Acceptance, Truth, and Probability: The Case of Conditionals[1]

Ginger Schultheis
January 2023

1 Introduction

In her classic paper *On Conditionals*, Dorothy Edgington reminds us of the importance of testing the predictions that a semantic theory of conditionals makes about what she calls our *less-than-certain judgments*—that is, our non-extreme credences—in conditionals. If we find that our credences in conditionals are radically different from those we would expect to have if we were forming credences in accordance with a certain theory, then that is evidence against the theory. If, moreover, there is an alternative theory that makes better predictions about our less-than-certain judgments, then that is reason to prefer the alternative.

This is, more or less, the state of play as I see it in current theorizing about conditionals. There are two leading theories—the *strict theory* and the *variably strict theory*. Our ordinary credences in conditionals are far out of line with what the strict theory recommends.[2] And there's a variably strict theory that does better: Stalnaker's variably strict theory.[3]

In light of this, I recommend rejecting the strict theory in favor of a Stalnakerian variably strict theory. Nevertheless, certain inference patterns that are valid on the strict theory and invalid on the variably strict theory—such as Transitivity, Contraposition, and Antecedent Strengthening—do share certain good-making features with classically valid arguments. Transitivity and Contraposition certainly seem like excellent principles. Although Antecedent Strengthening may seem less obvious at first, strict theorists have convincingly argued that the full range of data surrounding this principle is well explained by contextualist or dynamic strict theories. In particular, they observe that there is a striking asymmetry between the acceptability of so-called *forward Sobel sequences* and *reverse Sobel sequences*.

My task in this paper will be to develop a Stalnakerian variably strict theory on which the characteristically strict inference patterns are *pseudo-valid*:

[1]Thanks to Andrew Bacon, Cian Dorr, Melissa Fusco, Arc Kocurek, Justin Khoo, Paolo Santorio, and Snow Zhang for helpful conversations about this material. Special thanks to David Boylan and Matthew Mandelkern..

[2]Most of my observations about the strict theory's predictions about our credences in conditionals are not new. See Adams (1965), van Fraassen (1976), and Edgington (1995).

[3]See Stalnaker (1968) and Stalnaker (1975).

roughly, if the premises can be felicitously asserted in a given context, then if the premises are known in that context, the conclusion is also true in the context. In the first half of the paper, I will show that the pseudo-validity of these inference patterns for *indicative conditionals* falls out of Stalnaker's theory. This means that proponents of Stalnaker's theory have at their disposal a natural explanation of the apparent validity of Transitivity and Contraposition for indicatives, as well as a compelling account of the asymmetry between forward and reverse indicative Sobel sequences. Showing that the characteristically strict inference patterns are pseudo-valid for *counterfactuals* is a much more formidable task. Nevertheless, I will show that it can be done. In the second half of the paper, I develop a new Stalnakerian theory of counterfactuals—building on the work of Schulz (2014) and van Fraassen (1976)—and I show that the theory predicts that our three inference patterns are pseudo-valid for counterfactuals.

Here's the plan for the paper. I begin in §2 with a presentation of the strict theory and the variably strict theory. §3 presents the credence-theoretic argument against the strict theory. §4 explains–in an informal way—why Stalnaker's theory does better. §5 motivates a version of Stalnaker's theory for indicatives and then shows that the theory predicts that Transitivity, Contraposition, and Antecedent Strengthening are pseudo-valid. §6-7 presents and motivates a new theory of counterfactuals and then shows that the theory predicts that our three inference patterns are pseudo-valid for counterfactuals.

2 Strict and Variably Strict Conditionals

According to the strict theory, natural language conditionals have the truth conditions of necessitated material conditionals. The flavor of necessity depends on the kind of conditional. For indicatives, the necessity is epistemic necessity. An indicative with antecedent A and consequent B is true if and only if the material conditional ⌜not A or B⌝ is true throughout the closest epistemically possible worlds. For counterfactuals, the necessity is usually taken to be some form of historical or metaphysical necessity. A counterfactual with antecedent A and consequent B is true just in case the material conditional ⌜not A or B⌝ is true throughout the closest historically or metaphysically possible worlds.

To make this more precise, let W be a non-empty set of worlds. Let R_c be a contextually-supplied accessibility relation over W. We assume that R_c is reflexive. The strict theory runs as follows.[4]

[4]See Warmbrod (1981), Veltman (1985), von Fintel (2001), and Gillies (2007) for defenses of strict theories.

Strict Theory
$[\![\text{If A, then B}]\!]^{c,w} = 1$ if and only $\mathsf{R}_c(w) \cap [\![\text{A}]\!]^c \subseteq [\![\text{B}]\!]^c$

The strict theory retains many of the classical properties of the material conditional. Consider:

Transitivity
If A, then B, and if B, then C $\vDash$ If A, then C

Contraposition
If A, then B $\vDash$ If not B, then not A

Antecedent Strengthening
If A, then B $\vDash$ If A and C, then B

It is easy to check that all three inferences are validated by the strict theory.

Before moving on, a qualification. Contemporary strict theorists say that conditionals carry a *compatibility presupposition*. A conditional with antecedent A presupposes that there are accessible A-worlds. Often this presupposition is formalized in a trivalent framework; sentences whose presuppositions are not satisfied are said to be neither true nor false. On these theories, our three inferences are not classically valid; they are Strawson-valid. (Whenever the premises are true and the conclusion is either true or false, the conclusion is true.[5]) Other theorists prefer a *multidimensional* treatment of presupposition according to which sentences are always true or false, and presupposition is an independent dimension of meaning.[6] (We thus have a four-fold classification: true and presupposition satisfied, true and presupposition not satisfied, false and presupposition satisfied, and false and presupposition not satisfied.) On this theory, our three inferences are indeed classically valid. In what follows I'll assume a multidimensional treatment of presupposition. This assumption is made purely for ease of exposition.

To state the variably strict theory, we begin with a contextually-supplied *selection function* f_c that takes a world and an antecedent proposition to a set of worlds such that:

$$f_c(w, [\![\text{A}]\!]^c) \subseteq \mathsf{R}_c(w) \cap [\![\text{A}]\!]^c \tag{1}$$

Then the variably strict theory runs as follows.

[5] See von Fintel (2001).
[6] Herzberger (1973) and Kartunnen and Peters (1979).

Variably Strict Theory
$[\![\text{If A, then B}]\!]^{c,w} = 1$ if and only if $f_c([\![\text{A}]\!]^c, w) \subseteq [\![\text{B}]\!]^c$

The variably strict theory invalidates many classical inference patterns, including Transitivity, Contraposition, and Antecedent Strengthening.[7]

3 The Strict Theory and Less-than-Certain Judgments

Begin by observing that we often find ourselves uncertain of a conditional even though we know that the corresponding necessitated material conditional is false.[8] Consider a few examples.

There are twenty bikes at my gym. Almost always, if I get to the gym by three, at least one is free. I say:

(1) I'm confident that if I get to the gym by three, one of the bikes will be free.

Here I express high confidence in the proposition expressed by the indicative conditional:

(2) If I get to the gym by three, one of the bikes will be free.

This attitude is puzzling from the perspective of the strict theory. On the most natural way of filling out the example, I know that the material conditional

(3) Either I do not make it to the gym by three or one of the bikes will be free.

is not epistemically necessary—I know it's *possible* that I'll make it to the gym by three but find no bikes available.

Another example. My friend John and I are talking about the 2022 NBA Finals. Last year the Warriors faced the Celtics. The Warriors won, and it wasn't close. John asks:

(4) What's the chance that the Warriors would have won if they had faced the Bucks?

I reply:

(5) About 20-30%, I'd say.

Here I express low, but not zero, confidence in the counterfactual:

(6) If the Warriors had faced the Bucks in the 2022 NBA Finals, they would have won.

[7] For variably strict theories, see Stalnaker (1968) and Lewis (1973).
[8] To my knowledge this objection is due to Edgington (1995).

Once again this attitude is puzzling from the perspective of the strict theory. On the most natural way of filling out the example, I know that the material conditional

(7) Either the Warriors did not face the Bucks in the 2022 NBA Finals or they won.

is not historically or metaphysically necessary—I know that the Warriors *could have* lost to the Bucks.

One final example. My partner and I are deciding whether to enter a puzzle contest. If we enter, one of us will try to complete a difficult puzzle in five hours. I tell her that I don't care whether we enter, but that if we do, I should be the one to play. She asks why, and I reply:

(8) Because I'm more confident that I would finish in five hours if I tried than I am that you would finish in five hours if you tried.

This is a natural explanation of why I think I should be the one to play. But once again the psychological state that I'm reporting is mysterious from the perspective of the strict theory. I know that my partner could try to finish in five hours and fail; I know that I could try to finish in five hours and fail.

It is easy to generate more examples like this. We often rely on facts about how the probabilities of certain counterfactuals compare to the probabilities of others in our deliberation about what to do, and in our explanations of why we act the way we do. Since we usually know that the corresponding strict conditionals are false, strict theorists will have difficulty explaining these facts.

We turn now to a second feature of our less-than-certain judgments that the strict theory has a hard time explaining. Our degrees of confidence in conditionals tend to conform to the principle of *Conditional Excluded Middle*.

Conditional Excluded Middle
$\vDash$ If A, then B, or if A, then not B

To see this, return to our first example. Observe that insofar as I have high confidence in the proposition expressed by

(2) If I get to the gym by three, one of the bikes will be free.

I should have low confidence in the proposition expressed by

(9) If I get to the gym by three, none of the bikes will be free.

Suppose, for example, that I'm around 70% confident in (2). Then it seems I should be about 30% confident in (9). Furthermore, I assign zero credence to the conjunction:

(10) If I get to the gym by three, one of the bikes will be free, and if I get to the gym by three, none of the bikes will be free.

This means that I'm certain of the disjunction:

(11) Either, if I get to the gym by three, one of the bikes will be free, or if I get to the gym by three, none of the bikes will be free.

Moreover, if I learn something that decreases my confidence in (2)—say, that the cycling team is at the gym today—then I will presumably increase my confidence in (9) so that I remain certain of the disjunction.

Similar things can be said about our second case. Insofar as I have low confidence in the proposition expressed by

(6) If the Warriors had faced the Bucks in 2022 NBA Finals, they would have won.

I should have high confidence in the proposition expressed by

(12) If the Warriors had faced the Bucks in the 2022 NBA Finals, they would have lost.

Since I'm around 20-30% confident in (6), I should be around 70-80% confident in (12). Once again I assign no credence to the conjunction:

(13) If the Warriors had faced the Bucks in the Finals, they would have won, and if they had faced the Bucks in the Finals, they would have lost.

This means that I am certain of the disjunction:

(14) Either, if the Warriors had faced the Bucks in the Finals, they would have won, or if they had faced the Bucks in the Finals, they would have lost.

This pattern is pervasive. Our credences in the conditionals ⌜If A, then B⌝ and ⌜If A, then not B⌝, whether indicative or counterfactual, tend to sum to one. Since we ordinarily assign no credence to their conjunction, it follows that we are conforming to the principle of Conditional Excluded Middle.[9]

The problem is that strict theorists must reject Conditional Excluded Middle.[10] To see why, notice that Conditional Excluded Middle corresponds to the following constraint on the set of accessible worlds.

At Most One
$\mathsf{R}_c(w)$ contains at most one world.

[9] To my knowledge, this observation is originally due to Adams (1965). See also Bacon (2015), Goodman (2015), and Mandelkern (2018).

[10] This argument is from Cariani and Goldstein (2018).

If we want to secure Modus Ponens, we also need:

Reflexivity
$\mathsf{R}_c(w)$ contains w.

Putting At Most One and Reflexivity together, it follows that for any world w, $\mathsf{R}_c(w) = \{w\}$, and so the conditional ⌜if A, then B⌝ is true at w just in case the material conditional ⌜not A or B⌝ is true at w. This consequence is unacceptable; the overwhelming consensus of contemporary philosophers and linguists is that the material conditional analysis of natural language conditionals is inadequate.[11] Strict theorists must therefore reject Conditional Excluded Middle, and so they will have a hard time explaining why our credences tend to conform to the principle.

Finally, our credences do not seem to conform to the characteristically strict inference patterns of Antecedent Strengthening, Transitivity, or Contraposition. Consider Antecedent Strengthening. There are ten balls in an urn. They come in two colors (white and black) and two patterns (dots and stripes). Five balls in total are white, and five are striped. Among the striped balls, one is white. You are about to draw a ball. Consider the conditional:

(15) If I draw a ball, it will be white.

Since there are ten balls and five are white, you are 50% confident that (15) is true. Now consider:

(16) If I draw a ball and the ball has stripes, it will be white.

Since there are five striped balls and one of them is white, you are 20% confident that (16) is true. This combination of attitudes seems entirely reasonable. But if Antecedent Strengthening is valid, then the proposition expressed by (15) entails the proposition expressed by (16). And so if you were forming credences in accordance with the strict theory, you would have to be *more* confident in (16) than you are in (15).

Similar things can be said about Contraposition and Transitivity, but for reasons of space, I won't go through detailed examples.[12] (Note that given minimal background assumptions both Contraposition and Transitivity entail Antecedent Strengthening, so if we reject Antecedent Strengthening, we must reject these principles, too.)[13]

[11] See Jackson (1992) and Williamson (2020) for two exceptions.

[12] See Adams (1965) for more examples.

[13] Before moving on, it's worth taking a moment to discuss a reply to the credence-theoretic argument on the strict theorist's behalf. Some theorists, such as Bennett (2003), say that a sen-

4 Stalnaker's Theory and Less-than-Certain Judgments

This credence-theoretic objection to the strict theory would have little force if it turned out that *no* theory was capable of making sense of our credences in conditionals. But over the last several decades, a number of authors have shown that there is a variably strict theory that can make reasonable predictions: Stalnaker's variably strict theory.

For Stalnaker, a conditional with antecedent A and consequent B is true if and only either there are no accessible A-worlds, or else the closest accessible A-world is a B-world. (Formally: we state Stalnaker's theory using the variably strict semantics introduced in §2 and we require $f_c(w, [\![A]\!]^c)$ to contain at most one world. I will often refer to this world as the *selected A-world*.) Stalnaker's theory validates Conditional Excluded Middle.[14] When it comes to our less-than-

---

tence like (17)

(17) It's likely that if it rains, the picnic will be cancelled.

is conflated with

(18) If it rains, it's likely that the picnic will be cancelled.

If we apply this theory to conditionals under 'confident that' operators, then we should say that a sentence like (1)

(1) I am confident that if I get to the gym before three, one of the bikes will be free.

is conflated with (19):

(19) If I get to the gym before three, I am confident one of the bikes will be free.

There are several problems with this approach. For one, it is not at all obvious how the strategy can be extended to account for sentences like:

(20) Yesterday I was confident that if I went to the parade tomorrow, I would have have fun.

(Suppose today is Monday and the parade is on Tuesday.) In particular, (20) does not seem equivalent to:

(21) ? If I went to the parade tomorrow, yesterday I was confident that I would have fun.

Second, it is not obvious how to extend the strategy to counterfactuals. Consider (22):

(22) I'm confident that if I had died in the crash, you would have died, too.

(22) is obviously *not* equivalent to:

(23) ? If I had died in the crash, I would have been confident that you would have died, too.

Instead, I take it that the proponent of the conflation strategy wants to say that (22) is conflated with:

(24) If I had died in the crash, I am confident that you would have died, too.

But if that's right we need an explanation of why the attitude is not marked with past tense morphology.

[14] If there are accessible A-worlds, the selected A-world is either a B-world or a not-B-world. In the first case, ⌜If A, then B⌝ is true; in the second, ⌜If A, then not B⌝ is true. If there are no accessible A-worlds, both conditionals are true.

certain judgments in conditionals, this is already significant progress. If Conditional Excluded Middle is valid, then we should expect competent speakers to form credences in accordance with this principle, and as we saw in the previous section, that is exactly what we find.

But we can go further. We can give a much more general credence-theoretic argument in favor of Stalnaker's theory. To see this, consider:

> **Stalnaker's Thesis**
> The probability of a conditional is equal to the probability of its consequent given its antecedent.

It is widely accepted that our credences in indicative conditionals tend to conform to Stalnaker's Thesis. Consider some examples. My credence that one of the stationary bikes will be free, on the supposition that I get to the gym by three, is high. Correspondingly, I am confident in the conditional:

(2) If I get to the gym by three, one of the bikes will be free.

Suppose I am holding a fair, six-sided die in my hand. Consider:

(25) If I roll the die, it will land on one.

Intuitively, the probability of (25) is 1/6. That is also the conditional probability that the die lands on one, given that I roll the die.[15]

Building on the work of van Fraassen (1976), various authors, such as McGee (1989), Kaufman (2009), and Bacon (2015), have proven *tenability results* showing that certain versions of Stalnaker's Thesis are tenable on certain versions of Stalnaker's theory of conditionals.[16] Though the details of these results are complex, we can give an informal explanation of how they work. For Stalnaker, there are four constraints on the selection function. First: if there are epistemically accessible A-worlds, then the selected A-world is always an A-world. Second: if w is an A-world, then the selected A-world, at w, is w. Third: if there are epistemically accessible A-worlds, then the selected A-world is epistemically accessible.

[15] There are alleged counterexamples to Stalnaker's Thesis—see McGee (2000) and Kaufman (2004)—which I do not have the space to discuss. I don't think these examples undermine my central point: that Stalnaker's Thesis seems to hold in a great many cases, and that it is a significant advantage of Stalnaker's theory that it predicts this fact.

[16] I say 'certain versions of' Stalnaker's Thesis because not all of these authors defend Stalnaker's Thesis as I have stated it. McGee defends a version that applies only to conditionals whose antecedents do not contain conditionals; Bacon defends a context-sensitive version. It is also worth noting that the credence-theoretic argument in favor of Stalnaker's theory does not depend on the assumption that even a limited version of Stalnaker's Thesis holds *without exception*.

Fourth: If there are no epistemically accessible worlds, then there is no selected A-world: $f_c(w, [\![A]\!]^c) = \emptyset$. These are the only constraints. So if we're at an A-world, the selected A-world is the actual world. Otherwise, we can think of the selection function as randomly selecting a world from the epistemically possible A-worlds. With this metaphor in mind, we can begin to see how Stalnaker's semantics can capture our less-than-certain judgments. For one, we can be confident that a randomly selected A-world is a B-world even if we know that not *all* epistemically possible A-worlds worlds are B-worlds. Moreover, our credences in conditionals can be sensitive to the distribution of consequent-worlds among the epistemically possible antecedent-worlds. If few epistemically possible A-worlds are B-worlds—the probability of B given A is low—then the randomly selected A-world will probably not be a B-world, and so the probability of the conditional will be low. If most epistemically possible A-worlds are B-worlds—the probability of B given A is high—then the selected A-world will probably be a B-world, and so the probability of the conditional will be high.[17]

The fact that Stalnaker's Thesis is tenable on Stalnaker's theory of conditionals is a significant point in favor of that theory. As we have seen, our less-than-certain judgments tend to conform to Stalnaker's Thesis. A theory that is capable of vindicating Stalnaker's Thesis is therefore capable of vindicating our less-than-certain judgments in indicatives.

There is also reason to be optimistic that a Stalnakerian theory of counterfactuals is capable of making reasonable predictions about our credences in counterfactuals. As many authors have observed, our credences in counterfactuals tend to conform to *Skyrms' Thesis*, which says, roughly, that one's credence in a counterfactual is equal to one's expectation of the chance, at a relevant past time, of the consequent given the antecedent. Note that our credences in counterfactuals do not *always* conform to Skyrms' Thesis—counterlegals are counterexamples. Still, Skyrms' Thesis seems right in a wide range of cases, and recent work on the probabilities of counterfactuals suggests that a Stalnakerian theory of counterfactuals can account for this fact. (For recent work on this topic, see Khoo (2022), Schulz (2017), and Schultheis (forthcoming).)

Let's take stock. Stalnaker's variably strict theory has a clear advantage over the strict theory when it comes to predicting and explaining facts about our less-than-certain judgments in conditionals. I therefore recommend rejecting the strict theory, and Transitivity, Contraposition, and Antecedent Strengthening along with it. Nevertheless, as I said at the start of the paper, these three inference forms do share certain good-making features with classically valid arguments. I

[17] See Schulz (2017) for an excellent overview.

propose to account for their good-making features in terms of pseudo-validity. My task in the remainder of this paper will be to develop a Stalnakerian variably strict theory on which all three of our inference patterns are pseudo-valid. I start with indicatives in §5 and turn to counterfactuals in §6.

5 Stalnakerian Indicative Conditionals

In §5.1 I motivate a version of Stalnaker's theory of indicatives that draws a close connection between indicatives and epistemic modals. In §5.2, I show that the resulting theory predicts that Antecedent Strengthening, Transitivity, and Contraposition are pseudo-valid.

5.1 The Semantics

To state Stalnaker's theory, we begin with a contextually-supplied epistemic accessibility relation E. I assume for simplicity that $\mathsf{E}(w)$ is the set of worlds consistent with what's known, in w, by the speaker at the time of utterance. I assume that E is reflexive. A *Stalnakerian selection function* f_{E} is a contextually-supplied function that takes a world and a proposition to a set containing at most one world. Then:

Stalnaker Semantics for Indicatives
$[\![\text{If A, then B}]\!]^{\mathsf{E},w} = 1$ if and only if $f_{\mathsf{E}}(w, [\![\text{A}]\!]^{\mathsf{E}}) \subseteq [\![\text{B}]\!]^{\mathsf{E}}$

We make four assumptions about f_{E}.

Success
$f_{\mathsf{E}}(w, [\![\text{A}]\!]^{\mathsf{E}}) \subseteq [\![\text{A}]\!]^{\mathsf{E}}$

Minimality
If $w \in [\![\text{A}]\!]^{\mathsf{E}}$, then $f_{\mathsf{E}}(w, [\![\text{A}]\!]^{\mathsf{E}}) \subseteq \{w\}$.

Accessibility Constraint
$f_{\mathsf{E}}(w, [\![\text{A}]\!]^{\mathsf{E}}) \subseteq \mathsf{E}(w)$

Non-Vacuity
If $\mathsf{E}(w) \cap [\![\text{A}]\!]^{\mathsf{E}} \neq \emptyset$, then $f_{\mathsf{E}}(w, [\![\text{A}]\!]^{E}) \neq \emptyset$

Success secures the validity of Identity, the principle that ⌜if A, then A⌝ is always true. Minimality secures the validity of Modus Ponens.

The Accessibility Constraint is motivated by the feeling that indicative conditionals talk about possibilities that are compatible with our knowledge. To give a more concrete argument, I'll appeal to the close connection between indicatives and epistemic modals.[18] In particular, I'll assume that indicatives are interpreted according to the same accessibility relation as epistemic modals, and I'll assume a standard relational semantics for epistemic modals.

Epistemic Modals
$[\![\text{Must A}]\!]^{\mathsf{E},w} = 1$ if and only if $\mathsf{E}(w) \subseteq [\![\text{A}]\!]^{\mathsf{E}}$

I assume that ⌜Might A⌝ is the dual of ⌜Must A⌝ and is therefore true if and only if A is true at some accessible world. With the assumption that indicatives and epistemic modals are interpreted uniformly in place, we can now show that the Accessibility Constraint (when combined with Success and Non-Vacuity) secures the validity of the following two principles.[19]

Boxy Or-to-If
Must (A or B) ⊨ If not A, then B

Might-to-Might
Might A and if A, then B ⊨ Might B

We need Boxy Or-to-If to explain why the following (non-boxy) Or-to-If principle seems so compelling.

Or-to-If
A or B ⊨ If not A, then B

Suppose I say to you:

(26) Matt is either in Oxford or in London.

You conclude:

(27) Therefore, if Matt is not in Oxford, he's in London.

[18] The idea that indicatives and epistemic modals should receive a uniform interpretation is common, especially in dynamic treatments of conditionals. See, for example, Gillies (2004). See Santorio (2022) for extensive discussion about the interaction between 'if' and 'might'. See Dorr & Hawthorne (ms) for a defense of Boxy Or-to-If and Might-to-Might.

[19] *Proof of Boxy Or-to-If*. Suppose $[\![\text{Must (A or B)}]\!]^{\mathsf{E},w} = 1$. Then $[\![\text{not A}]\!]^{\mathsf{E}} \cap \mathsf{E}(w) \subseteq [\![\text{B}]\!]^{\mathsf{E}}$. By Success and the Accessibility Constraint, $f_{\mathsf{E}}(w, [\![\text{not A}]\!]^{\mathsf{E}}) \subseteq [\![\text{not A}]\!]^{\mathsf{E}} \cap \mathsf{E}(w)$. So $f_{\mathsf{E}}([\![\text{not A}]\!]^{\mathsf{E}}, w) \subseteq [\![\text{B}]\!]^{\mathsf{E}}$. By Stalnaker's Semantics, $[\![\text{If not A, then B}]\!]^{\mathsf{E},w} = 1$.

Proof of Might-to-Might. Assume (1) $[\![\text{Might A}]\!]^{\mathsf{E},w} = 1$ and (2) $[\![\text{If A, then B}]\!]^{\mathsf{E},w} = 1$. Suppose, for contradiction, that $[\![\text{Might B}]\!]^{\mathsf{E},w} = 0$. Then $\mathsf{E}(w) \subseteq [\![\text{not B}]\!]^{\mathsf{E}}$ and so $\mathsf{E}(w) \cap [\![\text{A}]\!]^{\mathsf{E}} \subseteq [\![\text{not B}]\!]^{\mathsf{E}}$. By the Accessibility Constraint and Success, $f_{\mathsf{E}}(w, [\![\text{A}]\!]^{\mathsf{E}}) \subseteq \mathsf{E}(w) \cap [\![\text{A}]\!]^{\mathsf{E}}$ and so $f_{\mathsf{E}}(w, [\![\text{A}]\!]^{\mathsf{E}}) \subseteq [\![\text{not B}]\!]^{\mathsf{E}}$. By (2) and Stalnaker's Semantics it follows that $f_{\mathsf{E}}(w, [\![\text{A}]\!]^{\mathsf{E}}) \subseteq [\![\text{B}]\!]^{\mathsf{E}}$. Since $f_{\mathsf{E}}(w, [\![\text{A}]\!]^{\mathsf{E}}) \subseteq [\![\text{not B}]\!]^{\mathsf{E}}$ and $f_{\mathsf{E}}(w, [\![\text{A}]\!]^{\mathsf{E}}) \subseteq [\![\text{B}]\!]^{\mathsf{E}}$, it follows that $f_{\mathsf{E}}(w, [\![\text{A}]\!]^{\mathsf{E}}) = \emptyset$. But that contradicts Non-Vacuity since by (1) $\mathsf{E}(w) \cap [\![\text{A}]\!]^{\mathsf{E}} \neq \emptyset$.

The inference seems flawless. But we can't say that Or-to-If valid. For if we did, it would be a short step to the view that indicative conditionals are logically equivalent to material conditionals.[20] If we want to deny this, we must deny that Or-to-If is valid, and find some other way of explaining why the inference from (26) to (27) seems valid. That's where Boxy Or-to-If comes in. Suppose I assert (26) and that I say something true. If you accept my assertion, then (26) becomes part of our shared information. That is to say, (28) becomes true in our context.

(28) Matt must either be in Oxford or in London.

If Boxy Or-to-If is valid, then (28) entails (27). So when you go on to assert (27) you say something true in our context.[21]

Might-to-Might also seems like an excellent principle. Consider:

(29) The fridge might stop working.

(30) If the fridge stops working, the food will spoil.

(31) Therefore, the food might spoil.

This reasoning is impeccable.

Finally, I will assume that that an indicative with antecedent A presupposes that there are accessible A-worlds.[22] Given our assumption that indicatives and epistemic modals are interpreted uniformly, we can state this compatibility presupposition as follows.

> **Might Presupposition**
> An indicative ⌜If A, then B⌝ presupposes in a given context what is expressed by ⌜Might A⌝ in that context.

Might Presupposition is well-motivated. In general, it is not appropriate to assert

[20] Suppose Or-to-If were valid. Then ⌜not A or B⌝ would entail ⌜if not not A, then B⌝. If we assume that ⌜not not A⌝ is logically equivalent to A, it follows that the material conditional ⌜not A or B⌝ entails the indicative conditional ⌜if A, then B⌝. The other direction—that ⌜if A, then B⌝ entails ⌜not A or B⌝ is a consequence of Minimality.

[21] Stalnaker (1975) makes use of the notion of a *context set*—a set of worlds consistent with what is *commonly presupposed* by the speakers—to capture the epistemic flavor of indicatives. Where S_c is the context set, Stalnaker assumes:

> **Stalnaker's Constraint**
> For all $w \in S_c$, if $S_c \cap [\![A]\!]^c \neq \emptyset$, then $f_c(w, [\![A]\!]^c) \subseteq S_c$

I prefer the theory in the main text because it preserves the connection between epistemic modals and indicatives. This of course assumes a relational semantics for epistemic modals. We might restore the connection between epistemic modals and indicatives on Stalnaker's theory by adopting a *domain semantics* following Yalcin (2007). For reasons discussed in Ninan (2018) and Mandelkern (2019), I believe that this approach to epistemic modality is not well-motivated.

[22] See Stalnaker (1975) and von Fintel (1999).

an indicative conditional if it is known that the antecedent is false. For example, if I assert (32), I can continue with (33), but I can't continue with (34).

(32) David is not at the party.

(33) ✓ If David had come to the party, he would have brought Matt.

(34) # If David is at the party, he brought Matt.

We can also argue for Might Presupposition by appeal to so-called *projection tests* for presupposition.[23] The presuppositions of a sentence are preserved when the sentence is embedded in certain environments. For example, the presuppositions of a sentence 'S' tend to be inherited by 'S?', 'I doubt that S', and 'maybe, S'. We can test Might Presupposition by examining whether the presupposition survives when indicatives are embedded in these environments. Might Presupposition seems to pass these tests. Suppose I'm wondering whether I can safely put my mug in the dishwasher. I ask my partner:

(35) Will it be ruined if it's hand-painted?

If she hadn't considered that the mug might have been hand-painted, it would be natural for her to respond:

(36) Wait, you think it might have been hand-painted?

Similar things can be said about indicatives under 'doubt' and 'maybe'. Consider:

(37) Maybe if it's hand-painted, it will be ruined.

(38) I doubt it will be ruined if it's hand-painted.

Once again, if my partner hadn't considered that the mug might have been hand-painted, it would be natural for her to respond with something like (36).

5.2 Pseudo-Validity and the Inferences

We turn now to the inferences of Transitivity, Contraposition, and Antecedent Strengthening. I propose to account for the good-making features of these inferences in terms of *pseudo-validity*.

Pseudo-Validity
The argument from $A_1, A_2, \ldots, A_n$ to C is pseudo-valid if and only if, whenever the premises can be felicitously asserted and ⌜Must A_1⌝, ⌜Must A_2⌝, . . . , and ⌜Must A_n⌝ are all true, C is also true.

[23] Chierchia & McConnell-Ginet (2000).

To say that an argument is pseudo-valid is to say, roughly, that if the premises can be felicitously asserted, and the premises are entailed by the information in a given context—that is to say, the premises are true in all epistemically possible worlds—then the conclusion is also true in that context. Pseudo-valid arguments tend to strike us as good arguments. Here's an example from Bledin (2015).

(39) Either Mrs. White did it or Miss Scarlet did it.

(40) Miss Scarlet didn't do it.

(41) Mrs. White must have done it.

Here's another example that is closely related to Or-to-If. I say:

(26) Matt is either in Oxford or London.

You conclude:

(42) Matt might be in Oxford, and if he isn't, he's in London.

Neither of these arguments is classically valid. In each case, the argument can have true premises and a false conclusion if the speaker doesn't know that the premises are true. Still, both arguments seem excellent. A natural explanation of this is that both arguments are pseudo-valid. Consider the second case, for example. On any plausible way of filling in the details of the example, the premise is not just true—it is also known by the speaker. If I assert (26), then ordinarily I am doing so because I know that Matt is either in Oxford or London and I don't know that Matt is not in Oxford. (In general, it is not appropriate to assert a disjunction unless you think both disjuncts are possible.[24]) If you accept my assertion, then you will come to know that Matt is either in Oxford or London, and you will also come to know that Matt might be in Oxford. If the inference from (26) to (42) is pseudo-valid, then it follows from these facts that (42) is true in our context. This means that when you assert (42) you say something that is true. In short, pseudo-valid arguments tend to strike as as excellent arguments because, ordinarily, if we come to know the premises *on the basis of a successful assertion of the premises*, the conclusions must be true, too.[25]

[24] This constraint can be justified on familiar Gricean grounds as Stalnaker (1975) suggests.

[25] Note that Or-to-If is also pseudo-valid. (Remember, we can't say that it is classically valid on pain of collapsing the indicative into the material conditional.) As we have seen, arguments conforming to Or-to-If are compelling; we are liable to mistake them for classically valid inferences. Pseudo-validity is also closely related to Stalnaker's notion of a *reasonable inference*. The inference from premises $A_1, \ldots, A_n$ to conclusion C is a reasonable inference if and only, for any context in which the premises can be felicitously asserted, if the context accepts the premises, then the context accepts the conclusion. There are two important differences between pseudo-validity and reasonable inference. The first is that reasonable inference is not defined in terms of epistemic modality. The second difference is that pseudo-validity does not say that the conclu-

According to the Stalnakerian theory of indicatives that I have presented, Transitivity, Contraposition, and Antecedent Strengthening are pseudo-valid. To see this, consider Transitivity. Suppose ⌜Must (If A, then B)⌝ and ⌜Must (if B, then C)⌝ are true. It then follows that ⌜Must (not A or B)⌝ and ⌜Must (not B or C)⌝ are both true. (This follows from Modus Ponens and our semantics for epistemic modals.) It follows that ⌜Must (not A or C)⌝ is true. Finally, by Boxy Or-to-If, it follows that ⌜If A, then C⌝ is true. The proofs for Antecedent Strengthening and Contraposition are similar.

We can appeal to the pseudo-validity of Transitivity, Contraposition, and Antecedent Strengthening to explain much of the data that is normally taken to motivate strict theories. Consider Transitivity first. Suppose my friend is having a party tonight. I say:

(43) If John goes to the party, Matt will go.

(44) If Matt goes to the party, David will go.

If you trust me, then you will and should conclude:

(45) If John goes to the party, David will go.

Here's another example. I say:

(46) If the power goes out out, the fridge will stop working.

(47) If the fridge stops working, the food will spoil.

If you trust me, you will and should conclude:

(48) If the power goes out, the food will spoil.

We can explain these facts in terms of pseudo-validity. Suppose I felicitously assert (43) and (44) and speak truly. If you accept my assertions, then (43) and (44) become part of our shared knowledge—(49) and (50) become true in our context.[26]

(49) It's gotta/has to/must be that if John goes to the party, Matt will go.

(50) It's gotta/has to/must be that if Matt goes to the party, David will go.

Since Transitivity is pseudo-valid, it follows from these facts that (45) is also true. Thus, when you assert (45) you say something that is true in our context.

Turn to Contraposition. Consider:

(51) If the power went out, the fridge stopped working.

---

sion is always *known* when the premises are assertable and known.

[26] I'm assuming that 'gotta' and 'has to' are synonymous with epistemic 'must' here.

(52) So, if the fridge didn't stop working, the power didn't go out.

Once again, this strikes us as an excellent argument, and once again, I propose to explain this appearance of validity using pseudo-validity. Suppose I felicitously assert (51) and that I say something true. If you accept my assertion, then (51) becomes part of our shared information, and so (53) will be true in our context.

(53) It's gotta/has to/must be that if the power went out, the fridge stopped working.

Since Contraposition is pseudo-valid, it follows from these facts that (52) is also true in our context. Thus, when you assert (52), you say something that is true in our context.[27]

Finally, turn to Antecedent Strengthening. Early advocates of the variably strict theory argued that Antecedent Strengthening is subject to counterexample. David Lewis (1973) appeals to so-called *Sobel sequences*, such as:

(54) If Alice comes to the party, it will be fun.

(55) But if Alice and Bob come to the party, it won't be fun.

This sequence of counterfactuals is unremarkable. But if Antecedent Strengthening were valid, Lewis argued, (54) and (55) would be inconsistent. Strict theorists are not convinced. They say that the alleged counterexamples exhibit signs of illicit context-shifting. For example, they observe that Sobel sequences sound much worse when their order is reversed.

(55) If Alice and Bob come to the party, it won't be fun.

(54) ? But if Alice comes to the party, it will be fun.

When we encounter (54) after an assertion of (55), we find it much harder to accept (54). (We're tempted to ask: What if Bob also comes?)

As strict theorists observe, this sort of order sensitivity is exactly what we should expect to find if the apparent counterexamples involved an illicit context shift.[28] Remember that an indicative carries a compatibility presupposition: ⌜if A, then B⌝ presupposes that there are accessible A-worlds. With this in mind, suppose I assert (54). If I say something true, then, according to the strict theory, all epistemically accessible worlds where Alice goes to the party are worlds where the party is fun. If Bob does not go to the party in any of these worlds, then the presupposition of (55) is not satisfied. In general, speakers strive to avoid ut-

[27] Stalnaker (1975) gives a similar explanation of the apparent validity of Transitivity and Contraposition for indicatives. He says that both are *reasonable inferences* in the sense defined in footnote 25. Stalnaker does not extend this story to counterfactuals.

[28] My explanation follows von Fintel (2001) and Gillies (2007).

tering sentences whose presuppositions are not satisfied, and hearers strive to interpret speakers in ways that do not leave the presuppositions of their utterances unsatisfied. Asserting (55), then, will tend to force the context to change; hearers will interpret the conditional relative to an expanded set of worlds, one that includes worlds where Alice and Bob go. If the party is not fun in any of these worlds, then (55) will come out true in our new context.

Now consider what happens when I assert (55) first. If I say something true, then all epistemically accessible worlds where Alice and Bob go to the party are worlds where the party is not fun. If you accept my assertion, the compatibility presupposition will ordinarily be satisfied. This means that the presupposition of (54) will also be satisfied. You have no reason to choose a different accessibility relation to interpret (54). The strict theory says that (55) and (54) cannot both be true when interpreted relative to the same accessibility relation. Therefore, it will typically not be acceptable for me to continue by asserting (54).

I propose to explain the contrast between forward and reverse Sobel sequences in terms of pseudo-validity. The explanation is similar to the strict theorist's. Forward Sobel sequences are often felicitous because the context tends to change midway through the sequence. Suppose I assert (54) and say something true. If you accept my assertion, then all epistemically accessible worlds where Alice goes to party will be worlds where it is fun. If Bob does not go to the party in any of these worlds, the presupposition of (55) won't be satisfied, and so asserting (55) will usually change the context. Reverse Sobel sequences are typically infelicitous because the context tends not to change when the sentences are uttered in reverse order. Consider any context in which (54) can be felicitously asserted. If Antecedent Strengthening is pseudo-valid, then (54) and (55) cannot both be entailed by our information in any such context. Suppose, then, that I assert (55) in such a context and that I say something true. If you accept my assertion, then the presupposition of (55) must be satisfied, which in turn means that the presupposition of (54) is satisfied. You have no reason to choose a different accessibility relation to interpret (54). Thus, we should expect that when I go on to assert (54), you will be puzzled—after all, I am asking you to accept something that you cannot accept if you have accepted (55).

We have seen that Transitivity, Contraposition, and Antecedent Strengthening are pseudo-valid on Stalnaker's theory of indicatives. In fact, it is easy to show that all characteristically strict inference patterns are pseudo-valid on this theory.[29] This means that if we want to argue for the strict theory on the basis of one

[29] Suppose that one or more strict conditionals ⌜Must(not A_1 or B_1)⌝, . . . , ⌜Must (not A_n or B_n)⌝ together entail another strict conditional ⌜Must (not C or D)⌝. Then the Stalnakerian theory of

or more of these characteristically strict principles, we will have to show that we have reason to prefer the strict theorist's explanation of the apparent validity of the principle—that the principle is valid—over the explanation I have suggested—that the principle is merely pseudo-valid. But we should be pessimistic that this can be done. For one, strict theorists already have to say that at least some apparently excellent inferences—such as Or-to-If—have some status, like pseudo-validity, that falls short of classical validity and that we tend to mistake for validity. Why should the characteristically strict principles of Transitivity, Contraposition, and so forth be treated differently? For another, the probabilistic facts discussed in §3 tell strongly in favor of the hypothesis that these inference patterns are merely pseudo-valid. Classically valid arguments do not just preserve knowledge; they also preserve credence. And we have seen that characteristically strict inference patterns are not credence-preserving.

6 Counterfactuals

To secure the pseudo-validity of Transitivity, Contraposition, and Antecedent Strengthening for indicative conditionals, we appealed to the close connection between indicatives and epistemic modality. To secure the pseudo-validity of these principles for counterfactuals, I will appeal to the close connection between counterfactuals and the modals 'had to' and 'could have'. In particular, I will develop a Stalnakerian theory of counterfactuals that validates both of the following principles.

Counterfactual Or-to-If
Had to be that (A or B) $\vDash$ Had not been that A, would have been that B

Qualified Counterfactual If-to-Or
Might not A, Must (A $\Box\!\!\rightarrow$ B) $\vDash$ Couldn't have been that (A and not B)

If we assume, as I will, that 'had to' and 'could have' are duals, then these two principles suffice to secure the pseudo-validity of our three inference patterns. To see this, consider Transitivity. Suppose A $\Box\!\!\rightarrow$ B and B $\Box\!\!\rightarrow$ C can be felicitously asserted in a given context.[30] In general, a counterfactual can be felicitously asserted only if its antecedent is not known to be true. (I will say more to justify this assumption in §7.1.) If that's right, then ⌜Might not A⌝ and ⌜Might not B⌝ are both true. Suppose further that ⌜Must (A $\Box\!\!\rightarrow$ B)⌝ and ⌜Must (B $\Box\!\!\rightarrow$ C)⌝ are

indicatives will say that the inference from ⌜if A_1, then B_1⌝, . . . , ⌜if A_n, then B_n⌝ to ⌜if C, then D⌝ is pseudo-valid.

[30] I use 'A $\Box\!\!\rightarrow$ B' as short for the counterfactual with antecedent A and consequent B.

both true in the context. By Qualified Counterfactual If-to-Or, it then follows that ⌜Couldn't have been (A and not B)⌝ and ⌜Couldn't have been (B and not C)⌝ are both true. This entails that ⌜Couldn't have been (A and not C)⌝ is also true. Since 'could have' and 'had to' are duals, it follows that ⌜Had to be that (not A or C)⌝ is true. Finally, by Counterfactual Or-to-If, the counterfactual $A \mathrel{\Box\!\!\rightarrow} C$ is true. The proofs for Antecedent Strengthening and Contraposition are similar.[31]

Here's the plan for the next two sections. I'll start in §6.1 by rehearsing some facts about the grammatical ingredients of counterfactuals. We will see that counterfactuals contain past tense morphology that appears not to be interpreted in the usual temporal way. I'll discuss the two main approaches to the semantics of this so-called *fake past*—the *past-as-past* approach and the *past-as-modal* approach. In §6.2-6.3, I'll present my own theory following in the past-as-modal tradition. In §7.1 I highlight a shortcoming of the theory as it stands. In particular, I will show that the theory does not yet secure Qualified Counterfactual If-to-Or. In§7.2, I suggest that the problem is a close relative of a familiar problem concerning the probabilities of indicatives. In §7.3, I suggest a solution that mirrors the standard solution to the problem for indicatives. I then show that the resulting theory validates Counterfactual Or-to-If and Qualified Counterfactual If-to-Or, which means that the theory also secures the pseudo-validity of Transitivity, Contraposition, and Antecedent Strengthening. In §7.4 I appeal to pseudo-validity to explain the good-making features of these three inference patterns for counterfactuals.

6.1 Fake Tense

Consider:

(56) If Daniel is not teaching this quarter, he will teach next quarter.

(57) If Daniel <u>were not</u> teaching this quarter, he <u>would</u> be teaching next quarter.

Notice that the antecedent and consequent of (57) take a morphologically past tense form. The antecedent contains 'were' instead of 'is', and the consequent contains 'would' instead of 'will'. Yet (57) does not seem to be about events in the

[31] You might be wondering why I'm not relying on the following stronger principle:

Counterfactual If-to-Or
Must $(A \mathrel{\Box\!\!\rightarrow} B) \vDash$ Couldn't have been that (A and $\neg$B)

I will explain why in §7.1.

past; instead, it makes a claim about a hypothetical scenario in which Daniel is not teaching this quarter—namely, that in that scenario, he teaches next quarter.

Iatridou calls this extra layer of past tense morphology *fake past*. She and others observe that, across a wide range of languages in different families, we find that past tense markers in certain constructions—such as conditionals—do not seem to be interpreted as referring to the past; instead, they seem to be interpreted *modally*.

How, exactly, does the past tense morphology in counterfactuals give rise to this apparent modal interpretation? The first theory—the *past-as-past* theory—says that the past tense markers are, contrary to appearances, interpreted in the usual temporal way. Proponents of this theory say that in counterfactuals like (57), the past tense takes wide scope over the conditional, modifying a covert modal; roughly, the past takes us back to a time when the antecedent was still open and conveys what was necessary or possible then. (57), for example, says that it *was* necessary that if Daniel were not teaching this quarter, he would be teaching next quarter.[32] The second theory—the *past-as-modal* theory—says that the past tense does indeed have a modal interpretation in certain constructions, such as counterfactuals. One intriguing idea, due to Iatridou (2000), is that the past tense has a core schematic meaning expressing some kind of *distance*. In its temporal interpretation, the past conveys *temporal distance* from a designated time. In its modal interpretation, the past conveys *modal distance* from a designated modal parameter.[33]

We find similar tense morphology in modal constructions. My focus will be on the modals 'could have' and 'had to'. To take an example, suppose our roommate David comes home, and leaves his wet umbrella at the door. You and I are heading out for a walk. Just as we step into the rain, I notice that you aren't wearing your raincoat.

(58) Why didn't you wear your raincoat? It had to be raining out—you saw David's umbrella.

Or suppose instead that it turns out not to be raining by the time we get outside. Surprised that you didn't prepare for rain, I say:

(59) Why didn't you wear your raincoat? It could have been raining out right now.

In both (58) and (59), we find past tense morphology on the modals. We have 'had to be raining' instead of 'has to be raining' in the case of (58) and 'could have

[32] See Ippolito (2013), Khoo (2015), and Khoo (2022) for past-as-past theories.
[33] See Iatridou (2000), Schulz (2014), and Mackay (2019) for past-as-modal theories.

been raining' instead of 'could be raining' in the case of (59). (58) seems to be saying, roughly, that we knew, or should have known, that it would be raining when we stepped outside; likewise, (59) seems to be saying, roughly, that rain was compatible with an impoverished knowledge state—in this case, our knowledge state before stepping outside.

Discussions about the interaction between modals and tense tend not to assimilate sentences like (58) and (59) to fake tense. The reason for this is perhaps that these sentences seem on their face to be more amenable to a past-as-past treatment than counterfactuals. But in the next section I will argue that we have reason to believe that counterfactuals and 'could have' and 'had to' are interpreted uniformly. If we're assuming a uniform treatment, our approach to counterfactuals should inform our approach to these modals; in particular, if we decide on a past-as-modal approach to counterfactuals, as I will suggest, it is natural to extend this past-as-modal approach to 'had to' and 'could have'.

The theory of fake tense that I favor follows in the past-as-modal tradition. I propose that the past tense, in its modal interpretation, shifts the accessibility relation relative to which we interpret the embedded conditional or modal. Roughly, a counterfactual with antecedent A and consequent B is true relative to our actual information if and only if the corresponding indicative is true relative to a contextually-determined *weakening* of our information. Likewise, I will say that the modal claim ⌜It had to be that A⌝ is true relative to our current information just in case ⌜It has to be that A⌝ is true relative to a contextually-determined weakening of our information. In the next two sections, I present a sketch of this theory.

6.2 A Past as Modal Theory

I begin with a referential theory of tense, according to which tenses, like pronouns, refer to particular entities.[34] Just as the pronoun 'she' is used to talk about a particular, contextually-salient individual, the past tense is used to talk about a particular, contextually-salient time. On this view, tenses are treated as free variables whose values are determined by a contextually-supplied assignment function g. ⌜Past_i A⌝ says that A is true at the time picked out by $g(i)$. I will also assume, following Heim (1994), that the past, in its temporal interpretation, carries a presupposition of *temporal precedence*: the past presupposes that the time picked out by g precedes the time of utterance.

The modal interpretation of past replaces variables over times with variables

[34] See Partee (1973).

over accessibility relations. We say that one accessibility relation E_1 is *at least as informed* as another accessibility relation E_2 (in symbols: $\mathsf{E}_1 \geq \mathsf{E}_2$) if and only if $\mathsf{E}_1(w) \subseteq \mathsf{E}_2(w)$ for all w. And we say that E_1 is *more informed* than E_2 (in symbols: $\mathsf{E}_1 > \mathsf{E}_2$) if and only if $\mathsf{E}_1(w) \subset \mathsf{E}_2(w)$ for all w. To interpret the modal past, we assume that context supplies a set of accessibility relations $\mathcal{E} = \{\mathsf{E}_1, \mathsf{E}_2, \ldots \mathsf{E}_n\}$ totally ordered by $\geq$ and a variably assignment g that assigns to each free variable an accessibility relation in $\mathcal{E}$. To interpret conditionals, we continue to assume that context supplies an accessibility relation E representing the knowledge of the speaker (and we assume that $\mathsf{E} \in \mathcal{E}$). We have the following semantic entry for the modal past.

Semantics for Modal Past
$[\![\text{Past}_i\ \text{A}]\!]^{w,g,\mathsf{E},\mathcal{E}} = 1$ if and only if $[\![\text{A}]\!]^{w,g,g(i),\mathcal{E}} = 1$

The presupposition of temporal precedence is replaced by a presupposition of *informational precedence*. In its modal interpretation, $\ulcorner$Past$_i$ A$\urcorner$ presupposes that the accessibility relation $g(i)$ is less informed than E (the accessibility relation representing the speaker's knowledge).[35]

I assume that a counterfactual involves a past tense operator scoping over the conditional.

(61) Past$_i$ (if A, then B)

I will continue to assume Stalnaker's semantics for indicative conditionals, restated below.

Stalnaker's Semantics
$[\![\text{If A, then B}]\!]^{w,g,\mathsf{E},\mathcal{E}} = 1$ if and only if $f_{\mathsf{E}}(w, [\![\text{A}]\!]^{g,\mathsf{E},\mathcal{E}}) \subseteq [\![\text{B}]\!]^{g,\mathsf{E},\mathcal{E}}$

[35]This proposal is partly inspired by Heim (1992)'s discussion of presuppositions in counterfactual antecedents. Consider:

(60) If John had attended too, we would have had a great time.

Since (60) can be felicitous, we must be interpreting the presupposition of 'too' relative to an information state that contains some, but not all, of our information. Heim writes:

> These examples suggest that the antecedent of a counterfactual is not really added to an 'empty' context, but to one which is in some sense a *revision* of the common ground *c*. It results from *c* by suspending some of the assumptions in *c*. (page 204)

My account also builds on Schulz (2014). However, there's an important difference between my account and hers. For Schulz, the modal past does not shift the information state relative to which we interpret the embedded expression to a *weakening* of our actual information state, but to a set of worlds that is *inconsistent* with our actual information. This has bad consequences—for example, Modus Ponens can fail. (See Mackay (2015).) observes. My proposal is also similar in spirit to the proposals of Mackay (2019) and Kaufman (2013), though the formal implementation differs from both.

Combining our semantics for modal past with Stalnaker's Semantics for conditionals gives us:

Semantics for Counterfactuals
$[\![\text{Past}_i \text{ (if A, then B)}]\!]^{w,g,\mathsf{E},\mathcal{E}} = 1$ if and only if $f_{g(i)}(w, [\![\text{A}]\!]^{g,g(i),\mathcal{E}}) \subseteq [\![\text{B}]\!]^{g,g(i),\mathcal{E}}$

This says that a counterfactual ⌜If had been that A, would have been that B⌝ is true at a world w, relative to our information state E, just in case the corresponding indicative conditional ⌜If A, then B⌝ is true at w, relative to a contextually-determined, less informed accessibility relation $g(i)$.

On this theory, all of the differences between indicatives and counterfactuals are grounded in differences in what information is held fixed when we evaluate the conditional. When we evaluate an indicative, we hold fixed everything we know: the selected antecedent-world must be epistemically possible. When we evaluate a counterfactual, we hold fixed only some of what we know: the selected antecedent-world must be consistent with $g(i)(w)$. Take an example. Suppose we know Matt made it on time to the dinner at six. We're told he caught the bus at five. Doubting that leaving at five left him enough time to get here by six, you say:

(62) If Matt had left at five, he wouldn't have made it to the dinner on time.

To evaluate (62), we temporarily suspend some of our knowledge—our knowledge of the fact that Matt made it to the dinner on time, among other things—but we hold much of what we know fixed. In particular, we tend to hold fixed much of our knowledge about what happened before the time of the events described in the antecedent. For example, we hold fixed facts about when Matt started to get dressed, facts about which buses were running at that time, and so forth. What we do and do not hold fixed is represented by the accessibility relation $g(i)$. If we're holding fixed facts about when Matt started to get dressed, then $g(i)$ takes each world w to a set of worlds consistent with what we know, in w, about when he started getting dressed. If we're holding fixed facts about the bus schedules, then $g(i)$ takes each world w to a set of worlds that is consistent with what we know, in w, about the bus schedules.

As many authors have noted, this pattern of holding the past history of the world fixed is pervasive. Many authors go so far as to say that when we're evaluating a counterfactual whose history concerns a particular period of time, there is some time *t*—usually not too long before the period of time in question—such that we hold fixed *all* facts about history before *t*.[36] Following Dorr (2016), I do

[36] See Lewis (1979).

not insist on this principle.[37] Nevertheless, I agree with Dorr and others that the principle is a good heuristic in the sense that, typically, when evaluating a counterfactual whose antecedent concerns a particular period of time, we hold fixed a *broad range* of facts about history before that time.

There is of course a great deal more to say about these issues, but it is beyond the scope of this paper to give a detailed account of what is and what is not held fixed. I take it that we have a a rough-and-ready understanding of the notion and that is enough for my purposes. Note that I do not assume that it is possible to give a reductive analysis of what is held fixed. It may be that we need counterfactuals to distinguish what we may rely on from what we may not rely on: roughly, the propositions that we hold fixed are those that *would still have been true* if the antecedent had been true.

Remember the four constraints on the selection function. First we have the principle of Success, which says that the selected A-world is an A-world. Second, we have the principle of Minimality, which says that if w is an A-world, then the selected A-world, at w, is w. The difference between indicatives and counterfactuals primarily concerns the final two constraints: the Accessibility Constraint and Non-Vacuity. For indicatives, these are stated in terms of what we know. For counterfactuals, they are stated in terms of what we're holding fixed for the purposes of evaluating the counterfactual.

Accessibility Constraint
$f_{g(i)}(w, [\![\mathrm{A}]\!]^{g,g(i),\mathcal{E}}) \subseteq g(i)(w)$

Non-Vacuity
If $g(i)(w) \cap [\![\mathrm{A}]\!]^{g,g(i),\mathcal{E}} \neq \emptyset$, then $f_{g(i)}(w, [\![\mathrm{A}]\!]^{g(i)}) \neq \emptyset$

Recall that I motivated the Accessibility Constraint for indicative conditionals by appealing to the close connection between indicatives and epistemic modals. In a similar spirit, I will motivate the Accessibility Constraint for counterfactuals by appeal to the close connection between counterfactuals and the modals 'could have' and 'had to'.

To begin, I will assume that 'has to' has an epistemic interpretation with the same meaning as epistemic 'must'. I will continue to assume a standard relational semantics for epistemic modals, restated below.

[37] If we say that all facts about history are held fixed in what Lewis (1979) calls *standard* contexts, then there is strong pressure to say that the laws of nature are *not* held fixed in these contexts. But—as Dorr persuasively argues—the hypothesis that the laws of nature are not held fixed in standard contexts has unpalatable consequences.

Epistemic Modals
$[\![\text{Has to be that A}]\!]^{w,g,\mathsf{E},\mathcal{E}} = [\![\text{Must A}]\!]^{w,g,\mathsf{E},\mathcal{E}} = 1$ if and only if $\mathsf{E}(w) \subseteq [\![\text{A}]\!]^{g,\mathsf{E},\mathcal{E}}$

I will assume that the modal claim ⌜It had to be that A⌝ has the following much-simplified logical form.[38]

(67) Past$_i$ (has to be that A)

Combining our semantics for the modal past with our relational semantics for epistemic modals gives us:

Semantics for 'Had to'
$[\![\text{Past}_i \text{ (has to be that A)}]\!]^{w,g,\mathsf{E},\mathcal{E}} = 1$ if and only if $g(i)(w) \subseteq [\![\text{A}]\!]^{g,g(i),\mathcal{E}}$

This says that ⌜It had to be that A⌝ is true relative to our current information just in case ⌜It has to be that A⌝ is true relative to a contextually-determined weakening of our information. Return to the raincoat example. As we step into the rain, I say:

(58) Why didn't you wear your raincoat? It had to be raining—you saw David's umbrella.

I say that (58) is true in our context if and only if the epistemic modal claim

(68) It has to be raining.

is true relative to a contextually-determined accessibility relation $g(i)$ that is less informed than E, the accessibility relation representing our knowledge. In

[38] Hacquard (2006) claims that English epistemic modals never occur under the past tense. She is aware of apparent counterexamples to this claim, such as the following example due to von Fintel and Gillies (2008). Sophie opens the freezer looking for ice cream. She doesn't find any. Asked why she opened the freezer, she replies:

(63) Because there might have been ice cream in the freezer.

Hacquard argues that this sentence contains an elided *I thought that* and that it is the *attitude verb* that shifts the time of the modal. Thus we have:

(64) Because ~~I thought that~~ there might have been ice cream in the freezer.

I don't see how Hacquard's story can account for all uses of sentences like (63). Suppose I lost my keys. We search the house. I decide not to search my office. You find them in the house. Later I tell you that I didn't search the office. You remind me that I have lost my keys there before. I reply:

(65) You're right. I guess they might very well have been in the office.

This does *not* mean:

(66) ? You're right. I guess I thought they might very well have been in the office.

See also Boylan (2020) for further arguments against this strategy.

the context most naturally evoked by (58), this less informed accessibility relation represents what we were in a position to know just before stepping outside. Thus, while (68) says that all worlds compatible with our present information—which includes, among other things, the fact that we now see raindrops—are worlds where it's raining, (58) says that all worlds compatible with our earlier information—which does not include the fact that we see raindrops, but does include the fact that David was carrying an umbrella—are worlds where it's raining.

I assume that 'could have' is the dual of 'had to' and therefore that the modal claim ⌜It could have been that A⌝ is true in our context just in case some $g(i)$-accessible world is an A-world.

Now that we have a semantics for 'had to' and 'could have' on the table, we can show that the Accessibility Constraint (when combined with Success and Non-Vacuity) secures the validity of the following counterfactual analogues of Boxy Or-to-If and Might-to-Might.[39]

Counterfactual Or-to-If
Had to be that (A or B) ⊨ Had not been that A, would have been that B

Could-to-Could
Could have been that A, and A $\square\!\!\rightarrow$ B ⊨ Could have been that B

Counterfactual Or-to-If looks just as plausible as its indicative counterpart. Consider the following case, adapted from Edgington (2008). We're hunting for a treasure. The organizer says, "I'll give you a hint. It's either in the attic or the garden." Trusting him, I go to the attic and tell my partner to search the garden.

[39] *Proof of Counterfactual Or-to-If.* Suppose $[\![\text{Past}_\text{i}\ (\text{has to be that (A or B)})]\!]^{w,g,\mathsf{E},\mathcal{E}} = 1$. By the semantics for modal past, $[\![\text{Has to be that (A or B)}]\!]^{w,g,g(i),\mathcal{E}} = 1$. By the semantics for modals, $g(i)(w) \cap [\![\text{not A}]\!]^{g,g(i),\mathcal{E}} \subseteq [\![\text{B}]\!]^{g,g(i),\mathcal{E}}$. By Success and the Accessibility Constraint, $f_{g(i)}(w, [\![\text{not A}]\!]^{g,g(i),\mathcal{E}}) \subseteq g(i)(w) \cap [\![\text{not A}]\!]^{g,g(i),\mathcal{E}}$. It follows that $f_{g(i)}(w, [\![\text{not A}]\!]^{g,g(i),\mathcal{E}}) \subseteq [\![\text{B}]\!]^{g,g(i),\mathcal{E}}$. By Stalnaker's Semantics, $[\![\text{If not A, then B}]\!]^{w,g,g(i),\mathcal{E}} = 1$. Finally, by the semantics for modal past, $[\![\text{Past}_\text{i}\ (\text{if not A, then B})]\!]^{w,g,\mathsf{E},\mathcal{E}} = 1$.

Proof of Could-to-Could. Assume (1) $[\![\text{Past}_\text{i}\ (\text{Could be that A})]\!]^{w,g,\mathsf{E},\mathcal{E}} = 1$ and (2) $[\![\text{Past}_\text{i}\ (\text{If A, then B})]\!]^{w,g,\mathsf{E},\mathcal{E}} = 1$. Suppose, for contradiction, that $[\![\text{Past}_\text{i}\ (\text{Could be that B})]\!]^{w,g,\mathsf{E},\mathcal{E}} = 0$. Then $[\![\text{Could be that B}]\!]^{w,g,g(i)\mathcal{E}} = 0$. By the semantics for modals, $g(i)(w) \subseteq [\![\text{not B}]\!]^{g,g(i),\mathcal{E}}$ and so $g(i)(w) \cap [\![\text{A}]\!]^{g,g(i),\mathcal{E}} \subseteq [\![\text{not B}]\!]^{g,g(i),\mathcal{E}}$. By Success and the Accessibility Constraint, $f_{g(i)}(w, [\![\text{A}]\!]^{g,g(i),\mathcal{E}}) \subseteq g(i)(w) \cap [\![\text{A}]\!]^{g,g(i),\mathcal{E}}$. It follows that $f_{g(i)}(w, [\![\text{A}]\!]^{g,g(i),\mathcal{E}}) \subseteq [\![\text{not B}]\!]^{g,g(i),\mathcal{E}}$. Next observe that by (2) and the semantics for modal past, we have that $[\![\text{If A, then B}]\!]^{w,g,g(i),\mathcal{E}} = 1$. So by Stalnaker's Semantics, $f_{g(i)}(w, [\![\text{A}]\!]^{g,g(i),\mathcal{E}}) \subseteq [\![\text{B}]\!]^{g,g(i),\mathcal{E}}$. We now have that $f_{g(i)}(w, [\![\text{A}]\!]^{g,g(i),\mathcal{E}}) \subseteq [\![\text{B}]\!]^{g,g(i),\mathcal{E}}$ and that $f_{g(i)}(w, [\![\text{A}]\!]^{g,g(i),\mathcal{E}}) \subseteq [\![\text{not B}]\!]^{g,g(i),\mathcal{E}}$. It follows that $f_{g(i)}(w, [\![\text{A}]\!]^{g,g(i),\mathcal{E}}) = \emptyset$. By Non-Vacuity, $g(i)(w) \cap [\![\text{A}]\!]^{g,g(i),\mathcal{E}} = \emptyset$ and so $[\![\text{It could be that A}]\!]^{w,g,g(i),\mathcal{E}} = 0$. But then it follows that $[\![\text{Past}_\text{i}\ (\text{Could be that A})]\!]^{w,g,\mathsf{E},\mathcal{E}} = 0$, which contradicts (1).

I discover the treasure. “Why did you tell me to search the garden?” she asks. I reply:

(69) The treasure had to have been either in the attic or the garden. (The organizer told me it was in one of those places.)

My partner concludes:

(70) If it hadn’t been in the attic, it would have been in the garden.

This inference is flawless. Here’s one more example. You say:

(71) It couldn’t have been that Matt left at five and made it to the dinner by six.

I infer:

(62) If Matt had left at five, he wouldn’t have made it to dinner by six.

Once again this inference seems excellent.

Could-to-Could also looks like a good principle. Consider:

(72) It could have been in the garden.

(73) If it had been in the garden, you would have found it in the garden.

(74) Therefore, you could have found it in the garden.

This reasoning is impeccable.

6.3 The Compatibility Presupposition

We turn now to the presuppositions of counterfactuals. Remember that an indicative conditional with antecedent A presupposes that there are epistemically accessible A-worlds. Since indicatives and epistemic modals are interpreted uniformly, this means that an indicative with antecedent A presupposes in a given context what is expressed by ⌜Might A⌝ in that context. Before I can say more about the presuppositions of counterfactuals, I need to make note of an important feature of presuppositions.

In general, the presuppositions of an embedded clause are evaluated relative to the sentence’s local embedding environment–often called a *local context*.[40] We can think of this local information state as representing the information that is *already available* at a certain point in the course of processing the sentence.

[40] There are different ways to model local contexts. Some follow Heim (1983) and model them in a dynamic framework. Others follow Schlenker (2009) and model them in a static framework. If we wanted to give a complete account of the presuppositions of counterfactuals, we would need to say more about what local contexts are and how they are updated. This is not the place to give such an account. See Mandelkern (2019).

For example, the local context of the right conjunct of a conjunction includes information in the left conjunct; the local context of the consequent of a conditional includes information in the antecedent. Compare (75), on the one hand, with (76) and (77), on the other.

(75) Susie stopped smoking.

(76) If Susie used to smoke, she stopped smoking.

(77) John believes that Susie stopped smoking.

(75) presupposes that Susie used to smoke and is felicitous only if the speaker knows that Susie used to smoke. This is not the case for (76) or (77). These sentences can be felicitously asserted even when the speaker does not know that Susie used to smoke. The standard explanation is that, in both cases the presupposition of 'Susie used to smoke' need only be satisfied relative to the sentence's local context. In the case of (76), the local context includes the information in the antecedent; in the case of (77), the local context is John's belief state.

Like all presuppositions, Might Presupposition is also interpreted in a local manner. Consider:

(78) John thinks that he is going to teach during autumn, and he also thinks that if he teaches in autumn, he won't teach in winter.

(78) does not presuppose that John might teach in autumn. Instead, the presupposition of the indicative conditional is interpreted relative to John's belief state, and (78) presupposes only that John believes he might teach during autumn.[41] Observe that (79) is felicitous.

(79) John is not going to teach in autumn. But John thinks that he is, and he also thinks that if he teaches in autumn, he won't teach in winter.

We have said that a counterfactual consists of an an indicative conditional embedded under a past tense operator, and that the role of the past tense is to shift the information state relative to which we interpret the embedded conditional to a less informed accessibility relation $g(i)$. A natural hypothesis is that when a conditional occurs embedded under a past tense operator, we interpret the conditional's presuppositions relative to this less informed accessibility relation.[42] If that's right, then our theory will predict:

[41]Note that it will not suffice to simply combine our semantics for conditionals with the standard presupposition algorithm if we want to predict that (79) presupposes that John thinks that he might teach in autumn. But there's reason to believe that an alternative approach—specifically, the one developed in Mandelkern (2022)—will deliver the right predictions.

[42]This hypothesis is supported Heim's example discussed in footnote 35. You and I both know that Mary went to the party. I'm wondering whether her partner John went too. You're pretty sure

Could Have Presupposition
A counterfactual with antecedent A presupposes in a given context what is expressed by ⌜It could have been that A⌝ in that context.

Although the data is more delicate than it is with indicatives, the counterfactual compatibility presupposition is plausible. Consider:

(80) The treasure had to have been hidden in the attic.

(81) ? If it hadn't been hidden in the attic, it would have been hidden in the garden.

We can also motivate Could Have Presupposition by appeal to the tests for presupposition mentioned in §5. Consider counterfactuals embedded in yes-no questions. I ask my partner:

(82) Would you have come with me if I had taken a job in Hawaii?

If she didn't know that a job in Hawaii was on the table, it would be natural for her to respond with:

(83) I didn't know that you could have taken a job in Hawaii.

Similar things can be said about 'maybe' and 'doubt'.

(84) Maybe if I had taken a job in Hawaii, you would have come with me.

(85) I doubt you would have come with me if I had taken a job in Hawaii.

Once again, if my partner hadn't considered that I could have taken a job in Hawaii, it would be natural for her to respond with something like (83).

7 Counterfactuals and Sequence Semantics

We have a Stalnakerian semantics for counterfactuals in place. But the theory is not quite right as it stands. Recall:

Qualified Counterfactual If-to-Or
Might not A, Must (A $\Box\!\!\rightarrow$ B) $\vDash$ Couldn't have been (A and not B)

---

he didn't. You say:

(60) If John had attended too, I would have seen him.

(60) is felicitous. This suggests that we are interpreting the presupposition of the antecedent—that a salient individual (in this case Mary) attended the party—relative to an information state that contains some of our knowledge, but not all of our knowledge. For example, the information state must contain our knowledge of the fact that Mary went, but it must not contain our knowledge of the fact that John did not go.

I will argue that this is a good principle, but that as is stands, our theory does not validate it.

7.1 Counterfactual If-to-Or

To motivate Qualified Counterfactual If-to-Or, I'll begin by considering the following stronger principle.

> **Counterfactual If-to-Or**
> Must (had been A, would have been B) ⊨ Couldn't have been (A and not B)

This principle says that if a counterfactual with antecedent A and consequent B is known in a given context, then ⌜Couldn't have been (A and not B)⌝ is also true in that context. There are reasons to think this stronger principle is a good principle. Take the case of the dinner party. I am confident that leaving at five would not have left Matt enough time to get to the dinner by six. I say:

(86) It's gotta/has to/must be that if Matt had left at five, he wouldn't have made it on time.

Assume you trust me and accept (86). Then it seems you must also accept:

(87) He couldn't have left at five and made it by six.

Now suppose instead that you disagree with me. You think that an hour might have been enough time. You say:

(88) I'm not so sure. He could have left at five and made it by six.

Here you present (88) as a challenge to (86). And if I accept (88), I have to retract (86).

Consider a second example—the treasure hunt case. I say:

(89) It's gotta/has to/must be that if hadn't been in the garden, it would have been in the attic.

Suppose you trust me and accept (89). Then you must also accept:

(90) The treasure couldn't have been anywhere else.

Now suppose instead that you have reason to believe that the man who gave me the tip was not the real organizer, and that it was a fluke that the treasure happened to be in the attic. You do not accept (89) and you reply:

(91) I'm not so sure. I think the treasure could have been hidden in the kitchen.

Once again you seem to be presenting (91) as a challenge to (89), and if I accept (91), I must retract (89).

On the basis of these examples and others like them, we might be tempted to conclude that Counterfactual Or-to-If is valid. But we should not. Counterfactual If-to-Or is in tension with the principle of *Conjunctive Sufficiency*—the principle that a counterfactual with antecedent A and consequent B is true whenever A and B are true.[43] And Conjunctive Sufficiency follows from Conditional Excluded Middle together with Modus Ponens. Since our Stalnakerian semantics for counterfactuals validates both of these principles, it also validates Conjunctive Sufficiency. We have seen that there are strong reasons to accept Conditional Excluded Middle, and it goes without saying that there are strong reasons to accept Modus Ponens.

We should not say that Counterfactual If-to-Or is valid, then. But it turns out that we don't need the full strength of Counterfactual If-to-Or to account for the seeming reasonableness of inferences conforming to this principle. To see this, observe that counterfactuals cannot be felicitously asserted when their antecedents are known to be true. If you assert a conditional at all in a case like this, it must be the indicative.[44] This means that Qualified Counterfactual If-to-Or suffices to account for cases like the dinner case and the treasure hunt case, and others like them. In the dinner case, I don't know that Matt left at five when I assert (86); in the treasure hunt case, I don't know that the treasure was not in the attic when I assert (89). Qualified Counterfactual If-to-Or allows us to say that if a counterfactual ⌜If had been that A, would have been that B⌝ can be felicitously asserted in a given context, then if the counterfactual is known in that context, it follows that ⌜Couldn't have been that A and not B⌝ is true.

The problem is that, as it stands, our theory does not validate even this weaker principle. Consider a simple three-world model. In w_1, Matt does not flip the coin. In w_2, Matt flips the coin and it lands heads. In w_3, Matt flips the coin and

[43] Assume Conjunctive Sufficiency. Then ⌜Must (A and B)⌝ entails ⌜Must (A $\Box\!\!\rightarrow$ B)⌝. But ⌜Must (A and B)⌝ does not entail ⌜Couldn't have been (A and not B)⌝. So if ⌜Must (A and B)⌝ is true, then ⌜Must (A $\Box\!\!\rightarrow$ B)⌝ will also be true, but ⌜Couldn't have been (A and not B)⌝ may be false.

[44] Indicatives also tend to be inappropriate when their antecedents are not known to be true. Still there appear to be exceptions to this generalization in the case of indicatives. Consider what Dorr and Hawthorne (2013) call *echoing* uses of indicatives.

(92) As we previously established, the murder took place before noon. If the murder took place before noon, then the butler, who was on the 10 o'clock train from Paddington could not have been responsible.

Counterfactuals are not licensed even in echoing contexts.

(93) ? As we previously established, the murder took place before noon. If the murder had taken place before noon, then the butler, who was on the 10 o'clock train from Paddington could not have been responsible.

it lands tails. Assume that our information entails that Matt did not flip the coin: $\mathsf{E}(w_1) = \{w_1\}$. Let $g(i)(w_1) = \{w_1, w_2, w_3\}$. Then (94) is true in w_1.

(94) Matt might not have flipped the coin.

Conditional Excluded Middle entails that one of the following is true in w_1.

(95) If Matt had flipped the coin, it would have landed heads.

(96) If Matt had flipped the coin, it would have landed tails.

Suppose that it is (95) that is true. Then (97) is true:

(97) It's gotta/has to/must be that if Matt had flipped the coin, it would have landed heads.

But in w_3, the coin lands tails, and $w_3 \in g(i)(w_1)$. So (98) is false in w_1.

(98) It couldn't have landed tails.

7.2 The Need for Fine-Grained Contents

The problem I am describing is closely related to a familiar problem concerning the probabilities of indicative conditionals—the *wallflower problem*.[45] Here I present a simplified version of the problem.[46]

Suppose I roll a six-sided die, but I have not yet seen how it landed. You might have thought we can model this scenario with exactly six equiprobable worlds, one for each outcome of the die roll. But this won't work. To see why not, consider:

(99) If it did not land on six, it landed on one.

The probability of (99) should be 1/5. But in our simple six-world model, there is no proposition whose probability is equal to 1/5. With just six equiprobable worlds, the unconditional probability of any proposition—and so, in particular, any conditional proposition—is a multiple of 1/6.

In response to this challenge, we should say there's a flaw in our simple model. We assumed that there were only six possible worlds, but if we take conditional propositions seriously, there must be many, many more. Consider the proposition that the die landed on six. Given Conditional Excluded Middle, one of the following is true in any world where the die lands on six:

(100) If it landed on four or five, it landed on four.

(101) If it landed on four or five, it did not land on four.

[45] See Hájek (1989).
[46] My presentation of the problem and the response follows Bacon (2015).

Clearly I don't know which of these two conditionals is true if the coin lands on six. So there must be at least two epistemic possibilities compatible with die landing on six—one where (100) is true, and one where (101) is true. We can generate many more epistemic possibilities by considering other antecedents that are false when the die lands on six. In short, if we take conditional propositions seriously, we must countenance many more epistemic possibilities than our original six.

I think we should say that there is a similar flaw in our model of the coin flip. I started with just three worlds: w_1 (where the coin is not tossed), w_2 (where the coin lands heads) and w_3 (where the coin lands tails). But once again, if we're taking conditional propositions seriously, this can't be right. Consider w_1, where the coin is not tossed. Given Conditional Excluded Middle, either (95) or (96) is true in w_1.

(95) If Matt had flipped the coin, it would have landed heads.

(96) If Matt had flipped the coin, it would have landed tails.

Clearly I don't know which of (95) or (96) is true. So it seems that w_1 needs to be split into two epistemic possibilities—one where (95) is true, and one where (96) is true.

Following van Fraassen (1976), many theorists propose to model these more fine-grained possibilities with *sequences* of 'factual' worlds. To illustrate this idea, let us look at the example of rolling the die. We begin with a set of six 'factual' possibilities corresponding to the six possible outcomes of the roll. We said that there are many different ways of settling the conditional facts that are compatible with the proposition that the die lands on six. For example, it could be that if the die didn't land on six, it landed on three, or it could be that if the die didn't land six, it landed on two, and so forth. We model these possibilities with sequences such as:

$$\langle w_6, w_2, w_3, w_4, w_5, w_1 \rangle$$

This sequence represents one way of the settling all of the facts. The first world tells us how the non-conditional facts have been settled—in this case, it tells us that the die landed on six. The other worlds in the sequence tell us how the conditional facts have been settled. For example, the second world tells us what is true if we are not in the first world. Thus, this sequence tells us that if the die didn't land on one, it landed on two. The third world tells us what is true if we are not in the first or the second world. Thus, this sequence tells us that if it didn't land on one or two, it landed on three. And so on.

In the next section, I will suggest a simple generalization of van Fraassen's sequence semantics that allows us to give a uniform semantics for indicatives and counterfactuals. Roughly, I will say that an indicative conditional is true at a sequence just in case the first *epistemically possible* antecedent-world in the sequence is a consequent-world, and that a counterfactual is true at a sequence just in case the first antecedent-world *consistent with what we're holding fixed when we evaluate the counterfactual* is a consequent-world.

Before moving on, a word of warning. I will only be considering *simple conditionals* in the next section—conditionals whose antecedents and consequents do not themselves contain conditionals.[47]

7.3 A Uniform Sequence Semantics for Conditionals

We begin with a set of 'factual' worlds W. We assume for simplicity that W is finite. We will let $\mathbf{S}_W$ be the set of all permutations of W. Thus, where the elements of W represent all possible ways of settling the non-conditional facts, the elements of $\mathbf{S}_W$ represent all possible ways of settling *all* of the facts—the non-conditional facts and the conditional facts alike. Where s is a sequence in $\mathbf{S}_W$, we will write w_s for the first world in s. The semantics for non-conditional sentences is straightforward: a non-conditional sentence A is true at a sequence s just in case A is true at w_s.

We interpret epistemic modals and indicatives relative to an epistemic accessibility relation. Before we used *factual accessibility relations*—accessibility relations over W—to interpret conditionals. Now we will need to use *lifted accessibility relations*—accessibility relations over $\mathbf{S}_W$. Where E is any factual epistemic accessibility relation over W, E determines a lifted accessibility relation $\mathbf{E}$ over $\mathbf{S}_W$ as follows.

Definition of Lifted Accessibility
$\mathbf{E}(s) = \{s' \in \mathbf{S}_W : w_{s'} \in \mathsf{E}(w_s)\}$

This says that one sequence s_1 is $\mathbf{E}$-accessible from another sequence s_2 just in case the first world in s_1 is E-accessible from the first world in s_2.

We interpret the modal past using a set of accessibility relations and a variable assignment. Before we used sets of factual accessibility relations, and a variable assignment that assigned factual accessibility relations to free variables. Now that we're assuming that conditionals are true or false relative to sequences of

[47]There is a way to generalize the semantics that I am offering to all conditionals, but the details are complex and would take us too far afield.

worlds, we will need to use sets of lifted accessibility relations and variable assignments over lifted accessibility relations.

Let $\mathcal{E} = \{\mathsf{E}_1, \ldots, \mathsf{E}_n\}$ be any set of factual accessibility relations totally ordered by $\geq$. (Remember: $\mathsf{E}_1 \geq \mathsf{E}_2$ if and only if for all w, $\mathsf{E}_1(w) \subseteq \mathsf{E}_2(w)$.) $\mathcal{E}$ determines a corresponding set of lifted accessibility relations $\boldsymbol{\mathcal{E}} = \{\mathbf{E}_1, \ldots, \mathbf{E}_n\}$ where $\mathbf{E}_1$ is the lifted counterpart of E_1, $\mathbf{E}_2$ is the lifted counterpart of E_2, and so forth. I will continue to use '$\geq$' to say that one lifted accessibility relation is at least as informed as another: $\mathbf{E}_1 \geq \mathbf{E}_2$ means that for any sequence $s \in \mathbf{S}_{\mathrm{W}}$, $\mathbf{E}_1(s) \subseteq \mathbf{E}_2(s)$. (Note that I will be assuming for simplicity that epistemic accessibility relations are equivalence relations. I think we can do without this assumption, but it significantly simplifies the presentation.)

We turn now to variable assignments. Where g is a 'factual' variable assignment that assigns factual accessibility relations to free variables, g determines a *lifted variable assignment* $\boldsymbol{g}$ defined as follows: if $g(i) = \mathsf{E}_{\mathrm{i}}$, then $\boldsymbol{g}(i) = \mathbf{E}_{\mathrm{i}}$, where $\mathbf{E}_{\mathrm{i}}$ is the lifted counterpart of E_{i}.

Now that we have lifted accessibility relations and lifted variable assignments, we need to redefine our selection function. To do so, I need to introduce one piece of terminology. For any set of sequences $\mathbf{P}$, we define the *flattening* of $\mathbf{P}$ as follows.

Flattening
$\downarrow \mathbf{P} = \{w \in \mathrm{W} : w = w_s \text{ for some } s \in \mathbf{P}\}$

The flattening of $\mathbf{P}$ is the set of first worlds of the sequences in $\mathbf{P}$. We now define the selection function as follows.

$$f_{\mathbf{E}}(s, \mathbf{P}) = \begin{cases} \text{the singleton containing the first} \downarrow [\mathbf{E}(s) \cap \mathbf{P}]\text{-world in s} & \text{if} \downarrow [\mathbf{E}(s) \cap \mathbf{P}] \neq \emptyset \\ \emptyset & \text{otherwise} \end{cases}$$

We are now in a position to state the semantics for epistemic modals, conditionals, and the modal past. Let's start with modals and conditionals. Where s is a sequence, $\boldsymbol{g}$ is a lifted variable assignment, $\boldsymbol{\mathcal{E}}$ is a set of lifted accessibility relations totally ordered by $\geq$, and $\mathbf{E}$ is a lifted accessibility relation in $\boldsymbol{\mathcal{E}}$, we have:

Epistemic Modals
$[\![\text{Must A}]\!]^{s,\boldsymbol{g},\mathbf{E},\boldsymbol{\mathcal{E}}} = [\![\text{Has to A}]\!]^{s,\boldsymbol{g},\mathbf{E},\boldsymbol{\mathcal{E}}} = 1$ if and only if $\mathbf{E}(s) \subseteq [\![\text{A}]\!]^{\boldsymbol{g},\mathbf{E},\boldsymbol{\mathcal{E}}}$

Stalnaker's Semantics
$[\![\text{If A, then B}]\!]^{s,\boldsymbol{g},\mathbf{E},\boldsymbol{\mathcal{E}}} = 1$ if and only if $f_{\mathbf{E}}(s, [\![\text{A}]\!]^{\boldsymbol{g},\mathbf{E},\boldsymbol{\mathcal{E}}}) \subseteq \downarrow [\![\text{B}]\!]^{\boldsymbol{g},\mathbf{E},\boldsymbol{\mathcal{E}}}$

⌜Must A⌝ or ⌜Has to be that A⌝ is true at a sequence s just in case A is true at all sequences E-accessible from s. ⌜If A, then B⌝ is true at a sequence s if and only if either there are no A-worlds that are E-accessible from w_s or the first A-world that is E-accessible from w_s is a B-world.

Let us check that all of the principles governing the logic of indicative conditionals discussed in §5 continue to hold. First, observe that Conditional Excluded Middle is valid. Consider any sequence s. If there are A-worlds E-accessible from w_s, then either the first E-accessible A-world is a B-world or it's a not-B-world. In the first case, ⌜if A, then B⌝ is true at s; in the second, ⌜if A, then not B⌝ is true at s. If there are no A-worlds E-accessible from w_s, both conditionals are true. Second, we continue to validate Boxy Or-to-If and Might-to-Might. This is because epistemic modals and indicatives are interpreted uniformly and the selection function satisfies the Accessibility Constraint, Success, and Non-Vacuity. I leave the proofs to a footnote.[48]

Turn now to the modal past. Where s is a sequence, $\boldsymbol{g}$ is a lifted variable assignment, $\boldsymbol{\mathcal{E}}$ is a set of lifted accessibility relations totally ordered by $\geq$, and E is a lifted accessibility relation in $\boldsymbol{\mathcal{E}}$, we have:

Modal Past
$[\![\text{Past}_i\ \text{A}]\!]^{s,\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}} = 1$ if and only if $[\![\text{A}]\!]^{s,\boldsymbol{g},\boldsymbol{g}(i),\boldsymbol{\mathcal{E}}} = 1$

We continue to assume that the modal past carries a presupposition of informational precedence: ⌜Past_i A⌝ presupposes that $\boldsymbol{g}(i) < \mathsf{E}$. Combining the semantics for the modal past with our semantics for epistemic modals and indicatives, respectively, gives us:

Semantics for 'Had to'
$[\![\text{Past}_i\ (\text{Has to be that A})]\!]^{s,\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}} = 1$ if and only if $\boldsymbol{g}(i)(w) \subseteq [\![\text{A}]\!]^{\boldsymbol{g},\boldsymbol{g}(i),\boldsymbol{\mathcal{E}}}$

Stalnakerian Semantics for Counterfactuals
$[\![\text{Past}_i\ (\text{if A, then B})]\!]^{s,\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}} = 1$ if and only if $f_{\boldsymbol{g}(i)}(s, [\![\text{A}]\!]^{\boldsymbol{g},\boldsymbol{g}(i),\boldsymbol{\mathcal{E}}}) \subseteq\ \downarrow [\![\text{B}]\!]^{\boldsymbol{g},\boldsymbol{g}(i),\boldsymbol{\mathcal{E}}}$

---

[48] *Proof of Boxy Or-to-If*. Suppose $[\![\text{Must (A or B)}]\!]^{s,\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}} = 1$. By the semantics for epistemic modals, $\mathsf{E}(s) \cap [\![\text{not A}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}} \subseteq [\![\text{B}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}}$ and so $\downarrow [\mathsf{E}(s) \cap [\![\text{not A}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}}] \subseteq\ \downarrow [\![\text{B}]\!]^{\text{g},\mathsf{E},\boldsymbol{\mathcal{E}}}$. By the definition of the selection function, $f_{\mathsf{E}}(s, [\![\text{not A}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}}) \subseteq\ \downarrow [\mathsf{E}(s) \cap [\![\text{not A}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}}]$. It follows that $f_{\mathsf{E}}(s, [\![\text{not A}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}}) \subseteq\ \downarrow [\![\text{B}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}}$. By Stalnaker's Semantics, $[\![\text{if not A, then B}]\!]^{s,\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}} = 1$.

Proof of Might to Might. Suppose (1) $[\![\text{Might A}]\!]^{s,\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}} = 1$ and (2) $[\![\text{if A, then B}]\!]^{s,\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}} = 1$. Suppose, for contradiction, that (3) $[\![\text{Might B}]\!]^{s,\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}} = 0$. By the semantics for epistemic modals, it follows from (3) that $\mathsf{E}(s) \subseteq [\![\text{not B}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}}$ and therefore that $\mathsf{E}(s) \cap [\![\text{A}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}} \subseteq [\![\text{not B}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}}$. This means that $\downarrow [\mathsf{E}(s) \cap [\![\text{A}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}}] \subseteq\ \downarrow [\![\text{not B}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}}$. By the definition of the selection function, $f_{\mathsf{E}}(s, [\![\text{A}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}}) \subseteq\ \downarrow [\mathsf{E}(s) \cap [\![\text{A}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}}]$ and so it follows that $f_{\mathsf{E}}(s, [\![\text{A}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}}) \subseteq\ \downarrow [\![\text{not B}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}}$. By (2) and Stalnaker's Semantics we also have that $f_{\mathsf{E}}(s, [\![\text{A}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}}) \subseteq\ \downarrow [\![\text{B}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}}$. It follows that $f_{\mathsf{E}}(s, [\![\text{A}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}}) = \emptyset$. But that contradicts the definition of the selection function since by (1) and the semantics for epistemic modals we know that $\downarrow [\mathsf{E}(s) \cap [\![\text{A}]\!]^{\boldsymbol{g},\mathsf{E},\boldsymbol{\mathcal{E}}}] \neq \emptyset$.

This says that ⌜It had to be that A⌝ is true at a sequence just in case all $\boldsymbol{g}(i)$-accessible sequences are A-sequences, and that a counterfactual with antecedent A and consequent B is true at a sequence just in case the first $g(i)$-accessible A-world in that sequence is a B-world. Since counterfactuals are interpreted uniformly with ‘had to’ and ‘could have’, we continue to validate Counterfactual Or-to-If and Could-to-Could. I leave the proofs to a footnote.[49]

The final order of business is to check that we validate Qualified Counterfactual If-to-Or.

> **Qualified Counterfactual If-to-Or**
> Might not A, Must (had been A, would have been B) $\vDash$ Couldn’t have been that (A and not B)

I leave the proof to a footnote, but here’s an informal explanation of why Qualified Counterfactual If-to-Or is valid on our unified sequence semantics.[50] Remember that we defined the epistemically possible sequences as the set of all permutations of W beginning with a world that is compatible with our factual information. Suppose our factual information leaves open some world where A and B are true. Then that world will be the first A-world compatible with our our factual information in some epistemically possible sequence. As a result, the indicative ⌜if A, then B⌝ will be epistemically possible. Likewise, suppose our factual information leaves open some world where A is true and B is false. Then that world will be the first A-world compatible with our factual information in some epistemically possible sequence. As a result, the indicative ⌜if A, then not B⌝ will

[49] *Proof of Counterfactual Or-to-If.* Suppose $\llbracket \text{Past}_\text{i} \text{ (Has to (A or B))} \rrbracket^{s,\boldsymbol{g},\mathsf{E},\mathcal{E}} = 1$. By the semantics for modal past, $\llbracket \text{Has to (A or B)} \rrbracket^{s,\boldsymbol{g},\boldsymbol{g}(i)\mathcal{E}} = \llbracket \text{Must (A or B)} \rrbracket^{s,\boldsymbol{g},\boldsymbol{g}(i)\mathcal{E}} = 1$. By Boxy Or-to-If, $\llbracket \text{If not A, then B} \rrbracket^{s,\boldsymbol{g},\boldsymbol{g}(i),\mathcal{E}} = 1$. By the semantics for modal past, it follows that $\llbracket \text{Past}_\text{i} \text{ (if not A, then B)} \rrbracket^{s,\boldsymbol{g},\mathsf{E},\mathcal{E}} = 1$.

Proof of Could-to-Could. Suppose (1) $\llbracket \text{Past}_\text{i} \text{ (Could be that A)} \rrbracket^{s,\boldsymbol{g},\mathsf{E},\mathcal{E}} = 1$ and (2) $\llbracket \text{Past}_\text{i} \text{ (If A, then B)} \rrbracket^{s,\boldsymbol{g},\mathsf{E},\mathcal{E}} = 1$. By the semantics for modal past, $\llbracket \text{Could be that A} \rrbracket^{s,\boldsymbol{g},\boldsymbol{g}(i),\mathcal{E}} = \llbracket \text{Might A} \rrbracket^{s,\boldsymbol{g},\boldsymbol{g}(i),\mathcal{E}} = 1$ and $\llbracket \text{If A, then B} \rrbracket^{s,\boldsymbol{g},\boldsymbol{g}(i),\mathcal{E}} = 1$. By Might-to-Might, $\llbracket \text{Might B} \rrbracket^{s,\boldsymbol{g},\boldsymbol{g}(i),\mathcal{E}} = \llbracket \text{Could be that B} \rrbracket^{s,\boldsymbol{g},\boldsymbol{g}(i),\mathcal{E}} = 1$. By the semantics for modal past, it follows that $\llbracket \text{Past}_\text{i} \text{ (Could be that B)} \rrbracket^{s,\boldsymbol{g},\mathsf{E},\mathcal{E}} = 1$.

[50] *Proof.* Suppose (1) $\llbracket \text{Might not A} \rrbracket^{s,\boldsymbol{g},\mathsf{E},\mathcal{E}} = 1$ and (2) $\llbracket \text{Must (Past}_\text{i} \text{ (if A, then B))} \rrbracket^{s,\boldsymbol{g},\mathsf{E},\mathcal{E}} = 1$. Suppose, for contradiction, that (3) $\llbracket \text{Past}_\text{i} \text{ (Could (A and not B))} \rrbracket^{s,\boldsymbol{g},\mathsf{E},\mathcal{E}} = 1$. By (3) and the semantics for modal past, there’s an $s_1 \in \boldsymbol{g}(i)(s)$ such that $\llbracket \text{A and not B} \rrbracket^{s_1,\boldsymbol{g},\boldsymbol{g}(i),\mathcal{E}} = 1$. It follows that $w_{s_1} \in\downarrow [\boldsymbol{g}(i)(s) \cap \llbracket \text{A} \rrbracket^{\boldsymbol{g},\boldsymbol{g}(i),\mathcal{E}}]$ and $w_{s_1} \in\downarrow \llbracket \text{not B} \rrbracket^{\boldsymbol{g},\boldsymbol{g}(i),\mathcal{E}}$. Note that $\mathsf{E}(s)$ is the set of all permutations of W beginning with a world in $\mathsf{E}(w_s)$. Since $\llbracket \text{Might not A} \rrbracket^{s,\boldsymbol{g},\mathsf{E},\mathcal{E}} = 1$, there’s an $s_2 \in \mathsf{E}(s)$ such that (a) $w_{s_2} \notin\downarrow \llbracket \text{A} \rrbracket^{\boldsymbol{g},\boldsymbol{g}(i),\mathcal{E}}$ and (b) w_{s_1} is the second world in s_2. Since $w_{s_1} \in\downarrow [\boldsymbol{g}(i)(s) \cap \llbracket \text{A} \rrbracket^{\boldsymbol{g},\boldsymbol{g}(i),\mathcal{E}}]$, and $\boldsymbol{g}(i)$ is an equivalence relation, $w_{s_1} \in\downarrow [\boldsymbol{g}(i)(s_2) \cap \llbracket \text{A} \rrbracket^{\boldsymbol{g},\boldsymbol{g}(i),\mathcal{E}}]$. And since w_{s_1} is the second world in s_2 and $w_{s_2} \notin \llbracket \text{A} \rrbracket^{\boldsymbol{g},\boldsymbol{g}(i),\mathcal{E}}$, w_{s_1} is the first $\downarrow [\boldsymbol{g}(i)(s_2) \cap \llbracket \text{A} \rrbracket^{\boldsymbol{g},\boldsymbol{g}(i),\mathcal{E}}]$-world in s_2. So $f_{\boldsymbol{g}(i)}(s_2, \llbracket \text{A} \rrbracket^{\boldsymbol{g},\boldsymbol{g}(i),\mathcal{E}}) = \{w_{s_2}\}$. Since $w_{s_2} \in \llbracket \text{not B} \rrbracket^{g,g(i),\mathcal{E}}$, $f_{g(i)}(s_2, \llbracket \text{A} \rrbracket^{\mathbf{g},g(i),\mathcal{E}}) \not\subseteq \llbracket \text{B} \rrbracket^{\boldsymbol{g},\boldsymbol{g}(i),\mathcal{E}}$. By Stalnaker’s Semantics $\llbracket \text{If A, then B} \rrbracket^{s_2,\boldsymbol{g},\boldsymbol{g}(i),\mathcal{E}} = 0$. By the semantics for the modal past, $\llbracket \text{Past}_\text{i} \text{ (if A, then B)} \rrbracket^{s_2,\boldsymbol{g},\mathsf{E},\mathcal{E}} = 0$. By the semantics for epistemic modals, $\llbracket \text{Must (Past}_\text{i} \text{ (if A, then B))} \rrbracket^{s,\boldsymbol{g},\mathsf{E},\mathcal{E}} = 0$. But that contradicts (2).

be epistemically possible. In this way, all ways of settling the indicative conditional facts that are compatible with our factual information are represented as epistemically possible.

Similar things can be said about counterfactuals. Suppose that the factual information *that we're holding fixed for the purposes of evaluating the counterfactual* leaves open some world w_1 where A and B are true; that is to say, suppose that ⌜It could have been that A and B⌝ is true in our context. Then so long as we don't know A, w_1 will be the first A-world compatible with the factual information that we're holding fixed at some epistemically possible sequence. As a result, the counterfactual ⌜if had been that A, would have been that B⌝ will be epistemically possible. Likewise, suppose that the factual information *that we're holding fixed for the purposes of evaluating the counterfactual* leaves open some world w_2 where A is true and B is false; that is to say, suppose that ⌜It could have been that A and not B⌝ is true in our context. Then, so long as we don't know A, w_2 will be the first A-world compatible with the factual information that we're holding fixed at some epistemically possible sequence. As a result, the counterfactual ⌜if had been that A, would have been that not B⌝ will be epistemically possible. In this way, all ways of settling the counterfactual conditional facts that are compatible with the factual information that we're holding are represented as epistemically possible.

Take the case of the coin flip. I know that Matt did not toss the coin, and I hold this knowledge fixed when I evaluate what would have happened if Matt had tossed the coin. I also hold fixed my knowledge of the fact that the coin was fair. This means that (102) and (103) are both true:

(102) The could have landed heads.

(103) The coin could have landed tails.

That is to say, there are worlds in $g(i)$ where the coin is tossed and it lands heads, and there are worlds in $g(i)$ where the coin is tossed and it lands tails. It follows that there are epistemically possible sequences whose first $g(i)$-accessible world is a world where the coin is tossed and lands heads and there are epistemically possible sequences whose first $g(i)$-accessible world is a world where the coin is tossed and lands tails. And so we rightly predict that I don't know either of the following counterfactuals.

(95) If Matt had flipped the coin, it would have landed heads.

(96) If Matt had flipped the coin, it would have landed tails.

7.4 Pseudo-Validity and the Inferences

We now have in place a Stalnakerian theory of counterfactuals that validates both Counterfactual Or-to-If and Qualified Counterfactual If-to-Or. As we saw at the beginning of §6, this means that our theory predicts that Transitivity, Contraposition, and Antecedent Strengthening are pseudo-valid. We can use this fact to explain the good-making features of these three inference patterns for counterfactuals.

Start with Transitivity. Consider an example from §5. Suppose my friend is having a party tonight. I say:

(104) If John had gone to the party, Matt would have gone.

(105) If Matt had gone to the party, David would have gone.

If you trust me, then you will and you should conclude:

(106) If John had gone to the party, David would have gone.

We can explain why this inference seems compelling in terms of pseudo-validity. Suppose I felicitously assert (104) and (105) and that I say something true. Suppose that you accept my assertions. Then (104) and (105) become part of our shared information—(107) and (108) will be true in our context.

(107) It's gotta/has to/must be that if John had gone to the party, Matt would have gone.

(108) It's gotta/has to/must be that if Matt had gone to the party, David would have gone.

Since Transitivity is pseudo-valid, it follows from these facts that (106) is true. Thus, when I assert (106) I say something that is true in our context.

Turn to Antecedent Strengthening. Just as we saw with indicatives, there is a contrast between forward and reverse Sobel sequences for counterfactuals. I will explain this contrast in terms of pseudo-validity. And once again, the explanation is similar to the strict theorist's. Forward Sobel sequences are often felicitous because the context tends to change midway through the sequence. Consider:

(109) If Alice had come to the party, it would have been fun.

(110) But if Alice and Bob had come to the party, it wouldn't have been fun.

Suppose I assert (109). If I say something true, and if you accept my assertion, then all worlds compatible with what we're holding fixed where Alice goes to the party are ones where the party is fun. If Bob does not go to the party in any of these worlds, the presupposition of (110) is not satisfied, and so asserting (110) will tend to force the context to change. Reverse Sobel sequences are typically infelicitous because the context tends not to change when the sentences are uttered in reverse order. Consider any context in which (109) can be felicitously asserted. If Antecedent Strengthening is pseudo-valid, then (109) and (110) cannot both be entailed by our information in any such context. Suppose, then, that I assert (110) in such a context and that I say something true. If you accept my assertion, then the presupposition of (110) will ordinarily be satisfied, and so the presupposition of (109) will also be satisfied. You have no reason to choose a different accessibility relation to interpret (109). Therefore, it will usually not be appropriate for me to continue by asserting (109).

Contraposition is admittedly less natural for counterfactuals than it is for indicatives. Nevertheless, as Goodman (1947) observes, something *like* Contraposition does appear valid for counterfactuals. Here's one of Goodman's examples.

(111) If the piece of butter had been heated to 150°, it would have melted.

(112) So, since the piece of butter didn't melt, it wasn't heated to 150°.

Goodman uses 'since' but 'if' could also have been used:

(113) So, if the piece of butter didn't melt, it wasn't heated to 150°.

On the theory of counterfactuals and indicatives that I have offered, this inference is pseudo-valid. To see this, suppose that (114) is true and licensed.

(114) It's gotta/has to/must be that if the piece of butter had been heated to 150°, it would have melted.

(114) entails:

(115) It's gotta/has to/must be that either the butter melted or it was not heated to 150°.

By Boxy Or-to-If, (115) entails:

(116) If the butter didn't melt, it was not heated to 150°.

8 Conclusion

We have seen that our ordinary credences in conditionals are far out of line with what the strict theory of conditionals recommends. We have also seen that there's a variably strict theory that is capable of making reasonable predictions about our credences in conditionals: Stalnaker's variably strict theory. In light of this, I recommended rejecting the strict theory in favor of a Stalnakerian variably strict theory. Still, it cannot be denied that certain characteristically strict inference patterns—such as Transitivity, Contraposition, and Antecedent Strengthening—share certain good-making features with classically valid arguments. Transitivity and Contraposition seem like excellent principles, and strict theorists have convincingly argued that the full range of data surrounding Antecedent Strengthening is well explained by contextualist or dynamic strict theories. In this paper, I developed a Stalnakerian variably strict theory on which the characteristically strict inference patterns are pseudo-valid. In the first half of the paper, I showed that the pseudo-validity of these inference patterns for *indicative conditionals* falls out of Stalnaker's theory. This means that proponents of Stalnaker's theory have at their disposal a natural explanation of the apparent validity of Transitivity and Contraposition for indicatives, as well as a compelling account of the asymmetry between forward and reverse indicative Sobel sequences. In the second half of the paper, I developed a new Stalnakerian theory of counterfactuals—building on the work of Schulz (2014) and van Fraassen (1976)—and I showed that the theory predicts that our three inference patterns are pseudo-valid for counterfactuals.

9 References

Adams, Ernest (1975). *The Logic of Conditionals*. Dordrecht, Holland: D. Reidel.

Bacon, Andrew (2015). "Stalnaker's thesis in context." *Review of Symbolic Logic*, 8 (1):131-163.

Bledin, Justin (2014). "Logic Informed." *Mind*, 123 (490): 277-316.

Boylan, David (2020). "What the Future 'Might' Brings," *Mind*, 129 (515): 809-829.

Cariani, Fabrizio & Goldstein, Simon (2018). "Conditional Heresies." *Philosophy and Phenomenological Research*, (2):251-282.

Chierchia, Gennaro & Sally McConnell-Ginet (2000). *Meaning and Grammar: An introduction to semantics*, 2nd edition. Cambridge, MA: MIT Press.

Dorr, Cian (2016). "Against Counterfactual Miracles." *Philosophical Review*, 125 (2): 241-286.

Dorr, Cian & John Hawthorne (2018). *If...: A theory of conditionals*. Manuscript.

Edgington, Dorothy (1995). "On conditionals." *Mind*, 104 (414):235–329.

Gillies, Thony (2007). "Counterfactual Scorekeeping." *Linguistics and Philosophy*, 30 (3): 329–360.

Gillies, Thony (2004). "Epistemic Conditionals and Conditional Epistemics." *Noûs*, 30 (3): 329–360.

Goodman, Nelson (1947). "The problem of counterfactual conditionals." *Journal of Philosophy*, 44 (5):113-128.

Hacquard, Valentine (2006). *Aspects of Modality*. PhD thesis, MIT.

Hájek, Alan (1989). "Probabilities of conditionals—revisited." *Journal of Philosophical Logic*, 18 (4):423 - 428.

Heim, Irene (1992). "Presupposition projection and the semantics of attitude verbs." *Journal of Semantics*, 9(3): 183–221

Heim, Irene (1994). "Comments on Abusch's Theory of Tense." In Hans Kamp (ed.), *Ellipses, Tense, and Questions*: 143-170

Herzberger, Hans (1973). "Dimensions of truth." *Journal of Philosophical Logic*, 2 (4): 535 - 556.

Holguín, Ben (2021). "Indicative Conditionals Without Iterative Epistemology." *Noûs*, 55 (3): 560–80.

Karttunen, Lauri & Stanley Peters (1979). "Conventional implicature." *Syntax and Semantics*, 11: 1–56.

Ippolito, Michela (2013). 'Subjunctive Conditionals: A Linguistic Analysis.' *Linguistic Inquiry Monograph* (Series 65), Cambridge: MIT Press.

Kaufmann, Stefan (2004). 'Conditioning against the grain: Abduction and indicative conditionals.' *Journal of Philosophical Logic* 33, pp. 583-606.

Kaufmann, Stefan (2009). 'Conditionals right and left: Probabilities for the whole family.' *Journal of Philosophical Logic* 38, pp. 1-53.

Khoo, Justin (2015). 'On Indicative and Subjunctive Conditionals.' *Philosopher's Imprint* 15, pp. 1-40.

Khoo, Justin (2022). 'The meaning of *If*.' Oxford University Press.

Lewis, David (1973). *Counterfactuals*. Basil Blackwell Ltd., Malden, MA.

Lewis, David (1979). "Counterfactual Dependence and Time's Arrow." *Noûs*, 13 (4): 455-476.

Mackay, John (2015). "Actuality and Fake Tense in Conditionals." *Semantics and Pragmatics*, 8 (12): 1-12

Mackay, John (2019). "Modal Interpretation of Tense in Subjunctive Conditionals." *Semantics and Pragmatics*, 12 (2)

Mandelkern, Matthew (2018). "Talking about worlds." *Philosophical Perspectives*, 32 (1): 298-325.

Mandelkern, Matthew (2019). "Bounded Modality." *Philosophical Review*, 128 (1): 1-61.

Mandelkern, Matthew (2022). "Bounds." Manuscript

McGee, Vann (1989). 'Conditional probabilities and compounds of conditionals.' *Philosophical Review* 98, 485-541.

Partee, Barbara (1973). "Some structural analogies between tenses and pronouns in English." *Journal of Philosophy*, 70 (18): 601-609.

Santorio, Paolo (2022). "Path Semantics for Indicative Conditionals." *Mind*, 131 (521): 59-98.

Schulz, Katrin (2014). 'Fake tense in conditional sentences: A modal approach.' *Natural Language Semantics* 22, pp. 117 - 144.

Schulz, Moritz (2017). *Counterfactuals and Probability*. Oxford: Oxford University Press.

Stalnaker, Robert (1968). 'A Theory of Conditionals.' In Nicholas Rescher (ed.) *Studies in Logical Theory*(American Philosophical Quarterly Monographs 2), Oxford: Blackwell. pp. 98-112.

Stalnaker, Robert (1975). 'Indicative conditionals.' *Philosophia* 5, pp. 269-286.

von Fintel, Kai (1998). 'The presupposition of subjunctive conditionals.' In Orin Percus & Uli Sauerland (eds.), *The interpretive tract*, Vol. 25, pp. 25- 44. Cambridge, Massachusetts: MIT Working Papers in Linguistics.

von Fintel, Kai (2001). "Counterfactuals in a Dynamic Context." In Ken Hale: A Life in Language, Michael Kenstowicz, editor. MIT Press, Cambridge.

van Fraassen, Bas C. (1976). 'Probabilities of conditionals.' In *Foundations of probability theory, statistical inference, and statistical theories of science*, Springer, pp. 261 - 308.

Veltman, Frank (1985). "Logics for Conditionals." Doctoral Dissertation, University of Amsterdam.

Warmbrod, Ken (1981). "Counterfactuals and Substitution of Equivalent Antecedents." *Journal of Philosophical Logic*, 10: 267–289.

Williamson, Timothy (2020). *Suppose and Tell: The Semantics and Heuristics of Conditionals*. Oxford, England: Oxford University Press.