Nonclassical counterfactuals

Paolo Santorio

March 24, 2018

(Very rough draft)[1]

For NYU Mind and Language readers: This is a first draft; despite my best efforts, I expect that there will be typos and some amount of imprecision/unclarity. I also suspect that there might be ways of strengthening the main result in §3 that I haven't seen yet. There is some overlap with the main paper that I sent along (CEM in Expressivist Semantics). The new material is concentrated in section 3, which shows how a triviality result for epistemic modals can be generalized in a straightforward way to counterfactuals. Feel free to focus on that section.

1 Introduction

Kratzer's theory of modality (Kratzer 1977, 1981a, 2012) has been the standard framework for modeling the meaning of modals and conditionals in natural language for forty years. One striking feature of this theory is its generality. The variation between modals and conditionals of different flavors is reduced to differences in the settings of a small number of parameters, while the basic logical scaffolding is shared by all modal items. As a result, the semantics of modality in natural language is distilled down to the combined effect of a few basic logical features, together with contextual input.

Over the past twenty years, however, Kratzer's framework has been under attack. The challenges have concentrated mostly on one particular area of modal discourse, namely epistemic modality. The challenges divide into two broad classes. The first concerns modal and conditional logic. Several writers have argued that the logic of epistemic discourse is importantly different from standard logics and requires a nonclassical treatment. For example, Yalcin 2007 argues that sentences like (1) are semantically inconsistent:

(1) It's not raining and it might be raining.

Unfortunately, this inconsistency cannot be vindicated on any classical logic for modality, unless we allow for the unacceptable logical equivalence of ϕ and *might* ϕ. Conversely, as Yalcin shows (building on classical work in dynamic semantics, in particular Veltman 1996), we can vindicate the inconsistency of (1) on nonclassical semantic

[1]Thanks to Fabrizio Cariani and Simon Goldstein for discussion and very useful feedback.

systems, which depart from truth-conditional semantics and give rise to nonclassical notions of consequence.

The second class of challenges concerns the relationship between modality and probability. A number of formal results (starting with Lewis's 1976 so called triviality results) show that, on a truth-conditional view, it is impossible to assign probabilities to conditionals in a way that conforms to very basic constraints. Results of this sort have been strengthened in a number of ways (see e.g. Bradley 2000, 2007, who derives them via extremely weak assumptions) and have been recently extended to all modals in the epistemic domain (see e.g. Russell & Hawthorne 2016, Goldstein forthcoming).

The result of these challenges has been the emergence of several non-truth-conditional theories of epistemic modality. These approaches differ in a number of respects, but they agree on an overall view of epistemic modal discourse, which for convenience I will refer to as the 'informational view'. On the informational view, the primary function of epistemic discourse is specifying constraints on information states. One upshot of this is that epistemically modalized claims don't have ordinary truth conditions, and their logic, though quite well-behaved, is nonclassical.

This leaves us with a split picture of modality in natural language. Epistemic modals (and perhaps modals of related flavors) are given a nonclassical analysis. Conversely, there is still broad consensus that other flavors of modality—including counterfactuals, historical modals like *will*, and circumstantial modals—are adequately treated in a classical truth-conditional framework. Even theories of these modals that deviate from this consensus (such as the dynamic accounts of counterfactuals defended by von Fintel 2001 and Gillies 2007) are motivated in very different ways, and have different features from informational semantics for epistemic modality.

This short paper is a contribution to a general argument that this split model is incorrect, and that we should recover a unitary view of modal discourse—though within a framework that is different from the classical Kratzerian framework. In particular, I focus on counterfactual modality. I show that the standard challenges to truth-conditional accounts of epistemic modality and conditionals generalize fully to counterfactuals. In particular, I argue that: (i) the logic of counterfactual conditionals is nonclassical in ways that parallel the nonclassicality of the logic of epistemic discourse; (ii) a triviality result that parallels triviality results that hold for epistemic modals also hold for counterfactuals. §2 argues for the first claim, and §3 argues for the second.

Throughout the paper, I represent *would*-counterfactuals with the conditional corner $>$ and *might*-counterfactuals with the usual diamondarrow '$\Diamond\!\!\rightarrow$' symbol.

2 The nonclassicality of counterfactuals: conditional logic

First, I argue that the logic of counterfactuals is nonclassical in ways that parallel the nonclassicality of the logic of epistemic modality. I start from a now-standard puzzle in the the literature on epistemic modals, and show that that puzzle generalizes to counterfactuals.

2.1 Epistemic contradictions

Yalcin (2007) points out that $\neg\phi$ and $\Diamond\phi$ seem to be inconsistent by the lights of some plausible tests. In particular, their conjunctions are infelicitous both when asserted by themselves and in embeddings:

(2) # It's not raining and it might be raining.

(3) # Suppose that it's not raining and it might be raining.

(4) # If it's raining and it might not be raining...

From a logical point of view, Yalcin's puzzle is triggered by the tension between the principle suggested by (2)–(4), on the one hand, and the requirement that $\Diamond\phi$ should not be veridical, i.e. that it should not entail ϕ, on the other.

Epistemic Contradiction. $\neg\phi \wedge \Diamond\phi \vDash \bot$

Nonfactivity of Epistemic Modality. $\Diamond\phi \nvDash \phi$

Both principles seem plausible, yet on a classical notion of logical consequence they are inconsistent.[2] Yalcin's suggested solution is to move to a nonclassical semantics, which generates a nonclassical notion of consequence that can accommodate both principles.

In this section, I argue that there is a natural generalization of Yalcin's argument to counterfactuals. This generalization is based on a standard puzzle about counterfactual logic. Counterfactuals seem to conform to two logical requirements: Conditional Excluded Middle and the incompatibility between certain *would*- and *might*-counterfactuals. Unfortunately, these requirements are jointly incompatible on a classical notion of consequence. I show that the puzzle created by these two requirements is naturally framed as a generalization of the problem posed by Yalcin.

2.2 Conditional Excluded Middle

The first principle at stake is Conditional Excluded Middle (below).

Conditional Excluded Middle. (CEM) $\vDash (\phi > \psi) \vee (\phi > \neg\psi)$

The literature includes an impressive battery of arguments in favor of CEM. Here I focus on a classical compositional argument: counterfactuals appear to be scopeless with respect to a number of logical operators. This effect is expected on semantics that vindicate CEM, but puzzling on semantics that treat counterfactuals as having quantificational force (like the accounts in Lewis 1973a,b; Kratzer 1981b).

[2]Proof:

i.	$\Diamond\phi$	Assumption
ii.	$\quad\neg\phi$	Supposition for conditional proof
iii.	$\quad\neg\phi \wedge \Diamond\phi$	(i, ii, $\wedge$-Introduction)
iv.	$\quad\bot$	(iii, Epistemic Contradiction)
v.	$\neg\neg\phi$	(ii-iv, Reductio)
vi.	ϕ	(v, propositional logic)

The point is obvious with respect to negation. Negation can be imported inside and outside the scope of a conditional without affecting truth conditions. The sentences in (5) are equivalent.

(5) a. It's not the case that, if Frida had taken the exam, she would have passed.
b. If Frida had taken the exam, she would not have passed.

Notice also that the phenomenon persists with items that lexicalize negation, like *doubt* ($\approx$ *believe not*) and *fail* ($\approx$ *not pass*).

(6) a. I doubt that, if Frida had taken the exam, she would have passed.
b. I believe that, if Frida had taken the exam, she would have failed.

The scopelessness of counterfactuals is not confined to negation, but generalizes to a number of other operators. Scopelessness arguments have been run for quantifiers (Higginbotham 1986, von Fintel & Iatridou 2002), the adverb *only* (von Fintel 1997), and comparative constructions (Korzukhin 2014). Here I discuss the case of quantifiers.

The observation that counterfactuals are scopeless with respect to quantifiers was originally made by Stalnaker (1981, 1984). Stalnaker's argument exploits a comparison between these two dialogs:

(7) A: President Carter has to appoint a woman to the Supreme Court.
B: Who do you think he has to appoint?
A: He doesn't have to appoint any particular woman; he just has to appoint some woman or other.

(8) A: President Carter would have appointed a woman to the Supreme Court last year if there had been a vacancy.
B: Who do you think he would have appointed?
A: # He wouldn't have appointed any particular woman; he just would have appointed some woman or other.

The felicitousness of A's response in (7) suggests that indefinites like *a woman* seem to have a *de dicto* reading under *has to*. This is exactly what we expect on a standard quantificational semantics for modals. Conversely, a similar reading is unavailable under *would*. This suggests that *would* is scopeless with respect to indefinites.[3]

To reinforce Stalnaker's observation, let me point out that the lack of scope interactions concerns specifically *would*, and not counterfactual modality in general. For example, counterfactual *might* does display ordinary scope interactions with indefinites: consider the following variants on (8). (Assume that (9) and (10) are evaluated in a situation on which only individuals who identify as women or men are possible candidates for appointment.)

[3] The data can be generalized to other determiners: some examples include proportional determiners like *most*, and numeric determiners like *three*. This blocks attempts at explaining data like (7) and (8) as due to the peculiar scope properties of indefinites (see Schwarz 2011 for an overview).

(9) If there had been a Supreme Court vacancy, not even this president might have appointed a man. Though, of course, he might have still failed to appoint any particular woman.

(10) # If there had been a Supreme Court vacancy, even this president would have appointed a woman. Though, of course, he would not have appointed any particular woman.

(9) has a coherent reading, while (10) does not. This suggests that *might*- and *would* have different scopal interactions with indefinites.[4]

2.3 *would-might* incompatibility

Would- and *might*-counterfactuals of the form $\phi > \neg\psi$ and $\phi \diamondto \psi$ seem incompatible. Discourses that involve counterfactuals of both forms are standardly heard as inconsistent; moreover, pairs of counterfactuals of this form can be used to generate disagreement.

(11) # If Maria had passed, Frida would not have passed; but, even if Maria had passed, Frida might have passed.

(12) A: If Maria had passed, Frida would not have passed.
B: I disagree. Even if Maria had passed, Frida might not have passed.

This infelicity persists also in linguistic environments that screen off pragmatic clashes, like supposition contexts (see Yalcin 2007).

(13) # Suppose that, if Maria had passed, Frida would not have passed, and that, if Maria had passed, Frida might have passed.

Thus there is a straightforward empirical case for the following:

Would-Might Contradiction. (WMC) $(\phi > \neg\psi) \wedge (\phi \diamondto \psi) \vDash \bot$

One *caveat*: when presenting his theory of counterfactuals, Lewis (1973a) doesn't directly endorse WMC. Rather, he treats *would* and *might*-counterfactuals as duals.

Duality $(\phi > \psi) \Dashv\vDash \neg(\phi \diamondto \neg\psi)$

In Lewis's framework, WMC and Duality are equivalent (given the uncontroversial entailment from $\phi > \psi$ to $\phi \diamondto \psi$). But, the two can come apart (see e.g. the semantics for indicatives in Santorio 2018).

[4] This point sets apart counterfactual and epistemic conditionals. Epistemic modals of any quantificational force seem to display a kind of scopelessness with respect to quantifiers. To see this, notice that e.g. (i) doesn't appear to have two readings, despite the fact that it involves two scope-taking items.

(i) Maria must have talked to most semanticists at the conference.

Hence arguments to the effect that epistemic conditionals are scopeless are more involved (see Santorio 2018 for some of them) and cannot pass simply through their interaction with quantifiers.

2.4 Generalizing Yalcin's puzzle

Unfortunately, in all classical counterfactual logics, the combination of CEM and WMC allows us to prove that $\phi > \psi$ and $\phi \diamondrightarrow \psi$ are equivalent. The direction $\ulcorner \phi > \psi \vDash \phi \diamondrightarrow \psi \urcorner$ is uncontroversial. Here is a proof of the other direction:

i.	$\phi \diamondrightarrow \psi$	Assumption
ii.	$\quad \phi > \neg\psi$	Supposition for conditional proof
iii.	$\quad \phi > \neg\psi \wedge \phi \diamondrightarrow \psi$	(i, ii, $\wedge$-Introduction)
iv.	$\quad \bot$	(iii, IMC)
v.	$\neg(\phi > \neg\psi)$	(ii-iv, Reductio)
vi.	$\phi > \psi$	(v, CEM, Disjunctive syllogism)

Hence, if we endorsed both CEM and ICM in a classical truth-conditional semantics, the following two statements (say) would be declared equivalent:

(14) a. If Maria had passed, Frida would have passed.

b. If Maria had passed, Frida might have passed.

This result, of course, is unacceptable. For this reason, all classical counterfactual semantics uphold one of CEM and IMC and discard the other. Famously, Stalnaker retains CEM and rejects the incompatibility of *would*- and *might*-conditionals. Conversely, Lewis rejects CEM and validates Duality. Given the data above, either of these choices is costly from an empirical point of view.

For current purposes, I want to highlight that the puzzle generated by CEM and IMC is a generalization of Yalcin's puzzle. Here is a compact characterization of the problem: the puzzle is due to the tension between three plausible but jointly inconsistent principles.

Conditional Excluded Middle. (CEM)	$\vDash (\phi > \psi) \vee (\phi > \neg\psi)$
Would-Might Contradiction. (WMC)	$(\phi > \neg\psi) \wedge (\phi \diamondrightarrow \psi) \vDash \bot$
Nonfactivity of Might-Conditionals. (NMC)	$\phi \diamondrightarrow \psi \nvDash \phi > \psi$

On this characterization, the analogy with Yalcin's puzzle is glaring. WMC and NMC are conditional counterparts of, respectively, Yalcin's Epistemic Contradiction and Nonfactivity principles. In the nonconditional case, the two principles are sufficient to generate inconsistency. In the conditional case, we also need CEM to draw a crucial inference from $\neg(\phi > \psi)$ to $(\phi > \neg\psi)$. But, aside from the extra assumption, one can run exactly parallel proofs (see above and fn. 2).

Hence I suggest that we should reframe the classical Stalnaker/Lewis debate about CEM and *might*-conditionals. That debate is standardly taken to concern the quantificational force of conditionals, and (relatedly) the nature of the closeness relation involved in the semantics of conditionals. All throughout the debate, a classical notion of consequence is presupposed. I suggest instead that that debate concerns a basic tension for classical semantics and logics for conditionals.

3 The nonclassicality of counterfactuals: triviality

A second, standard line of argument against truth-conditional treatments of epistemic conditionals comes from probability. A number of results, starting from Lewis's (1976) first triviality results, show that it is impossible to assign probabilities to epistemic conditionals while respecting some very plausible constraints.[5] These results have been recently generalized, beyond conditionals, to a number of modal operators (Russell & Hawthorne 2016, Goldstein forthcoming).

Many (if not all) of these results are linked to the nonclassical features of the logic of epistemic modality and conditionals. For example, Lewis's original triviality result in (1976) relies on the inference from ψ to $\phi > \psi$ (where '>' is interpreted as the epistemic conditional), which is valid on informational notions of consequence. Now, in §2 I argued that the logic of counterfactuals is nonclassical in ways that are similar to the nonclassical features of epistemic modality. So we might expect that triviality might be generated, exploiting the nonclassical features of the logic, also for counterfactuals. This section argues that this is exactly what we find.

Starting from intuitive probabilistic principles, I show that the we can prove that, for any probability function that models rational credence, the probability of a *would*-counterfactual equals the probability of the corresponding *might*-counterfactual.

$$\textbf{Probabilistic Collapse.} \quad Pr(\phi > \psi) = Pr(\phi \diamondto \psi)$$

The result is a kind of triviality result. In principle it can be generalized to all conditionals; but it is of particular interest for counterfactuals, for which there is no uncontroversial triviality proof so far.[6]

3.1 A collapse result for epistemic *might*

I start by reviewing a triviality result for epistemic possibility modals originally presented by Russell & Hawthorne 2016, in the context of providing a battery of triviality results for epistemic modals. I then show how this result can be strengthened into a collapse result that shows that ϕ and $\Diamond\phi$ have the same probability. This collapse result is the nonconditional analog of the collapse result for counterfactuals that I present in the next section.

Russell and Hawthorne start from the following intuitive principle:

> **Might.** For any probability function *Pr* that models rational credence, if $Pr(\Diamond\phi) > 0$, then:
>
> $$Pr(\phi \mid \Diamond\, \phi) > 0$$

[5]See Hájek & Hall 1994 for an overview of the literature on triviality up to the mid-90s. See also Van Fraassen 1976 Kaufmann 2009, Bacon 2015 (among others) for attempts at producing a semantics for conditionals that overcomes these results.

[6]There are only a few attempts at deriving triviality for counterfactuals (Williams 2012, Briggs 2017), and there is no agreement that any of them is successful. For relevant criticism, of Williams' and Briggs' results, see Schwarz 2016.

Russell and Hawthorne show that, starting from **Might** and from the assumption that the class of rational credence functions is closed under conditionalization, we can prove that $\neg\phi$ and $\Diamond\phi$ are incompatible. Suppose for *reductio* that there is a probability function *Pr* on which $\neg\phi$ and $\Diamond\phi$ are compatible. Take $Pr_{\neg\phi}$, i.e. the result of conditionalizing *Pr* on $\neg\phi$. Via **Might**, we have:

i. $Pr_{\neg\phi}(\phi \mid \Diamond\, \phi) > 0$

By the definition of conditional probability, this means:

ii. $Pr(\phi \mid \Diamond\, \phi \wedge \neg\phi) > 0$

But via the probability calculus, we also have:

iii. $Pr(\phi \mid \neg\phi) = 0$

Lines (ii) and (iii) are inconsistent. So we conclude that there is no probability function on which $\neg\phi$ and $\Diamond\phi$ are compatible after all.

Russell and Hawthorne's result can be easily strengthened into a collapse result. For this step, assume the following principle:

> **Not-might.** For any probability function *Pr* that models rational credence:
>
> $$\text{if } Pr(\neg\Diamond\phi) = 1, \text{ then: } Pr(\neg\phi) = 1$$

Assuming that epistemic necessity and possibility operators are duals, and assuming a classical view of negation, **Not-might** is equivalent to the principle that assigning credence 1 to *must* ϕ requires assigning credence 1 to ϕ. This seems uncontroversial.[7] Now, we reason as follows. Take an arbitrary, rational probability function *Pr*. Via total probability, we have:

i. $Pr(\neg\phi) = Pr(\neg\phi \wedge \Diamond\phi) + Pr(\neg\phi \wedge \neg\Diamond\phi)$

Given the result of Russell and Hawthorne's proof, we know that the first term goes to zero. Using the definition of conditional probability, we rearrange the second term as follows:

ii. $Pr(\neg\phi) = Pr(\neg\phi \mid \neg\Diamond\phi) \times Pr(\neg\Diamond\phi)$

Now, since $Pr(\cdot \mid \neg\Diamond\phi)$ is a rational probability function (via closure under conditionalization), and since obviously $Pr(\neg\Diamond\phi \mid \neg\Diamond\phi) = 1$, via **Non-Might** we know that $Pr(\neg\phi \mid \neg\Diamond\phi) = 1$. As a result, (ii) simplifies to:

iii. $Pr(\neg\phi) = Pr(\neg\Diamond\phi)$

But (iii) is disastrous, since it immediately leads to collapse. Via the probability calculus, we have:

[7] Notice that the principle doesn't amount to taking epistemic necessity to be factive, which might indeed be controversial. It just requires that, if we are certain of *must* ϕ, we are also certain of ϕ.

iv. $Pr(\phi) + Pr(\neg\phi) = 1$

From here, via (iii):

v. $Pr(\phi) + Pr(\neg\Diamond\phi) = 1$

Via the probability calculus again:

v. $Pr(\phi) + 1 - Pr(\Diamond\phi) = 1$

From here, rearranging:

v. $Pr(\phi) = Pr(\Diamond\phi)$

This result, of course, is absurd. In the rest of this section, I show that an analogous result can be derived for counterfactuals, following analogous steps.

3.2 Overview of the result

I show that the we can prove that, for any probability function that models rational credence:

Probabilistic Collapse. $Pr(\phi > \psi) = Pr(\phi \Diamond\!\!\rightarrow \psi)$

The proof proceeds in two steps. First, I use two basic principles linking the probabilities of *would*- and *might*-counterfactuals to derive a probabilistic version of the principle of Duality between *would*- and *might*-counterfactuals.

Probabilistic Duality. $Pr(\phi > \psi) = Pr(\neg(\phi \Diamond\!\!\rightarrow \neg\psi))$

The last step of the proof combines Probabilistic Duality with a claim that links probabilities of counterfactuals to expected chances, and which is normally known as 'Skyrms' Thesis' (from Skyrms 1981). Skyrms' Thesis, combined with **Probabilistic Duality**, allows us to derive the absurd conclusion that the probabilities of *would*-counterfactuals are identical to the probabilities of *might*-counterfactuals.

Before proceeding, let me emphasize that, even though I am relying on an assumptions that mentions chance, nothing in the proof exploits any substantial claims about what chances are, aside from the basic assumption that chances are probabilistic. The problem is generated not by any particular assumption about chance.[8] Rather, its source is the tension generated by probabilistic principles about the interaction of *would*- and *might* counterfactual, on the one hand, and principles in the general form of Stalnaker's Thesis, which are logically related to Conditional Excluded Middle.

The result resembles the collapse result in conditional logic in §2.4. But the proof runs on much weaker assumptions. In particular, I use elementary probabilistic principles about the interaction of *would*- and *might* counterfactuals that even defenders of CEM should be open to endorsing.

[8]Compare Schwarz's (2016) criticism of Williams' result, which depends on assuming the Principal Principle rather than the New Principle.

3.3 Basic constraints.

Constraints about the interaction of *would*- and *might*-conditionals. I use the following constraints probability functions.

> **Nonzero.** For all ϕ, ψ, and for all probability functions *Pr* that model rational credence and that are such that $Pr(\phi \diamondto \psi) > 0$:
>
> $$Pr(\phi > \psi \mid \phi \diamondto \psi) > 0$$
>
> **Upper bound.** For all ϕ, ψ, and for all probability functions that model rational credence *Pr*:
>
> $$\text{If } Pr(\neg(\phi \diamondto \psi)) = 1, \text{ then } Pr(\phi > \neg\psi) = 1$$

I will also appeal to the following principle of conditional logic:

$$\textbf{Conditional Non-Contradiction (CNC)} \quad \phi > \neg\psi \vDash \neg(\phi > \psi)$$

Finally, I assume that the class of probability functions that model rational credence is closed under conditionalization. I.e., I assume that, if *Pr* is a probability function that models rational credence, and $Pr(\phi) > 0$, then $Pr(\cdot \mid \phi)$ is also a probability function that models rational credence.

Nonzero can be framed as a basic constraint connecting the probabilities of *would*- and *might*-counterfactuals. It says that the probability of *If* ϕ*, would* ψ, conditional on *If* ϕ*, might* ψ, has to be greater than zero. **Nonzero** captures the intuition that it seems irrational to express full credence in a conditional like (15), and yet assign zero credence to a conditional like (16).

(15) If Sarah had tossed the coin, it might have landed tails.

(16) If Sarah had tossed the coin, it would have landed tails.

Nonzero can be questioned by appealing to the idea that propositions that express live possibilities can still receive probability zero (see e.g. Hájek 2003). In particular, we might grant that, in some cases, a *might*-counterfactual is true while the corresponding *would*-counterfactual has probability zero. As a candidate example, consider:

(17) If Sarah picked a real number at random between 0 and 1, she might pick 0.5.

(18) If Sarah picked a real number at random between 0 and 1, she would pick 0.5.

Even granting that **Nonzero** might have limited applicability, notice that the result still holds for counterfactuals for which **Nonzero** holds. (15) and (16) are exactly cases of this sort.

Upper bound is a weak probabilistic version of a basic logical principle about the interaction between *might*- and *would*-counterfactuals.

$$\textbf{\textit{Not-might-to-if}} \quad \neg(\phi \diamondto \psi) \vDash (\phi > \neg\psi)$$

Not-might*-to-*if, in turn, is equivalent to the right-to-left direction of Duality (assuming that negation behaves classically).

$$\textbf{Duality} \quad (\phi > \psi) \dashv\vDash \neg(\phi \diamond\!\!\rightarrow \neg\psi)$$

Let me emphasize that, for the purposes of this proof, I am *not* assuming Duality. This would not be dialectically advisable, given the tension between the semantics for counterfactuals that vindicate Duality and those that vindicate CEM. At the same time, the right-to-left direction of the principle seems hard to question. If it's not the case that the coin might have landed tails, had I tossed it, then the following seems uncontroversial: the coin would not have landed tails, had I tossed it. Moreover, strictly speaking I am not assuming **Not-*might*-to-*if***. Rather, I'm assuming that, whenever the premise receives credence 1, the conclusion also receives credence 1. This seems supported by intuitive judgments. Notice that certainty in (19) seems to require also certainty in (20).

(19) It's not true that, if Sarah had tossed the coin, it might have landed tails.

(20) If Sarah had tossed the coin, it would not have landed tails.

Skyrms' Thesis. Skyrms' Thesis is a counterfactual counterpart of Stalnaker's Thesis. It says that one's level of credence in a counterfactual should be equal to the expectation of the conditional chance of the consequent, given the antecedent. (One's expectation of chance is given by the weighted average of all the candidates for the actual chance function.)

> **Skyrms' Thesis.** Let Ch_w be the chance function at w. For all ϕ, ψ, and for all acceptable c-probability functions Pr such that $Pr(\phi) > 0$:
>
> $$Pr(\phi > \psi) = \sum_{w_i \in W} Pr(w_i) \times Ch_{w_i}(\psi \mid \phi)$$

Let me point out that the only reason why I am assuming Skyrms's Thesis is that it provides a bridge between probabilities of counterfactual and a kind of conditional probability. Any other principle of the same form would be equally effective.[9]

3.4 The proof

Step 1: incompatibility of $\phi > \neg\psi$ and $\phi \diamond\!\!\rightarrow \psi$. The first step establishes that a counterfactual $\phi > \neg\psi$ and the *might*-counterfactual $\phi \diamond\!\!\rightarrow \psi$ are incompatible: i.e., their conjunction has probability zero. Assume for *reductio* that $\phi > \neg\psi$ and $\phi \diamond\!\!\rightarrow \psi$ are compatible and that hence some probability function Pr assigns positive probability to both of them. Via **Nonzero**, we know:

[9] In fact, the whole point of using Skyrms' Thesis in the proof is deriving the equation:

$$Pr(\phi > \psi) + Pr(\phi > \neg\psi) = 1$$

I.e. a probabilistic version of CEM for counterfactuals. Any other assumption that led to this conclusion would do.

i. $Pr(\phi > \psi \mid \phi \diamond\!\!\rightarrow \psi) > 0$

Assuming that the class of rational probability functions is closed under conditionalization, we also have:

ii. $Pr_{\phi > \neg\psi}(\phi > \psi \mid \phi \diamond\!\!\rightarrow \psi) > 0$

Via the definition of conditionalization, (ii) is equivalent to:

iii. $Pr(\phi > \psi \mid \phi \diamond\!\!\rightarrow \psi \wedge \phi > \neg\psi) > 0$

However, via **CNC**, we know that

iv. $Pr(\phi > \psi \mid \phi > \neg\psi) = 0$

Hence (iii) and (iv) contradict. We conclude that $\phi > \neg\psi$ and $\phi \diamond\!\!\rightarrow \psi$ are incompatible.

Step 2: equivalence of $\phi > \neg\psi$ and $\neg(\phi \diamond\!\!\rightarrow \neg\psi)$. Take any *Pr* such that $Pr(\neg(\phi \diamond\!\!\rightarrow \psi)) > 0$. Then we can derive that $Pr(\phi > \psi)$ is equal to $Pr(\neg(\phi \diamond\!\!\rightarrow \neg\psi))$. We first observe, via the law of total probability:

i. $Pr(\phi > \neg\psi) = Pr(\phi > \neg\psi \wedge \phi \diamond\!\!\rightarrow \psi) + Pr(\phi > \neg\psi \wedge \neg(\phi \diamond\!\!\rightarrow \psi))$

Via the previous proof, $Pr(\phi > \neg\psi \wedge \phi \diamond\!\!\rightarrow \psi) = 0$. Reorganizing the term on the right-hand side:

ii. $Pr(\phi > \neg\psi) = Pr(\phi > \neg\psi \mid \neg(\phi \diamond\!\!\rightarrow \psi)) \times Pr(\neg(\phi \diamond\!\!\rightarrow \psi))$

Since $Pr(\cdot \mid \neg(\phi \diamond\!\!\rightarrow \psi))$ is an admissible probability function (via closure under conditionalization), and since $Pr(\neg(\phi \diamond\!\!\rightarrow \psi) \mid \neg(\phi \diamond\!\!\rightarrow \psi)) = 1$, via **Upper Bound** we get that $Pr(\phi > \neg\psi) \mid \neg(\phi \diamond\!\!\rightarrow \psi)) = 1$. Hence (ii) simplifies to

iii. $Pr(\phi > \neg\psi) = Pr(\neg(\phi \diamond\!\!\rightarrow \psi))$

Assuming that negation is classical, we get:

Probabilistic Duality. $Pr(\phi > \psi) = Pr(\neg(\phi \diamond\!\!\rightarrow \neg\psi))$

I show below that **Probabilistic Duality**, supplemented with an appropriate version of Skyrms' Thesis, gives rise to a collapse result. But it's worth emphasizing that, even without that, **Probabilistic Duality** has unacceptable consequences. Consider the following pair of *might*-counterfactuals

(21) If Sarah had tossed the coin, it might have landed tails.

(22) If Sarah had tossed the coin, it might not have landed tails.

Intuitively, (21) and (22) are both true in a situation where Sarah declined to toss a fair coin. Hence they should both get probability one, or near-one. Via **Probabilistic Duality**, this means that both the corresponding *would*-conditionals should get probability zero, or near-zero. This obviously doesn't match any intuitive assignment of probabilities to counterfactuals.

Step 3: Collapse. Assume now **Skyrms' Thesis**. The latter and **Probabilistic Duality** lead to collapse. ('PC' stands for 'probability calculus'.)

i. $\sum_{w_i \in W} Pr(w_i) \times Ch_{w_i}(\psi \mid \phi) + \sum_{w_i \in W} Pr(w_i) \times Ch_{w_i}(\neg\psi \mid \phi) = 1$ (PC)

ii. $Pr(\phi > \psi) + Pr(\phi > \neg\psi) = 1$ (i, Skyrms' Thesis)

iii. $Pr(\phi > \psi) + Pr(\neg(\phi \diamondsuit\!\!\rightarrow \psi)) = 1$ (ii, Probabilistic Duality)

iv. $Pr(\phi > \psi) + 1 - Pr(\phi \diamondsuit\!\!\rightarrow \psi) = 1$ (iii, PC)

v. $Pr(\phi > \psi) = Pr(\phi \diamondsuit\!\!\rightarrow \psi)$ (iv, algebra)

3.5 Discussion

Let me briefly survey the options for the truth-conditional theorist to resist this result.

Denying one of the starting constraints. Denying **Upper Bound** doesn't seem a realistic strategy. **Upper Bound** (as well as the right-to-left direction of **Duality**) holds on any combination of plausible truth-conditional semantics for *would*- and *might*-counterfactuals. As for the semantics of *might*-counterfactuals: the only available accounts seem to be variants on a Krazter-style semantics, on which *might*-counterfactuals are existential quantifiers over a set of (closest or close) worlds. As a result, $\neg(\phi \diamondsuit\!\!\rightarrow \psi)$ has roughly the truth conditions:

$\ulcorner \neg(\phi \diamondsuit\!\!\rightarrow \psi) \urcorner$ is true iff all closest ϕ-worlds are not ψ-worlds

These truth-conditions are bound to entail the counterfactual $\phi > \psi$, on any plausible semantics.[10]

Denying **Nonzero** might seem more promising. In fact, standard semantic theories that vindicate **Duality**, including Lewis's and Kratzer's own theory, invalidate **Nonzero** (given plausible assumptions about credence). These theories predict that, in some scenarios, it can be that both the *might*-counterfactuals in (23) are true and both the *would*-counterfactuals in (24) are false.

(23) a. If Sarah had tossed the coin, it might have landed tails.

b. If Sarah had tossed the coin, it might have landed heads.

(24) a. If Sarah had tossed the coin, it would have landed tails.

b. If Sarah had tossed the coin, it would have landed heads.

[10] One account that can elude this constraint is, in principle, the arbitrary selection account in Schulz 2014. If we let $\phi > \psi$ select an absolutely random world, with no further constraints, and constraint the domain of quantification of $\phi \diamondsuit\!\!\rightarrow \psi$ to a set of close worlds, **Upper Bound** fails. But this version of the view fails to vindicate the obvious principle that *would*-counterfactuals entail *might*-counterfactuals ($\phi > \psi \vDash \phi \diamondsuit\!\!\rightarrow \psi$), and hence is independently implausible.

If we assume that a reasonable subject can be certain of this, then we get a failure of **Nonzero**. However, embracing a semantics that predicts this will be of no help in blocking the collpase proof. If we have a semantics that vindicates **Duality**, **Probabilistic Duality** will also automatically hold, hence the proof will go through.

Denying **Skyrms' Thesis** seems the best option. This blocks the result, but it is empirically costly. It seems intuitive that, at least in some cases, the probabilities of counterfactuals should mirror the expectations of the relevant conditional chances. Examples involving dice and coins, like e.g. those in (24), seem cases in point. To block the conclusion that the probabilities of these counterfactuals are identical to the probabilities of the corresponding *might*-counterfactuals, we have to deny the corresponding instances of **Skyrms' Thesis**. This is obviously counterintuitive.

Invoking context dependence? One longstanding strategy for blocking triviality involves using the context-dependence of epistemic discourse to disrupt triviality proofs.[11] A number of triviality proofs involve assigning probability to a conditional both unconditionally and under supposition. (For example, in Step 2 in the proof in §3, $\phi > \neg\psi$ is assigned both an unconditional probability and a probability conditional on $\neg(\phi \diamond\!\!\rightarrow \psi)$.) One way to respond to the proof is to claim that shifts in what information is assumed involve shifts in the interpretation of the conditional, and that hence the relevant proofs involve equivocation.

I won't assess the merits of this response for epistemic conditionals here. But I want to notice that this strategy doesn't carry over in a straightforward way to counterfactuals. On all standard truth-conditional semantics for epistemic modality, the interpretation of epistemic claims is relativized to the knowledge or belief state of an agent, or group of agents. As a result, it is natural to think that, when the claim is assessed under supposition, it receives a different interpretation then when it is assessed on its own. But this consideration does not extend to counterfactuals. Counterfactuals are not interpreted relative to the knowledge or belief state of an agent or a set of agents, and in general their interpretation is not expected to shift when they are interpreted under epistemic supposition. Hence the context-dependent move cannot be straightforwardly replicated for counterfactuals.

4 Conclusion

I have argued that two of the standard challenges for truth-conditional accounts of epistemic modality generalize to counterfactuals. This is part of a broader argument to the effect that we need to develop a general, unified account of all flavors of modality in a nonclassical framework. A positive account of this form is a task for a different occasion.

[11]For a classical theory that involves this style of response, see Van Fraassen 1976 and related work.

References

Bacon, Andrew (2015). "Stalnaker's Thesis in Context." *Review of Symbolic Logic*, 8(1): pp. 131–163.

Bradley, Richard (2000). "A Preservation Condition for Conditionals." *Analysis*, 60(3): pp. 219–222.

Bradley, Richard (2007). "A Defence of the Ramsey Test." *Mind*, 116(461): pp. 1–21.

Briggs, R.A. (2017). "Two Interpretations of the Ramsey Test." In C. H. Helen Beebee, and H. Price (eds.) *Making a Difference: Essays in Honour of Peter Menzies*, Oxford University Press.

von Fintel, Kai (1997). "Bare plurals, bare conditionals, and only." *Journal of Semantics*, 14(1): pp. 1–56.

von Fintel, Kai (2001). "Counterfactuals in a Dynamic Context." *Current Studies in Linguistics Series*, 36: pp. 123–152.

von Fintel, Kai, and Sabine Iatridou (2002). "If and When If -Clauses Can Restrict Quantifiers." Unpublished draft, available at http://web.mit.edu/fintel/fintel-iatridou-2002-ifwhen.pdf.

Gillies, Anthony S (2007). "Counterfactual scorekeeping." *Linguistics and Philosophy*, 30(3): pp. 329–360.

Goldstein, Simon (forthcoming). "Triviality Results For Probabilistic Modals." *Philosophy and Phenomenological Research*.

Hájek, Alan (2003). "What Conditional Probability Could Not Be." *Synthese*, 137(3): pp. 273–323.

Hájek, Alan, and N. Hall (1994). "The Hypothesis of the Conditional Construal of Conditional Probability." In Ellery Eells, Brian Skyrms, and Ernest W. Adams (eds.) *Probability and Conditionals: Belief Revision and Rational Decision*, Cambridge University Press, p. 75.

Higginbotham, James (1986). "Linguistic Theory and Davidson's Program in Semantics." In E. LePore (ed.) *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Cambridge: Blackwell, pp. 29–48.

Kaufmann, Stefan (2009). "Conditionals Right and Left: Probabilities for the Whole Family." *Journal of Philosophical Logic*, 38(1): pp. 1–53.

Korzukhin, Theodore (2014). "Dominance Conditionals and the Newcomb Problem." *Philosophers' Imprint*, 14(9).

Kratzer, Angelika (1977). "What 'Must' and 'Can' Must and Can Mean." *Linguistics and Philosophy*, 1(3): pp. 337–355.

Kratzer, Angelika (1981a). "The Notional Category of Modality." In H. J. Eikmeyer, and H. Rieser (eds.) *Words, Worlds, and Contexts: New Approaches to Word Semantics*, Berlin: de Gruyter.

Kratzer, Angelika (1981b). "Partition and Revision: The Semantics of Counterfactuals." *Journal of Philosophical Logic*, 10(2): pp. 201–216.

Kratzer, Angelika (1991). "Modality." *Semantics: An international handbook of contemporary research*, pp. 639–650.

Kratzer, Angelika (2012). *Modals and Conditionals: New and Revised Perspectives*, vol. 36. Oxford University Press.

Lewis, David (1976). "Probabilities of Conditionals and Conditional Probabilities." *Philosophical Review*, 85(3): pp. 297–315.

Lewis, David K. (1973a). *Counterfactuals*. Cambridge, MA: Harvard University Press.

Lewis, David K. (1973b). "Counterfactuals and Comparative Possibility." *Journal of Philosophical Logic*, 2(4): pp. 418–446.

Russell, Jeffrey Sanford, and John Hawthorne (2016). "General Dynamic Triviality Theorems." *Philosophical Review*, 125(3): pp. 307–339.

Santorio, Paolo (2018). "Conditional Excluded Middle in Expressivist Semantics." Draft, University of California, San Diego.

Schulz, Moritz (2014). "Counterfactuals and Arbitrariness." *Mind*, 123(492): pp. 1021–1055.

Schwarz, Bernhard (2011). "Long distance indefinites and choice functions." *Language and Linguistics Compass*, 5(12): pp. 880–897.

Schwarz, Wolfgang (2016). "Subjunctive Conditional Probability." Forthcoming in *Journal of Philosophical Logic*.

Skyrms, Brian (1981). "The prior propensity account of subjunctive conditionals." In W. Harper, R. C. Stalnaker, and G. Pearce (eds.) *Ifs*, Reidel.

Stalnaker, Robert (1981). "A Defense of Conditional Excluded Middle." In W. Harper, R. C. Stalnaker, and G. Pearce (eds.) *Ifs*, Reidel, pp. 87–104.

Stalnaker, Robert (1984). *Inquiry*. Cambridge University Press.

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 (1996). "Defaults in Update Semantics." *Journal of Philosophical Logic*, 25(3): pp. 221–261.

Williams, J. Robert G. (2012). "Counterfactual Triviality." *Philosophy and Phenomenological Research*, 85(3): pp. 648–670.

Yalcin, Seth (2007). "Epistemic Modals." *Mind*, 116(464): pp. 983–1026.