

Bounds for metaphysical modality

Paolo Santorio

University of Maryland, College Park

paolosantorio@gmail.com

Abstract

Epistemic modals exhibit a kind of nonclassical behavior, exemplified by epistemic contradictions like (1). Conversely, metaphysical modals seemingly conform to the predictions of semantics based on classical modal logics, as witnessed by the consistency of (2).

- (1) # Al is tall and he might not be. (2) Al is tall and he might not have been.

I argue that this difference is illusory. Metaphysical modals display nonclassical behavior mirroring that of epistemics. This nonclassicality is manifested in puzzles concerning the interaction of *would* and *might*. I show how these facts be captured, while still preserving the contrast between (1) and (2), building on Mandelkern’s recent work on bounds.

1 Introduction

Epistemic modals notoriously exhibit a kind of nonclassical behavior, which is hard to capture on semantic theories based on Kripkean accessibility relations (Kratzer 1981, 2012 a.o.). This nonclassicality is typically exemplified by so-called epistemic contradictions, i.e. sentences of the form $\ulcorner A \wedge \Diamond \neg A \urcorner$ (Veltman 1985; Yalcin 2007; Willer 2013; Mandelkern 2019; Mandelkern 2024 a.o.). These sentences are consistent by the lights of classical theories, but they are infelicitous. For a simple example, consider (1).

- (1) #Ada is tall and she might not be.

Epistemic modality is taken to be unique in this behavior. Seemingly, other modal flavors are better behaved and conform to the predictions of a modal semantics based on standard modal logics. In particular, the flavor of modality that is often labeled ‘metaphysical’, and that includes both *would*-conditionals and contrary-to-fact *might*-claims, is usually considered a paradigm of classicality. At first sight, this is supported by empirical evidence. The metaphysical counterparts of epistemic contradictions, such as (2), are perfectly felicitous.

- (2) Ada is tall and she might not have been.

This short paper argues that, contrary to first appearances, the divergence between epistemic and metaphysical modality is largely illusory. Metaphysical modals display a kind of nonclassical behavior that mirrors that of epistemic modals. This kind of nonclassicality is manifested in puzzles concerning the interaction of *would* and *might*, some well-known and some new. Moreover, this nonclassicality can be captured in a theory that generalizes current theories of epistemic modals. For current purposes, I build on Mandelkern’s (2019, 2024) recent proposal about epistemic modality, which ties the domain of epistemic modals to the notion of local context. I state the motivating data in §2 and the positive account in §3. §4 concludes by discussing some general theoretical morals.

2 Nonclassicality for metaphysical modals: two puzzles

By ‘metaphysical modality’ I mean the modal flavor expressed by *would*-conditionals, non-epistemic *might*-conditionals, and non-epistemic *might* claims. Some examples are in (3).

- (3) a. If the die had landed odd, it would not have landed on two.
 b. If the die had landed odd, it might have landed on three.
 c. The die might have landed on three.

As often noticed, this flavor of modality involve characteristic tense and aspectual marking, which has come to be called ‘X-marking’ (Iatridou 2000, von Stechow and Iatridou 2023, a.o.).¹

Given restrictions of space, here I cannot make a full case for the claim that metaphysical modals are nonclassical. But I present two puzzles. Taken together, they are suggestive of the idea that there is an analogy between epistemic and metaphysical modals.

2.1 Puzzle #1: *would-might* interactions

The first puzzle I discuss is a variant on a puzzle recently presented by Greenberg 2021.² To start, consider sentence (4), as uttered in the following scenario.

A die was tossed and landed 1. You don’t know anything about whether the die was fair or biased in any way. I say:

- (4) #The die might have landed on 2, and if it had landed even, it would not have landed on 2.

(4) sounds infelicitous. Notice also that this infelicity is generated by the fact that the two conjuncts clash with each other. Given the information available in the scenario, each of the two conjuncts of (4) might be true, and in any case could be uttered felicitously. But, once we conjoin them, we obtain something that is clearly infelicitous.

The first conjunct of (4) is equivalent to *The die might have landed even and 2*. Hence (4) suggests that conjunctions of the form $\ulcorner \Diamond(A \wedge B) \wedge (A \Box \rightarrow \neg B) \urcorner$ are, in some sense, inconsistent.

Unconditional Might. (UM) $\Diamond(A \wedge B) \wedge (A \Box \rightarrow \neg B) \models \perp$

To support this, notice that, like (4), all instances of $\ulcorner \Diamond(A \wedge B) \wedge (A \Box \rightarrow \neg B) \urcorner$ are infelicitous.

- (5) a. #Ada and Bashir might have both come to the party, and, if Ada had come, Bashir would not have come.
 b. #Coins A and B might have both landed heads, and, if coin A had landed heads, coin B would not have landed heads.
 c. #Xintong and Yimei might have both been in DC, and, if Xintong had been in DC, Yimei would not have been in DC.

Interestingly, the infelicity of (4) is unexpected on many standard theories. Assume a simple Lewis-style semantics for counterfactuals with limit assumption.³

- (6) $\llbracket A \Box \rightarrow B \rrbracket^{w,g} = \text{true}$ iff $\forall w' \in \text{BEST}_{w,g}(\llbracket A \rrbracket_g), \llbracket B \rrbracket^{w',g} = \text{true}$

Assume also a simple semantics for metaphysical *might*: *might* is an existential quantifier over a contextually determined domain of metaphysically accessible worlds. Then the infelicity of (4) is

¹For current purposes, I set aside *could* and *could have*. While they are obviously related to *would* and *might*, it is unclear that they express unambiguously metaphysical modality. *can*, which is the non-X-marked counterpart of *could*, is notoriously polysemous, and has a habitual, circumstantial, epistemic, and ability interpretations.

²Greenberg’s original puzzle involves patterns of entailment between the conjuncts of sentences like (4). Using conjunctions makes things clearer for my purposes, and actually improves the quality of the data.

³For discussion of the limit assumption see, among many, Lewis 1973, Stalnaker 1981, Kaufmann 2017. I assume the limit assumption throughout my discussion. Also, a word about notation: ‘*g*’ is a variable for orderings over worlds. I use the notation $\llbracket A \rrbracket_g$ to denote the set of worlds $\{w : \llbracket A \rrbracket^{w,g}\}$, i.e. essentially the possible worlds proposition expressed by *A*, relative to ordering *g*.

unexpected. It can be that the closest *A*-worlds in the domain are non-*B*-worlds, but there are some *A*-and-*B* worlds further off.

To be sure, some versions of classical theories do predict the infelicity of (4). Assume that metaphysical *might* invariably quantifies over a set of ‘maximally similar’ worlds.⁴ Then the first conjunct of (4) says that one of the most similar worlds to the actual world is a die-landing-two worlds. This is indeed incompatible with standard variably strict truth-conditions. But, as Greenberg points out, making *might* strong in this way creates trouble elsewhere. Sentences of the form $\lceil \neg \Diamond(A \wedge B) \wedge (A \Box \rightarrow B) \rceil$, such as (7), are wrongly predicted to be consistent.

- (7) # The die mightn’t have landed 2, and if it had landed even, it would have landed 2.

2.2 Puzzle #2: CEM and Conditional Might

The second puzzle is familiar from the classical debate on counterfactuals (Stalnaker 1968, Stalnaker 1981, Lewis 1973). Counterfactuals seem to validate two plausible principles.

Conditional Excluded Middle. (CEM) $\models (A \Box \rightarrow B) \vee (A \Box \rightarrow \neg B)$

Conditional Might. (CM) $(A \Box \rightarrow B) \wedge (A \Diamond \rightarrow \neg B) \models \perp$

CEM has been subject to extensive empirical discussion; at this point, the case for it is very strong (see, a.o., Klinedinst 2011, Marty, Romoli, and Santorio 2021, Ramotowska et al. 2024). The core evidence for CEM concerns interactions between counterfactuals and other logical operators. In particular, semantics that conform to CEM predict that counterfactuals can commute with a number of logical operators, including negation, without affecting meaning. This prediction is borne out. For example, assuming that *lose* is equivalent to *not win*, the sentences in (8) sound equivalent.

- (8) a. Each of these tickets would have lost, if it had been bought. $\forall x(Buy(x) \Box \rightarrow \neg Win(x))$
 b. None of the tickets would have won, if it had been bought. $\forall x \neg (Buy(x) \Box \rightarrow Win(x))$

As for **CM**, the evidence for it is simple: conjunctions like (9) sound straight-up contradictory.

- (9) #If the die had landed even, it would have landed 2; and if it had landed even, it mightn’t have landed 2.

Unfortunately, **CEM** and **CM** are jointly untenable against a classical background, as they entail that *would* and *might*-counterfactuals are equivalent ($A \Box \rightarrow B \models A \Diamond \rightarrow B$). And in fact, classical theories for counterfactuals choose one between **CEM** and **CM**, and reject the other. Famously, Stalnaker validates **CEM** and rejects **CM** (Stalnaker 1968, Stalnaker 1981), while Lewis takes the opposite route (Lewis 1973).

3 Bounded semantics for metaphysical modality

I suggest that the puzzles of §2 show that metaphysical modals behave nonclassically. Moreover, the relevant kind of nonclassicality parallels the nonclassical behavior that motivates dynamic theories of epistemic modals. Consider again epistemic contradictions like (1):

- (1) #Ada is tall and she might not be.

⁴On some interpretations, this is the view defended by Kratzer (1981, 2012, a.o.). It is unclear to me whether Kratzer would want to endorse the assumption that all occurrences of *might* quantify over closest worlds, without allowing context to relax this constraint.

Theories in the dynamic tradition vary widely, but they share a basic explanation for the infelicity of (1). Roughly, in (1) the right conjunct is evaluated against a set of worlds that has been updated with the other conjunct. As a result, the possibility modal quantifies over a set of worlds where Ada is tall, and the possibility claim is false. This despite the fact that there is a clear sense in which the conjuncts of (1) are consistent. It is consistent that Ada is tall, and that the set of worlds representing speakers' information includes a world where she is not.

I suggest that something structurally analogous happens for metaphysical modality. Consider the problem conjunction (4):

- (4) #The die might have landed on 2, and if it had landed even, it would not have landed on 2.

I suggest that the basic truth conditions of the two conjuncts of (4) are simple. The first says that there is a metaphysically accessible world where the die landed on 2. The second says that, in the closest world where the die lands even, it lands on 2. Hence, in a sense, they are consistent. But, in the conjunction, the quantificational domains of the modals are constrained in a way that makes the sentence infelicitous. In particular, the first conjunct is evaluated against a set of worlds that is updated with the second conjunct, i.e. the counterfactual $\lceil \text{even} \Box \rightarrow \neg \text{two} \rceil$. Given basic logical facts about counterfactuals⁵, this has the effect of removing all even-and-two-worlds from the set. Hence the possibility modal quantifies over a set of worlds where the die did not land on 2, and is guaranteed to be false.

The remainder of the paper sketches a formal implementation of the idea. The system builds on the 'bounded theory' developed by Mandelkern (2019, 2024), *modulo* some simplifications.⁶ Aside from the technical analogies, the proposal involves the introduction of a new parameter, i.e. what I call 'metaphysical domain'. I first discuss metaphysical domains, and then present the formal system.

3.1 Metaphysical domains

The literature on presupposition introduces the notion of a local context (Stalnaker 1970, Karttunen 1974, Schlenker 2009 a.o.). Roughly, a local context is a set of epistemic possibilities that are available in the compositional computation, and are partly determined by surrounding sentential material. For example, the local context of the second conjunct in (10) is determined by the initial context set, updated with the proposition that France has a king.

- (10) France has a king, and the king of France is a snob.

Mandelkern (2019, 2024) has recently shown that the notion of local context is linked in interesting ways to the semantics of epistemic modality.⁷ In particular, the domain of quantification of epistemic modals appears to be constrained by the local context.

I propose that the interpretation of non-epistemic modality is similarly constrained. I introduce a non-epistemic counterpart of local contexts, which I call 'metaphysical domain' (MD for short). The notion of MD can be linked to the notion of openness, which is employed in the semantics for historical modals (see Copley 2009, Cariani and Santorio 2018 a.o.). At every point in time, we can individuate a set of worlds that represent 'open possibilities': roughly, ways that the future course of history might evolve. The set of open worlds at a context, OPEN_c ,

⁵In particular, the fact that counterfactuals (at least, non-nested counterfactuals) validate Weak Centering, i.e. they entail the corresponding material conditional ($A \Box \rightarrow B \models A \supset B$).

⁶I use Mandelkern's system simply because it is currently the best architecture we have to capture the behavior of epistemic contradictions and related phenomena. In principle, the idea can be implemented in other frameworks.

⁷Strictly speaking, Mandelkern's bounded theory does not exploit a notion of local context, but rather a closely related notion of a bound. Conceptually, the two notions are very closely related.

provides the starting value for the MD parameter at c . This parameter is then shifted by various operator in the course of the compositional computation.

Crucially, some items can expand the metaphysical domain. In particular, several authors have noticed that X-marked conditionals exploit a broader domain of quantification than non-X-marked ones (see e.g. Stalnaker 1975, Stalnaker 1988, von Fintel and Iatridou 2023). I assume that this expansion operation is tied to PAST, and represent its effects via a dedicated operator ‘ \uparrow ’. $\uparrow D$ is the domain we obtain by expanding D .⁸

The expansion operation introduces a complication. To track both whether the domain has been expanded and what information we have from the surrounding bits of the sentences, we cannot track the MD as a single set of worlds. Rather, we need to use a pair $\langle D, I \rangle$. The first member, D (the ‘background domain’), tracks whether expansion has happened. The second member, I , represents the information added during composition. The MD is determined by the intersection of the two sets. For a simple example, consider:

(11) The die will land odd and the coin will land even.

I assume that the MD of each conjunct is updated with the information provided by the other conjunct (see below). Hence the MD of the right conjunct is the pair consisting of D , and I augmented with the proposition **odd**. Assuming that the initial values of D and I are, respectively, $OPEN_c$ and the set of all worlds, the MD of the right conjunct is $OPEN_c \cap \mathbf{odd}$, i.e. the set of open worlds at the context where the die will land odd.

3.2 Formal system

The formal system assigns three layers of meaning to each sentence: (i) classical truth-conditionals; (ii) definedness conditions; (iii) an update function for the MD parameter.

Logical Forms. I assume that, at LF, both *would*- and *might*-claims involve an expansion operator ‘ \uparrow ’. (Arguably, the expansion operator is realized by ‘PAST’; see Santorio 2024 for a theory of X-marking along these lines). For current purposes, I assume that the LFs of (12-a) and (13-a) are in (12-b) and (13-b). I will also use the shorthand in (12-c) and (13-c).⁹

- | | | |
|------|--|--|
| (12) | a. If die had landed even, it would have landed 2.
b. \uparrow [the die ₁ land even $\Box \rightarrow$ it ₁ land 2]
c. the die ₁ land even $\Box \rightarrow \uparrow$ it ₁ land 2 | a. The die might have landed 2.
b. \uparrow [\Diamond (the die land 2)]
c. $\Diamond \uparrow$ (the die land 2) |
|------|--|--|

Layer 1: truth-conditions. At the basic layer, the semantics exploits just basic truth conditions. (For simplicity, I ignore the truth-conditional role of tense.)

Counterfactuals are selectional (Stalnaker 1968): a selection function s selects an antecedent-world within a modal base $f(w)$.

(14) $\llbracket A \Box \rightarrow \uparrow B \rrbracket^{c,s,f,w,D} = \text{true}$ iff (i) it is defined and (ii) $\llbracket B \rrbracket^{c,s,f,s(\llbracket A \rrbracket \cap f(w),w),D} = \text{true}$

I assume that the selection function obeys Centering: whenever $w \in \llbracket A \rrbracket$, $s(\llbracket A \rrbracket, w) = w$.

Possibility modals work simply as a existential quantifiers over a modal base $f(w)$.

(15) $\llbracket \Diamond \uparrow A \rrbracket^{c,s,f,w,D} = \text{true}$ iff (i) it is defined and (ii) $\exists w' \in f(w)$ s.t. $\llbracket B \rrbracket^{c,s,f,w',D} = \text{true}$

⁸For current purposes, it doesn’t matter exactly what the expanded set is. For convenience, we can just take it to be the set of all metaphysically possible worlds. (This assumption entails that iterating the expansion operator doesn’t lead to further expansions of the domain. So far as I can see, this seems a desirable consequence.)

⁹For the case of conditionals, this is a simplification: there is evidence that sentences like (12-a) involve two semantically active occurrences of PAST, which are interpreted locally in the antecedent and the consequent (see Santorio 2024). But this is irrelevant here.

Layer 2: definedness conditions. The second layer assigns recursively stated definedness conditions to sentences. (Definedness conditions can, but need not, be equated with presuppositions.) Following Mandelkern’s proposal about epistemic modals (Mandelkern 2019, Mandelkern 2024), the definedness conditions for modals requires that the modal base be a subset of the MD. The expansion operator maps D to an expanded domain $\uparrow D$.

Connectives $\lceil \neg A \rceil$ is defined at $\langle D, I \rangle$ iff A is defined at $\langle D, I \rangle$

$\lceil A \wedge B \rceil$ is defined at $\langle D, I \rangle$ iff A is defined at $\langle D, I \rangle + B$ and B is defined at $\langle D, I \rangle + A$

Exp. op. $\lceil \uparrow(A) \rceil$ is defined at $\langle D, I \rangle$ iff A is defined at $\langle \uparrow D, I \rangle$

Modals $\lceil \Diamond(A) \rceil$ is defined at $\langle D, I \rangle$ iff, for all $w \in D \cap I$, $f(w) \subseteq D \cap I$, and A is defined at $\langle D, I \rangle$

$\lceil A \Box \rightarrow B \rceil$ is defined at $\langle D, I \rangle$ iff (i) for all $w \in D \cap I$, $f(w) \subseteq D \cap I$; (ii) for all $w \in D \cap I$, $s(\llbracket A \rrbracket, w) \in D \cap I$; (iii) A and B are defined at $\langle D, I \rangle$.

Layer 3: updates. Let $\llbracket A \rrbracket_c$ be the proposition expressed by A at c . We have:

(i) For any A that is not of the form ‘ $\uparrow A$ ’: $\langle D, I \rangle + A = \langle D, I \cap (\overline{D} \cup \llbracket A \rrbracket_c) \rangle$

(ii) $\langle D, I \rangle + \uparrow A = \langle \uparrow D, I \rangle + A$

Update with regular sentences is essentially intersective update, with a minor twist.¹⁰ Update with a sentence involving an expansion operator triggers expansion of the background domain.

Consequence. I define consequence as preservation of definedness and truth.

(16) $A_1, \dots, A_n \models B$ iff, for all c s.t. A_1, \dots, A_n are def. and true at c , B is def. and true at c .

Predictions. Let us consider now some predictions.

(i) Instances of **Unconditional Might**, i.e. sentences of the form $\lceil \Diamond_{\uparrow}(A \wedge B) \wedge (A \Box \rightarrow_{\uparrow} \neg B) \rceil$ are never both defined and true. This explains the infelicity of the sentences in (4) and (5).

(ii) Conversely, sentences of the form $\lceil A \wedge \Diamond_{\uparrow} \neg A \rceil$ like (2) can be true and defined. The key reason (and the key difference with the case of epistemic modals) is that the expansion operator widens the metaphysical domain beyond the initial set OPEN_c .¹¹

(iii) **Conditional Excluded Middle** and **CM** both come out valid on this system (all instances of **CEM** are true, while instances of **CM** are always true if defined). This vindicates the motivating data in §2.2.

4 Conclusion

Nonclassical accounts of epistemic modals seem to require a major bifurcation in modal semantics: modals of different flavors work in very different ways. The current theory shows that we can recover the unity of modal semantics by generalizing mechanisms used to account for epistemic contradictions. Independently of whether the current system is ultimately correct, the fact that this can be done offers an important proof of concept, and enriches the space of options for our semantics for modality.

¹⁰The twist is that we update the information parameter with the proposition expressed by A disjoined with the complement of D . This is to prevent the update from affecting non- D worlds when the domain is expanded.

¹¹The computation of the MD of $\lceil A \wedge \Diamond_{\uparrow} \neg A \rceil$ works as follows. (i) The initial MD of the sentence is OPEN_c . (ii) Via the conjunction clause, the MD of the right conjunct is $\langle D, I \rangle + A$, i.e. (via the definition of update) $\langle D, I \cap (\overline{D} \cup \llbracket A \rrbracket_c) \rangle$. At this point, the MD consists of only A -worlds. (iii) But the expansion operator in the second conjunct expands the MD to $\langle \uparrow D, I \cap (\overline{D} \cup \llbracket A \rrbracket_c) \rangle$, hence to $\uparrow D \cap (I \cap (\overline{D} \cup \llbracket A \rrbracket_c))$. This set can include non- A worlds.

References

- Cariani, Fabrizio and Paolo Santorio (2018). “Will Done Better: Selection Semantics, Future Credence, and Indeterminacy”. In: *Mind* 127.505, pp. 129–165.
- Copley, Bridget (2009). *The Semantics of the Future*. New York: Routledge.
- von Fintel, Kai and Sabine Iatridou (2023). “Prolegomena to a theory of X-marking”. In: *Linguistics and Philosophy* 46.6, pp. 1467–1510.
- Greenberg, Gabriel (2021). “Counterfactuals and modality”. In: *Linguistics and Philosophy* 44.6, pp. 1255–1280.
- Iatridou, Sabine (2000). “The grammatical ingredients of counterfactuality”. In: *Linguistic inquiry* 31.2, pp. 231–270.
- Karttunen, Lauri (1974). In: *Theoretical Linguistics* 1.1-3, pp. 181–194. DOI: doi:10.1515/thli.1974.1.1-3.181. URL: <https://doi.org/10.1515/thli.1974.1.1-3.181>.
- Kaufmann, Stefan (2017). “The Limit Assumption”. In: *Semantics and Pragmatics* 10.18. DOI: 10.3765/sp.10.18.
- Klinedinst, Nathan (2011). “Quantified Conditionals and Conditional Excluded Middle”. In: *Journal of Semantics* 28.1, pp. 149–170. DOI: 10.1093/jos/ffq015.
- Kratzer, Angelika (1981). “Partition and Revision: The Semantics of Counterfactuals”. In: *Journal of Philosophical Logic* 10.2, pp. 201–216.
- (2012). *Modals and Conditionals: New and Revised Perspectives*. Vol. 36. Oxford University Press.
- Lewis, David K. (1973). *Counterfactuals*. Cambridge, MA: Harvard University Press.
- Mandelkern, Matthew (2019). “Bounded Modality”. In: *Philosophical Review* 128.1, pp. 1–61. DOI: 10.1215/00318108-7213001.
- (2024). *Bounded Meaning: The Dynamics of Interpretation*. Oxford University Press.
- Marty, Paul, Jacopo Romoli, and Paolo Santorio (2021). “Counterfactuals and Undefinedness: Homogeneity vs. Supervaluations”. In: *Proceedings of SALT*. Vol. 30, pp. 603–623.
- Ramotowska, Sonia et al. (2024). “Undefined Counterfactuals: Experimental Evidence for Selection Semantics”. draft.
- Santorio, Paolo (2024). “Modalizing the Past: A Theory of C-Marking”. draft, University of Maryland, College Park.
- Schlenker, Philippe (July 2009). “Local Contexts”. In: *Semantics and Pragmatics* 2.3, pp. 1–78. DOI: 10.3765/sp.2.3.
- Stalnaker, Robert (1968). “A Theory of Conditionals”. In: *Studies in Logical Theory*. Ed. by N. Recher. Oxford.
- (1970). “Pragmatics”. In: *Synthese* 22.1-2, pp. 272–289. DOI: 10.1007/bf00413603.
- (1975). “Indicative Conditionals”. In: *Philosophia* 5.
- (1981). “A Defense of Conditional Excluded Middle”. In: *Ifs*. Ed. by William Harper, Robert C. Stalnaker, and Glenn Pearce. Reidel, pp. 87–104.
- (1988). “Belief Attribution and Context”. In: *Contents of Thought*. Ed. by Robert H. Grimm and Daniel Davy Merrill. Tucson, pp. 140–156.
- Veltman, Frank (1985). “Logic for Conditionals”. PhD thesis. University of Amsterdam.
- Willer, Malte (2013). “Dynamics of Epistemic Modality”. In: *Philosophical Review* 122.1, pp. 45–92.
- Yalcin, Seth (2007). “Epistemic Modals”. In: *Mind* 116.464, pp. 983–1026.