A Semantics for Conceivability and Possibility

Micah Newman

September 2, 2009

A spectrum of modal logics have been developed that are available to model a range of interpretations of the logic of possibility and necessity. These involve an accumulation of axioms of increasing strength, allowing different modal entailments, characteristic of each particular logic, to be valid. The choice of modal logic depends on the application at hand, and what sorts of entailments are wanted in it. In this paper, I present a somewhat gerrymandered, but I think well-motivated, modal semantics that is tailored to express the distinction between conceivability and possibility. The distinction involved has numerous applications for the project of making sense of modal discourse without assuming that one can go straight from conceivability to genuine possibility.

1 Accessibility and Possibility

Many have adopted the logics S4 or S5 as correctly expressing a fundamental sort of modality because they allow iterated formulas, where modal symbols are "stacked" onto each other, to be collapsed into one. Due to the interdefinability of $\Box$ and $\Diamond$ via negation, the characteristic axioms of S4 and S5 can be put a number of equivalent ways. The most well-known of each is shown below.

(S4) $\Box p \Vdash \Box\Box p$

(S5) $\Diamond p \Vdash \Box\Diamond p$

Formulas with iterated modalities, while well-formed, can be somewhat puzzling to interpret; thus, it is attractive to many to adopt a logic where

these are superfluous. Application of the S4 axiom allows one to replace strings of the same modal operator with one of that operator. The S5 axiom amounts to the ability to handle strings of different modal operators by reducing them to the innermost operator. So, for example, due to the minimal formal meaning of "$\Box$," (S4) entails its own converse, allowing "boxes" to be collapsed. In that logic, then, something being said to be "necessarily necessary" just comes to it being necessary, period. In S5, furthermore, one can proceed straight from the somewhat perplexing "$\Diamond\Box p$" to just "$\Box p$".

The notion of an *accessibility* relation between worlds, introduced by Kripke in his modal semantics that came to be widely-adopted, is a very common way to interpret differences between modal logics. The accessibility relation allows a world to "see" other worlds in its domain, with respect to truth-values in neighboring worlds as they are relevant to truths of possibility (at *some* accessible world) and necessity (at *all* accessible worlds). For some modal systems weaker than S4 or S5, the accessibility relation is at least reflexive. In S4, the relation is transitive as well as reflexive. S5 adds symmetry to transitivity. Thus, the accessibility relation in its various permutations is an ingenious and useful heuristic for interpreting the differences between modal systems.

2 Conceivability and Possibility

As noted above, it is widely thought that in the most basic sense of modality, maximal accessibility should be allowed throughout the domain of worlds. I offer here a framework for modality at odds with that, one in which restrictions on accessibility can be used model some basic distinctions in thinking about modality itself.

This meshes with another cardinal philosophical conviction of mine, which is that conceivability by itself does not entail possibility, and that there ought to be a modal framework that allows for less-than-complete accessibility that will capture a separation between actual possibilities and mere conceivabilities. For example, we might want to say that it's conceivable that water not be H_2O, even if it's not possible, since *water* in fact rigidly designates just that stuff that is H_2O. One might want to say it "could have turned out" that water is not H_2O, and that such is a "conceptual possibility." (For more on such locutions, see Yablo (2002: §§9–10).)

Stephen Yablo has proposed to think of conceptual possibilities in a way

that is as metaphysically untendentious as possible. This is to consider them as *counteractual*—a weaker, purely epistemic notion than the metaphysically-loaded *counterfactual* (see Yablo 2002: §§11–12). The notion of the counteractual is wanted so as to not assume that conceptual necessity and apriority coincide, which adherents to "modal rationalism," who place heavy emphasis on conceivability arguments for metaphysical conclusions, tend to tacitly assume. But when examined, this equivalence turns out not to be correct, nor is an entailment in either direction valid, as is clear when one thinks through conceptual predicates that are, say, evaluative, recognitional, or theoretical. So, one way to think of conceivability without smuggling in possibility through the back door is to think of nonactual conceivabilities as the counteractual. That way, it isn't assumed that we know exactly *what* we are thinking *about* in counteractual worlds given our this-worldly conceptual repertoire.

So, given the above, how should we relate the conceivable to the genuinely possible? One important thing to note, as George Bealer (2002) has pointed out, is that the epistemic and nonepistemic senses of any given modal expression "do not mark out two associated *genuses* [species?]—epistemic possibility and nonepistemic possibility—of some common still more general modal category *Possible*." (77) So given that we want to think of the genuinely possible as a proper subclass of the conceivable—so that inclusion in the latter does not entail inclusion in the former—we could think of the possible as a special case of the conceivable. Thus, counteractual worlds that we specify via exercises in conceivability might be modeled as the merely "possibly possible." In such a model, we would want the following:

(DD) $\Diamond\Diamond p \nvdash \Diamond p$

(BB) $\Box p \nvdash \Box\Box p$

But you can't get the pair (DD) and (BB) (which are equivalent, holding in all the same models) in logics with the S4 axiom: (S4) is equivalent to (S4p).

(S4p) $\Diamond\Diamond p \vdash \Diamond p$

So in modal logics as strong as S4 and S5, iterated formulas collapse, and so, too, does the wanted distinction between conceivability and possibility. So, if the suggested way to model mere conceivability is to work, what is needed is a system weaker than S4.

3 Non-normal Models

When we look to weaken S4, we find in the neighborhood of S2 and S3 the *non-normal* models, first proposed by Kripke (1965). As Veikko Rantala has put it, "[a] non-normal world is a sort of "auxiliary" possible world which need not obey the same logical laws—concerning the underlying propositional or predicate logic or the modal operator—as a normal world." (Rantala 1982b: 41) One way this framework has been interpreted is to think of the "non-normal worlds" as *impossible* worlds, as discussed by Jaakko Hintikka (1975).

The application of "impossible worlds" comes in usefully with respect to the possible-worlds treatment of propositional attitudes. The whole possible-worlds approach of modeling intensional functions like knowledge and belief may seem to founder if is taken to commit one to *logical omniscience*, "the assumption that everyone knows all the logical consequences of what he knows, and analogously for all the other propositional attitudes." (*ibid.,* 475) As Hintikka discusses, this obviously false assumption seems to follow from applying the whole space of logical possibilities to propositional attitudes, where the semantics would go basically thus:

(PAS) A sentence of the form "a knows that p" is true in a world w iff p is true in all the epistemically possible worlds which are compatible with everything a knows in w.

Then, adding the failure of logical omniscience would seem to undermine a principle foundational to any possible-worlds semantics, (LP):

(LP) A sentence is logically true iff it is true in every logically possible world.

But the failure of logical omniscience contradicts (LP) only with the further assumption (CP):

(CP) Every epistemically possible world is logically possible. (*ibid.*)

Hence, the need for "impossible worlds," which vitiate (CP) and thus allow one to take account of the epistemic limitations of normal agents. The notion of non-normal worlds as impossible worlds was further applied by Veikko Rantala (1982a, 1982b) to the formal semantics of propositional attitudes in a general way.

Conceivability might be thought of a species of propositional attitude; in particular, denying (CP) would be sufficient to get a failure of general entailment of possibility from conceivability. The wanted denial of logical omniscience is also tantamount to denying the equivalence of conceptual necessity and apriority, as discussed in the previous section and by Yablo (2002). So, let us see how a non-normal modal model of conceivability and possibility might work given the desiderata canvassed in the previous section.

The basic non-normal modal language $\mathcal{L}_{NN}$, with standard syntax, uses a structure $< \mathcal{W}, \mathcal{N}, \mathcal{R}, \nu >$, where $\mathcal{N} \subseteq \mathcal{W}$, and $\mathcal{N} \neq \emptyset$. The usual interpretation of a non-normal model is one in which the non-normal worlds $w_{NN} \in \mathcal{W} - \mathcal{N}$ are such that all diamond-claims are true and all box-claims are false. And, logical consequence is defined only over the normal worlds ($w \in \mathcal{N}$).

To the basic $\mathcal{L}_{NN}$ structure, we would want to add the stipulation that non-normal worlds are, in fact, present in the model: $\mathcal{W} - \mathcal{N} \neq \emptyset$. We would also, I presume, want the accessibility relations $\mathcal{R}$ between worlds to be maximal: the only motivation, for present purposes, of specifying any less free accessibility than that is to disallow such as (S4) and (S4p). So, $\mathcal{R} \subseteq \mathcal{W} \times \mathcal{W}$, giving an S5 accessibility.

Call this more-specified non-normal language $\mathcal{L}_{NN+}$. In $\mathcal{L}_{NN+}$, we get (DD) and (BB): Since all necessities are false at the non-normal worlds ($w \in \mathcal{W} - \mathcal{N}$), some necessities, such as logical truths, are such that it is not the case that $\Box\Box p$ even though $\Box p$. So, (S4) fails. Likewise, since some possibilities (diamond-claims), namely *impossible* "possibilities," are possible only at the non-normal worlds, there are propositions p such that $\Diamond\Diamond p$ but not $\Diamond p$. Therefore, (S4p) fails.

What does this much get us? We wanted conceivability as a species of possibility, albeit a more-distant kind from *bona fide* possibilities *simpliciter*: one that disallows a straightforward logical consequence of possibility from conceivability. $\mathcal{L}_{NN+}$ does this and may for that reason be thought to do the job overall. However, there are some serious shortcomings in $\mathcal{L}_{NN+}$ for the job it is here intended for. In that, it seems to me to do both too much and too little.

Too much, in that $\mathcal{L}_{NN+}$ is still too permissive in its valid formulas, in making absolutely any proposition turn out to be "conceivable." Since all diamond-claims are true at non-normal worlds, all double-diamond (that is, conceivability) claims, at the normal worlds that access the non-normal worlds, come out true. So, in addition to our being able to, with $\mathcal{L}_{NN}$, talk

about water's not turning out to be H_2O, we are also, it seems, "forced to conceive" of there being round squares and the square root of 2 being a rational number and such. This seems an awfully high price to pay.

In another way, $\mathcal{L}_{NN+}$ does too little for present purposes, in that even the conceivabilities that we might want come too cheaply. If counterpossibles like water's being XYZ and Hesperus $\neq$ Phosphorus are modeled automatically regardless of their content (let alone their being not distinguishable in the model from logical "impossibilities"), this hardly makes conceivability interesting.

Rather, it seems that interestingly conceivable propositions have content that relates to the content of other propositions: actually true propositions, as a subset of those, certainly do. A possible world is a maximal state of affairs, as Plantinga (1974: §4.1) has put it. Such a state of affairs excludes certain propositions while including others. Along the same lines, a maximal state of affairs includes sets of propositions that stand or fall together. Water cannot be H_2O without there existing the elements hydrogen and oxygen, which in turn cannot exist without there being protons, neutrons, and electrons, and so on. The conjunctive proposition that water is H_2O but that there are no protons I would like to consider as a logical impossibility, and likewise the proposition that there are, or even could be, round squares.

In sum, all this is to say that another desideratum that we want but that $\mathcal{L}_{NN+}$ cannot give us is:

(NPP) $\nVdash \Diamond\Diamond p$

$\mathcal{L}_{NN}$ by itself does yield (NPP), since for all that $\mathcal{L}_{NN}$ stipulates, not every normal world need access a non-normal world. But antecedently, we wanted transitive and symmetric access between worlds in the model. And if we add that to $\mathcal{L}_{NN}$, that is, to get $\mathcal{L}_{NN+}$, we once again lose (NPP): everything becomes "conceivable." Where to go from there?

4 $\mathcal{L}_{BC}$

What I propose is a language, which I will call $\mathcal{L}_{BC}$, that is inspired by the use of non-normal worlds; it contains worlds that behave like non-normal worlds. Like $\mathcal{L}_{NN+}$, $\mathcal{L}_{BC}$ has a structure $< \mathcal{W}, \mathcal{N}, \mathcal{R}, \nu >$, with $\mathcal{N} \subset \mathcal{W}$, $\mathcal{N} \neq \emptyset$, and $\mathcal{W} - \mathcal{N} \neq \emptyset$. $\mathcal{L}_{BC}$ has three main points of departure from $\mathcal{L}_{NN+}$.

Firstly, the set of normal worlds $\mathcal{N}$ is split up into a large number of *domains* $\mathfrak{D}$, nonoverlapping subsets, of worlds:

$$\mathcal{N} = \{\mathfrak{D}_1, \mathfrak{D}_2, \ldots, \mathfrak{D}_n\}$$

$$\mathfrak{D}_1 \cap \mathfrak{D}_2 \cap \ldots \cap \mathfrak{D}_n = \emptyset$$

Secondly, in place of non-normal worlds properly called, we have what I will call *buffer worlds:* $\mathcal{W} - \mathcal{N} = \mathcal{B}$. (The reason for calling them such will soon become apparent.)

Finally, and most importantly, the accessibility relation $\mathcal{R}$ is as follows.

$$\mathcal{R} = \mathfrak{D}_1 \times \mathfrak{D}_1 \cup \mathfrak{D}_2 \times \mathfrak{D}_2 \cup \ldots \cup \mathfrak{D}_n \times \mathfrak{D}_n \cup \mathfrak{D}_1 \times \mathcal{B} \cup \mathfrak{D}_2 \times \mathcal{B} \cup \ldots \cup \mathfrak{D}_n \times \mathcal{B} \cup \mathcal{B} \times \mathcal{B}$$

In words, every world within a domain accesses all other worlds within that domain, and all buffer worlds access all normal worlds and all other buffer worlds.

Admissible valuations ν, and semantics for $\Box, \Diamond, \supset, \wedge, \vee$ and so on, are as in standard modal languages. As in $\mathcal{L}_{NN}$, logical consequence is defined only over $w \in \mathcal{N}$. This way, we will have (DD) and (BB) since they will be defined only over the normal worlds. (DD) and (BB) would not hold in general if logical consequence were defined over $w \in \mathcal{B}$ as well.

With the characteristic accessibility relation R for $\mathcal{L}_{BC}$, the semantics for buffer worlds is similar to what it would be for non-normal worlds. Only, instead of the truth and falsity of diamond- and box-sentences, respectively, being a matter of definition as they are in non-normal worlds, their values are made so because of a truthmaker, or lack of one, in some world in some domain.

So, senseless propositions like there being round squares are not found to be true in any world in the model. They are not possible with respect to any buffer world. This reflects the fact that such propositions are not even conceivable, by anyone. Also, it is conceivable (alternatively, "could have turned out to be the case") that water be an element like some ancient Greeks thought, and conceivable that it be XYZ. But it is not even conceivable that it be both. So those two propositions are true only in worlds of different domains to each other.

With this structure in hand, plus the fact that as in the usual sort of non-normal model, logical consequence is defined only over normal worlds, we get the following.

(NCP) $\Diamond\Diamond p \wedge \Diamond\Diamond q \not\Vdash \Diamond\Diamond(p \wedge q)$

Recall that in $\mathcal{L}_{NN+}$, any double-diamond statement whatever comes out true. We want certain propositions to not even be considered as conceivable; hence, (NPP). The foregoing is another reason that we don't just any double-diamond statement coming out true: only certain sets of propositions are co-possible with one another; hence, (NCP).

The structure of $\mathcal{L}_{BC}$ as a whole may be pictured as an enormous grid on which domains of worlds lie, the domains each being separated from one another by a "layer" of buffer worlds; hence the name. (Any layer of buffer worlds being more than one world "thick" would be redundant.) A world in a domain accesses a world in some different domain only by the double-diamond function. Any buffer world accesses any buffer world on the "grid," and via that, any normal world in any domain.

The fact that the normal worlds are separated into domains reflects the metaphysical connections between possibilities: there can be possible variations in states of affairs, but only so much variation as long as we are still talking in terms of the same "stuffs" that constitute our actual world. We can talk about epistemic, or conceptual, possibilities in which everything looks just the same as in actuality but that, say, water, fire, air, and earth really were the elements, as occupying a different domain of possibilities. Thus, in the context of everyday discourse, *water* really rigidly designates *our* water, but we can talk about what might have been the case *if it had turned out that* water was actually an element: call it water$_e$. Water$_e$ is thus a different, and rigid, *name* for the "stuff" considered counterpossibly. So names can be used as rigid designators without requiring that, say, when actualizing a world with water in it, God had no choice but to make the water out of H_2O, which seems an unnecessarily strong form of essentialism: (BB) gets us out of that.

In other words, $\mathcal{L}_{BC}$ buys much of what motivates so-called "two-dimensional" semantics—possibilities considered differently from the perspective of worlds considered as actual from those of worlds considered as counteractual—but without the associated "logical omniscience" about consequences stemming from the *this*-worldly as considered counteractual. In its place, we simply have the failure of certain entailments, by the theorems (DD) and (BB), without having to be able to say we can truly conceive of what *is* the case at those other-domain worlds.

In modeling conceivability as noncoextensive with possibility, some *con-*

nection between conceivability and possibility is at least still wanted: many conceivabilities *are* possible, and most possibilites, it would seem, are conceivable. (Even possibilities of which *we* can't conceive could be conceivable to *some* epistemic agent, and hence may be in the true model even if we don't know it!) $\mathcal{L}_{BC}$ gives us that connection. But genuine possibilities, whether accessible to this world or not, should at least be able to have truthmakers for them. These truthmakers must be part of some maximal state of affairs: a possible *world*. $\mathcal{L}_{BC}$ gives us as many of these as we might want. But not too many.

Systems with maximal interworld accessibility, like S5, are usually taken to be those that model "broadly logical" modality, while special weakenings are invoked for special applications, like deontic or temporal logics. The present system offers an alternative means of interpretation: the model of nonoverlapping domains delimited by buffer worlds is purported to represent broadly logical modality itself. Thus, no form of "logical omniscience" is assumed, whether that comes to epistemic closure under entailment, or of knowing the essential nature of a thing or stuff just by reflection. Rather, logical omniscience is relegated to the buffer worlds, of which ours is not one.

It should be pointed out that while $\mathcal{L}_{BC}$ gets us (BB) and (DD) instead of (S4) and (S4p), the converses of each of (S4) and (S4p) *do* hold. This seems natural for the converse of (S4), as a statement about what is conceptually necessary at all worlds in all domains. But the converse of (S4p), given the current semantics, says that whatever is (this-domainly) possible is conceivable. But this is as it should be, I think. p being conceivable, in the present sense, requires not that anyone actually does conceive of p, but that it cannot be ruled out. And if p is genuinely possible, no epistemic agent will encounter anything that will enable her to rule out that p—unless she is mistaken, of course, in which case she will be mistaken about what she is conceiving of. So this kind of error can be accounted for in this system, just as can the fallibility of possibility judgments from conceivability, and the possibility of modal error. But where there are true modal facts coming from this-worldly and this-domainly entailments, when we come to know these, they form the basis of our genuine modal knowledge. So, we can make true modal judgments about this-worldly, this-domainly things about which we are sufficiently acquainted. If water = H_2O, it is necessarily so. Yet there are other ways things could have turned out, and we can talk about those things, too, without assuming we know more than what the actual world can tell us.

References

Bealer, George, "Modal Epistemology and the Rationalist Renaissance," in Tamar Szabó Gendler and John Hawthorne, eds., *Conceivability and Possibility,* OUP, 2002

Hintikka, Jaakko, "Impossible Possible Worlds Vindicated," *Journal of Philosophical Logic* **4** (1975) 475–484

Kripke, Saul A., "Semantical Analysis of Modal Logic II. Non-Normal Modal Propositional Calculi," in J.W. Addison, Leon Henkin, and Alfred Tarski, eds., *The Theory of Models,* Amsterdam: N. Holland Publishing Co., 1965

Plantinga, Alvin, *The Nature of Necessity,* OUP, 1974

Rantala, Veikko, "Impossible Worlds Semantics and Logical Omniscience," *Acta Philosophica Fennica* **35** (1982a) 106–115

Rantala, Veikko, "Quantified Modal Logic: Non-Normal Worlds and Propositional Attitudes," *Studia Logica* **41** (1982b) 41–66

Yablo, Stephen, "Coulda, Woulda, Shoulda," in Tamar Szabó Gendler and John Hawthorne, eds., *Conceivability and Possibility,* OUP, 2002