

Counterfactuals and Modality

Gabriel Greenberg | October 17, 2019

This essay calls attention to a set of linguistic interactions between counterfactual conditionals, on one hand, and possibility modals like *could have* and *might have*, on the other. These data, I argue, present a serious challenge to the popular *variably strict* semantic analysis of counterfactual conditionals. The same evidence instead supports a new version of the *strict conditional* semantics, which I call the *unified strict analysis*. According to the unified strict analysis, counterfactuals and possibility modals share a single, contextually determined quantificational domain. I'll argue that pragmatic explanations of the same data which support the unified analysis are not available to the variable analysis. And putative counterexamples to the unified analysis, on careful inspection, in fact support it. Ultimately, the semantics of conditionals and modals must be linked together more closely than has been assumed, and, I argue, the unified analysis is the only way to fully achieve this.

Section 1 first outlines the standard semantics for modals and the rival strict and variably strict semantics for conditionals, then introduces the unified strict analysis. Sections 2 and 3 describe two patterns of modal-conditional interaction respectively and examine their semantic implications; I argue that the unified analysis explains both patterns directly, but no such semantic explanation is available to the variably strict analysis. Section 4 considers possible pragmatic explanations of the data on behalf the variably strict analysis, but finds that these too fall short. Section 5 discusses variations of the interaction data which at first appear to counterexample the unified analysis, but ultimately end up supporting it. Section 6 takes stock of the essay's conclusions.

1 Counterfactuals and Modals

Counterfactual (or *subjunctive*) *conditionals* belong to a broader linguistic class of counterfactual (or subjunctive) *modals*, including *might-counterfactuals*, as well as *bare possibility* and *bare necessity* modals. My focus in what follows is a set of interactions between counterfactual conditionals and bare possibility modals.

For present purposes, COUNTERFACTUAL BARE POSSIBILITY MODALS (for short, MODALS) will be represented symbolically as $\diamond\phi$, corresponding to sentences of the form:

- (1) It could/might have been the case that ϕ .

Here I'll follow the practice of referring to ϕ as the PREJACENT of the sentence $\diamond\phi$. Note that the

Acknowledgements: This essay has benefited from years of conversations with teachers and classmates at Rutgers University and students and colleagues at UCLA, as well as the comments of anonymous reviewers, an audience at the 2009 CEU Summer School on Conditionals, and the METH Philosophy of Language Group. Special thanks to Sam Cumming, David Kaplan, and Jessica Rett for detailed notes on recent drafts.

intended counterfactual reading of (1) can be distinguished from an alternative past-epistemic reading, because only the counterfactual reading is compatible with denial of the prejacent: “It could have happened, but it didn’t.”¹ (Starr 2014, p. 1023)

I’ll assume a widely accepted approach to the semantic analysis of modals that understands them to contribute a quantifier over possible worlds with existential force (Kripke 1963; Kratzer 1977; Kratzer 2012).² Roughly, $\diamond\phi$ is true if and only if there is some ϕ -world among the set of ACCESSIBLE worlds. Here I’ll let k be a contextually-selected function from worlds of evaluation to sets of accessible worlds. Then, for any world i and context c :

Standard semantics for possibility modals

$\lceil \diamond\phi \rceil$ is **true** _{i,c} iff $\exists w \in k(i) : w \in \llbracket \phi \rrbracket$

There is a world w among the worlds accessible from i , $k(i)$, such that w is a ϕ -world.

Next I’ll represent COUNTERFACTUAL CONDITIONALS (for short, CONDITIONALS) symbolically as $\phi > \psi$, corresponding to sentences of roughly the following form:³

- (2) If it had been/were the case that ϕ , it would have been/would be the case that ψ .

Two semantic theories of conditionals have distinguished themselves in recent years. The dominant view is the VARIABLY STRICT ANALYSIS, for short, the VARIABLE ANALYSIS (Stalnaker 1968; Lewis 1973; Kratzer 1981), while the chief alternative is the STRICT CONDITIONAL ANALYSIS, for short, the STRICT ANALYSIS (Lowe 1995; von Stechow 2001; Gillies 2007).⁴

For the semantics of the variable analysis, I introduce a contextually determined SELECTION FUNCTION f which takes as arguments an evaluation world and a sentence, and returns the set of worlds “closest” to the evaluation world in which the sentence is true. Such a function is defined relative to a closeness ordering \geq_i , a partial ordering over worlds, with the world of evaluation i as a closest element. Meanwhile, the strict analysis employs the more familiar notion of a set of accessible worlds. Here I use a contextually-determined accessibility function s , which maps worlds to sets of accessible worlds.

¹Another diagnostic is that, on the epistemic reading, synonymy is preserved if *could have/might have* is replaced by *may have*, but not so on the counterfactual reading.

²By stating the semantics in terms of truth-conditions here and throughout, I do not mean to exclude presuppositional or dynamic aspects of modal semantics, if any. I’ll discuss these, along with additional pragmatic and discourse-based amendments, as the need arises.

³I assume that there is some relatively systematic way of identifying counterfactual conditionals, and that, at some legitimate level of abstraction, counterfactual conditionals can be translated by a sentence of propositional logic containing two sub-clauses and a single logical connective. For discussion, see e.g. Kratzer 1986 and Lycan 2001, ch. 1.

⁴The strict view has ancient roots, perhaps originating with the Stoic philosopher Chrysippus in the third century B.C.E. (Sanford 2003); it was first defended in the 20th century by C.I. Lewis (1914) and C.S. Pierce (1933). Contemporary, context-sensitive versions of the theory have been developed by Ken Warmbröd (1981a; 1981b), William Lycan (1984; 2001), E.J. Lowe (1983; 1995), Kai von Stechow (2001), and Thony Gillies (2007). The subtle difference in truth-conditions between the strict and variably strict semantics give rise to significantly different logics. For discussion, see Stalnaker (1968), Lewis (1973), Fine (1975), Creary and Hill (1975), Loewer (1976), Ellis, Jackson, and Pargetter (1977), McKay and Van Inwagen (1977), Nute (1980), Warmbröd (1981b), Lowe (1995), and Lycan (2001).

The strict and variable analyses may now be stated, for any world i and context c :⁵

The Strict Analysis

$$\lceil \phi > \psi \rceil \text{ is true}_{i,c} \quad \text{iff} \quad \forall w : w \in s(i) \cap \llbracket \phi \rrbracket \rightarrow w \in \llbracket \psi \rrbracket$$

For every world w among the ϕ -worlds accessible worlds from i , w is a ψ -world.

The Variable Analysis

$$\lceil \phi > \psi \rceil \text{ is true}_{i,c} \quad \text{iff} \quad \forall w : w \in f(\phi, i) \rightarrow w \in \llbracket \psi \rrbracket$$

For every world w among the closest ϕ -worlds relative to i , w is a ψ -world.

The two analyses differ because the set of accessible antecedent-worlds comes apart from the set of *closest* antecedent-worlds. Even if we restrict the ordering of worlds to accessible worlds, the strict conditional always quantifies over the full set of accessible antecedent-worlds, while the variable analysis would only quantify over the closest antecedent-worlds among the accessible ones. It is this flexible departure from the set of accessible worlds which is distinctive of the variable analysis, but also the feature I will contest in the discussion to come.

As I have presented the two analyses of conditionals here, there's no assumed connection between the semantics of modals and conditionals. But in what follows I will highlight linguistic data which reveals substantive coordination between conditionals and modals. The view that I'll be defending, the UNIFIED STRICT ANALYSIS (OR UNIFIED ANALYSIS), is a version of the strict analysis in which modals and conditionals quantify over the same domain of counterfactual possibility. Formally, the domain for conditionals, s , is simply identified with the modal domain k , $k = s$. Thus conditionals express restricted universal quantifiers over the set of accessible worlds, while possibility modals express existential quantifiers over the same set of accessible worlds.⁶

The Unified Strict Analysis

$$\lceil \phi > \psi \rceil \text{ is true}_{i,c} \quad \text{iff} \quad \forall w \in k(i) : w \in \llbracket \phi \rrbracket \rightarrow w \in \llbracket \psi \rrbracket$$

$$\lceil \Diamond \phi \rceil \text{ is true}_{i,c} \quad \text{iff} \quad \exists w \in k(i) : w \in \llbracket \phi \rrbracket$$

In this analysis, conditionals and modals are conceived as devices of quantification that are keyed to a common domain. In this respect, they follow the pattern of the variety of explicit quantifiers in English, like nominal quantifiers (*a(n), some, no, all*), temporal quantifiers (*always, sometimes, never*), or personal quantifiers (*everbody, somebody*). Each of these involves a class of quantificational expressions, each of which is semantically tuned to a common, contextually-selected domain of quantification. According to the unified analysis, counterfactual modals are

⁵The definitions of the strict and variable analyses offered here constitute the semantic core of each view, but neither is limited to this core. A variety of additional constraints are commonly imposed, for example in order to validate *Modus Ponens*. The definition of the variable analysis in the text combines features of Stalnaker 1968 and Lewis 1973; it denies *Uniqueness* (with Lewis) and accepts the *Limit Assumption* (with Stalnaker).

⁶von Fintel (2001) briefly considers a view of this kind in footnote 15.

no different.

In defending the unified strict analysis, I assume that modals and conditionals are embedded in a system of contextual update in discourse, along the lines of von Fintel (2001) and Gillies (2007). As von Fintel and Gillies have shown, such a system allows a strict analysis to accommodate much of the challenging data which originally motivated the variable analysis over the strict analysis. For present purposes it doesn't matter if this system is built into the modal semantics, pragmatics, or some other means of discourse management.

2 *Impossibility Above*

I begin with a class of cases where the quantificational domain of modals appears to put constraints on that of conditionals. After reviewing the data, I'll show how both the unified and variable analyses have compelling semantic explanations of the evidence available to them. In the next section, I turn to more challenging phenomena.

2.1 *Impossibility Above: the evidence*

Consider the inference exhibited below.⁷

- (3) a. Though she stayed home, Isabel could have gone sailing Tuesday afternoon.
- b. But she couldn't have gone sailing and not brought a hat.
- c. So if she had gone sailing, she would have brought a hat.

Line (a) introduces the scene and ensures that the subsequent modals receive a counterfactual, rather than epistemic, reading. The object of interest is the transition from (b) to (c), that is, from a sentence of the form $\neg\Diamond(\textit{sail} \wedge \neg\textit{enjoy})$ to one of the form $\textit{sail} > \textit{enjoy}$. It seems that (c) can be legitimately inferred from (b), and this impression is reinforced by the felicitous use of the inferential marker *so* at the opening of (c). Other inferential markers like *therefore*, *thus*, *hence*, and *it follows that* work just as well.

I propose that this example conforms with a general pattern of linguistic inference, which I call *Impossibility Above* (IA), displayed below:⁸

$$\textit{Impossibility Above (IA)} \quad \frac{\neg\Diamond(\phi \wedge \neg\psi)}{\phi > \psi}$$

My claim for now is that IA is simply an INFERENCE PATTERN, a schema for sequences in which

⁷Unless otherwise noted, the reader should assume that all examples in this essay are spoken by the same individual, and that this individual is speaking in some kind of deductive context – either reasoning to herself or trying convince an interlocutor of her conclusions.

⁸Williamson (2008, p. 156) discusses this inference pattern, considered as a logical principle, under the name "NECESSITY".

the second sentence seems to “follow from” the first. By this I don’t mean, necessarily, that IA is a *logical entailment*, though I will ultimately come to that conclusion. The inferential force in question may be deductive, ampliative, some type of pragmatic inference, or a purely epistemic inferential relation. What the linguistic evidence most immediately reveals, I believe, is that a speaker who asserts the premise is thereby *committed* (in some sense) to the truth of the conclusion, and one who denies the conclusion is similarly committed to the denial of the premise. I will say that, according to IA, the premise *IMPLIES* the conclusion, in a similarly neutral sense of *imply*.

Further evidence that IA corresponds to an genuine inference pattern is revealed in a second set of data. These are cases which involve a stark *infelicity* when IA is violated. When reading these examples, it is important to imagine the speaker using a relatively flat affect, and to resist the temptation to substitute the contrastive discourse marker *but* for the discourse conjunctions *and* or *whatsmore* used in the examples below. The significance of this constraint is discussed in Section 5; the point, for now, is to avoid unintentional shifts to the relevant modal context.

- (4) a. Norbert could have gone home over the holidays.
- b. But he couldn’t have gone home and met the deadline.
- c. * What’s more, if he had gone home, he would have met the deadline.

- (5) a. Norbert could have gone home over the holidays.
- b. If he had gone home, he would have met the deadline.
- c. * And he couldn’t have gone home and met the deadline.

This type of example suggests that it is infelicitous to utter sentences of the form $\neg\Diamond(\phi \wedge \psi)$ and $\phi > \psi$ in conjunction. This is directly explained by IA: for by IA, $\neg\Diamond(\phi \wedge \psi)$ implies $\phi > \neg\psi$; further, assuming the antecedent is not contradictory, then on either semantics of the conditional, $\phi > \neg\psi$ entails $\neg(\phi > \psi)$; consequently $\neg\Diamond(\phi \wedge \psi)$ implies $\neg(\phi > \psi)$. But (4) and (5) above involve asserting the premise and the negation of the conclusion in conjunction, which we should expect to yield infelicity. I provisionally conclude that the IA is a genuine inference pattern.

2.2 *Impossibility Above*: semantic explanations

What explains the inferential force of IA? The unified analysis has a straightforward explanation, for according to the unified analysis, IA is simply a logical entailment. For if it is true at a world i that $\neg\Diamond(\phi \wedge \neg\psi)$, then there are no accessible $\phi \wedge \neg\psi$ -worlds. Thus all accessible ϕ -worlds (if any) are ψ -worlds; so, by the unified semantics of conditionals, $\phi > \psi$ must be true at i as well.⁹ The patterns of felicitous inference highlighted above reflect logical entailment; the

⁹A qualification: recent advocates of the strict analysis have proposed that conditionals presuppose the compatibility of their antecedents with the possibilities made available by the context of evaluation (von Stechow 1998, p. 2; von Stechow

patterns of infelicitous assertion reflect that fact that it is generally infelicitous to assert contradictions (Stalnaker 1978).

What of the variable analysis? As I presented the variable analysis in Section 1, there is no built-in alignment between the semantics for possibility modals and the semantics for conditionals. In particular, there are no constraints relating the values of the selection function f and the accessibility function k . The absence of constraint here gives rise to models which straightforwardly violate IA.

That said, there is a sympathetic amendment one can make to the variable analysis which results in the validation of IA. It is natural to think that all the worlds relevant to the evaluation of a conditional must be considered possible in the context in which it is expressed. Formally, we may express this idea as the constraint that every world in the domain of the closeness ordering must also be among the set of accessible worlds. I call this constraint *Accessibility*, defined below.

Accessibility

For any world i , comparative closeness function \geq , and modal accessibility function k :

The set of worlds ordered by \geq_i is a subset of $k(i)$.

IA is thereby validated on the variable analysis. For according to the premise of IA, there are no $\phi \wedge \neg\psi$ -worlds in $k(i)$, for an arbitrary i . Thus every ϕ -world in $k(i)$ is a ψ -world. By *Accessibility*, the closest ϕ -worlds to i must be in $k(i)$. So they must be ψ -worlds. So $\phi > \psi$ is true at i , on the variable analysis.

Since this way of validating IA is so readily available to variable theorists, and because it is independently plausible, I'll assume in what follows that the variable analysis takes *Accessibility* on board; there is good reason to accept it and none to reject it.

3 Impossibility Below

I turn next to linguistic evidence which cannot be so easily accommodated by the variable analysis. Whereas, in the IA-cases, the domain of the modal appears to constrain that of the conditional, the next set of cases suggest the reverse effect. As before, I begin with the apparent inference pattern exhibited in the data, then turn to possible semantic explanations.

2001, pp. 15-20; Gillies 2007, pp. 333-4). In the framework of this essay, this means that conditionals presuppose that their antecedents are compatible with the set of accessible worlds. If this principle is correct, and assuming presupposition failure leads to change in semantic value, then only a slightly weaker principle than IA is viable, namely:

$$\text{Weak Impossibility Above} \quad \frac{\neg\Diamond(\phi \wedge \neg\psi) \quad \Diamond\phi}{\phi > \psi}$$

But this does not disrupt the argument of the essay. The data which I cited as conforming with IA supports Weak IA just as well, and the unified analysis explains the inferential force of both by making both semantically valid.

3.1 *Impossibility Below: the evidence*

Consider the following inferential discourses:

- (6) a. Amira didn't play in the match yesterday, but she could have.
- b. If she had played, she would have won.
- c. So she couldn't have played and lost.
- (7) a. The season finale aired Saturday when Katie was busy, but it could have aired Tuesday.
- b. If the finale had aired Tuesday, Katie would have watched.
- c. Thus the finale couldn't have aired Tuesday without Katie watching.

Focusing on (6) for illustration, it seems to be appropriate to infer from (b) to (c), that is, from a sentence of the form $play > win$ to one of the form $\neg\Diamond(play \wedge \neg win)$. And again, this impression is reinforced by the felicitous use of *so* (or *thus, therefore, etc.*) at the beginning of (c).

These observations suggest the existence of a second inference pattern, this one the mirror image of *Impossibility Above*. I call this, *Impossibility Below* (IB).¹⁰

$$\text{Impossibility Below (IB)} \quad \frac{\phi > \psi}{\neg\Diamond(\phi \wedge \neg\psi)}$$

A range of further data supports the inferential force of IB. Consider the following dialogues involving *denial*:

- (8) a. A: If Talia had played, she would have won.
- b. B: No, she might have played and lost.
- (9) a. A: Talia might have played and lost.
- b. B: No, if she played, she would have won.

The felicitous use of the B-sentences to deny the A-sentences in each case suggest that sentences of the form $\phi > \psi$ and $\Diamond(\phi \wedge \neg\psi)$ are considered incompatible in context. And this is explained directly by IB, since by IB, $\phi > \psi$ implies $\neg\Diamond(\phi \wedge \neg\psi)$.

In a similar manner, IB can be used to explain the following *infelicities* as the assertion of incompatible propositions. (In evaluating these and other cases presented in this section, it is crucial to maintain a "flat" or "neutral" tone, since there is a temptation to put focus on the modal, which tends to shift the interpretation in a way that eliminates the proffered infelicity. I return to this effect below, and in Section 5.)

- (10) a. Yesterday Ivan decided to drive to work.

¹⁰This inference pattern came to me via John Hawthorne's 2008 metaphysics seminar at Rutgers University; he presented the inference in its inverted form, reasoning from $\Diamond(\phi \wedge \neg\psi)$ to $\neg(\phi > \psi)$.

- b. If he had stopped for coffee, he would have arrived late.
 - c. * And he might have stopped for coffee and arrived on time.
- (11)
- a. Talia missed the ping-pong match last Tuesday.
 - b. But she might have played and won.
 - c. * What's more, if she had played, she would have lost.

The same explanation seems to apply here: if IB is a genuine inference pattern, and $\phi > \psi$ implies $\neg\Diamond(\phi \wedge \neg\psi)$ then it should be infelicitous to assert $\phi > \psi$ and $\Diamond(\phi \wedge \neg\psi)$ in conjunction, just as we see in the examples above.

Vivid variants of these cases involve infelicitous dialogues, where the second interlocutor explicitly signals agreement with the first through the use of the affirmative *yes*.

- (12)
- a. A: If Talia had played, she would have won.
 - b. * B: Yes, and she might have played and lost.
- (13)
- a. A: Talia might have played and lost.
 - b. * B: Yes, and if she played, she would have won.

Another type of example relies on the inversion of IB, derived by reasoning from the negation of its conclusion to the negation of its premise, from $\Diamond(\phi \wedge \neg\psi)$ to $\neg(\phi > \psi)$. This inference is exhibited directly in the following script:

- (14)
- a. John wanted to take Organic Chemistry. But in all fairness, it's a very difficult class.
 - b. He could have taken it and failed.
 - c. So it isn't true that if he had taken it, he would have passed.

A final case is aimed at those (I count myself among them) who accept the duality of *would*-conditionals and *might*-conditionals, represented as $\phi \Diamond\rightarrow \psi$. (Bennett 2003, §73) That is:

(15) $\phi > \psi \Leftrightarrow \neg(\phi \Diamond\rightarrow \neg\psi)$

To add to this, I submit that the following inference is generally impeccable:

$$\text{Might Inference} \quad \frac{\Diamond(\phi \wedge \psi)}{\phi \Diamond\rightarrow \psi}$$

- (16)
- a. Isabel worked at the library all day yesterday.
 - b. But she might have gone into town and seen her friends.
 - c. So if she had gone into town, she might have seen her friends.

But if one assumes duality and *Might Inference*, then IB follows directly. For one can infer from $\phi > \psi$ to $\neg(\phi \diamond \rightarrow \neg\psi)$, by duality; and one can infer from $\neg(\phi \diamond \rightarrow \neg\psi)$ to $\neg\diamond(\phi \wedge \neg\psi)$, by *Might Inference*. So this is additional evidence for the inferential force of IB.

Stepping back from the evidence, my experience is that speakers vary in the degree to which they feel the force of the judgments reported here. Some, like myself, find the judgments in support of IB just as strong as those which support IA; others find the IB judgments of inference and infelicity in some way weaker, though nevertheless present. In any case, even the weaker judgments will do for my purposes. What is important to the present argument is not that the inferences intuitively feel like *validities*, or that the infelicities feel like *contradictions*— validity and contradiction are a matter of theory. Instead, what is important is that there are indeed judgments of some kind which conform with IB, even if they are less vivid. As I'll ultimately argue, on semantic grounds alone the variable semantics can only explain the judgments that do not conform with IB, but none of the judgments that support IB. Meanwhile, I'll argue that the unified strict analysis can explain both.

At this point readers will naturally be drawn to consider whether there are counterexamples to IB. These tend to take the form of felicitous conjunctions like the following:

- (17) a. If Kiara had played, she would have won.
b. But of course, she *might* have played and lost.

I'll discuss this and parallel examples in greater detail in Section 5. Cases like (17) seem at first blush to counter-example IB. But they notably include a cluster of distinctive linguistic markers, including focal stress (on *might*), discourse intensifiers (*of course, in fact*), and discourse contrast markers (*but, however*). In Section 5 I'll argue that, collectively, these devices have the effect of shifting the modal context between the first and second sentences. As a consequence, discourses like (17) do not represent challenges to the classical validity of IB, which assumes the fixity of context.

Importantly, the infelicitous conjunctions presented above, (10) and (11), did not invoke this cluster of markers. What's more, when these are stripped from (17), the result is again infelicitous, lending credence to the hypothesis that the markers in question works as context-shifters. Once again, this phenomenon and related effects will be examined in Section 5.

- (18) a. If Kiara had played, she would have won.
b. * And she might have played and lost.

I mention these points here to forestall the impression that there are *obvious* counterexamples to IB. Support for IB comes from sentences which do not change context. So assessment of IB must carefully control for shifts in context, and this in turn will require closer attention to certain kinds of linguistic detail, such as discourse markers, than has been the custom in debates about

the semantics of conditionals. To anticipate, I will treat separately the types of case which *support* IB and the types of case which *challenge* IB. In this section I've presented data of the first type, and I will consider semantic explanations of this data in what follows. I'll take up data of the second type in Section 5.

More generally, there are no decisive examples or counter-examples to either the unified or variable analyses at the linguistic level; there is only data to be explained. Readers will note that apparent counter-examples to IB are also apparent counter-examples to the unified analysis, and support the variable analysis; meanwhile, as I'll argue below, the data that supports IB counts in favor of the unified analysis and against the strict analysis. For each theory, there is data which seems to support it and data which seems to challenge it. Any adjudication must holistically weigh the full suite of evidence. My final argument for the unified analysis is that it represents the best explanation of both the IB-supporting and IB-challenging data. By contrast, the variable analysis, gives a reasonable explanation of the IB-challenging data, but a poor explanation of the IB-supporting data. It is the supporting data which is my immediate focus in this section and the next.

3.2 The logic of the unified analysis

The semantic significance of IB is brought out by the logical consequences of conjoining IA and IB, when it is assumed that both are validities. In short, if both are valid, then the unified analysis must be true, and variable analysis false.

To begin, suppose that both IA and IB are logically valid:

$$(19) \text{ IA: } \neg\Diamond(\phi \wedge \neg\psi) \models \phi > \psi$$

$$(20) \text{ IB: } \phi > \psi \models \neg\Diamond(\phi \wedge \neg\psi)$$

From which it follows that the two formulae are equivalent:

$$(21) \quad \phi > \psi \Leftrightarrow \neg\Diamond(\phi \wedge \neg\psi)$$

But by a few short steps, $\neg\Diamond(\phi \wedge \neg\psi)$ is equivalent to $\Box(\phi \supset \psi)$:

$$\begin{aligned} & \neg\Diamond(\phi \wedge \neg\psi) \\ \Leftrightarrow & \neg\Diamond\neg(\neg\phi \vee \psi) && \text{[DeMorgan's Law]} \\ \Leftrightarrow & \neg\Diamond\neg(\phi \supset \psi) && \text{[Definition of } \supset \text{]} \\ \Leftrightarrow & \Box(\phi \supset \psi) && \text{[Definition of } \Box \text{]} \end{aligned}$$

It follows that:

$$(22) \quad \phi > \psi \Leftrightarrow \Box(\phi \supset \psi)$$

In other words, the counterfactual is equivalent to the strict conditional. And since the \Box in question is defined as the dual of \Diamond , it follows that this strict conditional is also coordinated with

the \diamond in the manner of the unified analysis.

The preceding argument shows that the unified analysis follows from the validity of IA and IB. Intuitively, the argument works because IA and IB each restricts the quantificational domain of the conditional in opposite directions. On one hand, IA implies that the range of worlds quantified by the conditional are limited to those quantified by the modal. On the other, IB holds that the worlds quantified by the modal are limited to those quantified by the conditional. Putting these together, it follows that in any context, the domains quantified by the conditional and modal are the same, just as the unified analysis requires.

The significance of this result is that any approach that makes sense of IA and IB by making them both semantic entailments will result in a theory which is truth-conditionally equivalent to the unified analysis. As a consequence, semantic explanations of both IA and IB are simply not available to the variabilist. Since the adoption of *Accessibility* is a simple and plausible way for the variabilist to validate IA, I'll assume it for the variable analysis it going forward. Instead, the variabilist must search for an account of IB which does not render it logically valid.

3.3 *Impossibility Below: semantic explanations*

As the preceding argument shows, the inferential force of IB is directly explained by the semantics of the unified analysis. In explicitly semantical terms: the premise of IB, $\phi > \psi$, is only true at a world i if every accessible ϕ -world is a ψ -world. Hence no accessible ϕ -world is a $\neg\psi$ -world. And so, by the unified semantics for modals, the conclusion, $\neg\diamond(\phi \wedge \neg\psi)$ must also be true at i . With little elaboration, this explains the full range of data, from felicitous inferences, to denials, infelicitous conjunctions, and interaction with might-conditionals.

In Section 3.1 I noted that some find that the data supporting IB, while empirically valid, is weaker than the data supporting IA. A plausible explanation of this asymmetry, from the perspective of the unified analysis, is that subjunctive modals have (at least) two standard or default readings. On one reading, the domain of the modal is the same as that as the counterfactual; this explains the judgments I reported above. On another reading, it naturally assumes a wider domain than that of the counterfactual; this would explain the instinct to adjudicate in the opposite direction. Consider the infelicitous cases like (10) and (11). The unified-domain reading predicts both to be contradictory, but the wider-domain reading predicts both to be consistent. My conjecture is that readers who experience the weak IB judgments are, perhaps unconsciously, divided between these two readings. The same facts would not weaken the IA-judgments, thus explaining the alleged asymmetry between the evidence for IA and for IB. Crucially, this ambiguity hypothesis is perfectly consistent with the unified analysis, since the theory requires only that *one* reading of the modal be unified with the conditional.

As for the variable analysis, IB is not validated on the variable semantics as I originally presented it. This is to be expected from the argument presented above. Semantically, the invalidity

results from the lack of coordination between the accessibility function k and selection function f .¹¹

At this point, a variety of strategies suggest themselves to aid the variable analysis. All pursue the natural idea that the domain of accessible worlds for modals and the selection function for conditionals should be put into some kind of correspondence. One idea is to require, in the model, that the accessible worlds be a subset of the set of closest worlds *simpliciter*. Another idea is to build the connection into the semantics of the modal itself. Following Kratzer (1981; 2012, pp. 39-41), one such proposal is that a modal is true just in case there is at least one prejacent-world among the closest worlds *simpliciter*. A related but distinct idea, due to Stalnaker and Thomason (1970), is that the modal is true just in case there is at least one closest prejacent-world.¹²

Despite interesting differences between these analyses, all have the same effect for the variable analysis. If *Accessibility* is imposed, as I have assumed, each either fails to validate IB, or imposes constraints on the closeness ordering which make the variable analysis truth-conditionally equivalent to the unified analysis. This is the inevitable consequence of the argument described in the last subsection: for if IA and IB are validated, the truth-conditions of the conditional are equivalent to those given by the unified analysis. Since any purely semantic explanation of IB will, by design, validate it, such semantic explanations of IB are foreclosed for the variable analysis. I'll consider alternative pragmatic and discourse-based explanations on behalf of the variable next.

4 Counter-explanations?

In this section I ask whether the variable analysis might explain the inferential force of IB with the help of pragmatic or discourse-based resources. I'll consider two apparently plausible accounts culled from recent work on conditionals in discourse, but argue that, in each case, the proposed explanation are not supported by the linguistic evidence. The prospects for a pragmatic account of IB, I conclude, are dim.

I wish to begin by recalling the most straightforward positive evidence for IB. These included felicitous inferences as well as infelicitous monologues. (There was other evidence, which I don't focus on here, including cases of denial, and interactions with might-conditionals.) I reproduce the inference and infelicity cases in compact form below.

¹¹It is not valid even if we impose the natural constraint expressed by *Accessibility*. For *Accessibility* requires that the ordered worlds simply be *among* the accessible worlds. It could still be true that the closest ϕ -worlds are all ψ -worlds (the premise of IB), even while *some* accessible ϕ -worlds are $\neg\psi$ -worlds (the negation of the conclusion).

¹²Let $\min(\geq_i)$ be the set of worlds which are closest to i relative to \geq_i . Then the first proposal, the constraint on modal accessibility, can be stated as $k(i) \subseteq \min(\geq_i)$. The Kratzerian semantics holds that $\diamond\phi$ is $\text{true}_{i,c}$ iff $\exists w : w \in \min(\geq_i) \cap \llbracket\phi\rrbracket$. By contrast, according to the Stalnaker-Thomason semantics, $\diamond\phi$ is $\text{true}_{i,c}$ iff $\exists w : w \in f(\phi, i)$. Both the Kratzerian semantics and the constraint on modal accessibility have the effect of making the variable semantics equivalent to the unified strict analysis. The Stalnaker-Thomason semantics fails to validate IB.

- (23) a. If Amira had played, she would have won.
 b. So she couldn't have played and lost.
- (24) a. If Amira had played, she would have won.
 b. * And she might have played and lost.
- (25) a. Amira might have played and lost.
 b. * And if she had played, she would have won.

How might the variablist explain this sort of data? Inspiration comes from recent work by Moss (2012), Starr (2014), Lewis (2016; 2017), and Stojnić (2017; 2016) on the role of conditionals in discourse, particularly Sobel and reverse-Sobel sequences. Adopting the broad strategy of these authors, one might explain the inferential case (23) by claiming that sentence (a) induces some kind of systematic shift of context (either linguistic or epistemic) so that (b) is either entailed or naturally inferred from the new context. The inferential force of IB is explained as pragmatic, discourse-bound, or epistemic, rather than logical. I'll consider two ways of implementing this idea, ultimately arguing that neither will work.

In defending a variable analysis treatment of reverse-Sobel sequences, Lewis (2016; 2017) proposes that asserting a conditional tends to raise the set of nearest antecedent-worlds to salience. As a consequence, context shifts in such a way that these antecedent-worlds are now included among the closest worlds for subsequent conditionals. Adapted to the present case, we may imagine that modals too have the force of raising their prejacent-worlds to salience, such that subsequent conditionals treat those as among the closest worlds. This could explain the judgments about (25): the modal in (a) raises *play* \wedge *lose*-worlds to salience; these are treated as closest for the conditional in (b), which must thereby be false, since some closest *play*-worlds are *lose*-worlds.

But the same strategy cannot be extended to the other examples. Focusing on (23) for illustration, suppose that the conditional in (a) has the effect of raising the closest *play*-worlds to salience, all of which will also be *win*-worlds. Then, in the spirit of Lewis' suggestion, we can assume these must be included in the modal domain of (b). But this is not enough to ensure that (b) is entailed by the new context: the presence of closest *play* \wedge *win*-worlds among the accessible worlds doesn't settle the question of whether there are accessible *play* \wedge *lose*-worlds. Thus the general strategy fails to capture the full range of data that support IB. The stumbling block for this account, and others like it, is that it predicts strong order effects. For it only works when the modal precedes the conditional. By contrast, as (24) and (25) demonstrate, the evidence for IB is order invariant. This is part of why it is natural to explain the full range of evidence in terms of semantic entailment, a relation which is insensitive to the order of discourse presentation.

A second strategy, which circumvents this concern, is inspired by Starr's (2014) treatment of reverse-Sobel sequences and Stojnić's (2016; 2017) analyses of discourses involving indicative

conditionals. These analyses focus on the role of MODAL SUBORDINATION, the phenomenon in which one conditional or modal functions to provide comment on a possibility introduced into the discourse by an antecedent modal expression (Roberts 1989). Thus, in (26) below, (a) introduces a possibility, that a wolf might walk in, and (b) is modally subordinated to (a) since it says *of that possibility*, that the wolf would eat you. It does not say, for example, that a wolf *will* eat you, come what may.

- (26) a. A wolf might walk in.
b. It would eat you.

For concreteness, I'll focus on adapting Stojnić's detailed proposal to the data at hand. Following Stone (1997; 1999), Stojnić holds that modals and conditionals are covertly anaphoric, their quantificational domain restricted by an anaphorically selected proposition (set of worlds), just as in (26b) above. Stojnić makes two additions to this picture: first, the idea that modals and conditionals also promote the possibilities they describe to contextual prominence; and second, that in the right discourse context, subsequent modals are anaphorically linked to these prominent possibilities by default. The result is that, in the right discourse context, modals and conditionals tend to restrict the quantificational domain of subsequent modal expressions. Discourse context, in turn, is understood in terms of the available DISCOURSE COHERENCE RELATIONS, posited by linguists to help explain the informational relations between clauses in discourse (see e.g. Hobbs 1985; Kehler 2002; Asher and Lascarides 2003). Stojnić hypothesizes that some relations, like *Elaboration*, exhibited in (26), impose modal subordination, while others, like *Contrast*, do not. (To be clear, in Stojnić's analysis, all modals are anaphoric; what I am calling "modal subordination" is, in Stojnić's terms, modal anaphora to the most prominent possibility in the immediately preceding discourse.)

If this strategy could be sustained here, it would explain several key pieces of data. Here's how the story might go: in (23) the conditional in (a) shifts to prominence the set of closest *play*-worlds which are also *win*-worlds. If (b) is modally subordinated to this set, then of necessity (b) will be true, since no world in the set is a *play* \wedge *lose*-world. A parallel explanation applies to (24). As for (25): the modal in (a) raises to prominence the set of accessible *play* \wedge *lose*-worlds. The domain of the selection function in (b) is then restricted to this set, with the result that the conditional is necessarily false.

While this account offers an in-principle explanation of the available data, additional evidence tells against it. The problem is that the proposed explanation goes through only if the discourses in question (quite different from those originally discussed by Stojnić) are such that they impose modal subordination. Stojnić focuses only on the relation of *Elaboration* as a trigger of modal subordination, but there are many discourse relations which do *not* involve modal subordination.

The chief obstacle for the modal subordination account arises for felicitous *inferences* like (23). Here a relation of inference or conclusion is expressed by the discourse particle *so*. Yet the use of *so* actually seems to block modal subordination in certain contexts.¹³ Evidence comes from the following discourse, where the (a) and (b) sentences are close semantic analogues of those from (23).¹⁴ (I change “played” to “played using a new bat” here so as to allow for the possibility of playing and loosing.)

- (27) a. Amira didn’t play in the match yesterday, but she could have.
 b. If she had played using a new bat, she would have won.
 c. * So she couldn’t have lost.

Suppose (b) were subordinated to (a). Then the *play bat* \wedge *win*-worlds introduced by (a) would constrain the modal in (b). Since these worlds include no *lose*-worlds, (b) should be an obvious entailment, and the use of *so* should be felicitous as a result. Yet it is not, suggesting that (27) is not an instance of modal subordination, and by extension, neither is (23).

The same lesson is illustrated with other combinations of modals and conditionals. In the following discourses, modal subordination would predict that the conclusion is contextually entailed by the assertion of the premise, hence felicitous. Yet the use of *so* is once again infelicitous.

- (28) a. Amira might have played and won.
 b. * So if she had played, she would have won.
 (29) a. If Amira had played, she might have won.
 b. * So if she had played, she would have won.

As for the infelicitous conjunctions (24) and (25), the modal subordination account again requires that these cases be governed by *Elaboration* or some other modally subordinating relation. But the evidence here is mixed at best. In a comprehensive discussion of the discourse particle *and*, Txurruka (2003) argues that one function of *and* is to signal the class of *coordinating* discourse relations, which tend to introduce a new discourse topic, rather than offer further information on an existing one. While there are a variety of coordinating relations, including relations of narration and causation, *Elaboration* is specifically excluded. Whether some other relation, compatible with modal subordination is at work in these infelicitous conjunctions is a separate question. But absent specific evidence to that effect, the proposed explanation is merely promissory.

¹³I say that *so* blocks modal subordination in *certain* contexts, but its overall behavior is complex. The following discourse is felicitous: “A wolf might walk in. It would be hungry. So it would eat you.” As it this: “If a wolf walked in, it would be hungry. So it would eat you.” But this variant is not: “A hungry wolf might walk in. So it would eat you.” It seems that the modally anaphoric behavior of *so* may be modulated both by discourse context and lexical encoding. In the text I focus on cases that are relevantly similar to the target examples.

¹⁴Thanks to Una Stojnić (p.c.) for this example.

In sum, the modal subordination strategy fails for the inferential data, and is inconclusive for the infelicity data. Meanwhile, the unified analysis provides a clear explanation of both with little elaboration.

At this point, I've considered two possible pragmatic explanations of the phenomena at hand, both based on flexible, contemporary models of modal discourse. But neither one successfully explains the available evidence. Of course there may be some further pragmatic account compatible with the variabilist theory that I have not considered. But I know of no such proposal in the current literature. At the very least the explanatory burden has been moved to the variabilist's court.

While there is no proof that all variabilist pragmatic approaches will fail, such accounts face serious in-principle challenges. The evidence for IB covers a heterogeneous set of linguistic data, including inferences, denials, and infelicities. A semantic account of IB, like the unified analysis, provides the same kind of logic-based explanation across discursive contexts. But it is the nature of pragmatic and discourse-based phenomena to act differently in different discursive contexts. As a result, pragmatic accounts have trouble evoking stable results in these varied settings. We saw this with Lewis-based account, which could only predict IB-like effects in one discourse order, but not the other. And the Stojnic-based account could only predict IB-like effects with some discourse-relations, but not others. Any pragmatic account must overcome this basic challenge.

And the variabilists explanatory hurdles are not over. In the next section, we'll encounter a range of additional discourse phenomena which the unified analysis smoothly accounts for, but the variable analysis does not. These must also be counted as part of the total evidential burden facing each theory

5 Counter-examples?

I turn finally to a family of examples which seem to challenge the inferential force of IB, thus presenting apparent counterexamples to the unified analysis itself. I'll argue that these examples not only fail as counterexamples, they turn out to be part of a larger body of data which ultimately supports the unified analysis.

The examples in question have the following form, where the words written in italics are pronounced with emphasis:

- (30) a. If Kiara had played, she would have won.
b. But of course, she *might* have played and lost.
- (31) a. Kiara might have played and lost.
b. But in fact, *if* she had played, she *would* have won.

In each case, the (b) sentence is clearly felicitous. And this, on its face, is a challenge for the unified analysis, because according to the unified analysis the sentence with the logical form of (a) should be inconsistent with a sentence with the form of (b). Meanwhile, according to the variable semantics, the propositions expressed by the (a) sentences and (b) sentences are perfectly compatible, so no infelicity is predicted.

What is notable about these and similar cases is the presence of three distinctive linguistic signals: (i) focal stress; (ii) discourse intensifiers; and (iii) discourse contrast markers. FOCAL STRESS is the phenomenon in which a speaker puts special emphasis on a particular term as means of modulating its meaning. DISCOURSE INTENSIFIERS (as I call them) are clausal prefixes such as *of course*, *really*, or *in fact*. DISCOURSE CONTRAST MARKERS are clausal prefixes such as *but*, *however*, and *by contrast*. (It may be that *of course* belongs to this latter class as well.) Collectively, I shall call these SHIFT MARKERS, for reasons detailed below.

I'll return to the interpretation of shift markers shortly. What is striking is that, when they are removed, and the discourse particle *and* (or *what's more*) is introduced in their stead, the sense of contradiction and infelicity vividly reemerges:

- (32) a. If Kiara had played, she would have won.
b. * And she might have played and lost.
- (33) a. Kiara might have played and lost.
b. * What's more, if she had played, she would have won.

I'll call the alleged counterexamples, (30) and (31), MARKED cases, and the revised conjunctive examples, (32) and (33), UNMARKED cases. On semantic grounds alone (together with the assumption that contradictions are infelicitous), the variable analysis can explain the felicity of the marked cases, but not the infelicity of the unmarked cases. The reverse is true for the unified analysis. So each party must offer something like a pragmatic or discourse-based explanation of the opposing case.¹⁵ But this leaves the variabilist in a doubly bad position. First, as I argued in the last section, there seems to be no readily available variabilist explanation of the unmarked cases. But even if there were, the proffered explanation (whatever it may be) would seem to be of the wrong kind. It is the marked cases which we would normally expect to involve a shift in context and the unmarked cases where context should remain inert. I turn now to outline a discourse-based explanation of the marked cases on behalf of the unified analysis.

In the dynamic semantics outlined by von Stechow (2001) and Gillies (2007), context includes a MODAL DOMAIN (or a function from worlds to domains), which is the quantificational domain of conditionals. By default, the modal domain remains fixed as discourse progresses, so that

¹⁵It is a remarkable fact that nearly the entire corpus of linguistic data surrounding the variable/strict debate (including, for example, Sobel and reverse-Sobel sequences) has failed to control for the presence of discourse markers or test for their effects. It would be instructive to revisit these debates with this dimension of variation in mind— but this lies outside of the domain of this essay.

each successive sentence in a discourse expresses its content relative to the same contextual parameter.¹⁶ Only specific linguistic signals can trigger an expansion or contraction of the modal domain. For von Fintel (2001) and Gillies (2007), the primary case is domain expansion triggered by the possibility presupposition of the antecedent of a conditional. I propose to adopt this general framework, and to add into it a role for modals.¹⁷ Modals quantify over the same modal domain as conditionals, so that what I have been calling the *set of accessible worlds* I now identify with the modal domain.

But I take a more flexible attitude to shifts of the modal domain than von Fintel and Gillies. While they recognize standardized domain shifts driven by fixed semantic mechanisms, I propose that the modal domain liable to shift easily in response to a range of linguistic and extralinguistic factors (Kratzer 1977). I take inspiration from Lewis (1979, ex. 6), who suggests that the domain of modals can be expanded by mechanisms of ACCOMMODATION. When a sentence $\diamond\phi$ is asserted, if there are no accessible ϕ -worlds, then, under suitable circumstances, the context shifts to accommodate the assertion: the set of accessible worlds expands, so that, relative to the new context, $\diamond\phi$ expresses a true proposition which is compatible with preceding discourse. While there may be limits on what can be accommodated, Lewis argues that accommodation-driven shifts in context arise felicitously in normal conversation.

To Lewis's account, I make three modifications. First, I hold that, at least in the kinds of cases at hand, accommodation does not occur automatically; it must be triggered by an appropriate linguistic signal. In particular, the use of the contrastive discourse marker, in conjunction with the other shift markers, is sufficient to trigger an accommodation-driven change to the modal domain. Meanwhile, the use of *and*, in the absence of shift markers, tends to signal that the modal domain stay fixed.¹⁸ Second, whereas (D.) Lewis (1979, pp. 354-5) (as well von Fintel (2001, §8) and Gillies (2007, §9)) thought that only domain *expansions* are easy to achieve discourse, I join (K.) Lewis (2017, pp. 7-9) in concluding that both expansions and contractions of the modal domain occur smoothly in normal conversation. Third, I assume that accommodation is guided by something like a set of nested spheres of worlds, ordered by closeness, just as it is in the dynamic systems of von Fintel (2001, pp. 7, 19-22) and Gillies (2007, pp. 334-337). The idea is that accommodation-driven expansions and contractions are *conservative*, not arbitrary. Context

¹⁶To be clear, when I say here and below that the *modal domain* persists through discourse, I mean that, *relative to a world of evaluation*, the modal domain persists. What actually is carried forward in discourse is an accessibility function. I trust this way of speaking will cause no confusion.

¹⁷Once again, it is immaterial to the present proposal whether the mechanics of discourse are understood as dynamic semantics, as in the systems of von Fintel (2001) or Gillies (2007), or as a set of discourse rules that run over and above the semantics, roughly in the style of Lewis (1979).

¹⁸A potential simplification of this analysis might hold that the presence of the *Contrast* discourse relation is sufficient to trigger modal accommodation. Yet it seems to me that infelicitous readings of the marked cases are (more) available when only the contrast marker is used (which is sufficient to incur *Contrast*), but not available when all three shift markers are employed. Thus the discourse relation of *Contrast* is necessary for triggering a domain shift, but that is not the whole story about the discourse mechanisms at play here.

tends to change as little as possible to accommodate the target utterance, in accordance with the background ordering on worlds.

Consider again the roster of cases I began with. In the unmarked cases, the absence of any discourse marker except *and* signals the default persistence of the modal domain. Thus, in each case, the two sentences are evaluated at the same modal domain. By the unified analysis, the propositions expressed are logically incompatible, hence the second assertion in each case is infelicitous.

By contrast, in the marked cases (30) and (31), the cluster of shift markers in the second sentence signal an adjustment to the modal domain, which thereby expands or contracts to accommodate the main clause. In (30), accommodation of the modal results in an *expansion* of the modal domain. The conditional in (a) asserts that all *play*-worlds in the domain are *win*-worlds; then (b) triggers an expansion of the modal domain to include some *play* \wedge *lose*-worlds, so that the modal is true. In (31), accommodation of the conditional results in *contraction* of the modal domain. The modal in (a) asserts that there are some *play* \wedge *lose*-worlds in the modal domain; (b) triggers a contraction of the domain, restricting to *play*-worlds which are all *win*-worlds, thereby making the conditional true.

I maintain that this account of the linguistic evidence data is not at all *ad hoc*, but well motivated by a range of parallel phenomena. Here it is helpful to focus on the role of the contrastive discourse marker *but*, since it seems to be the one shift marker which is essential to trigger a shift. (Nevertheless, felicitous readings of the marked cases are clearer with the addition of intensifiers and focal stress.)

In the study of discourse structure, contrast markers are thought to signal a contrast, or violated expectation, between a preceding segment of discourse and the clause to follow. (Hobbs 1985, p. 32; Knott 1996) Yet the notion of contrast at work here is quite open-ended; the relevant contrast may be an inconsistency or difference in embedded clauses, predicates, presuppositions, implicatures, or merely associated expectations. (Hobbs 1985, pp. 21-23; Asher and Lascarides 2003, §1.1, §4.8.4) Plausibly, a contrastive particle could be used to signal a shift to a new context which differs from the previous context in a manner relevant for the discourse. In the alleged counterexamples, I contend, *but* is used to signal a shift from a wide modal domain that satisfies the possibility modal to a narrow modal domain that satisfies the conditional, or *visa versa*. On the other hand, *and* is thought to be a marker of coordination; it signals that the adjoined segments of discourse are in some sense parallel or equally situated in the information structure of the discourse (Hobbs 1985; Txurruka 2003). And it is further plausible that *and* has the opposite effect of *but*, actually requiring that the contextual domain to remain in its default fixed state.

As corroboration of these ideas, we find that the very same shift markers can be used to trigger a contextual shift in the domain of quantification for nominal quantifiers, where the semantics of the relevant expressions are relatively well established. For example, the standard analysis of

English quantifiers predicts that “there is (only) an α in β ” and “there is nothing in β ” are, in the same context, logically incompatible, hence infelicitous to assert in conjunction. Yet the combined use of contrast markers, intensifiers, and focal stress render these discourses acceptable:

- (34) a. My friend Thomas never goes shopping.
b. There is only a ketchup bottle and a can of beer in the fridge.
c. But really, there is *nothing* in the fridge.
- (35) a. My friend Thomas never goes shopping.
b. There is nothing in the fridge.
c. Though of course, there *is* a ketchup bottle and a can of beer in the fridge.

The context shifting mechanisms at work here appear to be precisely the nominal counterpart of that proposed for the modal case. Thus the use of shift markers in (34) appears to have the effect of admitting a narrowing of the domain of quantification, while those in (35) seem to have the effect of admitting a widening of the domain. Notably, the same pattern of infelicity that characterized the modal cases emerges when the shift markers are removed:

- (36) a. There is only a ketchup bottle and a can of beer in the fridge.
b. * And there is nothing in the fridge.
- (37) a. There is nothing in the fridge.
b. * And there is a ketchup bottle and a can of beer in the fridge.

Such observations lend credence to the account of signal-driven accommodation I’ve sketched on behalf of the unified analysis. They suggest that shift markers are a general way to signal accommodation of changes to the relevant quantificational domain, and discourse conjunctions a general way of blocking such domain change.

Still, the account of modal accommodation outlined here isn’t meant to be comprehensive. The use of shift markers to signal accommodation appears to be a partially pragmatic and defeasible process, potentially superseded by other linguistically encoded cues of accommodation. In particular, it is widely thought that conditionals presuppose the possibility of their antecedents as a matter of semantic encoding. And in the systems of von Stechow (2001, pp. 15-20) and Gillies (2007, pp. 333-4), this presupposition is automatically accommodated by dynamic semantic mechanisms. (Note that in the cases of accommodation I’ve considered so far, any antecedent possibility is already granted, and what is accommodated is the truth of the conditional as a whole.)

Such semantically-driven accommodation of the antecedent possibility appears to be an especially explicit and powerful way to push a possibility into the modal domain. Indeed, it may

force an expansion of the modal domain even in the face of contrary discourse signals. This can be seen in the relative felicity of Sobel sequences conjoined by *and*:

- (38) a. Mary took the bus to work and arrived late.
b. If Mary had driven to work she would have arrived on time.
c. And if she had driven to work and got into an accident, she would have arrived late.

Here the modal domain in (c) is expanded from that of (b), thereby making both (b) and (c) true and felicitous in their respective contexts. It is allowed to shift despite the presence of *and*, which normally signals the fixity of modal context, because the driver of accommodation is the conditional antecedent in (c). (Still, it is notable that a contrastive version with *but* is preferable.)

Furthermore, once the antecedent of a conditional is used to embed a possibility into the modal domain, it is difficult to eliminate it merely via signal-driven accommodation. This is illustrated by the case of reverse-Sobel sequences that employ shift markers.

- (39) a. Mary took the bus to work and arrived late.
b. If she had driven to work and got into an accident, she would have arrived late as well.
c. * But in fact, *if* Mary had driven to work, she *would* have arrived on time.

In this case, (c) is infelicitous, suggesting a failure of accommodation, despite the presence of shift-markers. It seems that the presuppositions of conditional-antecedents make a kind of indelible mark on the discourse record that can only be overcome through equally explicit means. This dominance of semantic presupposition over signal-driven accommodation is to be expected. It's one thing to assume an open ended change in context when the speaker invites you to do so. It's another to assume a change in context in the face of an explicit commitment on behalf of the speaker to the contrary. In general, explicit widenings of the domain cannot be implicitly clawed back through the use of shift markers.

I began this section with examples which appeared, at first glance, to be direct counterexamples to the unified analysis. On closer inspection, I've argued, just the opposite seems to be true: the unified analysis has at its disposal a simple, independently motivated account of the full range of relevant data.

6 Semantics and Modality

This essay has concerned itself with the close interactions between different linguistic expressions of modality— counterfactual conditionals and counterfactual possibility modals. I've presented evidence for two inference patterns, IA and IB, that establish clear links between modals

and conditionals. And I've shown that the unified analysis directly explains both, while the variable analysis falters on IB. I've argued that the variable analysis cannot explain IB as an *entailment* (on threat of collapse into the unified analysis), but nor can it explain it as a pragmatic or discourse-based inference. And I've argued that alleged counterexamples to the unified analysis only bolster its evidential position as compared to the variable analysis. Together, these considerations constitute my case for the unified analysis of both modals and conditionals, and by extension, for the strict analysis of conditionals, and against the variable analysis.

These conclusions offer a new perspective on the enduring debate between strict and variably strict theories of the conditional. The variable analysis was originally motivated by Stalnaker (1968) and Lewis (1973) with a range of non-monotonic data— like cases of antecedent strengthening and failures of transitivity— which, it was observed, could not be accounted for by the strict analysis on truth-conditional grounds alone. More recently, discourse-based treatments of strict conditionals by von Fintel (2001) and Gillies (2007) have offered principled and plausible answers to these challenges, and replied with their own set of data— reverse-Sobel sequences and their variants— which seem to pose a challenge to the variable analysis. And this data has, in turn, been met with counter-explanations on behalf of the variable analysis (Moss 2012; Starr 2014; Lewis 2017), and new challenges to the strict analysis (Lewis 2017; Nichols 2017). With plausible accounts of the data on both sides, as well as unanswered questions, it can look like the debate has reached an impasse. In this essay, I hoped to have tipped the balance in favor of the strict analysis by drawing attention to a new source of linguistic evidence, and a new set of arguments.

Beyond the particular debate at hand, the deeper lesson of this discussion is that modals and conditionals are not independent semantic devices, expressing merely loosely related modal facts. Instead, they are part of a unified suite of quantificational tools whose specialized task is the navigation of a common modal space. Mapping the geography of this space is the task for future elaborations of the unified approach to modality.

References

- Asher, N. and A. Lascarides (2003). *Logics of Conversation*. Cambridge University Press.
- Bennett, J.F. (2003). *A Philosophical Guide to Conditionals*. Oxford University Press, USA.
- Creary, L.G. and C.S. Hill (1975). "Review of David Lewis' Counterfactuals". In: *Philosophy of Science* 42.3, p. 341.
- Ellis, B., F. Jackson, and R. Pargetter (1977). "An objection to possible-world semantics for counterfactual logics". In: *Journal of Philosophical Logic* 6.1, pp. 355–357.
- Fine, K. (1975). "Critical notice: Counterfactuals". In: *Mind* 84, pp. 451–458.
- Gillies, A.S. (2007). "Counterfactual scorekeeping". In: *Linguistics and Philosophy* 30, pp. 329–360.
- (2009). "On truth-conditions for if (but not quite only if)". In: *The Philosophical Review* 118.3, p. 325. ISSN: 0031-8108.

- Hobbs, J. R. (1985). *On the Coherence and Structure of Discourse*. Center for the Study of Language and Information.
- Kehler, A. (2002). *Coherence, Reference, and the Theory of Grammar*. Center for the Study of Language and Information.
- Knott, Alistair (1996). "A Data-Driven Methodology for Motivating a Set of Coherence Relations". PhD thesis. University of Edinburgh.
- Kratzer, A. (1977). "What 'Must' and 'Can' Must and Can Mean". In: *Linguistics and Philosophy* 1, pp. 337–55.
- (1981). "The notional category of modality". In: *Words, worlds, and contexts*, pp. 38–74.
- (1986). "Conditionals". In: *Chicago Linguistics Society* 22.2, pp. 1–15.
- (2012). *Modals and conditionals: New and revised perspectives*. Vol. 36. Oxford University Press.
- Kripke, S. (1963). "Semantical considerations on modal logic". In: *Acta philosophica fennica* 16.1963, pp. 83–94.
- Lewis, C.I. (1914). "The Calculus of Strict Implication". In: *Mind* 23.1, p. 240.
- Lewis, D. (1973). *Counterfactuals*. Blackwell Publishers.
- (1979). "Scorekeeping in a Language Game". In: *Journal of Philosophical Logic* 8, pp. 339–59.
- Lewis, K. S. (2016). "Elusive counterfactuals". In: *Noûs* 50.2, pp. 286–313.
- (2017). "Counterfactual Discourse in Context". In: *Noûs*.
- Loewer, B. (1976). "Counterfactuals with Disjunctive Antecedents". In: *Journal of Philosophy* 73.16, pp. 531–537.
- Lowe, E.J. (1983). "A Simplification of the Logic of Conditionals". In: *Notre Dame Journal of Formal Logic* 24.3.
- (1995). "The truth about counterfactuals". In: *The Philosophical Quarterly*, pp. 41–59.
- Lycan, W.G. (1984). "A syntactically motivated theory of conditionals". In: *Midwest Studies in Philosophy* 9.1, pp. 437–455.
- (2001). *Real Conditionals*. Oxford University Press, USA.
- Mckay, T. and P. Van Inwagen (1977). "Counterfactuals with disjunctive antecedents". In: *Philosophical Studies* 31.5, pp. 353–356.
- Moss, Sarah (2012). "On the pragmatics of counterfactuals". In: *Noûs* 46.3, pp. 561–586.
- Nichols, Cory (2017). "Strict conditional accounts of counterfactuals". In: *Linguistics and Philosophy*, pp. 1–25.
- Nute, D. (1980). "Conversational scorekeeping and conditionals". In: *Journal of Philosophical Logic* 9.2, pp. 153–166.
- Peirce, CS (1933). *Collected Papers of Charles Sanders Peirce, Vol. III: Exact Logic*. Ed. by C. Hartshorne and P. Weiss. Cambridge, MA: Harvard University Press.
- Roberts, Craige (1989). "Modal subordination and pronominal anaphora in discourse". In: *Linguistics and philosophy* 12.6, pp. 683–721.
- Sanford, D.H. (2003). *If P, then Q: Conditionals and the Foundations of Reasoning*. Routledge.
- Stalnaker, R. (1968). "A theory of conditionals". In: *Ifs. Conditionals, Belief, Decision, Chance, and Time*. Dordrecht, D. Reidel Publishing Company, pp. 41–55.
- (1978). "Assertion". In: *Syntax and Semantics, Volume 9: Pragmatics*. Ed. by Peter Cole and Jerry Morgan. Academic Press, pp. 315–332.

- Stalnaker, R. and R.H. Thomason (1970). "A semantic analysis of conditional logic". In: *Theoria* 36.1, pp. 23–42.
- Starr, William B (2014). "A uniform theory of conditionals". In: *Journal of Philosophical Logic* 43.6, pp. 1019–1064.
- Stojnić, Una (2016). "Content in a Dynamic Context". In: *Notûs*.
- (2017). "One's Modus Ponens: Modality, Coherence and Logic". In: *Philosophy and Phenomenological Research* 95.1, pp. 167–214.
- Stone, M. (1997). "The Anaphoric Parallel between Modality and Tense". In: *IRCS Report* 97.
- (1999). "Reference to Possible Worlds". In: *RuCCS Report* 49.
- Txurruka, Isabel Gómez (2003). "The natural language conjunction and". In: *Linguistics and Philosophy* 26.3, pp. 255–285.
- von Fintel, K. (1998). "The Presupposition of Subjunctive Conditionals". In: *The Interpretive Tract, MITWPL* 25, pp. 29–44.
- (2001). "Counterfactuals in a Dynamic Context". In: *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press, pp. 123–52.
- Warmbröd, K. (1981a). "An Indexical Theory of Conditionals". In: *Dialogue: Canadian Philosophical Review* 20.4, pp. 644–664.
- (1981b). "Counterfactuals and substitution of equivalent antecedents". In: *Journal of Philosophical Logic* 10.2, pp. 267–289.
- Williamson, Timothy (2008). *The philosophy of philosophy*. John Wiley & Sons.