine-grained conception of meanings or propositions without believing that the fine-grainedness is to be cashed out in terms of what needs to be presupposed about w in order to show that the truth-conditions of the relevant sentence are satisfied at w.

The second potential source of misunderstanding is that the idea that one can get from the claim that every logical truth has empty truth-conditions to the conclusion that coming to know of a logical truth that it is true should be a trivial matter. But the conclusion only follows on the assumption that determining whether a sentence has empty truth-conditions was itself a trivial affair. And the assumption fails. Note, for example, that working out that sentences of the form '$\exists x(\phi(x) \supset \psi) \equiv (\forall x(\phi(x)) \supset \psi)$' are logically true is a non-trivial affair—to say nothing of, e.g. '$DP \supset GC$' (where 'DP' abbreviates the conjunction of an interesting subset of the first-order Dedekind-Peano Axioms and 'GC' abbreviates is a first-order statement of Goldbach's Conjecture). Learning of logical truths that they are true can be a non-trivial matter. But this is *not* because it is somehow a non-trivial matter to work out whether the world satisfies an empty requirement. It is because it can be a non-trivial matter to work out whether a sentence has empty truth-conditions.

Observation 2: Why-Closure

Knowing that a sentence has empty truth-conditions is not the same as understanding *why* it has empty truth-conditions. '$DP \supset$ there is no largest prime', for example, is logically true. But someone who knows that it is a logical truth on the basis of testimony need not understand why it is a logical truth. More interestingly, someone might have succeeded in following a proof of '$DP \supset$ there is no largest prime' and still have doubts which would be aptly expressed by saying 'I can see that it is a logical truth, but I still don't understand *why* it is a logical truth.' (Perhaps she didn't fully understand the proof, or she found that it wasn't sufficiently illuminating.) Such a person understands the meanings of lexical items in '$DP \supset$ there is no largest prime' and knows how they should be combined to generate truth-conditions. Since she knows that the sentence is a logical truth, she can see that the resulting truth-conditions will be empty. But she is unable to see *why* this is so. If she uses a why-question to voice this sort of concern, I shall say that the question should be understood as a *computational* question.

In general, when someone asks $\ulcorner$why is it the case that ϕ?$\urcorner$ she might be making requests of different kinds. In the most common case, the subject is clear about what would be required in order for the truth-conditions of ϕ to be satisfied, and wants to understand why the world is such as to satisfy this requirement. But someone might ask $\ulcorner$why is it the case that ϕ?$\urcorner$ and intend it as a computational question about ϕ, rather than as a question about ϕ's subject-matter: she might be able to see that that the meanings of ϕ's lexical items conspire to deliver the truth-conditions that ϕ actually has, but not fully understand why this is so.

Suppose someone asks: "why is it the case that ($DP \supset$ there is no largest prime)?" One can make sense of the question by reading it as a computational question. Accordingly, one might address it by offering a proof of '$DP \supset$ there is no largest prime', and making sure that the proof is illuminating and fully understood. But notice the following. It would be hard to make sense of one's interlocutor if she went on to reply: "I know full well that the truth-conditions of '$DP \supset$ there is no largest prime' are empty, and I fully understand why this is so. In addition, I can see as clearly as can be that nothing is required of the world in order for a sentence with empty truth-conditions to be true. What I want to know is *why* it is the case that ($DP \supset$ there is no largest prime)!"

Earlier I said that 'why is it the case that q?' is read as a grounding question when it is read either as requesting an elucidation of what it is for it to be the case that q or as requesting a justification for the claim that q (or both). I would now like to extend the definition as follows: $\ulcorner$why is it the case that ϕ?$\urcorner$ is read as a *grounding question* just in case it is read as either (a) requesting an elucidation of why ϕ has the truth-conditions that it has, (b) requesting an elucidation of what satisfying these truth-conditions would consist in, (c) requesting a justification for the truth of ϕ, or (d) some combination thereof. (As before, I say that ϕ is treated as *why-closed* if one is unable to make sense of $\ulcorner$why is it the case that ϕ?$\urcorner$ unless one reads it as a grounding question.)

With the updated terminology in place, I can set forth the following claim: when ϕ is a logical truth, one is unable to make sense of $\ulcorner$why is it the case that ϕ?$\urcorner$, unless one reads it as a grounding question. And if one does read it as a grounding question, one will often read it as a computational question about ϕ.

Observation 3: Intelligibility

Could one make sense of a story depicting a situation in which the truth-conditions of a logical truth fail to be satisfied? I suggest not. For such a story would have to depict a situation in which empty truth-conditions fail to be satisfied. But empty truth-conditions are trivially satisfied, so it would have to be a situation in which the relevant truth-conditions are both satisfied and unsatisfied, which is unintelligible.

Of course, if one is mistaken about ϕ's truth-conditions, one might believe that one is able to make sense of a story at which ϕ's truth-conditions fail to be satisfied, even if ϕ is, in fact, a logical truth. Say you mistakenly believe that '$DP \supset$ there is no largest prime' is false. Then you might also believe that you are able to make sense of a story that says '$\neg(DP \supset$ there is no largest prime)'. But if you learn that '$\neg(DP \supset$ there is no largest prime)' is a logical truth you will come to see the story as unintelligible. (Compare: someone who thinks that 'Mark Twain is Samuel Clemens' is false might believe that she is able to make sense of a *de re* story that says 'Mark Twain is not Samuel Clemens'. But if she learns that 'Mark Twain is Samuel Clemens' is true, she will come to see the story as unintelligible.)

The Upshot

The following sentences are both counted as metaphysical necessities by the Main Thesis:

$$\forall x(\text{Human}(x) \rightarrow \text{Mammal}(x)) \qquad\qquad \forall x(\text{Human}(x) \rightarrow \text{Human}(x))$$

But the sentences differ in an important respect: whereas the second has empty truth-conditions, the first does not.[11]

In order to get the result that the first sentence is true at every world in the canonical space, one needs to make sure that any worlds that fail to satisfy its (non-empty) truth-conditions are ruled out. In particular, one needs to rule out any worlds in which the extension of 'Human' is not a subset of the extension of 'Mammal'. One strategy for excluding the relevant worlds—the one I proposed in section 6—is to make a substantial assumption about metaphysical possibility: that any world incompatible with a true semi-identity statement should be counted as metaphysically impossible.

No assumptions of this kind are needed to ensure that the second sentence is true at every world in the canonical space. For its truth-conditoins are empty, and will be satisfied at a world regardless of what the world is like. This is *not* to say that one can get the result that the second sentence is true at every world without making substantial assumptions. I have made free use of classical logic in the metalanguage, presupposed a Kripke-style framework for theorizing about worlds and made classical assumptions about the meaning of the logical vocabulary. The point is that once these basic logico-semantic assumptions are in place, nothing *further* needs to be assumed about metaphysical possibility in order to get the result that sentences with empty truth-conditions are true at every possible world. There is no need to defend a special principle ruling out 'non-logical' worlds as metaphysically impossible. This is the main the take-home lesson of the present section, and will be important in the discussion to follow.

In sections 6.5 and 6.6 I suggested that there are tight links between possibility and intelligibility, and between necessity and why-closure. A secondary take-home lesson is that these links are not threatened by the claim that the modal status of sentences with empty and non-empty truth-conditions is accounted for in different ways. As illustrated by the case of logical truth, the connection between possibility and intelligibility is not threatened because one cannot make sense of a story depicting a situation in which empty truth-conditions fail to be satisfied, and the connection between necessity and why-closure is not threatened because sentences with empty truth-conditions are why-closed.

7.2 Arithmetic

The purpose of this section is to explain how a proponent of the account of possibility developed in earlier sections might secure the result that arithmetical truths are necessary.

Platonism is the view that the world contains numbers. On the most straightforward understanding of Platonism, there should be no problem making sense of a *de re* story

[11] It does, however, have what I elsewhere call *trivial* truth-conditions. See 'On Specifying Truth-Conditions'.

that says 'there are no numbers'. ("When God made the world, she began by creating spatiotemporal objects. She was tempted to go on to create abstract objects. But she changed her mind at the last minute. So numbers were never created.") The straightforward Platonist should not be inclined to treat 'there are numbers' as why-closed. ("Why are there numbers?" you ask. "Because after creating the concrete world, God went on to create numbers" someone might reply—a bad answer to be sure, but an answer all the same.) Accordingly, the straightforward Platonist can only claim that numbers exist necessarily by severing the connection between possibility and intelligibility ('there are numbers' is necessary, even though one can make sense of a *de re* that says 'there are no numbers'), and severing the connection between necessity and why-closure ('there are numbers' is necessary even though it is not why-closed).

There is, however, a subtler version of Platonism.[12] Suppose one thinks that for the number of the planets to be eight *just is* for there to be eight planets:

$$\text{The number of the planets} = 8 \ \equiv \ \exists!_8 x(\text{Planet}(x))$$

More generally, suppose one accepts every instance of the following schema:

$$\text{The number of the Fs} = n \ \equiv \ \exists!_n x(\text{F}(x))$$

Then one should find a *de re* story that says 'there are no numbers' unintelligible. (Trivially, it will be a story according to which there are no nonselfidentical things. But according to the subtle Platonist that is what it is for the number of the nonselfidentical things to be zero. So it must be a story according to which there are numbers. If it is also a story that says 'there are no numbers', it must be a story according to which there are numbers and there are no numbers, which is unintelligible.) I submit, moreover, that the subtle Platonist should be inclined to think that 'there are numbers' is why-closed. If this is right, then the subtle Platonist is in a position to claim that numbers exist necessarily without severing the connection between possibility and intelligibility, or the connection between necessity and why-closure.

It seems to me that subtle Platonism has a lot to recommend it, but two points are worth emphasizing. First, taking arbitrary instances of the above schema to be true is not enough to guarantee that every sentence of modal arithmetic gets the right truth-value. (In particular, one won't get every instance of the induction schema to count as necessarily true.) The problem can be remedied by supplementing the schema with a second-order statement such as the following:[13]

[12] Neofregeans might be read as endorsing a version of this view. See Wright (1983) and Hale and Wright (2001)—and, of course, Frege (1884).

[13] An alternative is to forego the schema altogether and set forth the following version of Hume's Principle (along with suitable definitions):

> the number of the Fs = the number of the Gs $\equiv_{F,G}$ the Fs are in one-one correspondence with the Gs.

Number
x is a number $\equiv_x \forall X((X(0) \wedge \forall y(X(y) \rightarrow X(y'))) \rightarrow X(x))$
(*read:* what it is to be a number is to be amongst any things that include zero and are closed under the successor operation).

So as long as subtle Platonists are prepared to use higher-order resources, all will be well. Without higher-order resources, however, matters become trickier. One can gain some ground without going beyond first-order languages by claiming that arbitrary instances of the following schema are true:

$$\top \ll \phi$$

where '$\top$' abbreviates a tautology and 'ϕ' is replaced by a true first-order sentence in the language of pure arithmetic. But the combined strength of the two schemas is still not enough to guarantee that every sentence of modal arithmetic gets the right truth-value.[14]

The second point worth emphasizing is that an identity-statement like 'for the number of the planets to be eight *just is* for there to be eight planets' is significantly different from an identity-statement like 'being water *just is* being H_2O'. For it relies on the idea that a realm of objects can be *thin*, in the sense that its existence and properties consist of no more than the existence and properties of a different realm objects. I am not myself troubled by this idea, but it would be disingenuous not to expect a certain amount of skepticism.

Fortunately, friends of the account of modality defended in earlier sections can use a different method to ensure that every sentence of modal arithmetic gets the right truth-value. Platonists are usually *committalists*: they believe that the world must contain numbers in order for the truth-conditions of typical arithmetical assertions to be satisfied. But committalism is not forced upon us. As I show in 'On Specifying Truth-Conditions', it is possible to specify a compositional assignment of truth-conditions to arithmetical sentences which yields the result that the world needn't contain numbers in order for the truth-conditions of typical arithmetical sentences to be satisfied. In particular, all that is required in order for the truth-conditions of 'the number of the planets is 8' to be satisfied is that there be eight planets, and *nothing* is required is required in order for the truth-conditions of 'there are infinitely many primes' to be satisfied.

When the noncommittalist assignment of truth-conditions is in place, the truth-conditions of any consequence of the axioms of (second-order) pure and applied arithmetic turn out to be *empty*. So in much the way that logical truths were shown to be metaphysically necessary in section 7.1 without setting forth a principle to rule out 'non-logical' worlds, the non-committalist can secure the result that any consequence of the axioms is metaphysically necessary without setting forth a principle to rule out 'non-mathematical' worlds. For an arithmetical truth's empty truth-conditions will be satisfied at a world regardless of what the world is like. This is *not* to say that noncommittalists make no substantial

[14]Suppose, for example, that each of the following is true: '$F(x) \ll_x \exists_1 y F(y)$', '$F(x) \ll_x \exists_2 y F(y)$', ... Then '$\Diamond(\exists x(\text{Number}(x) \wedge x$ is the number of the Fs))' should count as false, since there must be infinitely many Fs. But—in the absence of further constraints—it will be counted as true by the firstorderist.

assumptions. They make use of arithmetical reasoning in the metalanguage, presupposes a Kripke-style framework for theorizing about worlds and makes classical assumptions about the meaning of the logical vocabulary. The point is that—unlike their committalist rivals—noncommittalists need assume nothing *further* about metaphysical possibility once these basic logico-semantic assumptions are in place.

The noncommittalist line does not threaten the connection between possibility and intelligibility or the connection between necessity and why-closure. For we have seen that one cannot make sense of a story depicting a situation in which empty truth-conditions fail to be satisfied, and that sentences with empty truth-conditions are why-closed. When someone asks ⌜why is it the case that ϕ?⌝ for ϕ a truth of pure arithmetic, the question should typically be read as a computational question about ϕ, rather than as a question about ϕ's subject-matter.

Disagreement between a noncommittalist and a subtle Platonist who is also a committalist is more elusive than one might think. Notice, in particular, that they both agree that all it takes for the truth-conditions of, e.g. 'the number of the planets is 8' to be satisfied is that there be eight planets. One might reply that the views are nonetheless quite different. For whereas the subtle Platonist gets the result by relying on a standard semantics and setting forth a controversial identity statement—what it is for the number of the planets to be 8 is for there to be eight planets—the noncommittalist gets the result by relying on a non-standard semantics. But the extent to which there is a deep disagreement here, rather than a difference in bookkeeping, is not entirely clear to me. Fortunately, these are not issues that need to be resolved for present purposes.

7.3 Beyond Arithmetic

The noncommittalist line is not confined to the special case of arithmetic. Similar techniques can be used to account for the metaphysical necessity of other sentences involving abstract-object-vocabulary. It can certainly be done in the case of set-theory, higher-order quantification, and abstract-object-talk that is governed by abstraction principles (e.g. "A's age = B's age just in case A and B were born simultaneously'). But it is important to be clear that the availability of a noncommittalist line is not automatic. One can only develop a noncommittalist treatment for the sentences in a given class if one has succeeded in constructing the right sort of compositional semantics for those sentences. And this need not be a trivial task.[15]

8 Metaphysical Issues

We have been assuming so far that Realism is true: that there is good sense to be made of the question whether a semi-identity statement is *true*, over and above the question of whether it yields a congenial articulation of our methods of inquiry. The purpose of this section is to consider an alternative.

[15]For an extended treatment of such issues, see my 'On Specifying Truth Conditions'.

I would like to begin by contrasting three metaphysical myths:

1. *The Three-Day Myth*

   On the first day, God created space-time. On the second day, she decided which 'basic' properties were to be instantiated at each space-time region. "The total mass in this space-time region is to be so-and-so"—she commanded—"and the total charge in that one is to be thus-and-such." On the third day God rested.

2. *The Four-Day Myth*

   The first two days were as in the Three-Day Myth. But on the third day God took it upon herself to divide the world into objects. She said: "This mass-laden space-time region here is to be exactly occupied by an object; but not that one there—that space-time region has a subregion which is to be exactly occupied by an object and a subregion which fails to overlap any region which is to be exactly occupied by an object." (If God is a lover of promiscuous mereologies, she might go on to decree that every subregion of a region exactly occupied by an object is to be exactly occupied by an object; if she is a stickler for non-colocation, she might go on to decree that no space-time region is to be exactly occupied by more than one object.) On the fourth day God rested.

3. *The Five-Day Myth*

   The first three days were as in the Four-Day Myth. But on the fourth day God took it upon herself to endow her objects with *essences*. She proceeded in three stages. First she introduced a system of 'categories', and partially ordered them. "The category ELEPHANT"—she declared—"is to be a subcategory of the category MAMMAL, which is in turn is to be a subcategory of the category ANIMAL". Next, she decided which categories were to be *constitutive* of their bearers. "Falling under ELEPHANT"—she said—"is to be constitutive of its bearers: on the assumption that something falls under ELEPHANT, part of what it is to be that thing is to fall under ELEPHANT. GRAY, on the other hand, is not to be constitutive of its bearers: the assumption that something falls under GRAY is not enough to guarantee that part of what it is to be that thing is to fall under GRAY." Finally, God specified what it takes for an object to fall under each of the categories. "What it takes for an object to fall under the category GRAY"—she decreed—"is for it to exactly occupy a space-time region in which such-and-such 'basic' properties are instantiated. What it takes for an object to fall under ELEPHANT"—she continued—"is slightly more complicated. The object is to be the exact occupant of a space-time region in which such-and-such 'basic' properties are instantiated *and* be a member of a certain 'lineage': a finite sequence of objects starting with *that* object there, and such that consecutive members of the sequence are the exact occupants of regions in which thus-and-such 'basic' properties are instantiated and with thus-and-such relative space-time coordinates." Having done all that, God turned to a large, gray, elephant-shaped object, and announced: "you satisfy all the conditions for falling under ELEPHANT; so from this day forward part of what it is to be you is to fall under ELEPHANT—and since ELEPHANT is a subcategory of MAMMAL it is also part of what it is to be you to fall under MAMMAL". By the fifth day, all of God's creatures had been assigned essences. Only then did she rest.

A proponent of the Five-Day Myth might suggest that Realism is to be spelled out as follows:

GROUNDING

What makes 'Human(x) $\ll_x$ Mammal(x)' true is that 'Human' refers to HUMAN, 'Mammal' refers to MAMMAL, and God has decreed that HUMAN is to be a subcategory of MAMMAL. Similarly, what makes 'Human(x)' constitutive, is that 'Human' refers to HUMAN, and God has decreed that HUMAN is to be constitutive.

This is presumably *not* the best way of spelling out Realism. But it can be used as a prop to spell out a non-realist position. Consider a proponent of the Three-Day Myth who says the following. "Proponents of the Four- and Five-Day Myths make the mistake of deifying our conceptual framework. It is not *God* who divides the world into objects and assigns the resulting objects essences: that is something *we* do when we deploy our concepts. Other than that, the Four- and Five-Day Myths are basically right. And Grounding is not too badly off the mark when 'our conceptual framework' is substituted for 'God'."

"Let me be clear"—she continues—"I am not proposing a form of idealism. Whatever concepts we deploy, the external world will remain as God left it after the second day. The claim is rather that when we *represent* the world, our representations are mediated by concepts, and that these concepts play the role that God's purported third and fourth days of work are supposed to play in Grounding. More specifically, the view is as follows. It makes no sense to say of a sentence (in context, as used by a given linguistic community) that it is true or false *simplicter*. A sentence can only be said to be true or false relative to a conceptual framework. A conceptual framework does three things. Firstly, it specifies which space-time regions are to be regarded as exactly occupied by objects, and how many objects are to be regarded as doing the occupying. (It might specify, for example, that this space-time region here is to be regarded as exactly occupied by an object; but not that one there.) Secondly, a conceptual framework specifies a partially ordered system of concepts, and determines which of them are to be regarded as constitutive. (It might specify, for example, that the concept *elephant* is to be regarded as constitutive, and that it is to be regarded as subordinate in the ordering to the concept *mammal*.) Finally, a conceptual framework assigns satisfaction-conditions to concepts. (It might specify, for example, that an object is to be regarded as satisfying the concept *gray* just in case it is regarded as occupying a space-time region that in fact instantiates such-and-such 'basic' properties; it might also specify that an object is regarded as satisfying the concept *elephant* just in case it is regarded as occupying a space-time region that in fact instantiates such-and-such 'basic' properties *and* is regarded as being a member of a certain 'lineage': a finite sequence of objects starting with *that* object there, and such that consecutive members of the sequence are the exact occupants of regions that in fact instantiate thus-and-such 'basic' properties and in fact have thus-and-such relative coordinates.)"

"Let $\mathcal{F}$ be a conceptual framework, and let an $\mathcal{F}$-concept be a concept in $\mathcal{F}$. What it takes in order for '$\exists x\, \text{Elephant}(x)$' to be true relative to $\mathcal{F}$ is for the instantiation of 'basic' properties across space-time regions to be such that one of the objects postulated by $\mathcal{F}$ counts as satisfying the $\mathcal{F}$-concept that is expressed by 'Elephant'. What it takes for '$\text{Elephant}(x) \ll_x \text{Mammal}(x)$' to be true relative to $\mathcal{F}$ is for the $\mathcal{F}$-concept that is expressed by 'Elephant' to be counted as subordinate to the $\mathcal{F}$-concept that is expressed by 'Mammal'. And what it takes for '$\text{Human}(x)$' to be constitutive is for the by $\mathcal{F}$-concept that is expressed by 'Human' to be counted as constitutive."

"What does it take for a speaker to adopt a conceptual framework? Oversimplifying a bit, the story is this. Someone adopts a conceptual framework when she adopts a language and uses it to theorize about the world. The concepts in the subject's framework correspond to predicates in her language. The satisfaction-conditions of concepts are determined on the basis of broadly Davidsonian considerations: while limited by constraints such as compositionality, one does one's best to get the result that the subject is by and large reasonable, and by and large a speaker of the truth. Which objects are postulated by the framework is also determined by broadly Davidsonian considerations: if the speaker is prepared to assert 'there is an elephant over there', then—other things being equal—the mass-laden space-time region over there will be counted as being occupied by an object; if the speaker refuses to assert 'whenever there are apples on the table there are apple-halfs on the table', then—other things being equal—the framework will deliver a conservative mereology; if she is prepared to assert 'the statue = the clay', then—other things being equal—the framework will exclude colocation. Finally, the ordering amongst concepts and the status of concepts with respect to constitution is determined by the articulation of her methods of inquiry she finds most congenial and, more specifically, by the semi-identity statements she accepts. If she accepts $\ulcorner F(x) \ll_x G(x) \urcorner$, then—other things being equal—the concept corresponding to F will be counted as subordinate in the ordering to the concept corresponding to G, and if she treats F as constitutive, then—other things being equal—the concept corresponding to F will be counted as constitutive. (It is *not* a consequence of the Davidsonian approach that speakers cannot be mistaken, since one will often have to choose between counting the speaker as reasonable and counting her a saying something true.)"

"The picture that emerges is one whereby a sentence can only said to be true relative to a conceptual framework, but *not* one whereby assertions lack truth-conditions. For an assertion is counted as true just in case the sentence asserted is true relative to the framework adopted by conversational participants. When one asserts, e.g. 'there are elephants', the world must satisfy a non-trivial empirical condition in order for the assertion's truth-conditions to be satisfied. Which condition? By presupposing my own conceptual framework, I can state it as follows: the condition that there be elephants. (If one wanted to state the condition in a framework-neutral way one would have to say something like: the condition that the world be such that space-time regions instantiating such-and-such 'basic' properties are related in thus-and-such ways.) When one asserts 'Elephant$(x) \ll_x$ Mammal(x)', one's assertion also has truth-conditions. But in this case the truth-conditions are *trivial*: nothing is required of the world in order for them to be satisfied. For one's conceptual framework counts the concept *elephant* as subordinate to the concept *mammal*, and this guarantees that 'Elephant$(x) \ll_x$ Mammal(x)' will be true relative to one's framework regardless of what the world is like."

"Other things being equal, what it takes for one's conceptual framework to count the concept *elephant* as subordinate to the concept *mammal* is for one to accept 'Elephant$(x) \ll_x$ Mammal(x)'. So there is an important sense in which the truth of semi-identity statements is *conventional*: the semi-identity statements that are counted as true are those we choose to adopt.[16] But describing the proposal in this way is somewhat misleading. For it suggests that we are free to adopt whichever semi-identity statements we like, and this is not quite right. We adopt a family of semi-identity statements on the basis of whether it supplies—in conjunction with a family of theoretical claims—a congenial articulations of our methods of inquiry. But the world places substantial constraints on the methods of inquiry that will prove to be successful, and not every family of semi-identity statements delivers equally congenial articulations of a given methodology. So although it is true that the decision to accept a semi-identity statement is to some extent pragmatic, it does not float free of what the world is like."

As far as I can tell, this sort of picture supplies a genuine alternative to Realism. There is no need to take sides on the issue here, however. I noted earlier that the view I defend in this essay is compatible with Realism. It is also compatible with the alternative, since whenever I speak of true semi-identity statements I could be read as presupposing my own conceptual-framework.

9 Concluding Remarks

Throughout the paper I have defend the idea that the notions of semi-identity, possibility, intelligibility and why-closure are interconnected in significant ways: possibility is limited only by semi-identity; with lack of possibility comes unintelligibility; with necessity comes why-closure.

I have not, however, defended the view that one amongst these notions is, in some sense, *prior*. It seems to me that it is better to think of them as jointly supplying a picture of the modal landscape, with each of them placing constraints of the rest. Our views about what is possible are constrained by what we take to be intelligible; our views about what is intelligible are constrained by the semi-identity statements we accept; our acceptance of semi-identity statements is constrained by our ability to make sense of explanations; our ability to make sense of explanations is constrained by our views about possibility. And each of the points on this circle is constrained by its role in our theorizing about the world.

[16]For discussion of conventionalism in modality, see Sider (2003) and Sider's unpublished 'Reducing Modality'. For related discussion, see Cameron (forthcoming).

In order to get the result that ... is a true sentence of $L^\Diamond$	**it is enough to count ... as a true statement of L.**
Analyticity $\Box(\forall x(\mathrm{V}(x) \to \mathrm{F}(x)))$ (necessarily, every vixen is female)	 $\mathrm{V}(x) \ll_x \mathrm{F}(x)$ (part of what it is to be a vixen is to be female)
Analyticity $\Box(\forall x \forall y(\mathrm{S}(x,y) \to \exists z(\mathrm{P}(z,x) \wedge \mathrm{P}(z,y))))$ (necessarily, sisters share a parent)	 $\mathrm{S}(x,y) \ll_{x,y} \exists z(\mathrm{P}(z,x) \wedge \mathrm{P}(z,y))$ (part of what it is for objects to be sisters is for them to share a parent)
Determinates and determinables $\Box(\forall x(\mathrm{E}(x) \to \mathrm{M}(x)))$ (necessarily, every elephant is a mammal)	 $\mathrm{E}(x) \ll_x \mathrm{M}(x)$ (part of what it is to be an elephant is to be a mammal)
Cross-category prohibitions $\Box(\forall x(\mathrm{E}(x) \to \neg\mathrm{O}(x)))$ (necessarily, every elephant is not an octopus)	 $\mathrm{E}(x) \ll_x \neg\mathrm{O}(x)$ (part of what it is to be an elephant is to not be an octopus)
Supervenience $\Box(\forall x(\Phi(x) \to \Psi(x))$ (necessarily, if you have physical property Φ, you have psychological property Ψ)	 $\Phi(x) \ll_x \Psi(x)$ (part of what it is to have physical property Φ is to have psychological property Ψ)
Supervenience $\Box(\Phi \to M)$ (necessarily, if physical fact Φ obtains, moral fact M obtains)	 $\Phi \ll M$ (part of what it is for physical fact Φ to obtain is for moral fact M to obtain)
Identity $\Box(\exists y(y = \mathrm{h}) \to \mathrm{h} = \mathrm{p})$ (necessarily, if Hesperus exists, it is identical to Phosphorus)	 $\mathrm{h} = \mathrm{p}$ (Hesperus is Phosphorus)
Kind identity $\Box(\forall x(\mathrm{Water}(x) \leftrightarrow \mathrm{H_2O}(x))$ (necessarily, water is H_2O)	 $\mathrm{Water}(x) \equiv_x \mathrm{H_2O}(x)$ (what it is to be water is to be H_2O)

Figure 2: Examples of Metaphysical Necessities

In order to get the result that ... is a true sentence of $L^\Diamond$	**it is enough to count ... as a true statement of L.**
Essentiality of kind $\Box(\forall z(\mathrm{H}(z) \to \Box(\exists y(y = z) \to \mathrm{H}(z)))$ (necessarily, if you're human, you couldn't have failed to be human)	$\dfrac{\mathrm{H}(z)}{x = z \ \ll_x \mathrm{H}(x)}$ (assume z is human; then part of what it is to be z is to be human)
Essentiality of kind $\Box(\forall z(\mathrm{M}(z) \to \Box(\exists y(y = z) \to \mathrm{H}(z)))$ (necessarily, if you're a man, you couldn't have failed to be human)	$M(x) \ll_x H(x)$; $\dfrac{\mathrm{H}(z)}{x = z \ \ll_x \mathrm{H}(x)}$ (part of what it is to be a man is to be human; moreover: assume z is human; then part of what it is to be z is to be human)
Essentiality of origin $\Box(\forall z(\mathrm{B}(c, z) \to \Box(\exists y(y = z) \to \mathrm{B}(c, z)))$ (necessarily, if you have Charles as a biological parent, you couldn't have failed to have Charles as a biological parent)	$\dfrac{\mathrm{B}(c, z)}{x = z \ \ll_x \mathrm{B}(c, x)}$ (assume z has Charles as a biological parent; then part of what it is to be z is to have Charles as a biological parent)
Essentiality of origin $\Box(\forall x \forall y(\mathrm{B}(w, z) \to \Box(\exists y(y = z) \to \mathrm{B}(w, z)))$ (necessarily, if z has w as a biological parent, then z couldn't have failed to have w as biological parent)	$\dfrac{\mathrm{B}(w, z)}{x = z \ll_x \mathrm{B}(w, x)}$ (assume z has w as a biological parent; then part of what it is to be z is to have w as a biological parent)
Essentiality of constitution $\Box(\forall z(\mathrm{W}(z) \to \Box(\neg \mathrm{I}(z)))$ (necessarily, if you're made of wood, you couldn't have been made of ice)	$\dfrac{\mathrm{W}(z)}{x = z \ \ll_x \mathrm{W}(x)}$; $\mathrm{W}(x) \ll_x \neg \mathrm{I}(x)$ (assume z is made of wood; then part of what it is to be z is to be made of wood; moreover, part of what it is to be made of wood is to not be made of ice)
Essentiality of constitution $\Box(\forall x \forall y((\mathrm{C}(z) \land \mathrm{C}(w) \land \mathrm{P}(w, z)) \to \Box(\exists y(y = z) \to \mathrm{P}(w, z)))$ (necessarily, if z and w are portions of clay and w is part of z, then z couldn't have failed to have w as a part)	$\dfrac{\mathrm{C}(z) \land \mathrm{C}(w) \land \mathrm{P}(w, z)}{x = z \ll_x \mathrm{C}(x) \land \mathrm{C}(w) \land \mathrm{P}(w, x)}$ (assume z and w are portions of clay and w is part of z; then part of what it is to be z is to be made of clay, to have w as a part and for w to be made of clay)

Figure 3: Examples of Metaphysical Necessities (Continued)

In order to get the result that ... is a true sentence of $L^{\Diamond}$	**it is enough to count ... as a true statement of L.**
Reflexivity $\Box(\forall x(x < x))$ (necessarily, anything is part of itself)	 $x = x \ll_x x < x$ (part of what it is to be self-identical is to be be a part of oneself)
Antisymmetry $\Box(\forall x \forall y((x < y \wedge y < x) \rightarrow x = y))$ (necessarily, if x is part of y and y is part of x, then x is identical to y)	 $(x < y \wedge y < x) \ll_{x,y} x = y$ (part of what it is for x and y to be such that x is part of y and y is part of x is for x and y to be identical)
Transitivity $\Box(\forall x \forall y \forall z((x < y \wedge y < z) \rightarrow x < z))$ (necessarily, if x is part of y and y is part of z, then x is part of z)	 $(x < y \wedge y < z) \ll_{x,y,z} x < z$ (part of what it is for x, y and z to be such that x is part of y and y is part of z is for x to be part of z)
Strong Supplementation $\Box(\forall x \forall y(\neg(y < x) \rightarrow \exists z(z < y \wedge \neg \mathrm{O}(z, x)))$ (necessarily, if y is not a part of x, then y has a part that does not overlap with z)	 $\neg(y < x) \ll_{x,y} \exists z(z < y \wedge \neg \mathrm{O}(z, x))$ (part of what it is for x and y to be such that y is not a part of x is for y to have a part that does not overlap with x)
Unrestricted Fusions $\Box(\exists x(\phi(x)) \rightarrow \exists z \forall x(\mathrm{O}(x, z) \leftrightarrow \exists y(\phi(y) \wedge \mathrm{O}(x, y))))$ (necessarily, if there are any ϕs, then there is a sum of the ϕs: a z such that the things that overlap with z are precisely the things that overlap with some ϕ)	 $\exists x(\phi(x)) \ll \exists z \forall x(\mathrm{O}(x, z) \leftrightarrow \exists y(\phi(y) \wedge \mathrm{O}(x, y))))$ (part of what it is for there to be a ϕ is for there to be a sum of the ϕs: a z such that the things that overlap with z are precisely the things that overlap with some ϕ)

Figure 4: Examples of Metaphysical Necessities (Mereological Principles)

De dicto $\Diamond(\exists x\, \text{TinyElephant}(x))$ (*Read:* there might have been a tiny elephant)
De re *(with embedded possibility)* $\Diamond(\exists x\,(\text{Gray}(x) \wedge \Diamond(\text{TinyElephant}(x))))$ (*Read:* there might have been something gray that might have been a tiny elephant)
De re *(with embedded necessity)* $\Diamond(\exists x\,(\text{Gray}(x) \wedge \Box(\exists y(y = x) \rightarrow \text{TinyElephant}(x))))$ (*Read:* there might have been something gray that is essentially a tiny elephant)
De re *(with embedded necessity)* $\Diamond(\exists x \Box(\exists y(y = x) \rightarrow \forall y(y = x)))$ (*Read:* there might have been an essentially lonely object)
De re *(with double embeddings)* $\Diamond(\exists x \Diamond(\exists y \Box(x \neq x \vee y \neq y)))$ (*Read:* there might have been 'incompatible' objects: objects each of which could have existed but such that they couldn't have existed together.)
De re *(with multiple embeddings)* $\Diamond(\exists x_1 \Diamond(\exists x_2 \ldots \Diamond(\exists x_n(\Box(x_1 \neq x_2 \vee \ldots \vee x_n \neq x_n))) \ldots))$ (*Read:* there might have been 'n-incompatible' objects: n objects each of which could have existed but such that they couldn't have all existed together.)
De re *(with multiple embeddings)* $\Diamond(\exists x\ (\text{G}(x) \wedge \Diamond(\exists y\ (\text{S}(y,x)) \wedge \Box((\exists z(z = x) \wedge \exists z(z = y)) \rightarrow (\text{Y}(x,y) \wedge \Diamond(\exists z\ (\text{O}(z) \wedge \Box(\forall w\ (\text{H}(w) \rightarrow \text{M}(w))))))))))$ (*Read:* there might have been something gray that might have had a sister who couldn't have failed to be younger than it and is such that it might have been the case that there are octopuses and humans are necessarily mammals.)

Figure 5: Examples of metaphysical possibilities that are provably true in the canonical space of worlds just in case they are consistent with the core modal truths.

Appendix: Proofs

Preliminaries

If L is a first-order language, let an a-world (short for 'actualist-world') for $L^\Diamond$ be an ordered pair $\langle D, I\rangle$ such that:

- The domain D is a set of ordered pairs of the form $\langle x, \text{`actual'}\rangle$ (for $x$ an individual in the domain of discourse of $L$) or $\langle x, \text{`nonactual'}\rangle$ (for x an arbitrary object).
- The interpretation function I assigns a subset of D^n to each n-place predicate-letter of $L^\Diamond$, and a function from D^n to D to each n-place function-letter of $L^\Diamond$.
- If c is an individual constant of $L^\Diamond$ and x is its intended interpretation, I assigns the pair $\langle x, \text{`actual'}\rangle$ to c.

(A Kripke-semantics based on a-worlds can be characterized along familiar lines—see my 'An Actualist's Guide to Quantifying-In' for details. I assume that L contains no empty names, and that atomic formulas (including identity statements) fail to be satisfied by a pair at an a-world unless the pair is in the domain of the a-world.)

The easiest way of understanding how a-worlds are supposed to work is by comparing them to Lewisian worlds.[17] Like a-worlds, Lewisian worlds represent possibilities: a Lewisian world in which my counterpart has a sister, for example, represents a possibility whereby I myself have a sister. But representation works differently for a-worlds and Lewisian worlds. Whereas Lewisian worlds represent by analogy, a-worlds represent by satisfaction. A Lewisian world represents the possibility that I have a sister by containing a person who is similar to me in certain respects, and has a sister. An a-world, on the other hand, represents the possibility that I have a sister by satisfying the formula '$\exists x(\text{Sister}(x, \text{AR}))$', where 'Sister' is a predicate that expresses sisterhood on its intended interpretation and 'AR' is a name that refers to me on its intended interpretation. For instance, the a-world $\langle D_1, I_1\rangle$ represents a possibility whereby I have a sister who is a philosopher:

$$D_1 = \{\langle \text{Agustín}, \text{`actual'}\rangle, \langle \text{Socrates}, \text{`nonactual'}\rangle\}$$

$$I_1(\text{`Sister'}) = \{\langle\langle \text{Socrates}, \text{`nonactual'}\rangle, \langle \text{Agustín}, \text{`actual'}\rangle\rangle\}$$

$$I_1(\text{`Philosopher'}) = \{\langle \text{Socrates}, \text{`nonactual'}\rangle\}$$

$$I_1(\text{`AR'}) = \langle \text{Agustín}, \text{`actual'}\rangle$$

Say that two possibilities are *linked* if—as one is inclined to put it—they concern the same individual, even if the individual in question doesn't actually exists. Here is an example.

[17] See Lewis (1986), especially §4.1. Lewis takes the counterpart relation to be context-dependent, but here I shall treat it as constant for the sake of simplicity. Also for the sake of simplicity, I shall assume that the counterpart relation is an equivalence.

Although I don't have a sister, I might have had one. And had I had a sister, she might have been a philosopher. But her profession wouldn't have been essential to her: she might have been a cellist instead. Accordingly, there a is possibility whereby I have a sister who is a philosopher and a different possibility whereby *that very individual* is a cellist. So—as one is inclined to put it—the two possibilities concern the same individual, even though she doesn't actually exist.

Consider how linking gets addressed from the perspective of the Lewisian. l_1 and l_2 are Lewisian worlds: l_1 contains an individual a_1 who is my counterpart and an individual s_1 who is a_1's sister and a philosopher; l_2 contains an individual s_2 who is a cellist. Accordingly, l_1 represents a possibility whereby my sister is a philosopher, and l_2 represents a possibility whereby someone is a cellist. But nothing so far guarantees *linking* between the possibility represented by l_1 and the possibility represented by l_2. Nothing so far guarantees that—as one is inclined to put it—the individual l_1 represents as my sister is *the very individual* that l_2 represents as a cellist. What is needed for linking is that s_1 and s_2 be counterparts: that they be similar in the right sorts of respects.

A similar maneuver can be used to represent linking amongst a-worlds: just like the Lewisian uses *counterparthood* amongst representations to capture linking, we shall use *identity* amongst representations to capture linking. Here is an example. We have seen that the a-world $\langle D_1, I_1\rangle$ represents a possibility whereby I have a sister who is a philosopher. Now consider two additional a-worlds, $\langle D_2, I_2\rangle$ and $\langle D_3, I_3\rangle$:

$$D_2 = \{\langle \text{Socrates}, \text{`nonactual'}\rangle\}$$
$$I_2(\text{`Philosopher'}) = \{\}$$
$$I_2(\text{`Cellist'}) = \{\langle \text{Socrates}, \text{`nonactual'}\rangle\}$$

$$D_3 = \{\langle \text{Plato}, \text{`nonactual'}\rangle\}$$
$$I_3(\text{`Philosopher'}) = \{\}$$
$$I_3(\text{`Cellist'}) = \{\langle \text{Plato}, \text{`nonactual'}\rangle\}$$

Both of these worlds represent a possibility whereby someone is a cellist rather than a philosopher. But only $\langle D_2, I_2\rangle$ is linked to $\langle D_1, I_1\rangle$. For $\langle D_1, I_1\rangle$ and $\langle D_2, I_2\rangle$ both employ $\langle \text{Socrates}, \text{`nonactual'}\rangle$ as a representation, and it is this that guarantees that—as one is inclined to put it—the individual who $\langle D_1, I_1\rangle$ represents as my sister is *the very individual* that $\langle D_2, I_2\rangle$ represents as a cellist. On the other hand, since $\langle D_3, I_3\rangle$ represents a possibility whereby someone is a cellist by using $\langle \text{Plato}, \text{`nonactual'}\rangle$ rather than $\langle \text{Socrates`nonactual'}\rangle$, what one gets is that—as one is inclined to put it—the individual who $\langle D_1, I_1\rangle$ represents as my sister is *distinct* from the individual that $\langle D_2, I_2\rangle$ represents as a cellist.[18]

From the perspective of the Lewisian, an individual with a counterpart in the actual world represents its actual-word counterpart, and an individual with no counterpart in the

[18]One could, if one wanted, use counterparthood amongst representations rather than identity amongst representations to track a-world linking; but doing so would complicate the proposal, and bring no real benefit in the present context.

actual world represents a merely possible object. From the present perspective, a pair of the form '$\langle x, \text{'actual'}\rangle$' represents its first component, and a pair of the form '$\langle n, \text{'nonactual'}\rangle$' represents a merely possible object (even though the pair itself, and both of its components, are actually existing objects). Thus, $\langle D_2, I_2\rangle$ and $\langle D_3, I_3\rangle$ represent possibilities whereby there is an object that doesn't actually exist, and $\langle D_1, I_1\rangle$ represents a possibility whereby there is an object that doesn't actually exist and an object that does (i.e. me).

The Construction

I shall assume that the domain of L forms a set, and that L contains a name for every object in its domain. As in the main text, let S be the set of true semi-identity statements in $L^{\ll}$ (including conditional semi-identity statements). Say that an a-world w for $L^{\Diamond}$ is *good* just in case it meets the following conditions:

$G1$ For any $\ulcorner\phi(x_1, \ldots, x_n) \ll_{x_1,\ldots,x_n} \psi(x_1, \ldots, x_n)\urcorner$ in S, $\ulcorner\forall x_1 \ldots \forall x_n(\phi(x_1, \ldots, x_n) \rightarrow \psi(x_1, \ldots, x_n))\urcorner$ is true at w.

$G2$ Let $\phi(z, w_1, \ldots, w_n)$ be constitutive. If $\langle y, \text{'actual'}\rangle$, $\langle v_1, \text{'actual'}\rangle$, $\ldots$, $\langle v_n, \text{'actual'}\rangle$ are all in the domain of w, then $\langle\langle y, \text{'actual'}\rangle, \langle v_1, \text{'actual'}\rangle, \ldots, \langle v_n, \text{'actual'}\rangle\rangle$ is in the w-extension of $\phi(z, w_1, \ldots, w_n)$ just in case $\phi(z, w_1, \ldots, w_n)$ is, in fact, true of $y, v_1 \ldots, v_n$.

$G3$ (a) if $\langle x, \text{'actual'}\rangle$ is in the domain of w, then x is in the domain of discourse of L.

(b) if $\langle x, \text{'nonactual'}\rangle$ is in the domain of w, then x is a natural number.

Say that a *set* A of a-worlds is *good* just in case every world in w is good, and the following additional condition is satisfied:

$G4$ Let $\phi(z, w_1, \ldots, w_n)$ be constitutive, and let w and w' be a-worlds in A. If $\langle\langle y, e_0\rangle, \langle v_1, e_1\rangle, \ldots, \langle v_n, e_n\rangle\rangle$ is in the w-extension of $\phi(z, w_1, \ldots, w_n)$, and $\langle y, e_0\rangle, \langle v_1, e_1\rangle, \ldots, \langle v_n, e_n\rangle$ are all in the domain of w', then $\langle\langle y, e_0\rangle, \langle v_1, e_1\rangle, \ldots, \langle v_n, e_n\rangle\rangle$ is in the w'-extension of $\phi(z, w_1, \ldots, w_n)$.

If S is the set of true semi-identity statements, characterize the *canonical* set of a-worlds M as follows. If A is a good set of a-worlds, let the *indexing* A^* of A be the result of replacing each pair $\langle n, \text{'nonactual'}\rangle$ in the domain of some world in A by $\langle n_A, \text{'nonactual'}\rangle$. Let M be the union of A^*, for each good set A of a-worlds for $L^{\Diamond}$. (That M is small enough to form a set is guaranteed by $G3$.)

The Principle of Maximality

If every world in A satisfies $G1$, $G2$, $G3(a)$ and A satisfies $G4$, we shall say that A is *quasi-good*. If A is a quasi-good set of a-worlds, say that A^c is a *good copy* of L just in case A^c is a good set of a-worlds which is isomorphic to an elementarily equivalent submodel of A. (The Löwenheim-Skolem Theorem guarantees that every quasi-good set of a-worlds has a good copy.) If A is a quasi-good set, say that $A^{\#}$ is an *indexed* good copy of A just in case, for some good copy A^c of A, $A^{\#}$ is the result of replacing each pair $\langle n, \text{`nonactual'}\rangle$ in the domain of $A^c$ by $\langle n_A, \text{`nonactual'}\rangle$.

The Principle of Maximality can then be reformulated as follows:

> Principle of Maximality (Official Version)
> A sentence of $L^{\Diamond}$ is true just in case it is true in a space of worlds isomorphic to the union of all indexed good copies of quasi-good sets of a-worlds.

To see that the canonical space of worlds M meets this conditions, note: (1) that the result of indexing a good set is always an indexed good copy of a quasi-good set (since every good set is a quasi-good set), and (2) that every indexed good copy of a quasi-good set is isomorphic to the result of indexing a good set.

Proof that M verifies S

If A is a set of a-worlds, we shall say that a set of a-worlds A verifies S just in case: (*i*) every world in A satisfies $G1$, and (*ii*) A satisfies $G4$.

That every world in M satisfies $G1$ follows immediately from the construction of M. That M satisfies $G4$ can be verified as follows. Assume $\phi(z, w_1, \ldots, w_n)$ is constitutive. Let w and w' be a-worlds in M, and let $\langle\langle y, e_0\rangle, \langle v_1, e_1\rangle, \ldots, \langle v_n, e_n\rangle\rangle$ be in the w-extension of $\phi(z, w_1, \ldots, w_n)$. *Case 1:* e_i is 'actual' for every $i \leq n$. Then the result follows from the construction of M, together with the observation that every good a-world satisfies $G2$. *Case 2:* for some member of the sequence, $\langle\langle y, e_0\rangle, \langle v_1, e_1\rangle, \ldots, \langle v_n, e_n\rangle\rangle$, e_i is 'non-actual'. Because every good a-world satisfies $G3(b)$, the relevant pair must be of the form $\langle n_A, \text{`nonactual'}\rangle$, for some good A. If w' is not in A^*, the conclusion follows immediately (since $\langle n_A, \text{nonactual}\rangle$ is not in the domain of w'). If, on the other hand, w' is in A^*, then the result of removing indices from pairs in the domain of w' is in A, and the conclusion follows from the fact that A satisfies $G4$.

Further Results

We will verify that M meets each of the following conditions:

$M1$ Every basic modal truth is true in M.

$M2$ Every possibility statement consistent with the core of S is true in M.

The notions of *basic modal truth*, *possibility statement* and *core* will be defined in turn.

Basic Modal Truth. Say that a sentence ϕ is *basic* just in case it is equivalent to a sentence that can be built up from sentences containing no modal operators by using '$\Diamond$', '$\neg$' and '$\wedge$'. Say that a sentence is a *basic modal truth* just in case it is a basic sentence and is true in the Kripke-model based on the set of good worlds (with the 'actualized' a-world as a center).

Possibility Statement. Say that a formula of $L^\Diamond$ is in *normal form* if it is built up from atomic and negated atomic formulas using '$\wedge$', '$\vee$', '$\exists$', '$\forall$', '$\Diamond$' and '$\Box$', but without using '$\rightarrow$', '$\leftrightarrow$' or further applications of '$\neg$'. (One can transform any formula of $L^\Diamond$ into an equivalent formula in normal form.) Say that a formula of $L^\Diamond$ is *neutral* just in case it is equivalent to a formula ϕ with the following properties:

- ϕ is in normal form
- if $\Box(\psi)$ is a subformula of ϕ which does not itself fall within the scope of a box, then ψ is of the form $\ulcorner x_1 \neq x_1 \vee \ldots \vee x_n \neq x_n \vee \theta \urcorner$, where θ results from a basic sentence by uniformly substituting $x_1, \ldots, x_n$ for individual constants.[19]

A *possibility statement* is a sentence of the form $\ulcorner\Diamond(\phi)\urcorner$, for ϕ a neutral sentence.

Core. The *core* of S is the set consisting of: (a) every basic modal truth, (b) each sentence

$$\ulcorner\Box(\forall x_1, \ldots, x_n(\phi(x_1, \ldots, x_n) \rightarrow \psi(x_1, \ldots, x_n)))\urcorner$$

for $\ulcorner\phi(x_1, \ldots, x_n) \ll_{x_1,\ldots,x_n} \psi(x_1, \ldots, x_n)\urcorner$ in S, and (c) each sentence

$$\Box(\forall z \forall w_1, \ldots \forall w_n(\phi(z, w_1, \ldots, w_n) \rightarrow \Box(z = z \rightarrow \phi(z, w_1, \ldots, x_n))))$$

for $\phi(z, w_1, \ldots, w_n)$ constitutive.

Proof that M satisfies condition $M1$

If w is a good a-world, then $\{w\}$ is a good set. So we have the following: for any good w, w^* is in M (where w^* results from replacing each $\langle x, \text{nonactual}\rangle$ by $\left\langle x_{\{w\}}, \text{nonactual}\right\rangle$ in w). The definition of M guarantees that we also have the following: for any w in M, $w^\#$ is good (where $w^\#$ results from replacing each $\langle x_A, \text{`nonactual'}\rangle$ by $\langle x, \text{`nonactual'}\rangle$ in w). To show that every basic modal truth is true in M it is therefore sufficient to prove the following:

> Let G be the set of good worlds. Then, for any basic sentence ϕ of $L^\Diamond$, any w in G and any w' in M: (a) if ϕ is true at w in G, ϕ is true at w^* in M, and (b) if ϕ is true at w' in M, ϕ is true at $w'^\#$ in G.

The proof proceeds by induction on the number of modal operators in ϕ:

- ϕ contains no modal operators. Immediate.

[19] $\ulcorner x_i \neq x_i \urcorner$ is the negation of $\ulcorner x_i = x_i \urcorner$.

- ϕ is of the form $\ulcorner\Diamond(\psi)\urcorner$. Suppose ϕ is true at w in G. Then there is a world w_1 in G such that ψ is true at w_1 in G. But w_1^* is in M. By inductive hypothesis, ψ is true at w_1^* in M. So ϕ is true at w^* in M. Now suppose ϕ is true at w' in M. Then there is a world w_1' in M such that ψ is true at w_1' in M. But $w_1'^{\#}$ is in G. By inductive hypothesis, ψ is true at $w_1'^{\#}$ in G. So ϕ is true at $w^{\#}$ in G.

- ϕ is of the form $\ulcorner\neg\psi\urcorner$. Suppose ϕ is true at w in G. Then ψ is false at w in G. By inductive hypothesis, ψ is false at w^* in M. So ϕ is true at w^* in M. (The converse is analogous.)

- ϕ is of the form $\ulcorner\psi \wedge \theta\urcorner$. Suppose ϕ is true at w in G. Then ψ and θ are at w in G. By inductive hypothesis, ψ and θ are true at w^* in M. So ϕ is true at w^* in M. (The converse is analogous.)

Proof that M satisfies condition $M2$

Let ϕ be a possibility statement consistent with $\kappa(S)$, the core of S. Then there is a Kripke-model K of $\kappa(S) \cup \{\phi\}$. Moreover, the Löwenheim-Skolem Theorem guarantees that there is an elementarily equivalent countable submodel C of K.

We may assume with no loss of generality that C is a good set. For we may assume with no loss of generality that every object in the domain of some world in K is of the form $\langle x, \text{‘actual’}\rangle$ if it is in the domain of the ‘actual’ world of K, and of the form $\langle x, \text{‘nonactual’}\rangle$ otherwise. Accordingly, K may be thought of as a set of a-worlds with a designated center. That K satisfies $G1$ and $G4$ follows immediately from the fact that it is a model of $\kappa(S)$. That K satisfies $G2$ follows from the fact that $L^{\Diamond}$ has a name for each object in the domain of L, together with the fact that K is a model of $\kappa(S)$. And the construction of K guarantees that it also satisfies $G3(a)$. Since C is a submodel of K, C must also satisfy $G1$, $G2$, $G3(a)$ and $G4$. And since C is countable, it must be is isomorphic to a set of a-worlds satisfying $G1$–$G4$ (minus the designated center, which may be ignored from now on).

We know that C is a good set such that ϕ is true in C and every basic modal truth is true in C. And since ϕ is a possibility statement it is also a neutral formula. To prove that ϕ is true in M, it is therefore enough to show the following:

> Let ρ be a neutral formula, and A be a good set such that every basic modal truth is true in A. Let σ be an A-variable assignment and w be an a-world in A. Then if ρ is satisfied by σ at w in A, ρ is satisfied by σ^* at w^* in M (where w^* and σ^* result from replacing pairs $\langle n, \text{‘nonactual’}\rangle$ by $\langle n_A, \text{‘nonactual’}\rangle$ in w and σ, respectively).

The proof proceeds by induction on the complexity of formulas:

- ρ is an atomic or negated atomic sentence. Trivial.

- ρ is of the form $\ulcorner\psi \vee \theta\urcorner$, $\ulcorner\psi \wedge \theta\urcorner$. Follows immediately by inductive hypothesis.

- ρ is of the form $\ulcorner\exists x\psi\urcorner$. If $\ulcorner\exists x\psi\urcorner$ is satisfied by σ at w in A, there is a pair p in the domain of w such that ψ is satisfied by $\sigma[x/p]$ at w in A. By inductive hypothesis, there is a pair p^* in the domain of w^* such that ψ is satisfied by $\sigma^*[x/p^*]$ at w^* in M. So $\ulcorner\exists x\psi\urcorner$ is satisfied by σ^* at w in M.

- ρ is of the form $\ulcorner\forall x\psi\urcorner$. If $\ulcorner\forall x\psi\urcorner$ is satisfied by σ at w in A, every pair p in the domain of w is such that ψ is satisfied by $\sigma[x/p]$ at w in A. By inductive hypothesis, every p^* in the domain of w^* is such that ψ is satisfied by $\sigma^*[x/p^*]$ at w^* in M. So $\ulcorner\forall x\psi\urcorner$ is satisfied by σ^* at w in M.

- ρ is of the form $\ulcorner\Diamond(\psi)\urcorner$. If $\ulcorner\Diamond(\psi)\urcorner$ is satisfied by σ at w in A, there is a world w' in A such that ψ is satisfied by σ at w' in A. By inductive hypothesis, ψ is satisfied by σ^* at $(w')^*$ in M. So $\ulcorner\Diamond(\psi)\urcorner$ is satisfied by σ^* at w^* in M.

- ρ is of the form $\ulcorner\Box(x_1 \neq x_1 \vee \cdots \vee x_n \neq x_n \vee \psi)\urcorner$, where ψ results from a basic sentence by uniformly substituting $x_1, \ldots, x_n$ for individual constants. If $\ulcorner\Box(x_1 \neq x_1 \vee \cdots \vee x_n \neq x_n \vee \psi)\urcorner$ is satisfied by σ at w in A, every world w' in A is such that $\ulcorner x_1 \neq x_1 \vee \cdots \vee x_n \neq x_n \vee \psi\urcorner$ is satisfied by σ at w' in A. Let w^+ be a world in M. We show that $\ulcorner x_1 \neq x_1 \vee \cdots \vee x_n \neq x_n \vee \psi\urcorner$ is satisfied by σ^* at w^+ in M.

  Case 1: w^+ is $(w')^*$ for some w' in A. The result follows by inductive hypothesis.

  Case 2: $w^+ \neq (w')^*$ for every w' in A. *Case 2a:* for some $i \leq n$, $\sigma(x_i)$ is $\langle m, \text{‘nonactual’}\rangle$. So $\sigma^*(x_i)$ is $\langle m_A, \text{‘nonactual’}\rangle$. But $w^+ \neq (w')^*$ for every w' in A, so $\sigma^*(x_i)$ is not in the domain of w^+. It follows that that $x_i \neq x_i$ is satisfied by σ^* at w^+ in M, and therefore that $\ulcorner x_1 \neq x_1 \vee \cdots \vee x_n \neq x_n \vee \psi\urcorner$ is satisfied by σ^* at w^+ in M.

  Case 2b: $\sigma(x_1), \ldots, \sigma(x_n)$ are all pairs of the form $\langle z, \text{‘actual’}\rangle$. Then $\sigma(x_1) = \sigma^*(x_1), \ldots, \sigma(x_n) = \sigma^*(x_n)$. So $\ulcorner x_1 \neq x_1 \vee \cdots \vee x_n \neq x_n \vee \psi\urcorner$ is satisfied by σ^* at w^+ in M just in case ξ is satisfied by σ^* at w^+ in M, where ξ is the result of replacing each x_i ($i \leq n$) in $\ulcorner x_1 \neq x_1 \vee \cdots \vee x_n \neq x_n \vee \psi\urcorner$ with a name for the first component of $\sigma(x_i)$. But $\ulcorner\Box(\xi)\urcorner$ and its negation are both basic sentences, and one of them is a basic modal truth. Suppose the negation of $\ulcorner\Box(\xi)\urcorner$ is a basic modal truth. Since every basic modal truth is true in A, it follows that the negation of $\ulcorner\Box(\xi)\urcorner$ is true in A. So the negation of ξ is true at some world in A, contradicting the fact that every world w' in A is such that $\ulcorner x_1 \neq x_1 \vee \cdots \vee x_n \neq x_n \vee \psi\urcorner$ is satisfied by σ at w' in A. It follows that $\ulcorner\Box(\xi)\urcorner$ is a basic modal truth. Since every basic modal truth is true in M, it follows that $\ulcorner\Box(\xi)\urcorner$ is true in M, and therefore that ξ is true at w^+ in M. So $\ulcorner x_1 \neq x_1 \vee \cdots \vee x_n \neq x_n \vee \psi\urcorner$ is satisfied by σ^* at w^+ in M.

References

Block, N., and R. Stalnaker (1999) "Conceptual Analysis, Dualism, and the Explanatory Gap," *Philosophical Review* 108, 1–46.

Cameron, R. (forthcoming) "What's Metaphysical about Metaphysical Necessity?" *Philosophy and Phenomenological Research*.

Chalmers, D. (1996) *The Conscious Mind: In Search of a Fundamental Theory*, Oxford University Press, New York.

Chalmers, D., and F. Jackson (2001) "Conceptual Analysis and Reductive Explanation," *The Philosophical Review* 110, 315–360.

Davidson, D. (1967) "Truth and Meaning." In **?**.

Fine, K. (1994) "Essence and Modality," *Philosophical Perspectives* 8, 1–16.

Fine, K. (1995a) "The Logic of Essence," *Journal of Philosophical Logic* 24, 241–273.

Fine, K. (1995b) "Senses of Essence." In **?**.

Fine, K. (2000) "Semantics for the Logic of Essence," *Journal of Philosophical Logic* 29, 543–584.

Frege, G. (1884) *Die Grundlagen der Arithmetik*. English Translation by J.L. Austin, *The Foundations of Arithmetic*, Northwestern University Press, Evanston, IL, 1980.

Hale, B., and C. Wright (2001) *The Reason's Proper Study: Essays towards a Neo-Fregean Philosophy of Mathematics*, Clarendon Press, Oxford.

Jackson, F. (1998) *From Metaphysics to Ethics: A Defence of Conceptual Analysis*, Oxford University Press, Oxford.

Kment, B. (2006) "Counterfactuals and the Analysis of Necessity," *Philosophical Perspectives* 20, 237–302.

Lewis, D. (1986) *On the Plurality of Worlds*, Blackwell, Oxford and New York.

Sider, T. (2003) "Reductive Theories of Modality." In **?**, pp. 180-208.

Stalnaker, R. C. (2003) *Ways a World Might Be: Metaphysical and Anti-Metaphysical Essays*, Clarendon Press, Oxford.

Wright, C. (1983) *Frege's Conception of Numbers as Objects*, Aberdeen University Press, Aberdeen.

Yablo, S. (1993) "Is Conceivability a Guide to Possibility?" *Philosophy and Phenomenological Research* 53, 1–42.