One's *Modus Ponens*: Modality, Coherence and Logic*†

Forthcoming in *Philosophy and Phenomenological Research*

Una Stojnić
Department of Philosophy and Center for Cognitive Science,
Rutgers University

October 3, 2016

Abstract

Recently, there has been a shift away from traditional truth-conditional accounts of meaning towards non-truth-conditional ones, e.g., expressivism, relativism and certain forms of dynamic semantics. Fueling this trend is some puzzling behavior of modal discourse. One particularly surprising manifestation of such behavior is the alleged failure of some of the most entrenched classical rules of inference; viz., *modus ponens* and *modus tollens*. These revisionary, non-truth-conditional accounts tout these failures, and the alleged tension between the behavior of modal vocabulary and classical logic, as data in support of their departure from tradition, since the revisionary semantics invalidate some of these patterns. I, instead, offer a semantics for modality with the resources to accommodate the puzzling data while preserving classical logic, thus affirming the tradition that modals express ordinary truth-conditional content. My account shows that the real lesson of the apparent counterexamples is not the one the critics draw, but rather one they missed: namely, that there are linguistic mechanisms, reflected in the logical form, that affect the interpretation of modal language in a context in a systematic and precise way, which have to be captured by any adequate semantic account of the interaction between discourse context and modal vocabulary. The semantic theory I develop specifies these mechanisms and captures precisely how they affect the interpretation of modals in a context, and do so in a way that both explains the appearance of the putative counterexamples and preserves classical logic.

*Acknowledgments:** Earlier versions of this paper were presented at various conferences and departmental colloquia. I would like to thank all of these audiences. I am indebted to Chris Barker, Justin Bledin, Elisabeth Camp, Simon Charlow, Andy Egan, Graeme Forbes, Thony Gillies, Gil Harman, John Hawthorne, Ben Holguin, Nathan Howard, Angelika Kratzer, Barbara Partee, Jim Pryor, Craige Roberts, Timothy Williamson, and two anonymous referees for *Philosophy and Phenomenological Research* for comments on earlier drafts. I am especially indebted to Jeff King, Ernie Lepore and Matthew Stone for numerous discussions from the very early stages of this project and extensive comments on previous drafts.

†The paper is structured in two parts, followed by a formal appendix. In Part I (§ 1), I develop my account of modality. Part II (§ 2) describes the formal tools that allow us to build the semantics described in Part I. A formal appendix, (§ A) provides the complete formalization, along with a proof that the semantics preserves classical logic. Part II and the appendix should thus be seen as complementing Part I, by providing a formal implementation of the account developed in Part I.

0 Introduction

Recently, there has been a shift away from traditional truth conditional accounts of meaning towards non-truth-conditional ones, e.g., expressivism, relativism and certain forms of dynamic semantics. Fueling this trend are some puzzling behavior of modal discourse. One particularly surprising manifestation of such behavior is the alleged failure of some of the most entrenched classical rules of inference; viz., *modus ponens* (MP) and *modus tollens* (MT). Thus, several authors have independently touted counterexamples to MP and MT.[1] Because each challenge arises in the presence of modal language (typically involving embedded modals or conditionals), these critics believe they have uncovered a tension between the behavior of modal vocabulary and classical logic, with the moral being to revise the semantics for modality in order to invalidate certain classical patterns. These revisionary, non-truth-conditional accounts tout the alleged tension between the behavior of modal vocabulary in natural language and classical logic, as data in support of their departure from tradition, since the revisionary semantics invalidate some of these patterns.[2] I, instead, offer a semantics for modality with the resources to accommodate the puzzling data while preserving classical logic, thus affirming the tradition that modals express ordinary truth-conditional content. My account shows that the real lesson of the apparent counterexamples is not the one the critics draw, but rather one they missed: namely, that there are linguistic mechanisms, reflected in the logical form of a discourse, that affect the interpretation of modal language in a context in a systematic and precise way, which have to be captured by any adequate semantic account of the interaction between discourse context and modal vocabulary. I develop and defend a theory that captures this interaction: according to my account, context affects the interpretation of modals through mechanisms that are independently motivated, and are required to explain the effects of context on the interpretation of context-sensitive expressions quite generally, even in the most basic case—the case of pronouns. In particular, I shall argue, the relevant mechanisms are the ones that govern how individual utterances are organized to form a coherent discourse. While the impacts of these mechanisms explain (and predict) the appearance of counterexamples, the underlying logic, as I shall prove, is classical.[3]

In short: I defend a theory that captures how context affects the interpretation of modals, and does so in a way that reconciles classical logic with the semantics of modal language. My

[1] See in particular McGee (1985) and Willer (2010) for counterexamples to MP, and Yalcin (2012) and Veltman (1985) for counterexamples to MT. See also Kolodny and MacFarlane (2010) who reject MP as a reaction to certain puzzling behavior of deontic modals, and Cantwell (2008) who argues that MP, MT and reasoning by cases all fail when they involve modals in the consequents of conditionals. The revisionary accounts are typically non-truth-conditional, insofar as they deny that an utterance expresses a proposition that is true or false depending on the way the world is. In other words, these accounts deny utterances express propositional content. An exception to this is McGee (1985) , who rejects MP, and thus endorses a non-classical account, but one that is nevertheless not non-truth-conditional. In this respect, it differs significantly from these other accounts, which invalidate the relevant patterns of inference in part by rejecting truth-conditionality of modal vocabulary.

[2] This is the course explicitly taken by Yalcin (2012), who appeals to his counterexample as a basis for a rejection of MT. Some relativists also reject MP and/or MT based on the puzzling behavior of modal language (Kolodny and MacFarlane, 2010). And some dynamic semanticists, too, invalidate MT, where the diagnosis for rejection lies once again in the behavior of modal language (Gillies, 2010, 2004). I should note that, though the example I will focus on involves epistemic modality, the problem arises for other flavors of modality, as well; in particular, deontic modals, too, exhibit the behavior that *prima facie* violates classical patterns of inference. (See e.g. Kolodny and MacFarlane (2010).) The theory I will develop is not limited to epistemic modality, and my treatment of the apparent failure of classical patterns of inference naturally extends to other examples involving other types of modality, as well.

[3] To be precise, the underlying logic is a classical modal logic—a simple extension of classical logic with a modal necessity operator (in particular, system S_4.) This logic preserves a classical inference system, and is sound and complete. Importantly, unlike another famous attempt at a defense of classical logic, pioneered by H. P. Grice (1989), I shall not be arguing that the truth-conditions of the English indicative conditional are those of the material conditional, but rather those of the strict conditional.

strategy will be to focus on one particular case, due to Yalcin, but the account I develop naturally accommodates other known cases as well.[4] The paper divides into two parts. In part § 1, I argue that a systematic impact of discourse context on the interpretation of modals explains the apparent counterexample, through general mechanisms, while preserving the validity of MT, and I sketch an account of modality and context-change that captures this interplay between discourse context, and modal expressions. In part § 2, I demonstrate that we can formalize this systematic impact, in a way that explains the counterexample, and preserves classical logic.

1 Part I

1.1 The Challenge to *Modus Tollens*

The pattern [*If* p, q; $\neg q \therefore \neg p$] is known as *modus tollens*. Yalcin argues for its invalidity as follows:

> Take an urn with a 100 marbles. 10 of them are big and blue, 30 big and red, 50 small and blue, and 10 are small and red. One marble is randomly selected and hidden (you do not know which). Given this setup, (1) and (2) are licensed:

(1) If the marble is big, then it is likely red.

(2) The marble is not likely red.

But, surely, (3) does not follow:

(3) So, the marble is not big.

Since the inference from (1) and (2) to (3) looks to be an application of MT, we would seem to have a counterexample.[5]

Possible reactions to the apparent counterexample are:

i embrace the invalidity of MT, or

ii deny that (1)–(3) is a genuine instance of MT.

(i) requires modifying our semantics for modals in such a way so as to invalidate MT. This is Yalcin's option, one that embraces a "revisionary" semantics for modal language that denies that modal and conditional sentences express ordinary truth-conditional content.[6] It is important to note that these revisionary frameworks do not deny that, intuitively, the big premise in (1)–(3) is (in some sense) about conditional probability, i.e., that it concerns the probability of the marble being red, conditional on it being big, while the small premise is, intuitively, 'unrestricted' in this sense. Everyone in the debate concedes this. What the revisionists deny, however, is that

[4] E.g., the apparent counterexamples due to McGee (1985) and Kolodny and MacFarlane (2010).

[5] It is easy to construct other counterexamples along similar lines. See Veltman (1985) for an earlier counterexample to MT with right-nested conditionals. I shall focus on Yalcin, but my considerations extend to Veltman.

[6] See e.g. Yalcin (2012), Yalcin (2007) and Moss (2015) for an expressivist version of revisionary semantics, see e.g. Kolodny and MacFarlane (2010) for a relativist version, and see Gillies (2010, 2004) for dynamic semantics version. These semantics are 'revisionary' precisely insofar as they invalidate classical patterns of inference, and deny that modal discourse expresses standard truth-conditions. It is worth noting that, although Yalcin revises the standard compositional semantics for modals and conditionals, he defends expressivism as a *pragmatic* thesis. In particular, he does not hold that the semantic content of a modal utterance is the informational content the utterance expresses; in fact, he denies that modal utterances express any informational content. The semantics I will develop and defend will require a substantial departure from the traditional semantics for modal discourse (as in Kratzer (1977), and Kratzer (1983)), but one that allows us to vindicate the idea that modal utterances express truth-conditional content. Thus, my semantics is conservative, insofar as it preserves classical patterns of inference, and the classical truth-conditions for modal vocabulary.

this intuition can be adequately captured by saying that the *truth-conditional content* of the big premise *describes* the conditional probability, while the *truth-conditional content* of the small premise *describes* the unrestricted one. They all argue, in one way or another, that there is no plausible and systematic way of deriving the right truth-conditions given a context of utterance.[7]

Note that, though it might seem that the revisionary tendencies are localized to particular counterexamples to a particular inference pattern—MT, this reaction opens the floodgates: it's easy to devise similar counterexamples to numerous other inference patterns that are classically valid. The question then becomes which of the deductive rules of inference we should reject. To illustrate, consider the same scenario as in Yalcin's original counterexample: there's an urn with 100 marbles, etc. But then:

(4) Suppose that the marble is big.
(5) Then it is likely red.
(6) But the marble is not likely red.
(7) So, the marble is not big.

(4)–(7) is also horrible. If (1)–(3) provides grounds for rejecting MT, then by parity (4)–(7) provides grounds for rejecting MP or *reductio*. We might then reject both *modus ponens* and *modus tollens* (as in fact Kolodny and MacFarlane (2010) do). Or we might reject *reductio*. Should we reject all of these rules? Of course, we can derive MT through MP and *reductio*, on the assumption of the monotonicity of logical consequence. More generally, we know that rules of inference are holistic in this sense. But that is part of the problem—how do we isolate the culprit(s)? How do we choose?

None of this is, strictly speaking, an argument against option (i), since, the proponents of the revisionary views already accept that our modal vocabulary is at odds with classical logic.[8] But it does show that there is a lot more at stake than the loss of MT. Moreover, as we shall see shortly, contrary to what the proponents of the revisionary views assume, the problem is *not* tied specifically to modal vocabulary, but can be replicated with the most basic case of context-sensitive expressions—demonstrative pronouns. This suggests that the real problem may lie elsewhere. A theory of context-sensitivity resolution that preserves the aforementioned inference rules, while explaining away the appearance of counterexamples has a lot going for it. Since I am precisely interested in developing such a theory, I shall reject option (i).

One way to pursue option (ii) that has received some attention in the literature is to claim that epistemic modals—in particular, the modal operator 'likely'—takes obligatory wide-scope over the conditional, so that (1)–(3) is not a genuine instance of MT.[9]

If this were the only way to implement option (ii), we would face a hard choice indeed, for obligatory wide-scoping faces well-known problems. For one, it generates intuitively incorrect predictions. For example, on the (supposedly obligatory) wide-scope reading of the modal 'likely', we get the intuitively wrong interpretation for (8).[10]

[7] Again, assuming the standard notion of truth-conditions.

[8] A number of authors have suggested that the failure of MT is best explained by adopting the informational approach to logic and consequence relation. We can then provide a proof system for such logic, and study its relations to classical logic. See e.g. Yalcin (2007, 2012); Bledin (2014, forthcoming). See also Veltman (1985); Gillies (2004); Willer (2010) for a related approach to logical consequence. These authors point out that various classically valid inference patters—MT and *reductio* for instance—are valid within a restricted fragment that does not contain modal expressions. However, here, we are precisely interested in whether the presence of modal expressions gives rise to failures of classically valid patterns, and whether the best semantics for modality violates classical logic. In other words, the interesting question is whether MT or *reductio* are *unrestrictedly* valid. I shall argue, contrary to what these authors maintain, that the answer to this question is positive.

[9] This proposal has been discussed by e.g. Yalcin (2012), Dorr and Hawthorne (2014), and Kratzer (1983), though all of these authors reject it as unsatisfactory.

[10] Yalcin (2012) provides further discussion of this problem.

(8) If Bill comes to the party, then John will come and it is likely that Margaret will come, too.

(8) does not have a reading according to which it is likely that if Bill comes to the party, John and Margaret will come, too.

We can, alternatively, defend option (ii) by appeal to the context-sensitivity of modal operators. It is not particularly controversial that the interpretation of modal operators depends on context (although, exactly how is a matter of great controversy). A familiar view is that modals are quantifiers over possible worlds, but just which worlds depends on the context (Kratzer, 1977, 1981). We can exploit this to argue that the problematic counterexample can be explained away by maintaining that the modal 'likely' in (1) contributes a different semantic content than the one in (2), due to contextual effects on the interpretation of the two occurrences of the modal; and so, (2) and the consequent of (1) fail to contradict each other. Accordingly, (1)–(3) is not really an instance of MT.

This strategy captures the intuition that the consequent of (1) talks about a restricted (conditional) probability, while (2) talks about an unrestricted one.[11] The challenge is to explain exactly why and how the context secures different (and intuitively correct) interpretations for the two occurrences of the modal. To do so in a non-*ad hoc* way is notoriously difficult.[12]

To give more bite to the *ad hoc*-ness charge, note that it becomes even more pressing once we acknowledge that contextual effects are not freely available with many other uncontroversially context-sensitive expressions. To illustrate, consider the following example:

(9) If John ate the food from the fridge, then the fridge is empty.

(10) But the fridge is not empty.

(11) So, John didn't eat the food from the fridge.

One cannot freely shift the context so that 'empty' in (10) means *in the state of a vacuum*, and thus avoid the MT reading. So, why assume that context can freely shift between (1) and (2), but not between (9) and (10)?[13]

This apparent asymmetry raises a worry about the relation between language and logic on the contextualist accounts. These accounts typically leave the resolution of context-sensitivity to broadly open-ended pragmatic mechanisms (e.g. speaker's intentions, or non-linguistic contextual cues).[14] But, if logical forms are partly fixed by broadly unsystematic, open-ended mechanisms, then validity becomes partly a matter of such mechanisms, as well. As a result, on such accounts, there is no systematic, rule-governed way of determining whether (1)–(3), or (9)–(11), is expressing a valid argument, or not. For sure, given a set of fully-specified logical forms, one can determine a subset of valid ones; but given a surface form, like (1)–(3), one derives the fully-specified logical form only through open-ended, defeasible pragmatic processes. The rules of language alone do not determine whether a pattern like (1)–(3) is valid or not. Thus, logical

[11] However, as Yalcin (2012) notes, one should not be too quick to conclude, given this intuition, that the consequent of (1) and (2) express different contents, since there are non-truth-conditional, non-contextualist accounts available, Yalcin's included, that capture this intuition without positing a difference in content.

[12] See e.g King (2014), and Stanley (2000).

[13] Of course, one could try to modify (9)–(11) to achieve the context-shifting effect by, e.g. putting a focal stress on 'empty'. But focal stress is one of the linguistic mechanisms that systematically affect the interpretation of context-sensitive expressions, in a way that is predicted by the theory I shall defend. (More on this below.) The point is just that one has to additionally signal when there is a contextual shift via a mechanism such as, e.g., focal stress. The need for some such mechanism precisely shows that context *cannot* shift freely. So, the challenge of spelling out the mechanism(s) that would yield the interpretation of (1)–(3) that the proponent of option (ii) defends becomes only more pressing.

[14] This kind of an approach was initially advocated by Kaplan (1989a), as an account of the resolution of demonstrative pronouns. The idea is roughly that, with the possible exception of the so-called *pure indexicals*—expressions like 'I', whose linguistic meaning fully determines the interpretation given a context—context-sensitive expressions require pragmatic supplementation in order to be interpreted in a context.

inference becomes dependent on non-linguistic and psychological factors, like epistemic cues and communicative intentions. The link between grammar and logic is thus only indirect, mediated by pragmatic principles. This is worrisome if we cannot provide a systematic, yet non-*ad hoc*, account of when and how such mechanisms affect the resolution of context-sensitive expressions. And the problem is that the pragmatic mechanisms are too flexible to provide such systematic constraints: recall, if we want to claim that (9)–(11) is an instance of MT, while (1)–(3) is not, we need a principled story of why context affects the resolution in a particular way in one case, but not in the other. And we also need a story about why, by contrast with (9)–(11), we do not find contexts in which (1)–(3) expresses a valid logical form. It's hard to see what such a story would be, if it is to rely on speaker's intentions and general, non-linguistic contextual cues.

By contrast, I offer a systematic formal account of the effects of context on the interpretation of modals that is not *ad hoc*. I shall argue in what follows that we have *independent* evidence for context-change in (1)–(3). Moreover, I shall argue that we have evidence for the kind of context-change that is governed by mechanisms that affect the interpretation of context-sensitive vocabulary quite generally, and is resolved in a systematic, rule-governed way.[15] Once we see the import of these mechanisms, we will see that the puzzling behavior of modals is not in tension with classical logic. In particular, we will be able to devise a precise theory of context that predicts the appearance of examples in (1)–(3), while at the same time, provably preserves classical logic.

1.2 Modals and Pronouns

Consider the following example:

(12) If Jane is out, then she is having fun.

(13) She (pointing at Mary) is not having fun.

(14) So, Jane is not out.

(12)–(14) is obviously bad. But no one would invoke this argument to present a counterexample to MT. Obviously, (13) does not *really* contradict the consequent of (12). The occurrences of 'she' in (12) and in (13) resolve to different referents. This explanation works because it identifies non-*ad hoc* context-shifting in the inference. We can point to clear reasons why the reference is resolved in a particular way in the relevant cases. In (12), the pronoun 'she' occurs in an elaboration of a hypothetical scenario about Jane, and thus, is resolved to Jane. In (13), the pronoun occurs in tandem with a demonstrative gesture—a pointing at Mary—and thus is resolved to Mary. The *elaboration* in (12) and the pointing gesture in (13) render a certain referent the most prominent for the subsequent anaphora; the subsequent pronoun then, as a matter of its meaning—the character, in the sense of Kaplan (1989b)—picks out the most prominent referent.[16] As I have argued elsewhere (Stojnić, Stone, and Lepore, 2013, 2014), and will amplify on in the next section, we can capture these observations by rendering explicit the mechanisms

[15] As we shall see, moreover, these mechanisms do not merely affect the resolution of the context-sensitive items, but are also reflected in the logical from of an argument like (1)–(3). See also Stojnić (2016a) for further discussion.

[16] Note that the example in (12)–(14) is not merely trading on the difference between deictic and anaphoric (uses of) pronouns. Similar examples can easily be constructed where all the pronouns are interpreted anaphorically:

(i) Mary is upset because Jane is much luckier than she is.
   a. If Jane buys a ticket, she always wins.
   b. But, *she* does not always win.
   c. #So, Jane didn't buy a ticket.

No one would treat (i) as a serious threat to MT. We can point to reasons why a certain interpretation is naturally retrieved in the context: while 'she' in (i-a) is uttered in the course of an elaboration of a hypothetical scenario about Jane, and is thus resolved to Jane, the contrastive focus on the occurrence of 'she' in (i-b) requires that it refers to the contextually most prominent referent other than Jane—and this is Mary.

that systematically govern the resolution of the pronoun. We can pinpoint the mechanisms that, for our purposes, play the key role in governing the resolution of context-sensitivity by building on the resources of so called Discourse Coherence Theories. As we shall see in what follows, these mechanisms are reflected in the logical form of the discourse like (12)—(14), which will be particularly important insofar as we want to treat validity as a matter of logical form.[17]

1.3 Coherence

The key insight behind Coherence Theory is the simple but often neglected observation that a discourse is more than a random sequence of sentences. To flesh this out, we begin with an illustrative example from Hobbs (1979):

(15) John took the train from Paris to Istanbul. He has family there.

(16) John took the train from Paris to Istanbul. He likes spinach.

There is a stark contrast between (15) and (16). While (15) is a perfectly felicitous piece of discourse, (16) (out of the blue) is odd. What explains this contrast? Note that (15) does not merely report two random, unrelated facts about John. It signals that John took the train from Paris to Istanbul *because* he has family there; we understand the second sentence as providing an explanation of the events described in the first. Recognizing this explanatory connection between the two bits of discourse in (15) is part of understanding the contribution in (15)—unless we recognize it, we have simply failed to fully understand the discourse. Unless we understand how the two bits of discourse are related, we cannot fully understand the speaker's contribution. By contrast, due to difficulties in establishing such a connection, (16) seems off. We are left wondering: is Istanbul famous for its spinach? Does spinach cause a fear of flying? That such an interpretive effort is in play in an attempt to understand (16) suggests the requirement of a readily recoverable implicit organization of the discourse that renders it coherent.

Drawing on these observations, Coherence Theorists postulate an implicit organization of discourse that establishes inferential connections—*coherence relations*—among utterances (Hobbs, 1979; Kehler, 2002; Asher and Lascarides, 2003). This implicit organization arises from the communicative strategies that interlocutors exploit to convey and organize their ideas through an ongoing discourse. As demonstrated by the contrast between (15) and (16), successive contributions to a discourse must be linked by a recognizable network of interpretive relationships. The speaker must signal how she structures her contributions according to shared standards and conventions.[18] So, for instance, (15) is understood as connected by the coherence relation of *Explanation*.[19] Failure to establish such a connection in (16) makes it seem off.[20]

Crucial for us is that the task of establishing discourse coherence and resolving semantic ambiguities turn out to be mutually correlated processes. In particular, as has been confirmed

---

[17] See Hobbs (1979), Kehler (2002), and Asher and Lascarides (2003). I do not mean to suggest that mechanisms of discourse coherence are the only mechanisms affecting the resolution of either pronouns or modals. Other linguistic mechanisms, such as prosody, for example, likewise can affect the prominence in a discourse. See Stojnić, Stone, and Lepore (2013, 2014) for an account that incorporates the effects of various different linguistic mechanisms on the resolution of demonstrative pronouns.

[18] This is not to say that a discourse cannot be ambiguous with respect to coherence relations it harbors. We see one such example of ambiguity in (17) below. Part of the interpretive effort in understanding a discourse is in resolving such ambiguities.

[19] Coherence Theorists typically capture these observations by representing inferential relations—coherence relations—in the logical form of a discourse. Cf. Asher and Lascarides (2003). There is both good syntactic and good semantic evidence for representing coherence relations in the logical form of a discourse. See, in particular, Asher and Lascarides (2003) and Webber et al. (2003) for further discussion.

[20] Of course, we could make (16) felicitous, were we to provide a sufficiently rich context, which would allow us to establish the relevant relation—for example, if it were a part of the common ground that the best spinach is grown in Istanbul. This is just as expected.

by a number of empirical studies, pronoun resolution co-varies with the choice of coherence relation.[21] Here is an illustration:

(17) Phil tickled Stanley, and Liz poked him. (Smyth, 1994)

There are (at least) two ways we can understand (17). The second clause could be taken to describe the result of the event described by the first: Phil tickled Stanley, and *so*, Liz poked him (i.e. "Liz is avenging Stanley"); or, one could understand the two clauses as comparing and contrasting two parallel events: Phil tickled Stanley, and Liz poked him *as well* (i.e. "What happened to poor Stanley?"). Crucially, if the discourse is understood as connected by the Result relation, the pronoun refers to Phil; if it is organized around the Parallel relation, then 'he' is Stanley. The choice of a coherence relation guides the choice of pronoun resolution.

Moreover, the data suggest that the mutual constraints between these two tasks are both systematic and robust. That is, given the choice of a coherence relation, the interlocutors are radically constrained in the possible interpretation of a pronoun. To illustrate, consider the following example from Kehler (2002):

(18) Margaret Thatcher admires Ronald Reagan, and George W. Bush absolutely worships her.

Kehler reports that (18) is generally judged infelicitous by his subjects. The subjects expect the pronoun in the second sentence to resolve to Reagan, and intuitively feel the speaker has erred in uttering 'her' instead of 'him'. This is explained, again, by the interaction between the task of establishing coherence and that of resolving a pronoun. The discourse follows 'admires' by 'absolutely worships'—a stronger term in a scalar relationship—thus signaling that the discourse is organized by the coherence relation of Parallel. Parallel requires that the occurrence of a pronoun in the object position be resolved to Reagan (the object of the first clause). Given the gender mismatch, the utterance is judged infelicitous. This is surprising given the available referent for 'her' in the first conjunct, one that is a well-known subject of Bush's admiration. If the correlation were really a matter of mere general pragmatic defeasible reasoning, the perceived infelicity of (18) would be mysterious. The infelicity is, by contrast, expected if the effect of Parallel is a matter of an underlying convention—the convention determines that the referent has to be Reagan, and that is why we are stuck with infelicity, even in the presence of a nearby plausible interpretation.[22]

I take these (and other[23]) observations to show that coherence relations render certain referents prominent for subsequent anaphora resolution.[24] Here's one way to capture the constraints

[21] See e.g. Wolf, Gibson, and Desmet (2004), Kehler et al. (2008), and Kaiser (2009).

[22] Though Coherence Theorists agree that, due to the semantic effects of coherence relations on the truth-conditions of a discourse, as well as due to syntactic constraints on discourse structure, coherence relations need to be represented in the logical form of a discourse (cf. Asher and Lascarides (2003); Webber et al. (2003)), they typically take the process of pronoun resolution and establishing coherence to be merely mutually constraining. They understand the preference for a particular coherence relation to be an input to general holistic process of reasoning that attempts to find the overall most plausible interpretation of the discourse. By contrast, I have argued elsewhere that we should understand these dependencies not merely as mutual constraints between two related tasks, but rather as linguistic effects of coherence relations on pronoun resolution; coherence relations as a matter of linguistic convention make particular entities preferred candidates for subsequent anaphora. (For detailed defense of this view, see Stojnić, Stone, and Lepore (2013, 2014).) I advance considerations below that the effects of discourse coherence on modals is likewise conventionalized. For a more detailed defense of this position see Stojnić (2016b, chs. 1 and 2, and 5).

[23] It's worth noting, moreover, that languages differ with respect to the effects of coherence relations on the interpretation of pronouns, which suggests that, indeed, the effect is a matter of linguistic convention. See Stojnić, Stone, and Lepore (2014).

[24] The notion of prominence within a discourse that I'm relying on here should not be confused with the intuitive notion of real-world salience. That a referent is salient is not sufficient (or necessary) to make it prominent in my sense. Consider an utterance of "Jim came in. He sat down," while Bill is jumping up and down, making himself the most salient entity in the given situation. Unless the speaker points or somehow demonstrates that Bill is the

that the choice of coherence relation places on pronoun resolution. Suppose we rank candidate referents for anaphora in a discourse according to their relative prominence.[25] The import of discourse coherence is to affect this ranking, by making certain referents prominent for subsequent anaphora. The pronoun, then, according to its meaning, refers to the most prominent referent that satisfies the grammatical constraints encoded by the linguistic meaning of the pronoun (e.g. 'he' refers, roughly, to the top-ranked, third-person, singular, male candidate referent).[26] All of this supports our observations about (12)–(14). The consequent in (12) stands in an Elaboration relation with the antecedent; it is *because* 'she' in (12) occurs in an elaboration of the hypothetical scenario described by the antecedent that the pronoun refers to Jane. And it is *because* in (13) the pronoun occurs in a tandem with a pointing gesture, that it refers to the individual pointed at, namely, Mary. As a result, (12)–(14) is not an instance of MT.[27]

So far, I have demonstrated that even in the case of pronouns, if we fail to appreciate how they are resolved within a discourse, we would be misled to interpret examples such as (12)–(14) as "counterexamples". Moreover, I have argued that pronoun resolution is responsive to discourse structuring mechanisms, in particular, mechanisms of discourse coherence. Next, I shall argue that modals are analogous to pronouns in two crucial respects. First, like pronouns, their interpretation is an anaphoric process, by which I mean that they require a contextually available antecedent which is either linguistically introduced or available from a non-linguistic context.[28] And second, like the resolution of pronouns, the interpretation of modals is guided by the mechanisms that structure the information in discourse, in particular, mechanisms of discourse coherence. These mechanisms govern the interpretation of both types of expressions in a systematic, and rule-governed way. Once we incorporate the effects of these mechanisms on the resolution of context-sensitivity, we will see that we can devise a semantics for modals, the underlying logic of which remains classical.

1.4 Modals as Pronouns

That modals exhibit anaphoric-like behavior has been observed in the literature on modal subordination, as in (19), where a modal is interpreted relative to some other modal expression introduced earlier in a discourse (i.e. relative to a linguistically introduced antecedent):[29]

(19) A wolf might walk in. It would eat you first. (Roberts, 1989)

Perhaps the clearest argument for the full analogy between modals and pronouns with respect to the range of interpretive effects they permit has been offered by Stone (1997, 1999).[30] First, he

referent, the occurrence of the pronoun 'he' will be interpreted as Jim. This is not to say that prominence has nothing to do with the entity being salient in the pre-theoretic sense; rather, it is just to say that interlocutors tend to rely on a relatively narrow set of intra-discourse cues that language provides for guiding their attention towards particular referents in a discourse. See Stojnić, Stone, and Lepore (2014), and Stojnić (2016b).

[25] The idea of a ranking of referential candidates in a discourse comes from Centering Theory, which hypothesizes that the referential candidates in a discourse are ranked according to relative prominence, those ranked higher being preferred over those ranked lower as interpretations for subsequent anaphora. See Sidner (1983), Grosz, Joshi, and Weinstein (1995), and Bittner (2014).

[26] For a detailed development of this view, see Stojnić, Stone, and Lepore (2014).

[27] Note that according to this explanation, there are changes in context—making certain referents prominent as antecedents for subsequent pronouns at particular points in discourse—which explain why the pronouns in (12)–(14) are resolved the way they are. But these contextual changes are induced by linguistic mechanisms, reflected in the logical from of a discourse; in particular, they are governed by mechanisms of discourse coherence.

[28] Some theorists like to reserve the terms "anaphora" and "anaphoric" for cases where an item is bound by a linguistically introduced antecedent. The reader should bear in mind that I use these terms throughout in this broader sense specified in the text.

[29] See Roberts (1989). How exactly to account for modal subordination has been a matter of much debate, one that has generated a vast literature. The account defended in this paper shows how to model modal subordination.

[30] Stone's arguments for the parallel between modals and pronouns are analogous to Partee's famous arguments for the parallel between pronouns and tense. See Partee (1973, 1984).

observes that both can depend for their interpretation on an antecedent introduced linguistically (either by means of indefinite or definite reference) earlier in the discourse, as illustrated by (19). Just as 'he' in (20) is naturally understood to refer to the man introduced in the first sentence,[31] so 'would' in (19) is naturally understood restrictedly, as describing the hypothetical scenario introduced by the modal 'might' in the first clause.

(20) John owns a donkey. He beats it. (Based on Partee (1984).)

Second, like pronouns, modals allow for reference to specific entities from a non-linguistic context. In particular, just as (21), can be uttered out of the blue to refer to some significant woman available in the discourse context, so too (22), uttered out of the blue, can be understood to describe the hypothetical scenario that is salient in the discourse context (the scenario in which the speaker buys the stereo):

(21) (Referring to a certain significant female) She left me. (Partee, 1973)

(22) (Looking at a high-end stereo in an electronics store) My neighbors would kill me. (Stone, 1997)

Third, both types of expression allow for bound readings, where intuitively their semantic interpretation co-varies with the domain of some higher binding operator, as in (23) and (24):

(23) Every woman believes that she is happy.[32] (Partee, 1984)

(24) If a concert goer arrives late, he or she will not be permitted into the auditorium.[33] (Stone, 1997)

Finally, both types of expression allow for so called "donkey anaphora" readings, as witnessed by (25) and (26); crucially, just like 'it' in (25) co-varies with the indefinite NP 'a donkey' (without being within its syntactic scope), so in (26) the modal in the consequent 'will' co-varies with the sub-constituent of the antecedent clause 'if the enemy captures it'.

(25) If a man owns a donkey, he beats it. (Partee, 1984)

(26) If a submarine cannot self-destruct if an enemy captures it, the enemy will learn its secrets. (Stone, 1997)

These data strongly suggest an analogy between modals and pronouns with respect to the kind of interpretive dependencies they allow; both search for an antecedent either previously linguistically introduced, as in (19)–(20), and (23)–(26) or provided by the context, as in (21)–(22).[34]

Observe that conditionals, too, display this type of anaphoric behavior. For example, in (27), the second conditional depends on a scenario introduced by the first one; it is not simply evaluated against all epistemically accessible worlds, available in the context discourse initially. And the modal in the consequent of the second conditional is thus interpreted only relative to the hypothetical scenario introduced by the antecedent of the second conditional *relative to the hypothetical scenario described by the first conditional.*[35]

(27) If a wolf walks in, we will use the tranquilizer gun. If we manage to shoot it, we will be safe.

---

[31] I assume that there are no accompanying pointing gestures in (20).

[32] Similar examples are found in e.g. Partee (1973) and May (1977).

[33] I cite the original example that Stone provides, but the effect is easily replicated without 'will': "If a concert goer arrives late, he or she might not be permitted into the auditorium." The same goes for Stone's other examples involving 'will' (in particular, example (26)).

[34] Further support for the analogy between modals and pronouns with respect to their interpretive range is provided by examples from languages such as Warlpiri or American Sign Language (ASL) that allow for a single anaphoric expression to be ambiguous between pronominal and modal interpretation. See Bittner (2001) for data on Warlpiri, and Schlenker (2013) for data on ASL.

[35] Just as, as we shall see, the anaphoricity of modals plays a key role in explaining the counterexamples like the one in (1)–(3), so this anaphoricity of conditionals plays a key role in explaining the counterexamples to MP and MT involving right-nested conditionals, as McGee's (1985) original counterexample to MP, or Veltman's (1985) counterexample to MT.

One way we can think about the observed anaphoricity (in the sense of "anaphoricity" described above) of modals and conditionals is as follows. It is customary to treat modals as quantifier expressions, quantifying over possible worlds.[36] We know that quantifiers require a contextually supplied domain restriction—"Everyone had fun today" does not mean that everyone *in the universe* had fun today.[37] The same goes for modals—like other quantifiers, they also require that their domain of quantification be further contextually restricted.[38] "A wolf might walk in" does not convey the meaning that in at least one world *out of all possible worlds* a wolf walks in. Rather it conveys a more restricted meaning—at least one world out of the relevant, epistemically accessible worlds is such that in it a wolf walks in. Since in the case of the modals the domain of quantification is just a set of worlds, the restrictor on the domain will likewise be a set of worlds (a possibility, for short). I suggest that modal anaphora resolution is a matter of retrieving the possibility that serves as the restrictor on the domain of quantification of a modal operator. That is, modals (and conditionals) are anaphoric insofar as they require an anaphorically retrieved restrictor: just as with the antecedent of a pronoun, the restrictor can either be explicitly linguistically introduced in the discourse (e.g. by some modal utterance prior in the discourse), as in (19), (24), (26) and (27), or otherwise made prominent in the context, as in (22).[39]

How this anaphoric dependency is resolved—i.e. how the restrictor is retrieved in a context—brings us to our second analogy between pronouns and modals. Crucially, modals exhibit the same kind of responsiveness to mechanisms of discourse coherence as pronouns. The idea is that mechanisms of discourse coherence affect the prominence of possibilities that are candidates for the restrictor of a subsequent modal, just as they can affect the prominence of referents that are candidates for subsequent pronominal anaphora; they render certain possibilities prominent to serve as a restrictor of a subsequent modal. We see this already in (19), where the second sentence elaborates on the hypothetical possibility described by the first. It is because this Elaboration relation holds that the modal 'would' in the second sentence is understood as restricted by the possibility described by 'might' in the first— all of the hypothetical scenarios *out of those epistemically accessible ones in which the wolf walks in* are such that the addressee is eaten first. (Similarly, it is because there is an Elaboration relation between the antecedent and the consequent in (24) that the modal 'will not' in the consequent is understood as restricted by the possibility described by the antecedent.) The Elaboration relation is what makes this scenario prominent for the subsequent modal to pick up on.

The import of coherence is crucial in making a certain possibility prominent as a restrictor of a subsequent modal. Note that mere sequencing of modals is not sufficient for the correct interpretation. That is, it is not always the case that the modal will be restricted by a possibility introduced by the immediately preceding modal (when there is an immediately preceding modal). The fact that one modal follows another does not suffice to establish that the hypothetical scenario described by the second modal further elaborates the one introduced by the first. We easily see this with examples like (28).[40]

---

[36] Cf. Kratzer (1977, 1981, 2012).

[37] For more on quantifier domain restriction, see von Fintel (1994), and Stanley and Szabó (2000).

[38] Typically, in an ordinary discourse, the restrictor will eliminate at least some worlds from the domain of quantification, on pain of redundancy. In principle, however, the restrictor need not eliminate anything. This can happen, for instance, if the restrictor is a necessary proposition.

[39] Does this mean that quantifiers are also anaphoric in my sense? The answer is yes. Like modals they require a restrictor, provided either by the non-linguistic context, or the prior discourse. Moreover, though the details exceed the scope of the present paper, there are good reasons to hold that the way in which the restrictor of a quantifier is retrieved in a context is analogous to the way in which the restrictor of a modal and an antecedent of a pronoun is, i.e. that quantifier domain restriction is sensitive to discourse structuring mechanisms. (See Stojnić (2016b).)

[40] For similar examples, see Asher and McCready (2007).

(28) If a wolf walks in, it would eat you. But one probably won't walk in.

As before, the consequent of the first sentence in (28) elaborates on the information provided by the antecedent. The Elaboration relation between the antecedent and the consequent makes the hypothetical scenario introduced by the antecedent the most prominent one, and as a result, 'would' in the consequent further describes this scenario. Crucially, however, the modal 'won't' in the second sentence does not further elaborate upon scenario described by the two modals in the first sentence. The two sentences stand in a relation of Contrast, signaled by the discourse marker 'but', and are understood as contrasting two hypothetical scenarios—one in which a wolf walks in, and one in which one does not.

Intuitively, the Contrast relation requires that the first and the second sentence provide contrasting information about some body of information. A bit more precisely, the two bits of discourse provide contrasting information about some body of information regarding some common topic—in our example the topic is what is possible regarding a wolf's entrance.[41] The first sentence already sets the stage in determining the body of information the contrast has to be about—(assuming that the conditional is uttered discourse initially) it is interpreted relative to the set of epistemically accessible worlds determined by the context discourse initially, describing what might be the case if a wolf walked in, *given this overall body of knowledge.* The second sentence, then, has to provide a contrasting bit of information, regarding the possibility of a wolf's entrance, about this body of information available discourse initially. The Contrast relation makes this body of information prominent, and consequently, the modal in the second sentence selects it as its restrictor—*given this overall body of knowledge*, a wolf probably won't come in. This is the intuitively correct interpretation: *given the overall body of knowledge, if a wolf walks in*, it would eat the addressee, but *given the same body of knowledge*, one probably won't walk in.

What this sort of example establishes is that, just as with pronouns, the impact of discourse coherence on making a certain possibility prominent as a restrictor for a subsequent modal is crucial. As the contrast between Elaboration and Contrast shows, precise discourse mechanisms govern the prominence of possibilities in a discourse. We can capture this idea in a way similar to the way in which we captured the effects of coherence on the prominence of candidate referents for pronoun resolution. Here is a first pass proposal: let the context represent a ranking of sets of worlds (possibilities, for short) that are candidates for domain restrictors of modals in a discourse according to their relative prominence—the most prominent being the top-ranked one. A modal simply retrieves the most prominent epistemically live possibility as the restrictor for its domain of quantification.[42] The prominence ranking of candidate possibilities, in turn, is affected by a range of linguistic mechanisms, most notably, those of discourse coherence; coherence relations make certain possibilities prominent for subsequent modal anaphora.[43]

---

[41]The relevant topic is typically signaled by the cues in the information structure—the way the information is packaged together. One way of signaling this, in English, is by exploiting prosodic accents. For example, compare the following two utterances:

(i) John likes MARY.

(ii) JOHN likes Mary.

(i) is fine in the context in which we are wondering whom John likes, say, Mary or Sue, but not in the context in which we are wondering who likes Mary, say, Bill or John; the opposite is true of (ii). For more on sentential focus, see Rooth (1992), and for more on information structure, see e.g. Roberts (1996).

[42]I will use the term "epistemically live possibility" to denote the possibility that is not ruled out given the relevant body of knowledge, i.e., which is accessible given the relevant epistemic accessibility relation. For a more precise notion of epistemic accessibility, see § 2.1 and § A.2.

[43]Though limited space prevents me from going into details here, I have argued elsewhere that there are good reasons to treat these effects of coherence on the resolution of modal anaphora as linguistically encoded, rather then as byproduct of general reasoning. (See Stojnić (2016a,b).) For instance, as Asher and McCready point out, the direct translation of (19) in Japanese is infelicitous, unless there is an overt discourse marker

I shall assume that at the beginning of a conversation the top-ranked possibility is just a set of epistemically possible worlds—a set of worlds epistemically accessible from the actual world. The intuitive idea behind this assumption is the familiar idea that the ultimate goal of a conversation is to narrow down possible ways the actual world could be.[44] However, as the discourse progresses, the prominence ranking changes; this change in ranking is precisely what we want to model, since this is precisely the change relevant for the interpretation of modals. The changes arise as an effect of introducing novel possibilities (e.g. through utterances containing modals or conditionals), and through discourse structuring mechanisms that change and affect the prominence ranking, in the ways described earlier.

Connecting this with our previous observations, the anaphoricity of modals is captured by requiring that the restriction on the domain of quantification be retrieved in a way similar to how the antecedent of a pronoun is—either provided by the context, or explicitly by the prior discourse. The way anaphora is resolved, in both cases, is determined by discourse structuring mechanisms, in particular, mechanisms of discourse coherence. Just as the mechanisms of discourse coherence affect the prominence of candidate referents for the resolution of a pronoun, so too they affect the prominence of a candidate possibility for a restrictor on a subsequent modal.

To sum up: modals are like pronouns in two crucial respects: (a) they are anaphoric expressions, and (b) the resolution of modal anaphora is responsive to the same mechanisms that pronoun resolution is responsive to. We now have all the ingredients we need to tackle the original counterexample.[45]

---

signaling an Elaboration relation (Asher and McCready, 2006).) This would be surprising if the elaboration reading, where the second modal is understood as elaborating on the scenario described by the first one, were merely a byproduct of general reasoning. Moreover, as I argue in Stojnić (2016b,a) the mechanisms of discourse coherence sometimes force inconsistent interpretations of modal discourse, even when there are possible plausible alternative interpretations that could have been retrieved instead. That the inconsistent readings are retrieved even in the face of alternative common-sense interpretations is easily explained if the mechanisms of discourse coherence conventionally mandate such interpretations; however, this would be mysterious if the effects of these mechanisms were byproducts of pragmatic reasoning, since in such cases, we have overwhelming common-sense reasons to disprefer the inconsistent interpretations, and so override the effects of discourse coherence. These considerations, among others, suggest that the effects of the mechanisms of discourse coherence are a part of our linguistic repertoire, and should hence be reflected in the logical form of the discourse.

[44] Cf. Stalnaker (1978). More precisely, since typically there is uncertainty about which world is the actual one (given that our knowledge is limited), the initial set of epistemically live possibilities will be a set of worlds W that contains, for each world w that is a candidate for the actual world, the set of worlds accessible from w. Of course, if we are concerned with the interpretation of discourse initial uses of modals, it will matter a great deal whose body of knowledge is relevant for determining the set of epistemically accessible worlds. Since I shall not deal with this issue in the present paper, I shall simply assume that the relevant body of epistemically accessible worlds will be contextually provided discourse initially. However, *pace* Kolodny and MacFarlane (2010), Yalcin (2007), and Moss (2015), the resources of my account can show how this body of information is selected in a context, even discourse initially. (See Stojnić (2016b).) Relatedly, as I argue at length in Stojnić (2016b), the resources of my account naturally explain the patterns of intra- and inter-contextual (dis)agreement involving modal language, that have been argued to favor the revisionary theories precisely on grounds of the alleged failure of context to determine an adequate body of information in such cases. (For the revisionary arguments based on (dis)agreement patterns, see e.g. Egan, Hawthorne, and Weatherson (2005); MacFarlane (2014). For alternative contextualist responses to these arguments see, e.g., Dowell (2011) and von Fintel and Gillies (2009). These contextualist accounts (much like Kratzer's account) assume general pragmatic mechanisms of context-sensitivity resolution that are too unconstrained to systematically account for the apparent counterexamples to MT, as well as the puzzling behavior of modals in various embedding environments, including, but not limited to antecedents of conditionals. See Stojnić (2016b) for a detailed discussion.)

[45] Note that, according to the theory developed here, modals express truth-conditional content, and the explanation of the counterexample exploits the difference in the truth-conditional content between the consequent of the big premise, and the small premise in (1)–(3). Several authors have challenged the view on which modal vocabulary expresses truth-conditional content on grounds that are not directly related to the failure of classical inferences. (See e.g. Egan, Hawthorne, and Weatherson (2005), Yalcin (2007), and Moss (2015).) Though the theory developed here, given its systematic account of context-change, has means for accounting for those challenges as well, addressing them here is beyond the scope of the present paper. (I address these issues in detail in

1.5 Yalcin's Counterexample Explained Away

We can now explain what is going on with (1)–(3) as follows: in (1), the consequent of the conditional is understood as elaborating on the possibility introduced by its antecedent (which is just the set of epistemically accessible worlds in which the antecedent holds), and thus, the Elaboration relation renders this possibility the most prominent one for subsequent anaphora. Consequently, the modal 'likely' in the consequent, which is searching for the most prominent epistemically accessible possibility, selects this possibility as the restrictor for its domain of quantification. The consequent is thus understood as further describing the possibility introduced in the antecedent, providing the intuitively correct restricted reading—the marble is likely red, *given that it is big*. In turn, we naturally understand (2) as being linked to (1) by the relation of Contrast. This is seen even more clearly if we insert an explicit discourse marker 'but' in (2): 'But the marble is not likely red.' Note that some such way of signaling contrast is required, for the discourse consisting of (1) followed by (2) to be felicitous.

As always, the question of which relation holds (and between which relata) is a matter of disambiguation in a discourse, much as in the case of (17). There are often discourse-internal, linguistic cues that signal a particular relation (e.g. a discourse marker 'but'), but context can play a role in disambiguation as well. For example, here the initial context sets up a topic—the color of the marble, depending on a certain assumption about its size. (1) and (2) are then understood as providing contrasting bits of information about some body of information regarding this topic. As before, the first sentence already sets the stage in determining the body of information the contrast has to be about—since (1) is uttered discourse initially, it is interpreted relative to the overall body of knowledge available discourse initially. Thus, one understands (1) and (2) as providing two contrasting bits of information about *the overall body of knowledge or information available discourse initially*, regarding the likelihood of the marble being red, depending on a certain assumption about its size. Namely, given the overall available information discourse initially, *if the marble is big*, then it's likely red, but *additionally*, given the overall available information, the marble is not likely red (given no particular assumption about its size).

A bit more precisely, the effect of Contrast in (1) and (2) is the same as in (28). (1) and (2) are understood as contrasting two different bits of information about some initial overall

Stojnić (2016a) and Stojnić (2016b).) I shall just note that we can, for example, easily explain the problematic behavior of epistemic modals embedded in antecedents of conditionals—one of the key data points used by Yalcin to argue against the truth-conditional accounts (Yalcin, 2007, 2011). Yalcin notes the discrepancy between the following two conditionals:

(i) If it is raining and it might not be raining, ...

(ii) If it is raining and I/we don't know it, ...

The first one is odd, while the second perfectly fine, pointing to an apparent problem for the standard truth-conditional accounts of conditionals, that interpret modals as quantifying over some salient body of knowledge. Yet, the natural way to understand "it might not be raining" in the conditional above is as elaborating on the hypothetical raining scenario. This affects the resolution of modal anaphora—'might' is understood as quantifying over all the relevant epistemically accessible worlds in which it is raining, and so it is no surprise that the conditional winds up being bad. Note that my account does not predict that the conditional with the reversed order of conjuncts in the antecedent, i.e. of the form: "If it might p, and not p, then...," will automatically be bad. This is a desired result since, as Dorr and Hawthorne (2014) note, switching the order of the conjuncts in some cases (in particular, in Yalcin's original example) makes the conditional felicitous. Yet, my account does not predict that all such cases will automatically be felicitous either. This is because the badness of such a conditional will depend on which coherence relations, and other interpretive dependences, can be established between the two conjuncts in the antecedent, and between the conditional and the rest of the discourse in which it is embedded, for these factors can all affect the resolution of modal anaphora in a particular case. This is, again, a desired result. And the point holds more generally—embedding any sentence (including the original Yalcin's conditional) in a broader discourse might give raise to various interpretive dependences that might yield a different interpretation than the one we get when interpreting the same sentence-type out of the blue.

body of information, regarding the likelihood of the redness of the marble, given some or no assumption about its size. The first sentence already sets the stage in determining the body of information that the contrast has to be about—the conditional is interpreted relative to the set of epistemically accessible worlds determined by the context discourse initially,[46] describing what the likelihood of the redness of the marble is, *given this body of knowledge*, provided that it's big. The second sentence thus has to provide a contrasting bit of information about this same body of information (i.e. the set of epistemically accessible worlds, discourse initially) regarding the likelihood of the marble being red. Thus, this body of information is made prominent by Contrast. Consequently, the modal in (2), selects this body of information as its restrictor, conveying that the marble is not likely red *given this overall body of knowledge* (given no particular assumption about its size). Thus, we see that the two occurrences of 'likely' in (1) and (2) are interpreted differently, much like the two occurrences of 'she' are in (12) and (13). And thus, (1)–(3) is no more an instance of MT than (12)–(14) is.[47]

1.6 Conclusion to Part I

In § 1, I argued that the lesson of the alleged counterexamples to classically valid patterns of inference is *not* that we need a revisionary, non-truth-conditional semantics for modal vocabulary that invalidates these patterns, but rather that we need a different, more constrained picture of how context affects interpretation. In particular, I argued that modals are like pronouns not merely in that they are anaphoric expressions, but in that they are sensitive to some of the same interpretive mechanisms of anaphora resolution as pronouns are—in particular, the mechanisms of discourse structure. The "counterexamples" arise due to a failure to appreciate the import of these mechanisms on the interpretation of modals. But, a failure to appreciate how context changes with the evolution of a discourse would lead us to mistakenly treat even the examples like (12)–(14) as counterexamples. That the effect of context-change on modals is more subtle and less well explored only makes the failure to appreciate it that much more dangerous.[48]

The mechanisms of context-sensitivity that I have argued are relevant for the interpretation of modals, though pervasive and nuanced, are not unruly. We can devise a *formal* theory of context-change that models the import of these mechanisms on the interpretation of modals in a rule governed and systematic fashion, and indeed, this is what I do in § 2, and in the Appendix (§ A). Thus, the formalism developed there should be understood as an existence proof in support of the main argument of § 1. The theory accommodates the anaphoricity of modals, while preserving their standard truth-conditions, and maintaining the validity of the classical patterns of inference. Though the resulting semantics departs from the standard semantics for modal discourse, it does so in a way that allows us to capture the interplay between discourse structuring mechanisms and the interpretation of modal anaphora. It is precisely this departure that preserves classical logic, and permits modal utterances to express standard truth-conditional content.

It is important to note that the account developed here opens a promising line for further development. In particular, philosophers often identify context-sensitivity of various philosophi-

[46] Provided that the conditional is uttered out of the blue, which, by assumption, it is.

[47] Note that, while my account maintains that there is a change in context that transpires in (1)–(3), affecting the interpretation of modal expressions, the account does not maintain that such a change is somehow 'illicit'. In an important sense, there is no way of holding the context fixed throughout the course of (1)–(3). This is because (1)–(3) harbors linguistic elements (in particular, modals, antecedents of conditionals, and coherence relations) part of the meaning of which is to change the context in a way that affects the truth-conditions (by introducing certain possibilities, making them prominent, and demoting others).

[48] As I argue in more detail in Stojnić (2016b), the revisionary accounts in fact are not well suited to explain the full range of data concerning modal anaphora. For instance, they do not capture the modal anaphora data described in § 1.4. My account, in turn is precisely designed to account for these data. Thus, apart from preserving truth-conditions and classical inference patterns, the account is well motivated on independent grounds.

cally interesting terms—e.g., of implicit restriction on quantifier domains, knowledge ascriptions, vague predicates, normative terms—and use this context-sensitivity to motivate broad philosophical conclusions. But in doing so, they typically assume a model of context-sensitivity that is resolved by freely selecting one candidate resolution out of an open-ended list of potential ones, through general pragmatic mechanisms (e.g. speaker's intentions, and non-linguistic contextual cues). This predicts a level of flexibility that often fails to be born out in practice, and this flexibility in turn shapes the philosophical arguments that appeal to such context-sensitivity. Though arguing for this in full generality obviously exceeds the scope of the present paper, going beyond modals and pronouns, the kinds of tools developed here open a path for exploring the potential systematic constraints on other context-sensitive expressions. If the approach advocated here can be extended to capture the contextual-sensitivity of these other kinds of expressions, then that would show that contexts are much less powerful, and the resolution of context-sensitivity a much more constrained process than what philosophers typically assume in their arguments.[49]

2 Part II: A New Semantics for Modality

In what follows, we have two tasks: first, to modify the standard truth-conditional account, developed in Kratzer (1977, 1981, 1983, 2012), in order to accommodate the anaphoricity of modals and conditionals, and second, to develop an account of prominence and context-change that explains how modal anaphora is resolved in a discourse. I will undertake the first task in § 2.1, and the second task in § 2.2. As I shall demonstrate, the formal semantics that emerges assigns standard truth-conditions to modals and conditionals, and as I prove, in the Appendix (§ A), preserves classical logic.

2.1 Anaphoricity and Truth-Conditions

As mentioned earlier, the classic account treats modals as quantifying over a contextually supplied modal base, i.e. a set of contextually relevant worlds that comprise the domain of quantification. A modal then requires that a particular relation holds between the proposition expressed by the clause embedded under a modal (which is typically called *the prejacent*[50]) and the modal base.[51] For example "It might rain" requires that there is some possible world in the contextually supplied modal base in which it rains. Let 'q' denote the proposition expressed by the prejacent ϕ of '*might* ϕ' in a given context and 'R' an accessibility relation, that specifies the epistemically accessible worlds, i.e. the modal base.[52] Then the truth-conditions expressed by '*might* ϕ', relative to a context of utterance, are:

---

[49]Indeed, in Stojnić (2016b) I argue that the kind of account developed here extends to numerous other cases of context-sensitivity, and that the kind of mechanisms I described here are suited to explain context-sensitivity quite generally.

[50]So, in '*might* ϕ', ϕ is the prejacent.

[51]This captures the main insights from Kratzer's account, though I suppress formal details for simplicity. In particular, I suppress an ordering source parameter, which provides an ordering of worlds in a modal base according to some contextually provided standard, and the formal machinery that serves to derive a modal base parameter. (Cf. Kratzer (1977, 1981, 1983, 2012).) We could easily factor these elements back in. As Stone points out, modal base and ordering source parameters, as specified in Kratzer's account, cannot accommodate the anaphoricity of modals, since both are determined in complex ways, and neither provides a semantic parameter that can be contributed by prior discourse, so a modification of Kratzer's account is needed regardless (Stone, 1997).

[52]The relation R plays the role of a Kratzerian modal base. On Kratzer's account the modal base is contextually determined in complex ways. For our purposes, we can simplify even further, and let R be supplied by the model, since we are not dealing with discourse-initial uses of modals. See Stojnić (2016b) for details on how R is determined in context, discourse initially.

Definition 2.1.

$\{w \mid \exists w' : wRw' \;\&\; w' \in q\}$

That is, '*might* ϕ' is true at a world w (relative to a context c) just in case there is some world among the worlds epistemically accessible from w in which the proposition, q, expressed by the prejacent in c, holds. Anaphoric dependency is easily factored into the standard truth-conditions of modals explicitly as follows. (I shall use '$M(p, q)$' for the truth-condition expressed by an utterance of '*might* ϕ', where q is the proposition expressed by the prejacent ϕ of the utterance of '*might* ϕ', and p the proposition corresponding to an anaphorically retrieved restrictor. I omit the details about how context determines truth-conditions (and, in particular, the restrictor p) here. This will be the topic of § 2.2.)

Definition 2.2.

$M(p,\ q) = \{w \mid \exists w' : wRw' \;\&\; w' \in p \;\&\; w' \in q\}$

This gives us the resources to define other modal expressions: as is standard, '*must*' is the universal dual of '*might*'.[53]

For probability modals, such as '*likely*', we need a probability measure over the accessible worlds. Let $\mathcal{P}$ be a probability measure over the set of worlds in the universe W, that maps each subset of W to $[0, 1]$, satisfying the following constraints:

- $\mathcal{P}(W) = 1$
- $\mathcal{P}(p \cup q) = \mathcal{P}(p) + \mathcal{P}(q)$, when p and q are disjoint subsets of W.[54]

Then, where q is the proposition expressed by the prejacent ϕ of an utterance of '*likely* ϕ', and p the proposition corresponding to the anaphorically retrieved restrictor, the truth-conditions, $P(p, q)$, expressed by the utterance of '*likely* ϕ' are as follows:

Definition 2.3.

$P(p,\ q) := \{w \mid \mathcal{P}(\{w' \mid wRw' \;\&\; w' \in p \;\&\; w' \in q\}) \;/\; \mathcal{P}(\{w' \mid wRw' \;\&\; w' \in p\}) > .5\}$[55]

As expected, the truth-conditions expressed by an utterance of '*likely* ϕ' are the set of worlds such that, for each w in the set, the ratio of the probability that an R-accessible world from w be a p and q world to the probability that the R-accessible world from w be a p-world is greater than .5; i.e. an utterance of '*likely* ϕ' is true in w just in case given our modal base, the conditional probability of the prejacent, q, given the p-restricted modal base is greater than .5.[56]

[53] For the definition of truth-conditions of '*must* ϕ', see § A.2.

[54] I shall assume that the probability measure is supplied by the model. Alternatively, it could be provided by context. The choice is inessential for the context-sensitivity I'm aiming to model. For simplicity, I assume that W is finite. I also assume that $\mathcal{P}$ is a regular probability measure, i.e. assigning non-zero probability to all non-empty sets of worlds. Insofar as $\mathcal{P}(p)$ represents the prior, if $\mathcal{P}(p) = 0$, p wouldn't really be a possibility.

[55] It is typically assumed that the restrictor on the domain of quantification, p, is not empty. But if we wanted to allow for the possibility in which p is empty, we could modify our definition in a following way: $P(p,\ q) := \{w \mid \mathcal{P}(\{w' \mid wRw' \;\&\; w' \in p \;\&\; w' \in q\}) > 1/2\ \mathcal{P}(\{w' \mid wRw' \;\&\; w' \in p\})\}$. Also, one might wonder whether we should always impose a threshold of .5, or perhaps the threshold might vary with the context. For simplicity I choose the former option. The choice is inessential for our purposes.

[56] I depart from Kratzer (1991) in suggesting this quantitative characterization of the truth-conditional contribution of probability operators, rather than a qualitative one cashed out in terms of relative likelihood. There are well known problems with a purely qualitative account. For a detailed discussion, see Yalcin (2010); for a discussion of the prospects of basing the quantitative notion of probability on a qualitative one, see Kratzer (2012, ch 2.) and Holliday and Icard (2013). This particular choice I make here, though independently motivated, is inessential to the overall point of the paper. What matters for us is that the anaphoric potential of a probability modal is correctly captured. We can build the anaphoric dependency into the truth-conditions in the way suggested, regardless of what we think the correct account of the truth-conditions is. The point holds for the truth-conditions of other modals as well.

Finally, we need to specify the truth-conditional contribution of a conditional. We can easily factor in the anaphoric potential into the truth-conditions of a conditional, just as we did with modals, while otherwise preserving the standard truth-conditions. Where p, as before, is the anaphorically retrieved restrictor with respect to which the conditional is uttered, q corresponds to the proposition expressed by the antecedent, and r to the one expressed by the consequent, an utterance of a conditional expresses truth-conditions corresponding to a set of worlds such that for each w in the set, all the worlds w', R-accessible from w, that are p and q worlds, are r worlds as well; i.e. an utterance of a conditional is true in w if and only if all the p and q worlds in the modal base are r worlds as well:[57]

Definition 2.4.

Cond(p, q, r):= $\{w \mid \forall w' : wRw'$, if $w' \in p$ & $w' \in q$, then $w' \in r\}$

This in essence preserves the standard truth-conditions associated with a conditional, but factors in the fact that a conditional itself is always evaluated against some prominent body of information, that need not always correspond to the complete unrestricted set of epistemically live worlds.

This concludes my characterization of the truth-conditional contribution of modals and conditionals. According to the proposed account, the utterances containing modals and conditionals express the same truth-conditions one would expect given the canonical account,[58] except that the anaphorically retrieved restrictor is factored in separately. This allows us to flexibly track the way it is retrieved in a context. It is important to note that, provided that all anaphoric restrictors are resolved to the same set of worlds (e.g. to the set of all possible worlds), we get exactly the truth-conditions for modals we would expect in standard propositional modal logic (our conditional is a standard strict-conditional).[59] However, we have yet to specify the effects of context and context-change on the determination of truth-conditions of a given utterance. As I have argued, the most important impact of context for us will be in the resolution of anaphoric dependencies of modals and conditionals, i.e. in the role of context in determining the restrictor of the modal base of modals (and conditionals). In § 2.2, I lay out a formal theory of context-change, which will allow us to tackle the alleged counterexample we began with.

2.2 Truth-conditions and the Dynamics of Context-change

In order to capture the idea that the restrictor of a modal expression is the most prominent possibility, we need to have a way of modeling the prominence of possibilities in a discourse context.[60] As suggested in § 1.4, we can let a context represent a ranking of possibilities—sets of worlds, or propositions—according to relative prominence, the top-ranked possibility being the most prominent one. One way to think about such a context is as a sort of a conversational

[57]I ignore the ordering source, but we could easily factor this parameter in, and then state the truth-conditions by making a conditional true (in a world w and at a context c) just in case all the p and q worlds that are closest to w in the modal base are r worlds as well. Apart from factoring in the anaphora (and modulo the abstraction of the ordering source parameter), the truth-conditions I propose here match the ones developed in Kratzer (1983). Kratzer's account makes a conditional true in a world and at a context just in case all the (closest) antecedent worlds within a modal base are consequent words as well.

[58]Cf. Kratzer (1977), Kratzer (1981), Kratzer (1983) and Kratzer (2012)

[59]More precisely, on the assumption that R is reflexive and transitive, we would get the system $S4$. For a proof, see section § A.7. This is a common and natural assumption. Reflexivity, e.g., ensures that *must* p entails p, and also that p entails *might* p, and transitivity ensures, e.g., that *might(might p)* entails *might* p, again, provided that all anaphoric restrictors are resolved to the same set of worlds (e.g. to the set of all possible worlds).

[60]I sketch the key bits of the formal account here, but for the fully precise formal definitions and details, consult § A.

record, in the sense of Lewis (1979), an abstract "scoreboard" that tracks the moves and contributions interlocutors make in the flow of a discourse, and that comprises information relevant for interpretation, such as who's speaking, what the conversation is about, etc. Crucially, for us, the conversational record tracks propositions put into play in the course of a conversation as well as their relative prominence. Since this is the only aspect of the scoreboard that will matter for modal anaphora, we can abstract away from all other aspects that might be otherwise needed. In this spirit, we can see my context as modeling one aspect of a Lewisean scoreboard—the relative prominence of possibilities within a discourse.

We will let a context represent the ranking of possibilities according to their relative prominence. To model a context, we will exploit the idea of an assignment function, modeled as a stack, in the sense of theoretical computer science; an assignment function, or a stack, specifies an ordered sequence of worlds.[61] Let the context be a set of such stacks.[62] So, a context is just a set of sequences of worlds. For a given context G, let '$\mathbf{w}_i$' be a variable that stores a world at the i^{th} position of every stack in G.[63] Then, relative to G, '$\mathbf{w}_i$' stores a set of worlds that collects the worlds at the i^{th} position of every stack in G. This, in particular, will allow us to keep track of the propositions that have been introduced and promoted (as well as demoted), during the course of a discourse. I shall assume that every stack in G begins with the 0^{th} position, as the top-ranked position on the stack, and that for each position n, the position $n+1$ is one position lower in the ranking. Let 'G_i' denote the set of worlds that collects the worlds at the i^{th} position of every stack in G.[64] Then, G_0 is the top-ranked proposition in G, and for each n, G_{n+1} is a proposition one step lower in the ranking. This allows us to keep track of the relative prominence of possibilities for modal anaphora.

As noted in § 1.4, I shall assume that at the beginning of a conversation, the top-ranked possibility is just a set of epistemically accessible worlds. In turn, as the discourse progresses utterances affect the prominence ranking of candidate possibilities. As we have already seen in an informal way in § 1, utterances can promote novel possibilities and re-rank the ones introduced prior in the discourse. The simplest way to model this is to represent utterances as updates to the context that change the relative prominence of candidate possibilities. Formally, an update is represented as a relation between an input context and an output context where the update is true (relative to an input context and a world of evaluation w) just in case it relates an input context to some non-empty output context (relative to w).[65]

Thus, utterances have a two-fold contribution—they express truth-conditions (as per § 2.1), but they also contribute updates that change the context, by updating the ranking of possibilities. Both aspects of the interpretation are crucial. The updates associated with utterances capture the way in which these utterances change the context; in turn, such a dynamically maintained context

[61]Formally, a stack is an assignment function, mapping a finite convex subset of $\mathbb{N}$ to a set of worlds together with an undefined value $\perp$.

[62]This formalization, unlike the one exploiting a single stack ranking sets of worlds, allows us to keep track of not only sets of worlds, but also relations between individual members of different sets. (Cf. Bittner (2014), Brasoveanu (2010), and van den Berg (1996).) Though this is not strictly speaking crucial for our present purposes, it does become crucial once we want to develop an integrated semantics for anaphora both in the modal, and in the pronominal domain.

[63]In describing context change I am adopting the following strategy. I define a dynamic language that models the dynamics of prominence, and then provide a translation of a fragment of English into this language. The dynamic language has atomic expressions (propositional constants (p, q, r), and variables ($\mathbf{w}_i$ for $i \in \mathbb{N}$), conditions (propositional expressions comprising set of atoms closed under $\wedge$ and $\neg$), and update expressions, which we will define and describe shortly. For details see § A (in particular, § A.1, § A.4 and § A.7).

[64]Thus, G_i is the proposition stored at the i^{th} position in G.

[65]This is the standard notion of truth exploited in dynamic semantics. (See Dekker (2011).) Recall that on my account utterances also express ordinary truth-conditional content. The dynamic notion of truth defined here is exploited to capture the logic of context-change. We will see that once we have the logic of context-change in place, the ordinary truth-conditional content (and the ordinary notion of truth) falls out of it straightforwardly.

determines the relevant truth-conditional content of a given utterance. Thus, we effectively model the two-way interaction between a context and an utterance: on the one hand, an utterance changes the context, and on the other, such a dynamically evolving context determines its truth-conditional content.

Before laying down the updates associated with modals and conditionals, we need some preliminaries. We said that utterances contribute updates to a context by promoting and re-ranking propositions, i.e. possibilities, but they also express propositions. I am going to separate the role that propositions play in anaphora, from the one they play in semantic composition. To illustrate why we need this separation, take the example: "A wolf might eat Harvey". First, we need to compose the proposition expressed by the prejacent, *that a wolf eats Harvey*, with the modal '*might*'. But if this proposition automatically became the top-ranked possibility in the context, since the modal '*might*' selects the top-ranked epistemically live proposition as its restrictor, so long as there's some epistemically accessible world in which a wolf eats Harvey, we would predict that the restrictor for the modal in "A wolf might eat Harvey" is the proposition comprising all the epistemically live worlds in which a wolf eats Harvey. But, obviously, this proposition should not automatically be made a restrictor, as witnessed by examples like: "A wolf might walk in. It might eat Harvey."

There are several ways to get around this problem, but one elegant way is to store separately the propositions that enter into semantic composition, and the ones that are candidate restrictors for subsequent modals. That insures that the truth-conditional contribution does not automatically interfere with prominence ranking.[66]

To this end, I reserve a designated position on each stack in a context G that does not affect the ranking. Let '*comp*' (for compositional) denote the proposition that comprises all the worlds stored at this position in every stack in G.[67] We exploit '*comp*' to store, relative to G, each bit of propositional content that enters into semantic composition.[68] Here is how we do that. We will treat expressions that contain no proper propositional subparts as atomic formulae in our system. When ϕ is an atomic formula (and so, does not contain modals, or conditionals), its interpretation will just be the simplest update, defined below in 2.5, which stores the proposition expressed by that formula in the input context G as a new value of '*comp*'. This update relates the input context G to an output context H, (relative to a world of evaluation w) just in case H differs from G in at most the value of '*comp*', and the value of '*comp*' is the proposition that 'p' expresses in G and w. That is, where G is the input context, H the output context, w the world of evaluation, and $G \underset{n}{\sim} H$ just in case G differs from H in (at most) the n^{th} position, we can define this update as follows:

Definition 2.5.
$[\![\langle comp := \phi\rangle]\!](w, G, H)$ iff $G \underset{comp}{\sim} H$ & $H_{comp} = [\![\phi]\!]^{G,w}$.

When ϕ is non-atomic, it will be interpreted as a more complex update. However, crucially, its truth-conditional contribution determined by the update will still be stored as a value of '*comp*' of the output context, in a way to be specified presently.[69]

Observations in § 1.3 and § 1.4 show that modals require an anaphorically retrieved restrictor, but also introduce possibilities that can subsequently be picked up by other modals. I suggested

---

[66] This is not to say that one and the same proposition cannot play a role both in semantic composition and also be prominent for anaphora resolution. This is just to say that the mechanisms by which a proposition completes these two roles are best kept separate.

[67] So, now, formally, a stack is just a function from a finite convex subset of $\mathbb{N}$ plus *comp* to a set of worlds together with an undefined value, $\perp$.

[68] Similarly as before, 'G_{comp}' denotes the value of '*comp*' in G.

[69] In particular, 2.5–2.12 are clauses of a single recursive definition of update expressions in our dynamic language. See the Appendix (§ A) for details.

that the anaphorically retrieved restrictor will be the top-ranked epistemically live possibility. Let '@E' denote the top-ranked epistemically live possibility (the top-ranked possibility that is a subset of epistemically accessible worlds) in a given context.[70] Intuitively, we want the update associated with an utterance of '*might* ϕ' to have the following effect: first, (as with all other updates) its truth-conditions (as defined in 2.2) are stored as the value of '*comp*' of the output context. Second, it introduces a possibility comprising the top-ranked epistemically live worlds in which the prejacent holds. Here is how we can informally describe the effect on context that an update associated with '*might* ϕ' carries out. Where G is an input context, and H an output context, the update relates an input context G to an output context H just in case there are intermediate contexts G' and G'' and:

a. G' is a result of updating G with the update associated with prejacent of the modal (recall that since updates associated with both atomic and non-atomic formulae will store their truth-conditional content as the value of '*comp*', G'_{comp} will just be the truth-conditional contribution of the prejacent in G, i.e. $[\![\phi]\!]^{G,w}$),

b. G'' is just like G' except that it stores the top-ranked epistemically live possibility in which the prejacent holds (which is just $(G'_{comp} \cap [\![@E]\!]^{G,w})$) as a novel top-ranked possibility, and pushes all other values one position down, and

c. the final output context H is just like G'' except that it stores the truth-conditional content expressed by the utterance as a new value of '*comp*', which by 2.2, is the set that comprises all the worlds w such that there's some world epistemically accessible from w in which both the restrictor (the top-ranked epistemically live possibility in the input context G) and the prejacent hold (i.e. $M([\![@E]\!]^{G,w}, G'_{comp})$).

Putting all this together, we can now define the update associated with '*might* ϕ'. Let us define a relation between contexts, $\underset{n}{\approx}$, where for any two contexts G and H, this relation holds just in case H is obtained from G by storing a novel value for n and pushing all other values one position down in the ranking; more precisely, $G \underset{n}{\approx} H$ just in case H is identical to G up until n, it differs from G in (at most) the n^{th} position, and for all m, such that $m \geq n, G_m = H_{m+1}$.[71] Then, where K is an update associated with the prejacent ϕ, we define the update associated with an utterance of '*might* ϕ', that carries out the steps in (a.)–(c.), in the following way:

Definition 2.6.
$[\![\text{MIGHT}(@E, K)]\!](w, G, H)$ iff there is a G' and G'' such that $[\![K]\!](w, G, G')$ & $G' \underset{0}{\approx} G''$ & $G''_0 = G'_{comp} \cap [\![@E]\!]^{G,w}$ & $G'' \underset{comp}{\sim} H$ & $H_{comp} = M([\![@E]\!]^{G,w}, G'_{comp})$.

Let us see how this works through an example.

(29) A wolf might walk in.

Given 2.5 and 2.6, we can represent (29) as follows. Let 'p' stand for the prejacent ('a wolf walks in'); since atomic, by 2.5, it's just interpreted as $\langle comp := p\rangle$. Thus, we get:

- $[\![\text{MIGHT}(@E, \langle comp := p\rangle)]\!](w, G, H)$ which by 2.6 holds just in case there is a G' and G'' such that: $[\![\langle comp := p\rangle]\!](w, G, G')$ & $G' \underset{0}{\approx} G''$ & $G''_0 = G'_{comp} \cap [\![@E]\!]^{G,w}$ & $G'' \underset{comp}{\sim} H$ & $H_{comp} = M([\![@E]\!]^{G,w}, G'_{comp})$.

[70] Formally, we include in the basic vocabulary of our dynamic language unary predicates and a unary operator '@'. Where 'E' is a unary predicate and '@ a unary operator, '@E' is an atom. '@E' is interpreted as taking a property denoted by 'E' and delivering the top-ranked proposition satisfying it, denoted by '@E'.

[71] Even more precisely: $G \underset{n}{\approx} H$ iff $\{g_{0,n} + g_{n+1...} | g \in H\} = G$ and $G_{comp} = H_{comp}$. See section § A.3 of the Appendix for details.

- By 2.5, $[\![\langle comp := p\rangle]\!](w, G, G')$ holds just in case $G \underset{comp}{\sim} G'$ & $G'_{comp} = [\![p]\!]^{G,w}$. That is, it holds just in case G'_{comp} is the proposition expressed by p in G and w, namely, the proposition *that a wolf walks in.*

- Moreover, the possibility corresponding to a set of top-ranked epistemically live worlds in which a wolf walks in is introduced as a novel top-ranked possibility, pushing all other possibilities further down in the ranking (thus we get G'').

- Finally, the proposition expressed by the modal utterance is stored as the value of '*comp*' in the final output context H, which is otherwise just like G''; by 2.2, this proposition stored as the value of '*comp*' in H corresponds to a set of worlds R-related to some world in which both the anaphorically retrieved restrictor and the prejacent of the modal hold (i.e. $M([\![@E]\!]^{G,w}, G'_{comp})$). The anaphorically retrieved restrictor is the top-ranked epistemically live possibility in the input context G, (i.e. $[\![@E]\!]^{G,w}$) which, assuming that (29) is uttered out of the blue, just is the set of epistemically accessible worlds discourse initially. Thus, the proposition expressed by (29), the value of *comp* in the output context H, is the proposition that some of the epistemically accessible worlds discourse initially are such that in them a wolf might walk in.

Putting all this together, we get that (29) (a) expresses the proposition that it is compatible with what is known discourse initially that a wolf walks in, and (b) introduces a possibility comprising the top-ranked epistemically live worlds in which a wolf walks in, making it available for subsequent modal anaphora. This is just the desired result.

It is now also easy to see what the updates associated with utterances containing '*must*' and '*likely*' should look like. They will exactly parallel the update associated with utterances containing '*might*', the only difference being in the truth-conditions. So, the update associated with '*likely*' proceeds in exactly the same steps (a.)–(c.). The only difference will transpire in step (c.), where now the truth-conditional content stored as a value of '*comp*' of the final output context, (H), will be the truth-conditional content expressed by an utterance containing '*likely*', which, by 2.3, is $P([\![@E]\!]^{G,w}, G'_{comp})$, i.e. a proposition that requires that given our modal base, the conditional probability of the prejacent, G'_{comp}, given the appropriately restricted modal base, $[\![@E]\!]^{G,w}$, is greater than .5. So, the update associated with '*likely*' can be defined as follows:[72]

Definition 2.7.

$[\![\text{LIKELY}(@E, K)]\!](w, G, H)$ iff there is a G' and G'' such that $[\![K]\!](w, G, G')$ & $G' \underset{0}{\approx} G''$ & $G''_0 = G'_{comp} \cap [\![@E]\!]^{G,w}$ & $G'' \underset{comp}{\sim} H$ & $H_{comp} = P([\![@E]\!]^{G,w}, G'_{comp})$

Finally, we need to specify the update associated with a conditional. The update will proceed in a similar fashion as before, with one difference: now we have to first process the update associated with the antecedent, and then with the consequent. That is, where G is an input context, the update first stores the proposition expressed by the antecedent (in G, relative to the world of evaluation w), as the value of '*comp*' and introduces the top-ranked epistemically live worlds in which the antecedent holds as the top-ranked possibility (pushing all other possibilities one position down). Relative to thus obtained intermediate context (G''), it stores the proposition expressed by the consequent (in G'', relative to w), as the value of '*comp*' and introduces the top-ranked epistemically live worlds in which the consequent holds pushing all other possibilities one position down in the ranking, resulting in the intermediate context (G''''). Lastly, the final output

[72] For the definition of an update associated with '*must*', see § A.4.

context (H) differs from the intermediate context G'''' only insofar as it stores the propositional contribution of the conditional as the value of '*comp*': as per 2.4, it stores the proposition true at a world w just in case all the epistemically accessible worlds from w in which the anaphorically retrieved restrictor and the antecedent hold are such that the consequent holds in them as well.[73] More precisely, the update relates an input context G and an output H just in case there are some contexts G', G'', G''' and G'''', and:

a$'$. G' is a result of updating G with the update associated with the antecedent (thus storing the truth-condition expressed by the antecedent in G, as the value of '*comp*' in G'),

b$'$. G'' is just like G' except that it stores the top-ranked epistemically live possibility in which the antecedent holds, as a novel top-ranked possibility (G''_0), and pushes all other values one position down,

c$'$. G''' is the result of updating G'' with the update associated with the consequent (thus storing the truth-condition expressed by the consequent in G'', as the value of '*comp*' in G'''),

d$'$. G'''' is just like G''' except that it stores the top-ranked epistemically live possibility in which the consequent holds, as a novel top-ranked possibility (G''''_0), and pushes all other values one position down, and finally,

e$'$. the final output context H is just like G'''' except that it stores the truth-conditional content expressed by the whole utterance of the conditional as a new value of '*comp*', which by 2.4, just is $Cond([\![@E]\!]^{G,w}, G'_{comp}, G'''_{comp})$, i.e. a proposition that requires that in all the epistemically accessible worlds in which the antecedent (G'_{comp}) and the restrictor for the conditional ($[\![@E]\!]^{G,w}$) hold, the consequent (G'''_{comp}) holds as well.

Putting all this together, where K_1 and K_2 represent updates associated with the antecedent and the consequent, respectively, we define the update associated with the conditional as follows:

Definition 2.8.
$[\![\textsc{if}(@E, K_1, K_2)]\!](w, G, H)$ iff there is a G', G'', G''' and G'''' such that
$[\![K_1]\!](w, G, G')$ & $G' \underset{0}{\approx} G''$ & $G''_0 = G'_{comp} \cap [\![@E]\!]^{G,w}$ & $[\![K_2]\!](w, G'', G''')$ & $G''' \underset{0}{\approx} G''''$ & $G''''_0 = G'''_{comp} \cap [\![@E]\!]^{G'',w}$ & $G'''' \underset{comp}{\sim} H$ & $H_{comp} = Cond([\![@E]\!]^{G,w}, G'_{comp}, G'''_{comp})$

This completes the specifications of updates associated with modals and conditionals. As I have shown, we capture both aspects of their behavior—namely, we characterize the truth-conditions expressed by an utterance containing a modal or a conditional, and the way in which such an utterance changes the context in which it is uttered. The updates associated with utterances change the context, by updating the prominence ranking of possibilities that are candidate restrictors for subsequent modals and conditionals; the choice of a restrictor in turn affects the truth-conditions of an utterance containing a modal or a conditional. More generally, the updates associated with utterances affect the way in which the context evolves as the discourse progresses; the context in turn determines the truth-conditions expressed by the utterances. Both aspects of interpretation are crucial, and they are interrelated—unless we captured the change in the context prompted by an update associated with an utterance containing a modal or a conditional, we would not be able to tell how the modal can make a possibility available for subsequent anaphora; and unless we calculated in the anaphoric dependency of utterances

[73] As before, I abstract away from the contribution of the ordering source.

containing modals or conditionals, we would not be able to correctly predict which proposition a given utterance containing a modal or a conditional expresses, since the anaphorically retrieved restrictor crucially factors into its truth-conditions.

As we have seen, the updates in 2.5–2.8 all store their corresponding utterances' truth-conditional content as the value of '*comp*' of the output context. But, we also want to characterize what it takes to *assert* this content. Minimally, an assertion of an utterance requires that the proposition expressed holds at the world of evaluation. Plausibly, it also makes the possibility associated with the asserted content prominent. We can capture this by ensuring that an assertion promotes the set of top-ranked epistemically live worlds in which the asserted content holds as a novel top-ranked possibility,[74] and requires that the actual world be within that set. We can introduce a simple assertion update that achieves this effect:

Definition 2.9.

$[\![\textsc{Assert}(K)]\!](w, G, H)$ iff there is a G' such that $[\![K]\!](w, G, G')$ & $G' \underset{0}{\approx} H$ & $H_0 = G'_{comp} \cap [\![@E]\!]^{G,w}$ & $w \in H_0$.

We now have almost all the ingredients we need to capture MT. We need to introduce one last thing—negation. The truth-conditional contribution of '*not* ϕ' is simple—it is true (at a context and relative to a world w) just in case w is a non-ϕ world. We can let the update associated with an utterance of '*not* ϕ' simply store the complement of the truth-condition expressed by 'ϕ' in the input context (relative to w), as the value of '*comp*' of the output context. We define this update as follows:

Definition 2.10.

$[\![\textsc{not}(K)]\!](w, G, H)$ iff there is a G' such that $[\![K]\!](w, G, G')$ & $G' \underset{comp}{\sim} H$ & $H_{comp} = [\![\neg comp]\!]^{G',w}$, where $[\![\neg comp]\!]^{G,w} = \mathcal{D}_\omega \setminus [\![comp]\!]^{G,w}$, where $\mathcal{D}_\omega$ is a set of possible worlds provided by the model.

Finally, we can now tackle the task of specifying MT pattern of inference. *Prima facie*, we can represent the pattern of the form ['*if* ϕ, ψ', '*not* ψ'∴ '*not* ϕ'], as follows. Let $T_d(\phi)$ and $T_d(\psi)$ stand for whatever updates correspond to ϕ and ψ. *Prima facie*, then, the following seems to be the pattern we are after:

(30) $\textsc{Assert}(\textsc{if}(@E, T_d(\phi), T_d(\psi)); \textsc{Assert}(\textsc{not}(T_d(\psi))); \textsc{Assert}(\textsc{not}(T_d(\phi)))$[75]

However, the following is not yet the full logical form of (1)–(3) (indeed, it's not a fully specified logical form at all). First, we cannot know whether (30) has the form of MT or not, until we know what updates $T_d(\phi)$ and $T_d(\psi)$ are. Until this has been specified, the form is incomplete; not every instance of the schema (30) is an instance of MT. As we have seen, the updates that constitute a discourse affect which truth-conditions are expressed by it. Whether or not a discourse that is an instance of (30) will be an instance of MT (partly) depends on whether the truth-conditional content expressed by ψ in the context obtained after updating with the antecedent, and the truth-conditional content expressed by '*not* ψ' in the context obtained after updating with the big premise, indeed do negate each other.[76] Only if this is the case we'll have an instance of MT. And whether or not this condition obtains will depend on the respective

[74]Note that this will basically be a new set of candidate worlds for the actual world.

[75]I use the standard notation, representing the sequencing of updates with a semicolon. Thus, where K_1 and K_2 are updates, $K_1; K_2$ is also an update, that performs the update K_1 followed by K_2 (Muskens, 1996).

[76]I say 'partly' because, while validity is primarily a matter of logical form, it is also a semantic notion, capturing certain semantic patterns. As we shall see in what follows, we will be able to fully capture validity as a matter of logical form. Moreover, we will be able to prove that all classically valid patterns are associated with a valid logical form.

contexts that determine the truth-conditions expressed by 'ψ' and 'not ψ', which in turn will depend on the way the updates that result in these contexts proceed. This is partly a matter of what updates $T_d(\phi)$ and $T_d(\psi)$ are, but it is also a matter of which discourse structuring mechanisms organize the premises and the conclusion into a coherent discourse, since we have seen that those mechanisms also update the context in a way that affects the truth-conditions of modal discourse.

We can then state the following generalization: whenever the truth-conditional content expressed by the small premise negates the one expressed by the consequent of the conditional in the big one, the truth of the big and the small premise together will entail the falsity of the truth-conditional content expressed by the antecedent of the conditional in the big premise. (For a proof of the generalization see A.6.1.) Provided that the updates in (30) satisfy the constraint that the truth-conditional content expressed by the utterances they represent indeed conform to the pattern of MT, that is, that the proposition expressed by the small premise negates the proposition expressed by the consequent of the big one, this pattern is valid; whenever the premises describe a possible update, the conclusion does so as well.[77] Only those fully specified instances of (30) that preserve the adequate form to meaning mapping corresponding to MT will be genuine instances of MT; and all those are valid. This is exactly the same point, as the point that we cannot decide what the full logical form of (12)–(14) is until we know how the different occurrences of the pronoun are resolved. This is hardly a threat to MT.

Note that by characterizing MT in this way, we characterize it as a pattern depending partly on the truth-conditional content expressed, not merely on the underlying syntactic form, because only those ways of specifying (30) that ensure that the truth-conditions of the premises and the conclusion conform to the pattern of MT will count as MT. One might instead maintain that MT should be characterized exclusively in terms of a *unique* syntactic form. Yalcin offers this as an additional argument against MT (Yalcin, 2012). Namely, the standard Kratzerian "restrictor" analysis of conditionals does not recognize the English conditional as a binary operator, and crucially according to this analysis, what is in the scope of the negation in (2) would not even be a constituent of (1) (Kratzer, 1983); thus, Yalcin argues, since there is no single dyadic operator corresponding to the English language conditional (but rather just a multiplicity of different dyadic operators that correspond to different modals), MT, which he takes to be a generalization about this alleged dyadic operator, makes no sense. Since there's no "stable notion" of "antecedent" and "consequent", there is no MT.[78]

I find this line of argument unpersuasive; that a certain syntactic/semantic analysis eliminates a unique binary operator corresponding to the conditional should be independent of the fact that a certain *semantic pattern* is valid. Even if something like the Kratzerian analysis of a conditional turns out to be true, that will hardly constitute a demonstration that MT and MP actually do not exist. Perhaps, we should understand the intuitively valid patterns like MT and MP as precisely the patterns that reflect the behavior of modals in modally subordinated environments, but that does not mean that such patterns are merely an illusion. (Perhaps, it is useful to reflect on the fact that though we can formalize propositional calculus by means of, say, Sheffer stroke, it would be odd to argue on that basis that in such a system MT is somehow ill conceived.)

What, then, follows for the alleged counterexample we began with? *Prima facie*, the argument's structure is as follows. Where 'p' stands for "the marble is big", 'q' for "the marble is red", and '$@E$' as before denotes the top-ranked epistemically live possibility in the context, a

[77] Here, I'm appealing to the standard dynamic notion of validity: an inference pattern is dynamically valid just in case the sequential updates with the premises followed by the update with the conclusion lead to a non-empty context. (See § A.5 for a precise definition.) What we can prove is that dynamic system embed classical logic: whenever the truth-conditions associated with the premises classically entail the ones expressed by the conclusion, the inference pattern is dynamically valid. See § A.

[78] See (Yalcin, 2012).

first pass representation of (1)–(3) is as follows:

(31) $\text{ASSERT}(\text{IF}(@E, \langle comp := p\rangle, \text{LIKELY}(@E, q)); \text{ASSERT}(\text{NOT}(\text{LIKELY}(@E, q)));$
$\text{ASSERT}(\text{NOT}(p)))$

However, (31) is still not a full-blown logical form of (1)–(3), since it leaves out some of the relevant mechanisms that affect the truth-conditions.[79] An instance of MT requires that the truth-conditional content expressed by the consequent of the big premise is a negation of the truth-conditional content expressed by the small premise, and whether this is the case, depends on the way the relevant updates affect the context which determines the truth-conditions; in particular, here it depends on the way the anaphoric dependency (i.e., the value of '$@E$') of the modal in the big premise and the small one is resolved. We have seen earlier that discourse coherence plays a crucial role in resolving modal anaphora; thus, in order to determine whether (1)–(3) is an instance of MT or not, we need to take into account the contribution of these mechanisms, which are left out of (31). To get a full-blown logical form (1)–(3) we need to describe the way that mechanisms of discourse coherence update the context, as well.

I argued that mechanisms of discourse coherence change the context by updating the prominence ranking of possibilities that are candidates for anaphora resolution. Furthermore, I argued that (1)–(3) is not an instance of MT, because the Elaboration relation between the antecedent and the consequent in the big premise requires that the modal in the consequent further elaborates on the possibility made prominent by the antecedent, while the Contrast relation between the small premise and the big premise requires that the modal in the small premise quantifies over the body of information that both premises are about—i.e., the whole set of epistemically accessible worlds discourse initially. So far, I have specified the updates associated with modals and conditionals. Now we need to capture the effects of the mechanisms of discourse coherence on prominence.

We can capture the effect of these mechanisms on the prominence ranking by representing coherence relations as contributing prominence-affecting updates. Let us first characterize Elaboration. I argued that Elaboration promotes the possibility that is elaborated upon. Here's one way of capturing this idea: when an utterance elaborates on a possibility ϕ, a two-fold contribution is made—first, the possibility elaborated upon, ϕ, is promoted to prominence, pushing all other possibilities one position down in the ranking, and second, it is required that the propositional content expressed by the utterance in question, stands in an elaboration relation to ϕ. We can provisionally characterize an elaboration relation between propositions ϕ and ψ, $Elab(\phi, \psi)$, by requiring that it holds just in case ϕ and ψ are centered around the same event or entity, i.e. just in case the event or scenario described by ψ is a part of the event or scenario described by ϕ.[80] Putting all this together, where ϕ is a possibility, and K an update representing the utterance elaborating on ϕ, we can characterize the update associated with Elaboration as follows:

Definition 2.11.

$[\![\text{ELAB}(\phi, K)]\!](w, G, H)$ iff there are G' and G'' such that $G \underset{0}{\approx} G'$ & $G'_0 = [\![\phi]\!]^{G,w}$ & $[\![K]\!](w, G', G'')$ & $G'' \underset{0}{\approx} H$ & $H_0 = G''_{comp}$ & $Elab([\![\phi]\!]^{G,w}, H_0)$.

The contrast relation, as we have seen in § 1, has a different effect on the context than Elaboration. Its main effect is the following: the two bits of discourse provide contrasting

[79] As noted before, once we have the full-blown logical forms we can restore the idea of validity as a matter of logical form: in particular, a sequence of updates expressing classically valid truth-conditional pattern will be dynamically valid, as well.

[80] Cf. Hobbs (1979), Asher and Lascarides (2003). The provisional characterization suffices, because the exact characterization of Elaboration is not crucial for us; the only thing that matters is the way in which the relation affects prominence. Ditto for other coherence relations.

information about some body of information regarding some common topic. In turn, this body of information is made prominent. Thus, to characterize Contrast formally, we need to have a way of specifying the body of information that a given sentence is about—that is, a body of information it contributes information relative to. Here's one way of doing this. Where ϕ is a formula, we say that ϕ is about a set of worlds θ just in case, where G is an input context to ϕ, $\theta = [\![@E]\!]^{G,w}$. The idea is just, once more, that as a discourse progresses we are trying to narrow down the space of epistemic possibilities—thus, a sentence is just about a currently top-ranked epistemic possibility. I will use 'θ_ϕ' to denote the set of worlds ϕ is about. Then we can characterize Contrast as follows:

Definition 2.12.
$[\![\textsc{Contrast}(K_1, K_2)]\!](w, G, H)$ iff there is a G' and G'' such that $[\![K_1]\!](w, G, G')$ & $G' \underset{0}{\approx} G''$ & $G''_0 = [\![\theta_{K_1}]\!]^{G,w}$ & $[\![K_2]\!](w, G'', H)$ & $[\![\theta_{K_1}]\!]^{G,w} = [\![\theta_{K_2}]\!]^{G'',w}$ & $\mathit{Contrast}(G''_{comp}, H_{comp})$

According to 2.12, when two bits of a discourse contrast with each other, a two-fold contribution is made: first, the body of information they are about is made prominent, and second, the propositions expressed by them are required to stand in Contrast relation ($\mathit{Contrast}(G''_{comp}, H_{comp})$), i.e., to provide contrasting information about this body of information, regarding some common topic.[81]

Finally, now that we have specified the ways in which prominence ranking changes as the discourse evolves, putting all this together, we can return to the original counterexample. Where p stands for "the marble is big" and q for "the marble is red", we represent (1) and (2) as follows:

(32) $\textsc{Contrast}(\textsc{Assert}(\textsc{if}(@E, \langle comp := p\rangle, \textsc{Elab}(\mathbf{w}_0, \textsc{Likely}(@E, \langle comp := q\rangle)))),$
$\textsc{Assert}(\textsc{not}(\textsc{Likley}(@E, \langle comp := q\rangle))))$[82]

The following are the key steps in (32).[83] By 2.8, the conditional update introduces the possibility corresponding to the set of top-ranked epistemically accessible worlds in which the proposition expressed by the antecedent holds, i.e. the set of epistemically live worlds in which the marble is big. The consequent provides an elaboration of this possibility; as a result, this possibility is promoted to prominence (as per 2.11). Furthermore, it is required that the possibility introduced by the consequent stands in the Elaboration relation to the possibility introduced by the antecedent, which at this point is the possibility ranked at the position 0 (and, so, denoted by '$\mathbf{w}_0$' in (32)). Since the consequent contains an occurrence of the modal '*likely*', by 2.7, the proposition expressed by the consequent of the given utterance of the conditional corresponds to the proposition that the marble is likely red, given the top-ranked possibility, which due to the effect of Elaboration at this point is the set of epistemically accessible worlds in which the marble is big. Thus, to put it simply, the consequent expresses the proposition that, for all that is known, the marble is likely red, given that it is big. By 2.8 again, the whole conditional expresses the proposition that for all that is known, if the marble is big, then it is likely red, given that it is big. By 2.9, the assertion update requires that the conditional holds of the actual world and promotes the set of epistemically live worlds in which it holds. Due to the effect of Contrast, as

[81] I represent Contrast as operating on two updates, and Elaboration as operating on a proposition and an update. This is in line with a more general distinction between two classes of coherence relations, ones that select their arguments structurally (based on syntactic and structural constraints), and ones that select their arguments anaphorically (Webber et al., 2003). I presuppose this distinction here without defending it, due to considerations of space.

[82] Note that given this formalization, Contrast will not contribute any asserted content on its own. This is a welcome conclusion, but it is inessential. We could in principle make discourse relations a part of asserted content, by imposing additional constraints on the world of evaluation in the specification of the updates associated with coherence relations.

[83] For a detailed derivation, and a proof that (1)–(3) is not an instance of MT, see § A.8.

specified in 2.12, the body of information that the first utterance is about (which, given that the conditional is uttered discourse-initially, as by assumption it is, is just the set of epistemically accessible worlds discourse initially) is promoted to prominence. Then the modal in the small premise will be interpreted with respect to this body of information—given all that is known, the marble is likely red. By 2.10, negation then expresses the complement of this possibility; i.e. the small premise expresses the proposition that it is not the case that given all that is known the marble is likely big. The assertion (by 2.9, again) makes sure that this propositional content holds of the actual world, and promotes the set of epistemically accessible worlds in which the content holds. Finally, the propositions expressed by the two premises are required to stand in the Contrast relation—i.e to provide contrasting information regarding a common topic, about the body of information they are about, i.e. the set of epistemically accessible worlds discourse initially.

Crucially, (32) guarantees that the proposition expressed by the small premise, and the one expressed by the consequent of the big premise do *not* contradict each other. The information expressed is the following: *given all that is known, if the marble is big, then it's likely red, but, given all that is known, the marble is not likely red.* This pattern does not fit the pattern of the premises of MT. So, *a fortiori*, (1)–(3) is not a counterexample to MT.[84]

2.3 Conclusion to Part II

The formal semantics developed in § 2 demonstrates that the account argued for in § 1.2–§ 1.4 can be given a precise formal characterization. The semantics provided in § 2 models the two main aspects of the interpretation of modals and conditionals—the way in which utterances containing these expressions change the context, and the way in which the context determines their truth-conditions. We have seen that both aspects of interpretation are crucial, and that they are interrelated. Unless we captured the change in the context prompted by an update associated with an utterance containing a modal or a conditional, we would not be able to tell how the modal can make a possibility available for subsequent anaphora; and unless we calculated in the anaphoric dependency of utterances containing modals or conditionals, we would not be able to correctly predict which proposition a given utterance containing a modal or a conditional expresses, since the anaphorically retrieved restrictor crucially factors into its truth-conditions. Only once both aspects of the interpretation are taken into account can we give a full explanation of the apparent failure of classically valid patterns of inference.

The formal theory developed in § 2 thus provides a precise theory of context-change, and captures the way the mechanisms of discourse structure affect the interpretation of modals. As we have seen, these mechanisms are independently motivated, and their impact is systematic and rule governed; most importantly, the underlying logic is classical. Thus, the semantics achieves our two main goals: it preserves the validity of classically valid patterns, while at the same time explaining away the apparent counterexamples.

References

Asher, Nicholas and Lascarides, Alex. 2003. *Logics of Conversation.* Cambridge University Press.

[84]Recall, the account maintains that (1)–(3) harbors linguistic elements (modals, antecedents of conditionals, and coherence relations) part of the meaning of which is to change the context in a way that affects the truth-conditions (by introducing certain possibilities, making them prominent, and demoting others). These elements are reflected in the logical form of (1)–(3): it is because (1)–(3) harbors these elements that it is not associated with a valid logical from.

Asher, Nicholas and McCready, Eric. 2006. "Modal Subordination in Japanese: Dynamics and Evidentiality." *University of Pennsylvania Working Papers in Linguistics* 12:237–249.

—. 2007. "Were, Would, Might and a Compositional Account of Counterfactuals." *Journal of Semantics* 24:93–129.

van den Berg, Martin. 1996. *Some Aspects of the Internal Structure of Discourse: The Dynamics of Nominal Anaphora*. Ph.D. thesis, University of Amsterdam.

Bittner, Maria. 2001. "Topical Referents for Individuals and Possibilities." In Rachel Hastings, Brendan Jackson, and Zsófia Zvolenszky (eds.), *Proceedings of SALT XI*, 36–55. CLC Publications, Cornell University, Ithaca.

—. 2014. *Temporality: Universals and Variation*. Wiley-Blackwell.

Bledin, Justin. 2014. "Logic Informed." *Mind* 123:277–316.

—. forthcoming. "*Modus Ponens* Defended." *Journal of Philosophy* .

Brasoveanu, Adrian. 2010. "Decomposing Modal Quantification." *Journal of Semantics* 27:437–527.

Cantwell, John. 2008. "Changing the Modal Context." *Theoria* 74:331–351.

Dekker, Paul. 2011. "Dynamic Semantics." In Klaus von Heusinger, Claudia Maienborn, and Paul Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning, Volume 1.*, 923–945. Berlin: Mouton De Gruyter.

Dorr, Cian and Hawthorne, John. 2014. "Embedding Epistemic Modals." *Mind* 122:867–913.

Dowell, Janice L. 2011. "Flexible Contextualist Account of Epistmeic Modals." *Philosopers' Imprint* 11:1–25.

Egan, Andy, Hawthorne, John, and Weatherson, Brian. 2005. "Epistemic Modals in Context." In Gerhard Preyer and Peter Georg (eds.), *Contextualism in Philosophy*, 131–169. Oxford University Press.

von Fintel, Kai. 1994. *Restrictions on Quantifier Domains*. Ph.D. thesis, University of Massachusetts, Amherst.

von Fintel, Kai and Gillies, Anthony S. 2009. "'Might' Made Right." In Andy Egan and Brian Weatherson (eds.), *Epistemic Modality*. NY: Oxford Univeristy Press.

Gillies, Anthony S. 2004. "Epistemic Conditionals and Conditional Epistemics." *Noûs* 38:585–616.

—. 2010. "Iffiness." *Semantics and Pragmatics* 3:1–42.

Grice, H.P. 1989. "Indicative Conditionals." In *Studies in the Way of Words*, 58–85. Cambridge, MA: Harvard University Press.

Grosz, Barbara J., Joshi, Aravind K., and Weinstein, Scott. 1995. "Centering: A framework for modelling the local coherence of Discourse." *Computational Linguistics* 21:203–225.

Hobbs, Jerry R. 1979. "Coherence and Coreference." *Congitive Science* 3:67–90.

Holliday, Wesley H. and Icard, Thomas F. III. 2013. "Measure Semantics and Qualitative Semantics for Epistemic Modals." *Proceedings of SALT* 23:514–534.

Kaiser, Elsi. 2009. "Effects of Anaphoric Dependencies and Semantic Representations on Pronoun Interpretation." In *Anaphora Processing and Applications*, 121–130. Heidelberg: Springer.

Kaplan, David. 1989a. "Afterthoughts." In Almog Joseph, Perry John, and Wettstein Howard (eds.), *Themes From Kaplan*, 565–614. Oxford University Press.

—. 1989b. "Demonstratives." In Almog Joseph, Perry John, and Wettstein Howard (eds.), *Themes From Kaplan*, 481–563. Oxford University Press.

Kehler, Andrew. 2002. *Coherence, Reference and the Theory of Grammar*. Stanford, CA: CSLI Publications.

Kehler, Andrew, Kertz, Laura, Rohde, Hannah, and Elman, Jeffrey L. 2008. "Coherence and Coreference Revisited." *Journal of Semantics* 25:1–44.

King, Jeffrey C. 2014. "The Metasemantics of Contextual Sensitivity." In Alexis Burgess and Bret Sherman (eds.), *New Essays on Metasemantics*, 97–118. Oxford Univeristy Press.

Kolodny, Niko and MacFarlane, John. 2010. "Ifs and Oughts." *Journal of Philosophy* 107:115–143.

Kratzer, Angelika. 1977. "What *Must* and *Can* Must and Can Mean." *Linguistics and Philosophy* 1:337–355.

—. 1981. "Notional Category of Modality." In Hans-Jürgen Einkmeyer and Hannes Rieser (eds.), *Words, Worlds, and Contexts: New Approaches in Word Semantics*, 38 – 74. Berlin: de Gruyeter.

—. 1983. "Conditionals." In Anne M. Farley, Peter Farley, and Karl-Erik McCullough (eds.), *Papers From the Parasession on Pragmatics and Grammatical Theory*, 1–15. Chicago: Chicago Linguistic Society.

—. 1991. "Modality." In Arnim von Stechow and Dieter Wunderlich (eds.), *Semantics: An international handbook of contemporary research*, 639–50. Berlin: de Gruyter.

—. 2012. *Modals and Conditionals*. Oxford: Oxford University Press.

Lewis, David K. 1979. "Scorekeeping in a Language Game." *Journal of Philosophical Logic* 8:339–359.

MacFarlane, John. 2014. *Assessment Sensitivity: Relative Truth and Its Applications*. Oxford Univeristy Press.

May, Robert. 1977. *The Grammar of Quantification*. Ph.D. thesis, MIT.

McGee, Van. 1985. "A Counterexample to Modus Ponens." *Journal of Philosophy* 92:462–471.

Moss, Sarah. 2015. "On the Semantics and Pragmatics of Epistemic Modals." *Semantics and Pragmatics* 8:1–81.

Muskens, Reinhard. 1996. "Combining Montague Semantics and Discourse Representation." *Linguistics and Philosophy* 19:143–186.

Partee, Barbara. 1973. "Some Structural Analogies between Tenses and Pronouns in English." *The Journal of Philosophy* 70:601–609.

—. 1984. "Nominal and Temporal Anaphora." *Linguistics and Philosophy* 7:243–286.

Roberts, Craige. 1989. "Modal Subordination and Pronominal Anaphora in Discourse." *Linguistics and Philosophy* 12:683–721.

—. 1996. "Information Structure: Towards an integrated formal theory of pragmatics." In Jae Hak Yoon and Andreas Kathol (eds.), *OSUWPL Volume 49: Papers in Semantics*, 91–136. The Ohio State University Department of Linguistics.

Rooth, Mats. 1992. "A Theory of Focus Interpretation." *Natural Language Semantics* 1:75–116.

Schlenker, Philippe. 2013. "Temporal and Modal Anaphora in Sign Language (ASL)." *Natural Language and Linguistic Theory* 31:207–234.

Sidner, Candice. 1983. "Focusing in the Comprehension of Definite Anaphora." In Michael Brady and Robert C. Berwick (eds.), *Computational Models of Discourse*, 267–330. MIT Press.

Smyth, Ron. 1994. "Grammatical Determinants of Ambiguous Pronoun Resolution." *Journal of Psycholinguistic Research* 23:197–229.

Stalnaker, Robert. 1978. "Assertion." *Journal of Semantics and Syntax* 9:315–332.

Stanley, Jason. 2000. "Context and Logical Form." *Linguistics and Philosophy* 23:391–434.

Stanley, Jason and Szabó, Zoltán Gendler. 2000. "On Quantifier Domain Restriction." *Mind and Language* 15:219–261.

Stojnić, Una. 2016a. "Content in a Dynamic Context." *Manuscript, Rutgers University* .

—. 2016b. *Context-Sensitivity in A Coherent Discourse*. Ph.D. thesis, Manuscript, Rutgers Universtiy.

Stojnić, Una, Stone, Matthew, and Lepore, Ernest. 2013. "Deixis (Even Without Pointing)." *Philosophical Perspectives* 27:502–525.

—. 2014. "Discourse and Logical Form." Manuscript, Rutgers University.

Stone, Matthew. 1997. "The Anaphoric Parallel Between Modality and Tense." *IRCS Report 97–06, University of Pennsylvania* .

—. 1999. "Reference to Possible Worlds." *RuCCS Report 49., Rutgers University, New Brunswick* .

Veltman, Frank. 1985. *Logics for Conditionals*. Ph.D. thesis, University of Amsterdam.

Webber, Bonnie L., Stone, Matthew, Joshi, Aravind, and Knott, Alistair. 2003. "Anaphora and Discourse Structure." *Computational Linguistics* 29:545–587.

Willer, Malte. 2010. "New Surprises for the Ramsey Test." *Synthese* 176:291–309.

Wolf, Florian, Gibson, Edward, and Desmet, Timothy. 2004. "Discourse Coherence and Pronoun Resolution." *Language and Cognitive Processes* 19:665–675.

Yalcin, Seth. 2007. "Epistemic Modals." *Mind* 116:983–1026.

—. 2010. "Probability Operators." *Philosophy Compass* 5:916–937.

—. 2011. "Nonfactualism about Epistemic Modality." In Brian Weatherson and Andy Egan (eds.), *Epistemic Modality*, 295–332. NY: Oxford Univeristy Press.

—. 2012. "A Counterexample to Modus Tollens." *Journal of Philosophical Logic* 41:1001–1024.

A Appendix: Formal definitions

In this appendix, I first provide a dynamic language that formalizes my approach to modality with anaphora (§ A.1–§ A.5), and then prove that the system preserves the validity of MP and MT (§ A.6). Then I go on to provide a translation from a fragment of English into this language and into a simple classical language and prove that our dynamic interpretation preserves classical interpretation (§ A.7–A.7.4). Finally, I prove that Yalcin's counterexample is not an instance of MT (§ A.8).

A.1 Syntax:

In this section I specify the expressions of the language. We first start by listing the basic vocabulary:

- Propositional expressions: the elements of the set $\mathcal{C}$ of constants ($p, q, r...$), and the elements of the set $\mathcal{V}$ of variables (*comp* and $\mathbf{w}_n$ for $n \in \mathbb{N}$).
- Unary predicates: P, Q, R
- Unary operator: @
- Update expressions: K, H
- Connectives: $\wedge, \neg$
- Identity: $=$

The following are atomic formulae (atoms) in our language:

- Propositional expressions are atoms.
- $@P$ is an atom, where P is an unary predicate.
- Nothing else is an atom in our language.

These are the conditions in our language:

- All atoms are conditions.
- $\phi = \psi$ is a condition, where ϕ, ψ are conditions. (Stands for identity.)
- $\neg\phi$, where ϕ is a condition. (Stands for negation.)
- $\phi \wedge \psi$ is a condition where ϕ, ψ are conditions. (Stands for conjunction.)

These are the update expressions:

- $\langle comp := \phi \rangle$ is an update expression, where ϕ is an atom.
- If ϕ is a condition, then $[\phi]$ is an update expression.
- $K; K'$ is an update expression, if K is an update expression and K' is an update expression.
- MIGHT(ϕ, K) is an update expression, if ϕ is a condition and K an update expression.
- MUST(ϕ, K) is an update expression, if ϕ is a condition and K an update expression.
- LIKELY(ϕ, K) is an update expression, if ϕ is a condition and K an update expression.
- IF(ϕ, K_1, K_2) is an update expression, if ϕ is a condition and K_1 and K_2 are update expressions.
- AND(K_1, K_2) is an update expression, if K_1 and K_2 are update expressions.
- NOT(K) is an update expression, if K an update expression.
- ASSERT(K) is an update expression, if K is an update expression.
- ELAB(ϕ, K) is an update expression, if ϕ is a condition and K an update expression.
- CONTRAST(K_1, K_2) is an update expression, if K_1 and K_2 are update expressions.

A.2 Models:

I define frames and models in the usual way:

- **A Frame** is a tuple $\mathcal{F} = \langle \mathcal{D}_w, \mathcal{D}_t = \{0, 1\}, R, \mathcal{P} \rangle$ such that $\mathcal{D}_t$ is a domain of truth values ($\mathcal{D}_t = \{0, 1\}$), $\mathcal{D}_w$ is a finite domain of possible worlds, $\mathcal{D}_t \cap \mathcal{D}_w = \emptyset$, with R, a (transitive and reflexive) accessibility relation defined over $\mathcal{D}_w$, and $\mathcal{P}$, a probability measure over $\mathcal{D}_w$, that maps each subset of $\mathcal{D}_w$ to $[0, 1]$, satisfying the following constraints:
  - i $\mathcal{P}(\mathcal{D}_w) = 1$
  - ii $\mathcal{P}(p \cup q) = \mathcal{P}(p) + \mathcal{P}(q)$, when p and q are disjoint subsets of $\mathcal{D}_w$.
  - iii $\mathcal{P}$ is a regular probability measure: if $p \neq \emptyset$ then $\mathcal{P}(p) > 0$.
- **A Model** is a pair $\mathcal{M} = \langle \mathcal{F}, \mathcal{I} \rangle$, where $\mathcal{F}$ is a frame and $\mathcal{I}$ an interpretation function, which assigns to each propositional constant p a subset of $\mathcal{D}_w$ and each predicate constant P a set of subsets of $\mathcal{D}_w$.

A.2.1 Truth-conditional contributions of modals and conditionals:

Let us define some meta-language abbreviations that will help us state the truth-conditions associated with updates associated with modals and conditionals. These correspond to propositions expressed by modals and conditionals.

- Where R is the accessibility relation, and $\mathcal{P}$ the probability measure over sets of possible worlds provided by the model:

Definition A.1. (Definition 2.2 in the text)

$M(p, q) := \{w \mid \exists w' : wRw' \;\&w' \in p \;\&\; w' \in q\}$—*might* q, relative to some possibility, p.

Definition A.2.

$N(p,q) := \{w \mid \forall w' : wRw'$, if $w' \in p$ then $w' \in q\}$—*must* q, relative to some possibility p.

Definition A.3. (Definition 2.3 in the text)

$P(p,q) := \{w \mid \mathcal{P}(\{w' \mid wRw' \ \&\ w' \in p \ \&\ w' \in q\}) \ / \ \mathcal{P}(\{w' \mid wRw' \ \&\ w' \in p\}) > .5\}$—*probably* q, given p.

Definition A.4. (Definition 2.4 in the text)

$Cond(p,q,r) := M(p \ \&\ q, r) = \{w \mid \forall w' : wRw'$, if $w' \in p \ \&\ w' \in q$, then $w' \in r\}$—*if* q, r, relative to p.

A.3 Describing operations on stacks (sequences of worlds) and sets of stacks.

Here I define operations on stacks and sets of stacks, which I will use to define the semantics for our language later on. Formally, a stack is just a function from a finite convex subset of $\mathbb{N}$ plus *comp* to a set of worlds plus $\bot$, where '$\bot$' denotes an undefined value.[85] (I'll assume that '*comp*' is a designated position on the stack. Where the stack is intended to model prominence ranking, '*comp*' is not affecting the prominence ranking, as described in § 2.)

- Where $m \in \mathbb{N}$, and i is a stack, i_m is the m^{th} member of the stack if m is within the domain of i, and $i_m = \bot$ otherwise. (i_{comp} is the member of the stack stored at the designated position *comp*.)— Selecting a member of the stack.
  - Where G is a set of stacks (i.e. a 'context'), g a stack, and u a world, $G_m = \bigcup_{g \in G}\{u | g_m \neq \bot \ \&\ g_m = u\}$, for $m \in \mathbb{N}$ or $m = comp$.—Getting the m^{th} element in the set of stacks G.
- For m, $n \in \mathbb{N}$, and a stack i, $i_{m,n}$ is a stack j defined on the set $\{0, ..., n-m\} \cup \{comp\}$ such that for $k \in \mathbb{N}, j_k = i_{(m+k)}$ if j is defined on k, and $j_{comp} = i_{comp}$.
  - Where G is a context, and g and j are stacks, $G_{m,n} = \bigcup_{g \in G}\{j | j = g_{m,n}\}$ and for $H = G_{m,n}, H_{comp} = G_{comp}$.
- For $m \in \mathbb{N}$, and a stack i, $i_{m...}$ is the stack j defined on the set $\{k \in \mathbb{N} \mid i$ is defined at$(m+k)\} \cup \{comp\}$ such that, for $k \in \mathbb{N}, j_k = i_{(m+k)}$ and $j_{comp} = i_{comp}$.
  - Where G is a context, and g, j are stacks, $G_{m...} = \bigcup_{g \in G}\{j | j = g_{m...}\}$ and for $H = G_{m...}, H_{comp} = G_{comp}$.
- If i is a stack with a finite domain with maximal element $k-1$ then for a stack j, $i + j$ is a stack h where, for $x \in \mathbb{N}, h_x = i_x$ if i is defined at x, and $h_x = j_{(x-k)}$ otherwise (and $h_{comp} = i_{comp}$).
- Where u is a world and i is a stack, u, i is a stack j, such that $j_0 = u$, and for all $n \in \mathbb{N}$, such that $n > 0, j_n = i_{(n-1)}$ if i is defined on n, and $j_n = \bot$ otherwise and $j_{comp} = i_{comp}$.— Appending to a stack.
  - Where G is a context, u is a world, and g, j are stacks, $G_{u...} = \bigcup_{g \in G}\{j | j = u, g\}$ and for $H = G_{u...}, H_{comp} = G_{comp}$.
- $g[n]g'$ iff $g_m = g'_m$ for $m \neq n$ (where $m, n \in \mathbb{N} \cup \{comp\}$).

[85] A set of numbers S is convex just in case if $x \in S$, $y \in S$ and $x < m < y$ then m is in S.

- $G \underset{n}{\sim} G'$ iff $\{g'|g[n]g', g \in G\} = \{g'|g[n]g', g \in G'\}$ (where $n \in \mathbb{N} \cup \{comp\}$).
- $G \underset{n}{\approx} G'$ iff $\{g_{0,n} + g_{n+1...}|g \in G'\} = G$ and $G_{comp} = G'_{comp}$.

A.4 Semantics:

The Interpretation of Atoms: The interpretation of an expression e, relative to the interpretation function $\mathcal{I}$ a context G, and a world w:

- $[\![p]\!]^{G,w} = \mathcal{I}(p)$, if $p \in \mathcal{C}$.
  - Constants.
- $[\![\mathbf{w}_m]\!]^{G,w} = G_m$, if $\mathbf{w}_m \in \mathcal{V}$ and $m \in \mathbb{N}$
  - Variables.
- $[\![comp]\!]^{G,w} = G_{comp}$
  - A designated position on the stack.
- $[\![@P]\!]^{G,w} = \emptyset$ if $G_0 = \bot$, $[\![@P]\!]^{G,w} = G_0$, if $G_0 \in \mathcal{I}(P)$, and $[\![@P]\!]^{G,w} = [\![@P]\!]^{G_{1...},w}$ otherwise.
  - Find the top ranked entity in G, satisfying P.

The Interpretation of Conditions:

- $[\![\phi = \psi]\!]^{G,w} = \mathcal{D}_\omega$, if $[\![\phi]\!]^{G,w} = [\![\psi]\!]^{G,w}$; $[\![\phi = \psi]\!]^{G,w} = \emptyset$, otherwise.
  - Identity.
- $[\![\neg\phi]\!]^{G,w} = \mathcal{D}_\omega \setminus [\![\phi]\!]^{G,w}$.
  - Negation.
- $[\![\phi \wedge \psi]\!]^{G,w} = [\![\phi]\!]^{G,w} \cap [\![\psi]\!]^{G,w}$.
  - Conjunction.

The Interpretation of Update Expressions

- $[\![\langle comp := p\rangle]\!](w, G, H)$ iff $G \underset{comp}{\sim} H$ & $H_{comp} = [\![p]\!]^{G,w}$
- $[\![[\phi]]\!](w, G, H)$ if and only if $H = G$ and $w \in [\![\phi]\!]^{G,w}$
- $[\![K; K']\!](w, G, H)$ iff $\exists G' : [\![K]\!](w, G, G')$ and $[\![K']\!](w, G', H)$
- The following are updates that describe how propositional content (A.2.1) in context is determined. Where p is a proposition (an anaphorically retrieved restrictor) and '$@E$' denotes the top-ranked proposition that is the subset of the epistemically accessible worlds:[86]

[86]For generality, I let the restrictor in the definition be any proposition p. However, as argued above, epistemic modals and conditionals select the top-ranked possibility in a given context ('$@E$') as their restrictor.

- $[\![\text{MIGHT}(\phi, K)]\!](w, G, H)$ iff there is a G' and G'' such that $[\![K]\!](w, G, G')$ & $G' \underset{0}{\approx} G''$ & $G''_0 = G'_{comp} \cap [\![@E]\!]^{G,w}$ & $G'' \underset{comp}{\sim} H$ & $H_{comp} = M([\![\phi]\!]^{G,w}, G'_{comp})$

- $[\![\text{MUST}(\phi, K)]\!](w, G, H)$ iff there is a G' and G'' such that $[\![K]\!](w, G, G')$ & $G' \underset{0}{\approx} G''$ & $G''_0 = G'_{comp} \cap [\![@E]\!]^{G,w}$ & $G'' \underset{comp}{\sim} H$ & $H_{comp} = N([\![\phi]\!]^{G,w}, G'_{comp})$

- $[\![\text{LIKELY}(\phi, K)]\!](w, G, H)$ if and only if there is a G' and G'' such that $[\![K]\!](w, G, G')$ & $G' \underset{0}{\approx} G''$ & $G''_0 = G'_{comp} \cap [\![@E]\!]^{G,w}$ & $G'' \underset{comp}{\sim} H$ & $H_{comp} = P([\![\phi]\!]^{G,w}, G'_{comp})$

- $[\![\text{IF}(\phi, K_1, K_2)]\!](w, G, H)$ iff there is a G', G'', G''' and G'''' such that $[\![K_1]\!](w, G, G')$ & $G' \underset{0}{\approx} G''$ & $G''_0 = G'_{comp} \cap [\![@E]\!]^{G,w}$ & $[\![K_2]\!](w, G'', G''')$ & $G''' \underset{0}{\approx} G''''$ & $G''''_0 = G'''_{comp} \cap [\![@E]\!]^{G'',w}$ & $G'''' \underset{comp}{\sim} H$ & $H_{comp} = Cond([\![\phi]\!]^{G,w}, G'_{comp}, G'''_{comp})$

- $[\![\text{AND}(K_1, K_2)]\!](w, G, H)$ iff there is a G', G'', G''' and G'''' such that $[\![K_1]\!](w, G, G')$ & $G' \underset{0}{\approx} G''$ & $G''_0 = G'_{comp} \cap [\![@E]\!]^{G,w}$ & $[\![K_2]\!](w, G'', G''')$ & $G''' \underset{0}{\approx} G''''$ & $G''''_0 = G'''_{comp} \cap [\![@E]\!]^{G'',w}$ & $G'''' \underset{arg_0}{\sim} H$ & $H_{comp} = G'_{comp} \cap G''_{comp}$

- $[\![\text{NOT}(K)]\!](w, G, H)$ iff there is a G' such that $[\![K]\!](w, G, G')$ & $G' \underset{comp}{\sim} H$ & $H_{comp} = [\![\neg comp]\!]^{G',w}$

- $[\![\text{ASSERT}(K)]\!](w, G, H)$ iff there is a G' such that $[\![K]\!](w, G, G')$ & $G' \underset{0}{\approx} H$ & $H_0 = G'_{comp} \cap [\![@E]\!]^{G,w}$ & $w \in H_0$

In order to define the truth-conditions for updates associated with coherence relations, we assume the following abbreviations:

Definition A.5. $Elab(\phi, \psi)$ iff ϕ and ψ are centered around the same event or entity, i.e. iff the event or scenario described by ψ is a part of the event or scenario described by ϕ.

Definition A.6. A formula, ϕ, is about of body of information θ iff, where G is the input context to ϕ, $\theta = [\![@E]\!]^{G,w}$, where '$E$' is a predicate denoting the property of being an epistemically accessible proposition, and thus, '$@E$' denotes the top-ranked epistemically accessible proposition. I use 'θ_ϕ' to denote the body of information that ϕ is about.

Definition A.7. $Contrast(\phi, \psi)$ iff ϕ and ψ describe contrasting information about some body of information regarding a common topic.

- $[\![\text{ELAB}(\phi, K)]\!](w, G, H)$ iff there are G' and G'' such that $G \underset{0}{\approx} G'$ & $G'_0 = [\![\phi]\!]^{G,w}$ & $[\![K]\!](w, G', G'')$ & $G'' \underset{0}{\approx} H$ & $H_0 = G''_{comp}$ & $Elab([\![\phi]\!]^{G,w}, H_0)$.

- $[\![\text{CONTRAST}(K_1, K_2)]\!](w, G, H)$ iff there is a G' and G'' such that $[\![K_1]\!](w, G, G')$ & $G' \underset{0}{\approx} G''$ & $G''_0 = [\![\theta_{K_1}]\!]^{G,w}$ & $[\![K_2]\!](w, G'', H)$ & $[\![\theta_{K_1}]\!]^{G,w} = [\![\theta_{K_2}]\!]^{G'',w}$ & $Contrast(G''_{comp}, H_{comp})$

A.5 Truth, validity, entailment.

- K is true, relative to a context G, a world w, and a model $\mathcal{M}$, if there is some H, s.t. $H \neq \emptyset$ and $[\![K]\!](w, G, H)$. K is false (relative to a context G, a world w, and a model $\mathcal{M}$,) otherwise.
- K is valid iff it's true in all models.
- K_1 entails K_2 iff for any model $\mathcal{M}$, any context G, and any world w if there is a G' such that $G' \neq \emptyset$ and $[\![K_1]\!](w, G, G')$, then there is a G'' such that $G'' \neq \emptyset$ and $[\![K_2]\!](w, G', G'')$.

A.6 Modus Ponens (MP) and Modus Tollens (MT):

We want to prove that MP and MT are preserved within our system. In order to do that, we need to have a way of characterizing MT and MP correctly.

A.6.1 MT

Let us start with MT. We need to have a way of individuating instances of MT, first. Let us begin with the following discourse:

1. $\text{Assert}(\text{if}(@E, K_1, K_2)); \mathit{Assert}(\text{not}(K_3))$

- where '$@E$' denotes the top-ranked epistemically accessible proposition.

What we want to show is that when the propositions expressed by K_2 and $\text{not}(K_3)$ in their respective contexts contradict each other, then if the proposition expressed by $\text{if}(@E, K_1,$ $K_2)$ and $\text{not}(K_3)$ both hold (relative to the world of evaluation w), then the proposition corresponding to the intersection of the truth-conditional contribution of K_1 and $[\![@E]\!]^{G,w}$, where G is the input context, will be false (relative to the world of evaluation w). (Note that for our proof it does not really matter what '$@E$' is; the proof goes through regardless of what we take the anaphoric restrictor to be. I use '$@E$' because, as I argue in the paper, epistemic modals and conditionals select top-ranked epistemic possibility in the context as a restrictor; but this is orthogonal to our proof.)

Proof.

2. $[\![\text{Assert}(\text{if}(@E, K_1, K_2)); \text{Assert}(\text{not}(K_3)))]\!](w, G, H)$ iff there is a G', such that $[\![\text{Assert}(\text{if}(@E, K_1, K_2))]\!](w, G, G')$ and $[\![\text{Assert}(\text{not}(K_3))]\!](w, G', H)$.
3. Take such a G, G' and H.
4. $[\![\text{Assert}(\text{if}(@E, K_1, K_2))]\!](w, G, G')$ iff there is a G'', such that $[\![\text{if}(@E, K_1, K_2)]\!](w, G, G'')$ & $G'' \underset{0}{\approx} G'$ & $G'_0 = G''_{comp} \cap [\![@E]\!]^{G,w}$ & $w \in G'_0$.
5. Take such a G''.
6. $[\![\text{if}(@E, K_1, K_2)]\!](w, G, G'')$ iff H', H'', H''' and H'''', such that $[\![K_1]\!](w, G, H')$& $H' \underset{0}{\approx} H''$ & $H''_0 = H'_{comp} \cap [\![@E]\!]^{G,w}$ & $[\![K_2]\!](w, H'', H''')$ & $H''' \underset{0}{\approx} H''''$ & $H''''_0 =$ $H'''_{comp} \cap [\![@E]\!]^{H'',w}$ & $H'''' \underset{comp}{\sim} G''$ & $G''_{comp} = \mathit{Cond}([\![@E]\!]^{G,w}, H'_{comp}, H'''_{comp})$.
7. Take such H', H'', H''', and H''''.

8. $[\![\text{ASSERT}(\text{NOT}(K_3))]\!](w, G', H)$ iff there is a G''' such that $[\![\text{NOT}(K_3)]\!](w, G', G''')$ & $G''' \underset{0}{\approx} H$ & $H_0 = G'''_{comp} \cap [\![@E]\!]^{G,w}$ & $w \in H_0$.

9. Take such G'''.

10. $[\![\text{NOT}(K_3)]\!](w, G', G''')$ iff there is a G'''' such that $[\![K_3]\!](w, G', G'''')$ & $G'''' \underset{comp}{\sim} G'''$ & $G'''_{comp} = [\![\neg comp]\!]^{G'''',w}$.

11. Now, what is left to prove is that when $H'''_{comp} \cap G'''_{comp} = \emptyset$ & $w \in Cond([\![@E]\!]^{G,w}, H'_{comp}, H'''_{comp})$ & $w \in G'''_{comp}$, then $w \notin H'_{comp} \cap [\![@E]\!]^{G,w}$ (i.e. when the truth-conditional contribution of the small premise negates the truth-conditional contribution of the consequent, and both premises hold at a world w, then the antecedent does not hold at w).

12. *Per reductio*, suppose that $H'''_{comp} \cap G'''_{comp} = \emptyset$ & $w \in Cond([\![@E]\!]^{G,w}, H'_{comp}, H'''_{comp})$ & $w \in G'''_{comp}$, and also $w \in H'_{comp} \cap [\![@E]\!]^{G,w}$.

13. Then, w is such that $\forall w' : wRw'$, if $w' \in [\![@E]\!]^{G,w}$ & $w' \in H'_{comp}$ then $w' \in H'''_{comp}$, and also $w \in G'''_{comp}$.

14. Since, $w \in G'''_{comp}$, and $H'''_{comp} \cap G'''_{comp} = \emptyset$, then $w \notin H'''_{comp}$.

15. Suppose that $w \in [\![@E]\!]^{G,w}$ & $w \in H'_{comp}$. (Since R is reflexive and transitive, we know that wRw.)

16. Then, $w \in H'''_{comp}$.

17. Contradiction!

18. So, if $H'''_{comp} \cap G'''_{comp} = \emptyset$ & $w \in Cond([\![@E]\!]^{G,w}, H'_{comp}, H'''_{comp})$ & $w \in G'''_{comp}$, then $w \notin H'_{comp} \cap [\![@E]\!]^{G,w}$. □

- Note that whenever the proposition expressed by $\text{NOT}(K_3)$ negates the one expressed by K_2, and the one expressed by $\text{NOT}(K_4)$ negates the one expressed by the proposition expressed by the intersection of the truth-conditional contribution of K_1 and $[\![@E]\!]^G, w$, where G is the input context for the conditional, given (§ A.4) and (§ A.5), $\text{ASSERT}(\text{IF}(@E, K_1, K_2))$; $Assert(\text{NOT}(K_3))$ will entail $Assert(\text{NOT}(K_4))$.

A.6.2 MP

As with MT, we need to have a way of individuating instances of MP, first. Let us begin with the following discourse:

1. $\text{ASSERT}(\text{IF}(@E, K_1, K_2)); \text{ASSERT}(K_3)$

What we want to show is that when the propositions expressed by K_3 and the one corresponding to the intersection of the propositions expressed by K_1 and '$@E$' are identical, if the proposition expressed by $\text{IF}(@E, K_1, K_2)$ and K_3 hold (relative to a world of evaluation w), the proposition corresponding to the truth-conditional contribution of K_2 will be true (relative to a world of evaluation w, as well).

Proof.

2. $[\![\textsc{Assert}(\textsc{if}(@E, K_1, K_2); \textsc{Assert}(K_3))]\!](w, G, H)$ iff there is a G' such that $[\![\textsc{Assert}(\textsc{if}(@E, K_1, K_2)]\!](w, G, G')$ and $[\![\textsc{Assert}(K_3)]\!](w, G', H)$.

3. Take such a G, G' and H.

4. $[\![\textsc{Assert}(\textsc{if}(@E, K_1, K_2));]\!](w, G, G')$ iff there is a G'' such that $[\![\textsc{if}(@E, K_1, K_2)]\!](w, G, G'')$ & $G'' \underset{0}{\approx} G'$ & $G'_0 = G''_{comp} \cap [\![@E]\!]^{G,w}$ & $w \in G'_0$.

5. Take such a G''.

6. $[\![\textsc{if}(@E, K_1, K_2)]\!](w, G, G'')$ iff H', H'', H''' and H'''' such that $[\![K_1]\!](w, G, H')$& $H' \underset{0}{\approx} H''$ & $H''_0 = H'_{comp} \cap [\![@E]\!]^{G,w}$ & $[\![K_2]\!](w, H'', H''')$ & $H''' \underset{0}{\approx} H''''$ & $H''''_0 = H'''_{comp} \cap [\![@E]\!]^{H'',w}$ & $H'''' \underset{comp}{\sim} G''$ & $G''_{comp} = Cond([\![@E]\!]^{G,w}, H'_{comp}, H'''_{comp})$.

7. Take such H', H'', H''', and H''''.

8. $[\![\textsc{Assert}(K_3)]\!](w, G', H)$ iff there is a G''' such that $[\![K_3]\!](w, G', G''')$ & $G''' \underset{0}{\approx} H$ & $H_0 = G'''_{comp} \cap [\![@E]\!]^{G,w}$ & $w \in H_0$.

9. It is easy to show given § A.4 that if $[\![(K_3)]\!](w, G', G''')$, whatever form K_3 has, its truth-conditional bit will be stored as a value of G'''_{comp}.

10. Suppose $[\![(K_3)]\!](w, G', G''')$.

11. So, now we just have to prove that when $G'''_{comp} = [\![@E]\!]^{G,w} \cap H'_{comp}$, and $w \in Cond([\![@E]\!]^{G,w}, H'_{comp}, H'''_{comp})$, and $w \in G'_{comp}$, then $w \in H'''_{comp}$ (i.e. when the truth-conditional contribution of the small premise and the antecedent are identical, and both the antecedent and the small premise hold at w, then the consequent holds at w, as well).

12. *Per reductio*, suppose that $G'''_{comp} = [\![@E]\!]^{G,w} \cap H'_{comp}$, and $w \in Cond([\![@E]\!]^{G,w}, H'_{comp}, H'''_{comp})$ and $w \in G'''_{comp}$, and also $w \notin H'''_{comp}$.

13. Then w is such that $\forall w' : wRw'$, if $w' \in [\![@E]\!]^{G,w}$ & $w' \in H'_{comp}$ then $w' \in H'''_{comp}$, and also $w \in G'''_{comp}$.

14. By (12), $w \notin H'''_{comp}$.

15. Then, by (14) and reflexivity of R, $\exists w' : wRw'$, and $w' \in [\![@E]\!]^{G,w}$ & $w' \in H'_{comp}$ and $w' \notin H'''_{comp}$, namely, w.

16. Contradiction!

17. So, if $G'''_{comp} = [\![@E]\!]^{G,w} \cap H'_{comp}$ and $w \in Cond([\![@E]\!]^{G,w}, H'_{comp}, H'''_{comp})$ and $w \in G'''_{comp}$, then $w \in H'''_{comp}$. □

- Note that whenever the proposition expressed by K_3 is identical to the one corresponding to the intersection of the propositions expressed by K_1 and $@E$, and the one expressed by K_2 identical to the one expressed by K_4, given (§ A.4) and (§ A.5), $\textsc{Assert}(\textsc{if}(@E, K_1, K_2)); \mathit{Assert}(K_3)$ will entail $\mathit{Assert}(\textsc{not}(K_4))$.

A.7 Relation between the dynamic (A.7.2) and classical (A.7.3) interpretations

I shall now prove that my dynamic interpretation preserves a classical one. To this end, I shall first give a dynamic translation for a fragment of English, specifying the updates associated with utterances containing modals and conditionals. Then I will give a classical translation for the same fragment, and prove that the dynamic interpretation preserves the truth-conditions assigned by the classical interpretation. For ease of comparison between the two translations, I shall avail myself of abstract level of logical forms (ALFs) for the relevant fragment of English. The reader should bear in mind that we do not have to take a stand on the existence of a level of representation corresponding to ALFs. This level of representation is merely a dispensable convenience that helps us compare the two interpretations in a streamlined way.

A.7.1 Abstract Logical Forms (ALFs) for a Fragment of English:

Terms:

- Propositional constants from our base language in § A.1 (set $\mathcal{C}$).

Atoms:

- All terms are atoms, and nothing else is an atom.

ALFs:

- Atoms are ALFs.
- If ϕ and ψ are ALFs, then $might(\phi, \psi)$ is an ALF. (Stands for "it might be the case that ψ, given the restrictor ϕ".)
- If ϕ and ψ are ALFs, then $must(\phi, \psi)$ is an ALF. (Stands for "it must be the case that ψ, given the restrictor ϕ".)
- If ϕ and ψ are ALFs, then $likely(\phi, \psi)$ is an ALF. (Stands for "it's likely the case that ψ, given the restrictor ϕ".)
- If ϕ, ψ and γ are ALFs, then $if(\phi, \psi, \gamma)$ is an ALF. (Stands for "given the restrictor ϕ if ψ, then γ.")
- If ϕ and ψ are ALFs, then $and(\phi, \psi)$ is an ALF. (Stands for "ϕ and ψ.")
- If ϕ is an ALF, then $not(\phi)$ is an ALFs. (Stands for "Not ϕ.")
- If ϕ is an ALF, then $Assert(\phi)$ is an ALF. (Assertion operator—makes sure that the proposition ϕ is asserted.)

A.7.2 Dynamic Interpretation:

I first give a translation of the relevant fragment of English into our dynamically interpreted language (A.7.2.1), and then a translation of the same fragment into a classically interpreted language (A.7.3)

A.7.2.1 Dynamic Translations

In this section, I provide a translation of the relevant fragment of English, into our dynamically interpreted language defined in § A.1–§ A.4. I'll assume the ALFs for the relevant fragment of English defined in A.7.1, (e.g. $might(\phi, \psi)$ for "it might be the case that ψ", where the modal is anaphorically dependent on $[\![\phi]\!]^{G,w}$, for an input context G.)

(Base case, where $T_d(\phi)$ is a translation of a formula ϕ into our dynamic system.)

- If ϕ is an atom, then $T_d(\phi) = \langle comp := \phi \rangle$.

(Recursive case)

- $T_d(might(\phi, \psi)) = \textsc{might}(\phi, T_d(\psi))$
- $T_d(must(\phi, \psi)) = \textsc{must}(\phi, T_d(\psi))$
- $T_d(likely(\phi, \psi)) = \textsc{likely}(\phi, T_d(\psi))$
- $T_d(if(\phi, \psi, \gamma)) = \textsc{if}(\phi, T_d(\psi), T_d(\gamma))$
- $T_d(and(\phi, \psi)) = \textsc{and}(T_d(\phi), T_d(\psi))$
- $T_d(not(\phi)) = \textsc{not}(T_d(\phi))$
- $T_d(Assert(\phi)) = \textsc{Assert}(T_d(\phi))$

Next, I shall introduce the classical interpretation of modals and conditionals in A.7.3, and then go on to prove that our dynamic interpretation of modals and conditionals (A.7.2.1) preserves the classical interpretation in (A.7.4).

A.7.3 Classical Interpretation

I will now introduce a classical toy modal language, and provide a translation of the relevant fragment of English into this modal langauge, so I can compare our dynamic interpretation and classical interpretation. Assume the following modal language:

Terms:

- Propositional terms: propositional constants (p, q, r).

Atomic formulae:

- All terms are atoms, and nothing else is an atom.

Now we introduce well-formed formulae:

- All atoms are well-formed formulae.
- If ϕ and ψ are formulae, then $\Diamond(\phi, \psi)$ is a well-formed formula.
- If ϕ and ψ are formulae, then $\Box(\phi, \psi)$ is a well-formed formula.
- If ϕ and ψ are formulae, then $\phi \to \psi$ is a well-formed formula.
- If ϕ and ψ are formulae, then $\phi \wedge \psi$ is a well-formed formula.
- Nothing else is a well-formed formula.

A.7.3.1 Classical Semantics:

I now define classical semantics for the simple modal language. I assume models are defined as in § A.2. (Assuming the definition of models in § A.2, and sets of sequences in § A.3), where R is an accessibility relation provided by the model, and ϕ a restriction on the domain of quantification of a modal:

- $[\![p]\!]^{G,w} = w \in \mathcal{I}(p)$
- $[\![\Diamond(\phi,\psi)]\!]^{G,w} = \{w|\exists w'\colon wRw' \ \&\ w' \in [\![\phi]\!]^{G,w}, w' \in [\![\psi]\!]^{G,w}\}$
- $[\![\Box(\phi,\psi)]\!]^{G,w} = \{w|\forall w'\colon \text{if } wRw' \ \&\ w' \in [\![\phi]\!]^{G,w} \text{ then } w' \in [\![\psi]\!]^{G,w}\}$
- $[\![\phi \wedge \psi]\!]^{G,w} = [\![\phi]\!]^{G,w} \cap [\![\psi]\!]^{G,w}$
- $[\![(\phi \wedge \psi) \rightarrow \gamma]\!]^{G,w} = \{w|\forall w'\colon \text{if } wRw' \ \&\ w' \in [\![\phi]\!]^{G,w} \cap [\![\psi]\!]^{G,w} \text{ then } w' \in [\![\gamma]\!]^{G,w}\}$

A.7.3.2 Classical Translation:

Assuming the same ALFs for the relevant fragment of English as in A.7.1, I now define the following translations of the relevant bits of the fragment into our classical language.
Where T_c(p) is a translation of a formula p into classical system:

- Where ϕ is an atom $T_c(\phi) = [\![\phi]\!]^{G,w}$
- $T_c(might(\phi,\psi)) = [\![\Diamond(\phi,\psi)]\!]^{G,w}$
- $T_c(must(\phi,\psi)) = [\![\Box(\phi,\psi)]\!]^{G,w}$
- $T_c(if(\phi,\psi,\gamma)) = [\![(\phi \wedge \psi) \rightarrow \gamma]\!]^{G,w}$

This translation does not capture the systematic effects of context on the interpretation of modals: it doesn't control for where the restrictor comes from. But, note that we don't have to care about what the restrictor ϕ is. We can simply assume that it is the domain of all possible worlds ($\mathcal{D}_w$), the set of epistemically live worlds, or any other proposition. So long as all the anaphoric restrictors are denoting the same proposition, what the restrictor is won't matter, because by resolving the restrictor always to the same proposition we are making sure that the domain of quantification for modals is held constant throughout, as in a simple classical modal logic.

A.7.4 Proof.

- We want to prove that our dynamic interpretation (A.7.2) preserves the classical interpretation (in A.7.3). In particular, we prove that for any $T_d(p)$, if $[\![T_d(p)]\!](w,G,H)$, then $H_{comp} = [\![T_c(p)]\!]^{G,w}$, where $[\![T_c(p)]\!]^{G,w}$ is the corresponding translation of the formula p in classical system; $Assert(T_d(p))$ guarantees that $[\![T_c(p)]\!]^{G,w}$ is true at the actual world. We do a proof by induction.

* Base case. First prove that for an atom p, and translation $T_d(p)$, if $[\![T_d(p)]\!](w,G,H)$, then $H_{comp} = [\![p]\!]^{w}$.

Proof.

2. By (A.7.2.1), $T_d(p) = \langle comp := p \rangle$.
3. By (§ A.4), we have $[\![\langle comp := p\rangle]\!](w,G,H)$ iff $G \underset{comp}{\sim} H$ & $H_{comp} = [\![p]\!]^{G,w}$.

4. Suppose $[\![\langle comp := p\rangle]\!](w, G, H)$.

5. By (2)–(4), $H_{comp} = [\![p]\!]^{G,w}$, and $[\![p]\!]^{G,w}$ iff $w \in \mathcal{I}(p)$. Thus, $H_{comp} = [\![T_c(p)]\!]^{G,w}$, which we were set to prove. □

* Recursive case.

Proof.

- **IH:** Assume that for a formula ϕ of the depth k or less, if $[\![T_d(\phi)]\!](w, G, H)$, then $H_{comp} = [\![\phi]\!]^w$, where $[\![\phi]\!]^w$ is the corresponding classical interpretation of the formula ϕ.

- Consider a formula ϕ of the depth $k+1$. We prove that the **IH** holds for the possible ways of constructing ϕ:

i Let ϕ =*might*(χ, ψ). Then, by A.7.2.1, $T_d(\phi) = \textsc{might}(\chi, T_d(\psi))$. Suppose that $[\![\textsc{might}(\chi, T_d(\psi))]\!](w, G, H)$. We know by (§ A.4) that $[\![\textsc{might}(\chi, T_d(\psi))]\!](w, G, H)$ iff there is a G' and G'' such that $[\![T_d(\psi)]\!](w, G, G')$ & $G' \underset{0}{\approx} G''$ & $G''_0 = G'_{comp} \cap [\![@E]\!]^{G,w}$ & $G'' \underset{comp}{\sim} H$ & $H_{comp} = M([\![\chi]\!]^{G,w}, G'_{comp})$. Take such a G' and G''. We have that $[\![T_d(\psi)]\!](w, G, G')$; thus, by **IH**, $G'_{comp} = [\![\psi]\!]^{G,w}$. Then, since $H_{comp} = M([\![\chi]\!]^{G,w}, G'_{comp})$, given Definition (A.1) and (A.7.3.2), H_{comp} is equivalent to $[\![T_c(might(\chi, \psi))]\!]^{G,w}$, by simple math.

ii Let ϕ = *must*(χ, ψ). Then, by A.7.2.1, $T_d(\phi) = \textsc{must}(\chi, T_d(\psi))$. Suppose that $[\![\textsc{must}(\chi, T_d(\psi))]\!](w, G, H)$. So, by (§ A.4), there is a G' and G'' such that $[\![T_d(\psi)]\!](w, G, G')$ & $G' \underset{0}{\approx} G''$ & $G''_0 = G'_{comp} \cap [\![@E]\!]^{G,w}$ & $G'' \underset{comp}{\sim} H$ & $H_{comp} = N([\![\chi]\!]^{G,w}, G'_{comp})$. Take such a G' and G''. Since $[\![T_d(\psi)]\!](w, G, G')$, by **IH**, $G'_{comp} = [\![\psi]\!]^{G,w}$. Then, since $H_{comp} = M([\![\chi]\!]^{G,w}, G'_{comp})$, given Definition (A.2) and (A.7.3.2), H_{comp} is equivalent to $[\![T_c(must(\chi, \psi))]\!]^{G,w}$, by simple math.

iv Let χ = *if*$(\chi,\ \psi,\ \gamma)$. Then, by A.7.2.1, $T_d(\chi) = \textsc{if}(\chi, T_d(\psi), T_d(\gamma))$. Suppose that $[\![\textsc{if}(\chi, T_d(\psi), T_d(\gamma))]\!](w, G, H)$. So, by (§ A.4), we know that there is a G', G'', G''' and G'''' such that $[\![T_d(\psi)]\!](w, G, G')$ & $G' \underset{0}{\approx} G''$ & $G''_0 = G'_{comp} \cap [\![@E]\!]^{G,w}$ & $[\![T_d(\gamma)]\!](w, G'', G''')$ & $G''' \underset{0}{\approx} G''''$ & $G''''_0 = G'''_{comp} \cap [\![@E]\!]^{G'',w}$ & $G'''' \underset{comp}{\sim} H$ & $H_{comp} = Cond([\![\chi]\!]^{G,w}, G'_{comp}, G''_{comp})$. Take such a G', G'', G''' and G''''. Since $[\![T_d(\psi)]\!](w, G, G')$, by **IH** we know that $G'_{comp} = [\![\psi]\!]^{G,w}$ and since $[\![T_d(\gamma)]\!](w, G'', G''')$, we know by **IH** that $G'''_{comp} = [\![\gamma]\!]^{G,w}$. Then, since $H_{comp} = Cond([\![\chi]\!]^{G,w}, G'_{comp}, G'''_{comp})$ given Definition (A.4) and (A.7.3.2), H_{comp} is equivalent to $[\![T_c(cond(\chi, \psi, \gamma))]\!]^{G,w}$, by simple math. □

A.8 Yalcin's Counterexample:

I shall now formally represent Yalcin's counterexample in our system and prove that the counterexample is not an instance of MT (and, thus, is not a counterexample). Recall the two premises of the counterexample:

(1) If the marble is big, then it's likely red.

(2) The marble is not likely red.

Let $\mathcal{I}(p)$ be the proposition *that the marble is big*, and $\mathcal{I}(q)$ be the proposition *that the marble is red*. Then we represent the two premises of the counterexample as in (32), repeated below:

(32) $\textsc{Contrast}(\textsc{Assert}(\textsc{if}(@E, \langle comp := p\rangle, \textsc{Elab}(\mathbf{w}_0, \textsc{Likely}(@E, \langle comp := q\rangle)))),$
$\textsc{Assert}(\textsc{not}(\textsc{Likley}(@E, \langle comp := q\rangle))))$

Let 'K_1' stand for '$\textsc{if}(@E, \langle comp := p\rangle, \textsc{Elab}(\mathbf{w}_0, \textsc{Likely}(@E, \langle comp := q\rangle)))$' and '$K_2$' for '$\textsc{not}(\textsc{Likley}(@E, \langle comp := q\rangle))$'. Then, we show that (32) is not an instance of MT.

Proof. First, we shall calculate the truth-conditions expressed by the consequent of the big premise, then, we shall calculate the truth-conditions expressed by the small premise, and then we show that the two do not contradict each other.

1. $[\![\textsc{Contrast}(\textsc{Assert}(K_1); \textsc{Assert}(K_2))]\!](w, G, H)$ iff there is a G' and G'' such that $[\![\textsc{Assert}(K_1)]\!](w, G, G')$ & $G' \underset{0}{\approx} G''$ & $G''_0 = [\![\theta_{\textsc{Assert}(K_1)}]\!]^{G,w}$ & $[\![\textsc{Assert}(K_2)]\!](w, G'', H)$ & $[\![\theta_{\textsc{Assert}(K_1)}]\!]^{G,w} = [\![\theta_{\textsc{Assert}(K_2)}]\!]^{G'',w}$ & $Contrast(G'''_{comp}, H_{comp})$. (By (§ A.4).)

2. Take such a $G, G', G'' and H$. Then $[\![\textsc{Assert}(K_1)]\!](w, G, G')$ and $[\![\textsc{Assert}(K_2)]\!](w, G'', H)$. (By (1).)

3. First we calculate the truth-condition of the consequent of the big premise. $[\![\textsc{Assert}(K_1)]\!](w, G, G')$ iff there is a G''' such that $[\![K_1]\!](w, G, G''')$ & $G''' \underset{0}{\approx} G'$ & $G'_0 = G'''_{comp} \cap [\![@E]\!]^{G,w}$ & $w \in G'_0$. (By (§ A.4).)

4. Take such G'''. $[\![K_1]\!](w, G, G''')$ iff there is a H', H'', H''' and H'''' such that $[\![\langle comp := p\rangle]\!](w, G, H')$ & $H' \underset{0}{\approx} H''$ & $H''_0 = H'_{comp} \cap [\![@E]\!]^{G,w}$ & $[\![\textsc{Elab}(\mathbf{w}_0, \textsc{Likely}(@E, \langle comp := q\rangle))))]\!](w, H'', H''')$ & $H''' \underset{0}{\approx} H''''$ & $H''''_0 = H'''_{comp} \cap [\![@E]\!]^{H'',w}$ & $H'''' \underset{comp}{\sim} G'''$ & $G'''_{comp} = Cond([\![@E]\!]^{G,w}, H'_{comp}, H'''_{comp})$. (By (§ A.4) and the definition of K_1.) Note that by § A.4 and A.7.2.1, H'''_{comp} stores the truth-conditions of the consequent of the big premise.

5. Take such H', H'', H''' and H''''. $[\![\textsc{Elab}(\mathbf{w}_0, \textsc{Likely}(@E, \langle comp := q\rangle))))]\!](w, H'', H''')$ iff there are some J and J' such that $H'' \underset{0}{\approx} J$ & $J_0 = [\![\mathbf{w}_0]\!]^{H'',w}$ & $[\![\textsc{Likely}(@E, \langle comp := q\rangle)]\!](w, J, J')$ & $J' \underset{0}{\approx} H'''$ & $H'''_0 = J'_{comp}$ & $Elab([\![\mathbf{w}_0]\!]^{H'',w}, H'''_0)$. (By (§ A.4).)

6. Take such J and J'. $[\![\textsc{Likely}(@E, \langle comp := q\rangle)]\!](w, J, J')$ iff there is a J'' and J''' such that $[\![\langle comp := q\rangle]\!](w, J, J'')$ & $J'' \underset{0}{\approx} J'''$ & $J'''_0 = J''_{comp} \cap [\![@E]\!]^{J,w}$ & $J''' \underset{comp}{\sim} J'$ & $J'_{comp} = P([\![@E]\!]^{J,w}, J''_{comp})$. (By (§ A.4).)

7. Thus, by (A.3), $J'_{comp} = \{w | \mathcal{P}(\{w' \mid wRw'$ & $w' \in [\![@E]\!]^{J,w}$ & $w' \in J'_{comp}\})$ / $\mathcal{P}(\{w' \mid wRw'$ & $w' \in [\![@E]\!]^{J,w}\}) > .5\}$.

8. By 5 and the definition of '$\underset{n}{\approx}$', $J'_{comp} = H'''_{comp}$.

9. By (4)–(6), and (A.7.2.1), $H'''_{comp} = \{w | \mathcal{P}(\{w' \mid wRw'$ & $w' \in [\![p]\!]^{G,w} \cap [\![@E]\!]^{G,w}$ & $w' \in \mathcal{I}(q)\})$ / $\mathcal{P}(\{w' \mid wRw'$ & $w' \in [\![p]\!]^{G,w} \cap [\![@E]\!]^{G,w}\}) > .5\}$. These are the truth-conditions expressed by the consequent of the big premise. Now we calculate the truth-conditions of the small premise.

10. From (2) we have: $[\![\text{ASSERT}(K_2)]\!](w, G'', H)$.
By (§ A.4), $[\![\text{ASSERT}(K_2)]\!](w, G'', H)$ iff there is a I such that $[\![K_2]\!](w, G'', I)$ & $I \underset{0}{\approx} H$ & $H_0 = I_{comp} \cap [\![@E]\!]^{G'',w}$ & $w \in H_0$.

11. Take such I. By definition of K_2 and (§ A.4), we have: $[\![K_2]\!](w, G'', I)$ iff there is a I' such that $[\![\text{LIKLEY}(@E, \langle comp := q\rangle)]\!](w, G'', I')$ & $I' \underset{comp}{\sim} I$ & $I_{comp} = [\![\neg comp]\!]^{I',w}$.

12. Take such I'. Then, by (§ A.4), $[\![\text{LIKLEY}(@E, \langle comp := q\rangle)]\!](w, G'', I')$ iff there is a I'' and I''' such that $[\![\langle comp := q\rangle]\!](w, G'', I'')$ & $I'' \underset{0}{\approx} I'''$ & $I_0''' = I''_{comp} \cap [\![@E]\!]^{G'',w}$ & $I''' \underset{comp}{\sim} I'$ & $I'_{comp} = P([\![@E]\!]^{G'',w}, I''_{comp})$

13. By (11), (6), and (A.3), the truth-condition expressed by the small premise is as follows: $\mathcal{D}_w \setminus \{w | \mathcal{P}(\{w' \mid wRw' \text{ \& } w' \in [\![@E]\!]^{G'',w} \text{ \& } w' \in I''_{comp}\}) \,/\, \mathcal{P}(\{w' \mid wRw' \text{ \& } w' \in [\![@E]\!]^{G'',w}\}) > .5\}$, where $\mathcal{D}_w$ is the domain of possible worlds from the model.

14. By (1), (A.6), (§ A.4) and (A.7.2.1), we have: $\mathcal{D}_w \setminus \{w | \mathcal{P}(\{w' \mid wRw' \text{ \& } w' \in [\![@E]\!]^{G,w} \text{ \& } w' \in \mathcal{I}(q)\}) \,/\, \mathcal{P}(\{w' \mid wRw' \text{ \& } w' \in [\![@E]\!]^{G,w}\}) > .5\}$., where $\mathcal{D}_w$ is the domain of possible worlds from the model. This is the truth-condition expressed by the small premise.

15. From (9) and (13), we see that the truth-conditions expressed by the big premise and the consequent of the small one do not contradict each other.[87] Hence, (32) does not correspond to the premises of an instance of MT. □

[87] Yalcin's scenario with an urn with 100 marbles can be used to generate a model in which both propositions are true. In particular, suppose the domain of worlds $\mathcal{D}_w$ is partitioned according to a color-size distribution: into big and blue, small and blue, big and red and small and red worlds. Where $\mathcal{I}(p)$ is the proposition *that the marble is big*, $\mathcal{I}(q)$ the proposition *that the marble is red*, $\mathcal{I}(r)$ the proposition *that the marble is blue* and $\mathcal{I}(s)$ the proposition *that the marble is small*, let the probability measure assign the following probabilities: $\mathcal{P}(\mathcal{I}(p) \cap \mathcal{I}(q)) = .3, \mathcal{P}(\mathcal{I}(p) \cap \mathcal{I}(r)) = .1, \mathcal{P}(\mathcal{I}(s) \cap \mathcal{I}(q)) = .1$, and $\mathcal{P}(\mathcal{I}(s) \cap \mathcal{I}(r)) = .5$.