Noûs

NOÛS 00:0 (2015) 1–32
doi: 10.1111/nous.12111

*Triviality For Restrictor Conditionals**

NATE CHARLOW
University of Toronto

Abstract

I present two Triviality results for Kratzer's standard "restrictor" analysis of indicative conditionals (while also clarifying the sense in which Kratzer's semantics might avoid such results). I both refine and undermine the common claim that problems of Triviality do not arise for Kratzer conditionals since they are not strictly conditionals at all.

1. Overview

I present two Triviality results, adapted from David Lewis and Richard Bradley, for Kratzer's linguistically standard "restrictor" semantics for indicative conditionals. On Kratzer's semantics, a conditional $A \to C$ does not generally express a conditional relation (i.e. $\to$) between antecedent and consequent. Instead, conditionals are understood as devices for making explicit the domain restriction argument of some (possibly covert) quantifier. In the process, I clarify the sense in which this type of analysis of the conditional might avoid such Triviality results. The upshot is to both *refine* and *undermine* the claim—stated occasionally in the literature, but subscribed to, more or less routinely, by natural language semanticists familiar with the philosophical literature on conditionals—that problems of Triviality do not arise for "Kratzer conditionals,"[1] because Kratzer conditionals are not, strictly speaking, conditionals at all (and do not, moreover, involve probabilistic commitments of the sort that Lewis appears to reduce to absurdity, e.g., Stalnaker's Thesis[2]). Crucially, my argument, following the now well-known strategy of Bradley (2000, 2007), does not assume any such probabilistic commitments on behalf of the analysis of the conditional under consideration (though I do make use of other plausible assumptions that could, in principle, be denied).

The claim that Kratzer conditionals are not subject to problems of Triviality is only sporadically addressed in the literature, but is commonly assumed as part of the philosophical and linguistic folklore surrounding the conditional. It is, however, surprisingly hard to credit. The first part of this paper attempts, unsuccessfully, to

*Special thanks to Daniel Rothschild and Jonathan Weisberg, whose feedback was essential for this project. Additional thanks to Simon Charlow, Kai von Fintel, Thony Gillies, Arno Goebel, Alan Hájek, Jeremy Goodman, Angelika Kratzer, Brian Leahy, Eric Raidl, Mark Richard, Jason Stanley, H. Orri Stefánsson, Una Stojnic, Timothy Williamson, several anonymous referees for this journal, and an audience at the 2015 Belgrade Conference on Conditionals. Support from the Social Sciences and Humanities Research Council of Canada (Insight Grant #435-2015-0423) is gratefully acknowledged.

© 2015 Wiley Periodicals, Inc.

isolate a plausible reading of the claim. As we will see, Kratzer herself sometimes speaks as if conditionals could not have probabilities—indeed, as if it makes no sense to give an answer to the question *what is the probability of a conditional?* at all. We might be tempted to draw a comparison here to Austin (1961): "I can only answer a question of the form 'What is the meaning of *x*?' if *x* is some *particular* word you are asking about." Similarly, for Kratzer, '*the meaning of* $A \rightarrow C$' is non-referring absent a specification of the operator whose domain the conditional's *if*-clause is functioning to restrict. But this is easily controlled for, by restricting our focus to the proposition expressed, according to Kratzer, by a simple, unembedded conditional in a standard context (see §2.2). Showing that Kratzer's analysis of such conditionals falls prey to Triviality would be sufficient to upset the consensus that is this paper's target.

Kratzer also, as we will see, tends to speak as if judgments apparently concerning the probability of a conditional, in thought and language, are not properly taken as *evidence for theoretical claims about the probability of the conditional* within the probability calculus (e.g. Stalnaker's Thesis). §3 shows that, while this may be right, it is beside the point, so far as the most worrying aspect of Triviality—namely, trivialization of a probability distribution over a wide range of cases—is concerned. As Bradley (2000, 2007) has noted, Triviality tends to be narrowly (and misleadingly) interpreted as displaying a tension between a probabilistic and truth-conditional understanding of the meaning of the conditional—a tension that could be addressed by jettisoning the probabilistic understanding in favor of the truth-conditional or propositional understanding (as the reading of Kratzer under consideration suggests). If, however, we take Lewis' Triviality result (for example) at face-value, the choice is, I will argue, rather more stark than that: between a propositional understanding of the conditional and a probability calculus on which propositions can receive meaningful probabilities at all. To put it slightly differently: if we cannot assign the conditional a sensible probability within the probability calculus without trivializing the probability calculus, this can be taken as fairly good evidence that the probability calculus *should not apply* to the conditional, in which case the conditional should not be taken to express a proposition at all. The major goal of this paper is to show that, while the details may differ, Kratzer conditionals are caught up in this dialectic no less than any other propositional semantics for the conditional.

The remainder of the paper constructs proofs of two different Triviality results: the result of Lewis (1976) and (in an Appendix) the result of Bradley (2000). The proofs, I note, are threatened by the possibility that the proposition expressed by the conditional is—very roughly—sensitive to the "context" represented by a probability function. I make two points in reply: (i) appeals of this sort to context-sensitivity are a well-known strategy for dealing with problems of Triviality. In making such an appeal, an advocate of the Kratzer analysis would thus concede my main point: that Triviality applies the same theoretical pressures to the Kratzer conditional as it applies to any analysis of the conditional. Nevertheless, (ii) such appeals are not substantively plausible, given the most natural interpretation of the formal apparatus that this paper develops.

2. Background

This section fills in two pieces of essential background. First, it reviews the Triviality proof of Lewis (1976). Second, it offers a précis of Kratzer's analysis of the conditional.

2.1. Lewis-Triviality

The Triviality Result of Lewis (1976) makes use of two major assumptions.[3]

Stalnaker's (1970) Thesis (ST) $Pr(A \to C) = Pr(C|A)$

Conditionalization Thesis (CT) $Pr_A(C) = Pr(C|A)$

These say respectively that (i) the probability of an indicative conditional $A \to C$ equals the corresponding conditional probability, $Pr(C|A)$, and (ii) the probability of C once one conditionalizes on A equals the corresponding conditional probability. These are used to prove the "If-And" Lemma.

If-And Lemma $Pr((A \to C)|X) = Pr(C|(A \wedge X))$

Proof. By CT,

$$Pr((A \to C)|X) = Pr_X(A \to C)$$

By ST,

$$Pr_X(A \to C) = Pr_X(C|A) = \frac{Pr_X(C \wedge A)}{Pr_X(A)}$$

By the Conditionalization Thesis,

$$\frac{Pr_X(C \wedge A)}{Pr_X(A)} = \frac{Pr((C \wedge A)|X)}{Pr(A|X)}$$

By the definition of conditional probability,

$$\frac{Pr((C \wedge A)|X)}{Pr(A|X)} = \frac{\frac{Pr(C \wedge A \wedge X)}{Pr(X)}}{\frac{Pr(A \wedge X)}{Pr(X)}} = \frac{Pr(C \wedge A \wedge X)}{Pr(A \wedge X)} = Pr(C|(A \wedge X))$$

□

This lemma is then used to show:

Lewis-Triviality (LT) $Pr(A \to C) = Pr(C)$

Proof. By the Law of Total Probability:

$$Pr(A \to C) = [Pr((A \to C)|C) \cdot Pr(C)] + [Pr((A \to C)|\neg C) \cdot Pr(\neg C)]$$

By the If-And Lemma, this equals:

$$[Pr(C|(A \wedge C)) \cdot Pr(C)] + [Pr(C|(A \wedge \neg C)) \cdot Pr(\neg C)]$$

Finally, since $Pr(C|(A \wedge C)) = 1$ and $Pr(C|(A \wedge \neg C)) = 0$, this all reduces to $Pr(C)$. □

This seems unacceptable. Suppose conditionals have probabilities—and if they express propositions, it is hard to see why they would not (more on this below). According to Triviality, their probabilities are, in an informal sense, "trivial"—equal simply to the probabilities of their consequents.

2.2. *Kratzer Conditionals*

It is widely thought that Kratzer's analysis of conditionals avoids this sort of Triviality, on what amounts to a technicality. Rothschild (2013), for example, writes:

> Kratzer (1981, 1986) denies the syntactic parsing that is required to formulate this problem. She claims that the function of the 'if'-clauses is to restrict higher-up modal quantifiers in the sentence. (53)

To understand how this might bear on Triviality, some details about Kratzer's analysis will be needed. (I will add more below.) It will help first to separate the main idea of Kratzer's analysis into two parts: (i) Modal quantifiers are generalized quantifiers—quantifiers that express relations between the set specified by their *restriction* and the set specified by their *scope* (Barwise and Cooper, 1981); (ii) The compositional function of an 'if'-clause is to restrict the domain of a (by default, modal) quantifier. The range of data for which this analysis accounts is astonishing (an entrée: von Fintel, 2011; Gillies, 2010, 2012). So it is no surprise that it is simply presupposed by most work in semantics.[4]

What relation is expressed between the sets specified respectively by the quantifier's restriction and scope? The first thing to note is that it will depend on the nature of the quantifier. Consider:

(1) If my keys aren't in my coat, they must be on the table (epistemic necessity)
(2) If my keys aren't in my coat, they might be on the table (epistemic possibility)
(3) If it rains, you must cancel the match (deontic necessity)
(4) If it has stopped raining, you may go outside (deontic possibility)

Note also that, if the compositional job of the 'if'-clause is to restrict, then "[w]henever there is no explicit operator, we have to posit one" (Kratzer, 1986). Bare conditionals—conditionals like (5) whose matrix clauses lack an overt modal—are thus supposed to contain a covert modal expressing (ordinarily epistemic) necessity, so that (5) and (1) receive the same interpretation.

(5) If my keys aren't in my coat, they are on the table

This paper will generally concern the sort of quantification expressed by epistemic modals and unembedded "epistemic conditionals" like (1) and, implicitly, (5) that invoke these modals. Without this restriction, appeal to the notion of a "Kratzer conditional" would be in fact ill-founded. According to Kratzer, there is, for instance, no "semantic value of the conditional" whose surface form is $A \to C$, when $A \to C$ or C is embedded under an operator whose domain A functions to restrict (cf. Égré and Cozic, 2011, 23), as in:

(6) Probably, if my keys aren't in my coat, they are on the table
(7) If my keys aren't in my coat, they are probably on the table

Let $must(X)(Y)$ represent a Kratzerian dyadic (i.e., two-argument-place) modal expressing whatever sort of restricted epistemic necessity is routinely expressed by a simple, unembedded indicative conditional of the form $X \to Y$; the left argument is the modal's restriction and the right argument its scope; *might* is *must*'s dual: $might(X)(Y) := \neg must(X)(\neg Y)$. By *Kratzer Conditional*, I shall refer only to a conditional of surface form $A \to C$ or $A \to must(C)$ whose truth condition is, according to Kratzer, equivalent to that of a modal of the form $must(A)(C)$. Kratzer conditionals, then, are associated with modal "logical forms" in which their "antecedents" fill the modal's restriction and their "consequents" fill its scope—stripped of their modality if the conditional is explicitly epistemic, otherwise as is. If we let p be the sentence 'the keys are in my coat' and q the sentence 'the keys are on the table', the logical form of both (1) and (5) is a modal of the form $must(\neg p)(q)$.

3. Avoiding Triviality

This section will canvass some different reasons for thinking Kratzer conditionals *could not*—that is to say, in principle—be subject to problems of Triviality. These reasons are all found wanting.

3.1. No Probabilities of Conditionals

Rothschild (2013) notes that, for Kratzer, there is a sense in which the very notion of *the probability of a conditional* may be ill-founded: for Kratzer, "when we evaluate the probability of [a conditional] we do *not* think about how likely it is that it's true" (54). But what, exactly, could this mean? Of course there is first the fact that "a conditional" does not pick out a single kind of syntactic constituent, as just seen. But, since we have explicitly restricted our focus to conditionals that share a single type of logical form, I take it that we have controlled for this.

Still, there remains the issue that, for Kratzer, a claim of the form *$A \to C$ is x-probable* is—as indicated in the prior section—only properly interpreted as a claim of the form $x\text{-}probable(A)(C)$, with $x\text{-}probable$ a generalized probability operator, A its restriction, C its scope. Kratzer writes:

> [C]ontrary to appearance, quantificational operators, including adverbs of quantification, probability operators, and other modal operators, do not operate over "conditional propositions." The persistent belief that there could be such "conditional propositions"

> is based on a simple syntactic mistake. If-clauses need to be parsed as adverbial modifiers that restrict operators that might be silent and a distance away. (Kratzer, 2012: 107)

The suggestion seems to be that Triviality proofs assume essentially that the notion of a *conditional's probability* is a sensible one (most clearly in ST), when in fact it is not.[5] Though there is an ordinary proposition—an ordinary member of the relevant algebra over which *Pr* is assumed to be (totally) defined—that, for Kratzer, is expressed by the conditional, the present suggestion seems to be that *Pr* will be *undefined* for that proposition, thus rendering unformulable any step of the proof at which the probability of a conditional is mentioned.

This idea cannot be taken seriously. The suggestion that *Pr* simply has no value for conditionals seems, of course, ad hoc. Putting this aside, however, the suggestion is independently untenable, if we recall the dialectical context surrounding the Triviality results. We do not have very many solid diagnostics for assessing whether a sentence expresses a proposition—something evaluable for truth and fit for being the object of representational attitudes like belief—or not. One very clear diagnostic we do have is the following: ask whether it makes sense to regard probability functions as *being defined for* that sentence's semantic value (cf. Hájek, 2012, 147ff). In suggesting that it does not make sense to see probability functions as defined for the conditional, an advocate of Kratzer conditionals would seem to concede the very conclusion for which proponents of, e.g., No Truth-Value (NTV) views of the conditional have historically cited the Triviality results to argue.[6]

3.2. Against ST

What this does highlight, however, is that intuitive judgments of the sort commonly cited in motivating ST might not, in fact, bear on the motivation of ST at all (this point is made very clearly in Rothschild forthcomingb). In cases where it would be judged that $Pr(C|A) = x$, it may be that we do find ourselves readily disposed to accept a corresponding description of the probability of $A \to C$; the converse holds as well. It is a familiar point that such transitions tend to be judged impeccable:

(8) The probability that it will rain, given low pressure, is 30%
(9) ≈It is 30% likely that it will rain if there is low pressure

Given Kratzer's analysis of (9), it is possible to account for the fact that these transitions are licensed without committing to thinking that the probability of a conditional proposition is .3. (Indeed, for Kratzer, the transitions are trivially valid, since (8) and (9) have essentially the same logical form.) So it is possible generally to account for the fact that we accept overtly probabilized conditionals just when we accept corresponding claims expressing conditional probabilities—a fact which constitutes the bulk of the traditional motivation for ST—*without accepting ST*.

3.3. ACCEPT and REJECT

Whether Kratzer can account for the data that is standardly taken to motivate ST without taking on ST is, in my view, simply not relevant to the question of whether Kratzer conditionals are subject to problems of Triviality. First, there are a number of ways to Triviality, only some of which go through ST. (Similarly, there are a number of Triviality Results, distinct in various ways from what Lewis originally showed, all apparently threatening to a semantics which is committed to them; see Appendix A for one.) A remarkable feature of even Lewis' original proof is that neither the If-And Lemma nor ST seem to play essential roles. ST was used only in the proof of the If-And Lemma, and the If-And Lemma only to establish:

$$Pr((A \to C)|C) = 1 \text{ and } Pr((A \to C)|\neg C) = 0$$

It will be handy to give these conditions names. We will dub them:

ACCEPT $Pr((A \to C)|C) = 1$

REJECT $Pr((A \to C)|\neg C) = 0$

If we could establish ACCEPT and REJECT independently for some analysis of conditionals, LT would of course follow immediately for that analysis, whether or not it took on ST.

The trouble—pointed out in a slightly different way by Bradley (2000, 2007)—is that, quite independently of ST, ACCEPT and REJECT might plausibly be taken as *constraints of adequacy* on a semantic theory for the conditional, so long as that semantic theory assigns the conditional a proposition as its semantic value. Generally, we are in a position to fully accept the conditional whenever we are in a position to fully accept its consequent: *come what may*, the consequent must be true; *a fortiori*, if the antecedent is true, the consequent will be.[7] In such situations, *whatever proposition* the semantic theory assigns the conditional should, it would seem, receive probability 1. Similarly, we are in a position to fully reject the conditional whenever we are in a position to fully reject its consequent: *come what may*, the consequent must be false; *a fortiori*, if the antecedent is true, the consequent will not be. In such situations, *whatever proposition* the semantic theory assigns the conditional should, it would seem, receive probability 0.

Of course, these conditions on the indicative conditional's probability do not hold on every semantics for it—e.g. the material conditional analysis. Generally:

$$Pr((A \supset C)|\neg C) = \frac{Pr((A \supset C) \wedge \neg C)}{Pr(\neg C)} = \frac{Pr(\neg A \wedge \neg C)}{Pr(\neg C)} = Pr(\neg A|\neg C) > 0$$

Provisionally, however, this should be regarded as a strike against the material conditional analysis.[8]

Triviality raises a hard choice for *any* propositional semantics of the indicative conditional (including Kratzer's): defend an analysis (like the material conditional analysis) on which the probability of a conditional on update with its consequent (or its consequent's negation) is non-extremal, or else embrace LT for the conditional. The first option will invite suspicion that the semantics one is proposing is (like the material conditional analysis) revisionary, in the sense of departing from banal judgments about the conditional and the probability of the conditional. Let us suppose (only for the moment) that it will be regarded as undesirable.

3.4. *Deflating Triviality*

The other option is to *embrace LT* for the conditional: embrace the notion that whatever probabilities the propositions assigned to conditionals by the semantic theory might have, those probabilities will be Trivial, in the sense of being unsuitable for playing any interesting theoretical function. There is no attempt to *avoid* LT in this sort of posture—only an attempt to *deflate* it. This section will argue that this is not an available strategy for the advocate of Kratzer's analysis of the conditional (or, indeed, any analysis at all).

Why might this strategy be attractive? Kratzer's treatment of linguistic ascriptions of probabilities to conditionals could be used—with some plausibility—to deflate a Triviality result. Kratzer is consistently forthright about her non-interest in the project of assigning conditionals non-trivial probabilities: whenever a claim to the effect that $A \to C$ has non-trivial probability $x \in (0, 1)$ is intuitively acceptable, as we might imagine for (8), Kratzer just denies that this this is a judgment about the probability of the proposition expressed by $A \to C$. It is rather something like a judgment that half of the circumstances in which it doesn't snow are ones in which it rains. On Kratzer's account, then, there is no immediate need for *conditional propositions* to have non-trivial probabilities.[9]

3.5. *Restricted Triviality*

The deflationary response can only be satisfactory when accompanied with a blanket rejection of ST. Lewis (1976) notes that LT coupled with ST has a disturbing consequence:

$$Pr(A \to C) = Pr(C|A) = Pr(C)$$

Thus, arbitrary A and C are probabilistically independent, trivializing Pr. To avoid this conclusion, ST must be *positively rejected*, if LT is accepted. This applies generally: whether one is a proponent of Kratzer's analysis of the conditional is irrelevant. The major problem with triviality, of course, is not that it makes probabilities of conditionals uninteresting (because identical to the probabilities of their consequents). It is that it tends to trivialize the very notion of a probability distribution.[10]

This puts the Kratzer analysis in a worrying position. The Kratzer analysis—the restrictor portion anyway—is compatible with a wide range of specific truth-conditions for the conditional: anything from material conditional, to Variably Strict conditional (à la Stalnaker, 1968; Lewis, 1973; Kratzer, 1981), to full-blown strict conditional.[11] Once LT is granted, we need to look carefully at each version of a Kratzerian truth-condition for the conditional $A \to C$ and ask (i) in what sorts of cases the probability of the proposition expressed by $A \to C$ will be predicted to be the corresponding conditional probability (i.e. over what class of cases ST holds), (ii) evaluate whether we are happy, in these cases, viewing A and C as probabilistically independent (since this can be established using the relevant restriction of ST together with LT). If we are not happy (as seems likely), we have a *restricted trivialization* of *some class* of probability distributions. While this is in some sense a "less dire" situation than full-blown trivialization, it is to be avoided. Whether a particular analysis in the Kratzer framework manages to do so must be evaluated on a case by case basis.

Here is an illustration. Even for a semantics on which ST does not hold in full generality, it can be demonstrated to hold in a rather wide range of cases given fairly standard assumptions. For instance, Ellis (1978, 116–7) (noted by Rothschild, 2013, 67) shows that, for any analysis of the conditional where what is usually called Strong Centering holds, ST will be valid in any case in which a fairly banal Independence Condition is met.

Strong Centering $(A \wedge C) \equiv ((A \rightarrow C) \wedge A)$

Independence Condition $Pr((A \rightarrow C)|A) = Pr(A \rightarrow C)$

Proof. Suppose that the Independence Condition holds for $A \rightarrow C$ and Pr. Then:

$$\frac{Pr((A \rightarrow C) \wedge A)}{Pr(A)} = Pr(A \rightarrow C)$$

By Strong Centering:

$$\frac{Pr(A \wedge C)}{Pr(A)} = Pr(A \rightarrow C)$$

□

Given Strong Centering, then, ST holds restrictedly, but quite widely: namely whenever the probability of the conditional is independent of the probability of its antecedent. But if LT independently holds in full generality, then whenever A is probabilistically independent of $A \rightarrow C$, C must be probabilistically independent of A: $Pr(C) = Pr(C|A)$. What sort of a propositional semantics would, at least ordinarily, validate Strong Centering and the Independence Condition? One natural answer is: *any plausible propositional semantics.* Another is: *Kratzer's* semantics.

Strong Centering is a generally desirable constraint on the logic of the indicative conditional (and it is routinely assumed in the literature). Its right-to-left direction encodes modus ponens. Its left-to-right direction differentiates it from so-called Weak Centering, by supposing that the truth of C in addition to A suffices for the truth of $A \rightarrow C$—a pattern that seems to respected in ordinary speaker judgments of conditionals.[12]

Independence might seem a puzzling condition, but its use can be motivated in two ways: (i) citing its intuitive appeal in a range of cases; (ii) showing that it follows from a specific semantic treatment of the conditional. Regarding (i), it does frequently seem correct to say that the likelihood of a particular relationship—whatever relationship the conditional expresses—holding between A and C is independent of whether A is true. Learning that the gardener is innocent generally has no effect on one's willingness to accept:

(10) If the gardener is innocent, the butler is guilty

Of course there are many analyses of the conditional on which Independence ordinarily fails—see, again, the material conditional analysis, on which the probability of the antecedent is typically *inversely* related to the probability of the conditional. But, again, it seems reasonable to regard this as a strike against any such analysis.

Regarding (ii), consider Kratzer's analysis of the conditional. Kratzer, we have seen, treats $A \to C$ as equivalent to a restricted epistemic necessity, i.e., a proposition about what follows from the A-possibilities compatible with the relevant epistemic state. For such a conditional, the assumption that A is probabilistically independent of $A \to C$ seems easily and frequently met: whenever "we don't think [the relevant epistemic] state is [linked to] whether the antecedent is true... [I]f there is no such [link] then our learning the antecedent is true or false should have no effect on our confidence in the propositions about" what is necessary relative to the relevant epistemic state (Rothschild, 2013, 59). It may be that Independence will, according to the Kratzer analysis, fail in certain cases—in which A is a proposition bearing somehow (perhaps indirectly) on the relevant epistemic state.[13] In most cases, however, it would, for the Kratzer analysis, seem to hold.

The relevant question, then, is this: in *every* such case, is it sensible to think $Pr(C) = Pr(C|A)$? The apparent answer is: no. Consider a case in which raising one's confidence in A has no effect on one's (nonzero) confidence in the proposition expressed by $A \to C$, e.g., (10). Surely it is possible—indeed, one would think, ordinarily obligatory—that raising one's confidence in A will nevertheless raise one's confidence in C, if one does not utterly reject $A \to C$. Similar points will, it would seem, hold for any Kratzerian semantics for the conditional on which modus ponens (!) is upheld. While LT coupled with a restricted validation of ST does not fully trivialize the probability distribution, it is liable to trivialize it over a fairly wide range of cases.

Advocates of Kratzer's analysis of the conditional cannot, then, simply cite their proprietary understanding of the data underlying ST to justify shrugging off of Lewis' Triviality Results. LT seems relatively easily established without appeal to ST. So long as ST holds restrictedly over a range of cases, LT will tend to trivialize the probability distribution within that range of cases. It turns out, then, that LT must be positively resisted.[14]

3.6. Signpost

To emphasize the running theme: none of these problems turns on (nor can any be avoided by citing) the peculiarities of Kratzer conditionals. Problems of Triviality arise for any propositional semantics for the conditional, since any such semantics should, on the face of it, endorse (i) Strong Centering (ordinarily); (ii) Independence (ordinarily); (iii) ACCEPT; (iv) REJECT. Contrary to what is so often supposed, the Kratzer analysis does not escape this dialectic just by entering its (admittedly very alluring) treatment of the data ordinarily cited in favor of ST into the record.

4. Lewis-Triviality

The last section reviewed, largely informally, the Triviality dialectic for Kratzer conditionals. This section will give these reflections some formal focus. I will articulate a fairly minimal semantics for the Kratzer conditional and a fairly minimal integration of probability into that semantic framework. I will note that, given one or two fairly plausible additional assumptions, LT can be shown to hold in a wide range of cases.

4.1. Semantics for Must

We have seen that Kratzer treats the bare indicative conditional as expressing a species of restricted epistemic necessity, so that the truth-condition of $A \rightarrow C$ is generally equivalent to a modal of the form $must(A)(C)$. While semantics of the modal *must* is a complicated matter (see Kratzer, 1981; Swanson, 2008), we require only a few details. The domain of quantification—which it is the function of the 'if'-clause to restrict—is held to be determined in part[15] by a body of information relevant at a circumstance of evaluation (world, situation, world-time pair, ...). (Here, for concreteness, I will take circumstances of evaluation to be worlds.) Most important for present purposes is that: (i) these modals have as their domains of quantification elements of an independently specified Boolean algebra (e.g., as I will suppose for concreteness, sets of possible worlds), (ii) facts witnessed throughout the domain of quantification are correspondingly witnessed by necessity modals that quantify universally over that domain.[16]

Let s (an information state) be a function from a circumstance of evaluation into a domain of quantification (an element of the relevant Boolean algebra). All I will assume about the domain of quantification is that it is "well-behaved" in the sense of Gillies (2010), i.e., Reflexive and Euclidean:

Reflexivity $w \in s(w)$
Euclideanness $v \in s(w)$ implies $s(w) \subseteq s(v)$

Reflexivity means that the domain of quantification for epistemic modals is given by a body *of knowledge*: the actual world is never eliminated. Euclideanness means that what is *absolutely epistemically possible*—what is compatible with what is known—does not vary across possibilities compatible with what is known at w. It is a routine assumption in this literature; note, for instance, that if a domain of quantification is non-Euclidean, we permit the truth of apparently contradictory claims like *it might be raining but maybe it must not be*.[17]

From these assumptions it follows that the domain of quantification is Closed.

Closure $v \in s(w)$ implies $s(w) = s(v)$

In other words, what is *absolutely* epistemically necessary—entailed by what is known—or *absolutely* epistemically possible—compatible with what is known—does not vary across possibilities compatible with the domain of quantification.[18]

The only thing I will assume about the semantics of *must* is an Inheritance assumption. To state it, first some notation. As is customary, we will let $[\![\cdot]\!]$ be an interpretation function mapping sentences and information state-circumstance of evaluation pairs into truth values; $[\![P]\!]^{s,w}$ generally denotes either 1 (true) or 0 (false). We will use expressions without the index of evaluation superscript (i.e. expressions of the form $[\![P]\!]^{s}$) to denote functions from indices of evaluation into truth values, or, equivalently, the set of indices of evaluation at which P is true relative to s:

$$[\![P]\!]^{s} = \{w : [\![P]\!]^{s,w} = 1\}$$

According to the Inheritance assumption, then, if at each point $v \in s(w)$, Q is true at $\langle s, v\rangle$ (notation: $s(w) \subseteq [\![Q]\!]^{s}$), then $must(P)(Q)$ must be true at $\langle s, w\rangle$ (notation: $[\![must(P)(Q)]\!]^{s,w} = 1$).

Inheritance When $s(w) \cap [\![P]\!]^{s} \neq \emptyset$, $s(w) \subseteq [\![Q]\!]^{s}$ implies $[\![must(P)(Q)]\!]^{s,w} = 1$

Inheritance says nothing more than: what holds throughout a non-trivial domain of quantification D is necessary relative to D (and any restriction of D). It is difficult to imagine a treatment of necessity modals as universal quantifiers on which Inheritance does not hold (and all quantificational treatments of necessity modals of which I am aware validate it).

It will help to make things clearer if we have a schematic truth-condition for *must* on deck. Here is a natural one:

Schematic must $[\![must(P)(Q)]\!]^{s,w} = 1$ iff $f_s([\![P]\!]^{s}, w) \subseteq [\![Q]\!]^{s}$

It is assumed only that f_s is a selection function selecting from $s(w)$ some range of possibilities compatible with a restriction. (For simplicity, this schematic truth condition makes use of Lewis' (1973) "Limit Assumption". But, since it is only playing an illustrative role here, nothing will ultimately turn on this.) Given a conditional $A \to C$, $f_s([\![A]\!]^{s}, w)$ is some range of relevant A-possibilities compatible with $s(w)$. For a material conditional analysis, $f_s([\![A]\!]^{s}, w) = \{w\}$ (if $w \in [\![A]\!]^{s}$) or $\emptyset$ (otherwise). For a pure strict conditional analysis, $f_s([\![A]\!]^{s}, w) = s(w) \cap [\![A]\!]^{s}$. Given this definition, Strong Centering and Inheritance could be expressed as follows:

Strong Centering (1) If $w \in [\![A]\!]^{s}$, then $f_s([\![A]\!]^{s}, w) = \{w\}$

Inheritance (1) When $s(w) \cap [\![P]\!]^{s} \neq \emptyset$, $s(w) \subseteq [\![Q]\!]^{s}$ implies $f_s([\![P]\!]^{s}, w) \subseteq [\![Q]\!]^{s}$

It is important to note that, though Closure holds, possibly $f_s([\![P]\!]^{s}, w) \neq f_s([\![P]\!]^{s}, v)$. In other words, we do not assume the conditional has a world-invariant truth-value across $s(w)$, because we do not assume the domain of quantification for the conditional is world-invariant across $s(w)$.

4.2. Integrating Probability

Probabilistic claims, while not modal, display certain fairly systematic relationships to the modal. For instance: (11) entails (12) and (12) entails (13).

(11) The keys must be on the table
(12) The keys are probably on the table
(13) The keys might be on the table

Simple domains of quantification seem not to have the sort of structure that would allow us to account for either the meaning of probabilistic claims like (12) or the suggested relationships between epistemic and probabilistic claims. To accommodate this, we must think of the information relevant for evaluating modal and probabilistic claims as not just qualitative, but also probabilistic, in nature: we should *extend* our notion of information state to cover information of both qualitative and probabilistic character (cf. Yalcin, 2010).[19]

Let W be the universe (and the power-set of W be the relevant Boolean algebra). Then:

Definition 1. *$\sigma = \langle s, Pr\rangle$ is an **information state** iff for all w:*

i. $\emptyset \subset s(w) \subseteq W$
ii. *$Pr(w)$ is a function from subsets of W into* $[0, 1]$ *such that:*
  a. $Pr(w)(s(w)) = 1$ *(**Normalization**)*
  b. If $P \wedge Q = \bot$, $Pr(w)(P \vee Q) = Pr(w)(P) + Pr(w)(Q)$ *(**Additivity**)*

Given a circumstance w, think of $\sigma = \langle s, Pr\rangle$ as specifying whatever information is relevant at w for evaluating information-sensitive claims, with $s(w)$ providing the qualitative dimension of that information, $Pr(w)$ the probabilistic dimension. $Pr(w)$ is normalized to $s(w)$ to account, inter alia, for entailments like those between (11) and (12) and (12) and (13). Note: we will sometimes write $\sigma(w)$ to designate $\langle s(w), Pr(w)\rangle$; I will be careless about referring to *both* such extensional objects and their intensional counterparts as 'information states.'

For our purposes it is important to define notions of support and update—the informational analogue of conditionalization—for an information state.

Definition 2. *The **update** of an information state $\sigma = \langle s, Pr\rangle$ with P at w (notation: $\sigma(w)[P]$) is defined as: $\langle s(w) \cap P, Pr(w)(\cdot|P)\rangle$*[20]

Definition 3. *A state σ **supports** P at w (notation: $\sigma(w) \models P$) iff $\sigma(w)[P] = \sigma(w)$.*

Definition 2 says simply that the result of updating an information state $\langle s(w), Pr(w)\rangle$ with p is an information state $\langle s'(w), Pr'(w)\rangle$, where $s'(w)$ is the result of eliminating the $\neg p$-worlds from $s(w)$ and $Pr'(w)$ is a new probability function obtained by conditionalizing $Pr(w)$ on p (and, implicitly, normalizing to $s(w) \cap [\![p]\!]^{\sigma}$).

The fundamental lemma for our own discussion follows directly in this setup.

Reflection For $\sigma = \langle s, Pr\rangle$: if $s(w) \cap [\![P]\!]^s \neq \emptyset$ and $\sigma(w) \models Q$, then $Pr(w)\,(must(P)(Q)) = 1$.

Proof. Suppose $s(w) \cap [\![P]\!]^s \neq \emptyset$ and $\sigma(w) \models Q$. Then $s(w) \subseteq [\![Q]\!]^\sigma$ and $[\![must(P)(Q)]\!]^{\sigma,w} = 1$ (Inheritance). Since $s(w)$ is Closed, for any $v \in s(w)$, $s(w) = s(v)$. So, for any $v \in s(w)$, $[\![must(P)\,(Q)]\!]^{\sigma,v} = 1$. Then $s(w) \subseteq [\![must(P)\,(Q)]\!]^\sigma$. Since $Pr(w)(s(w)) = 1$ (Normalization), and since $s(w) \subseteq [\![must(P)\,(Q)]\!]^\sigma$, $Pr(w)\,(must(P)\,(Q)) = 1$. □

4.3. Proving LT

It is, given these assumptions, not difficult to show that Kratzer conditionals satisfy the following restatement of LT. (We will in this section generally omit the world index for readability.)

K-LT For $\sigma = \langle s, Pr\rangle : Pr(must(A)\,(C)) = Pr(C)$

As already noted, establishing ACCEPT and REJECT would be sufficient for establishing LT.

K-ACCEPT For $\sigma = \langle s, Pr\rangle : Pr\ (must(A)(C)|C) = 1$

K-REJECT For $\sigma = \langle s, Pr\rangle : Pr\ (must(A)(C)|\neg C) = 0$

But something quite close to these claims seems (i) correct and (ii) easily proved with Reflection. After all, by CT, the probability of P on update with Q equals the probability of P given Q. But notice that the following are consequences of Reflection.

When $\sigma \models C$ and $s \cap [\![A]\!]^s \neq \emptyset$, $Pr(must(A)\,(C)) = 1$ Corollary(a)

When $\sigma \models \neg C$ and $s \cap [\![A]\!]^s \neq \emptyset$, $Pr(must(A)\,(C)) = 0$ Corollary(b)

A fortiori, we have:

K-ACCEPT* For $\sigma = \langle s, Pr\rangle$: when $Pr_{[C]}(A) > 0$, $Pr_{[C]}(must(A)(C)) = 1$

K-REJECT* For $\sigma = \langle s, Pr\rangle$: when $Pr_{[\neg C]}(A) > 0$, $Pr_{[\neg C]}(must(A)(C)) = 0$

Now suppose we take on the following restatement of CT (according to which the probability of C on update with A at w—which I will write as $Pr_{[A]}(w)(C)$, or, more simply, as $Pr_{[A]}(C)$—equals the conditional probability assigned to C given A at w).

K-CT For $\sigma = \langle s, Pr\rangle : Pr_{[A]}(C) = Pr(C|A)$

We can now use K-ACCEPT* and K-REJECT* to derive (slightly restricted versions of) K-ACCEPT and K-REJECT and, thus, LT (Restated).

Proof. Consider $\sigma = \langle s, Pr \rangle$ such that $Pr_{[C]}(A) > 0$ and $Pr_{[\neg C]}(A) > 0$. By K-CT:

$$Pr(must(A)(C)|C) = Pr_{[C]}(must(A)(C))$$
$$Pr(must(A)(C)|\neg C) = Pr_{[\neg C]}(must(A)(C))$$

By K-ACCEPT* and K-REJECT*: $Pr_{[C]}(must(A)(C)) = 1$ and $Pr_{[\neg C]}(must(A)(C)) = 0$. But, by the Law of Total Probability,

$$Pr(must(A)(C)) = [Pr(must(A)(C)|C) \cdot Pr(C)] + [Pr(must(A)(C)|\neg C) \cdot Pr(\neg C)]$$

Given what we have just shown, this reduces to $Pr(C)$. □

4.4. Avoiding LT: First Pass

Rothschild (2013, 62) suggests a strategy for disrupting this sort of proof that amounts to rejecting K-REJECT*. His discussion centers around the following example:

(14) If the car crashes at 35mph, the airbag will go off

We imagine the car has a defect causing its airbag to go off in such a crash. Rothschild asks us to consider a sample analysis of the conditional on which $A \to C$ expresses a kind of epistemic strict conditional, and which Rothschild glosses as "X knows that $A \supset C$." Rothschild grants that $A \to C$ satisfies ST with respect to the probability function supplied by $\sigma = \langle s, Pr \rangle$ at w. The question he then asks is this: does this conditional satisfy Independence (hence ST) with respect to the probability function supplied by $\sigma[\neg C] = \langle s_{[\neg C]}, Pr_{[\neg C]} \rangle$ at w? In other words, is it the case that (omitting the world index for readability): $Pr_{[\neg C]}(A \to C|A) = Pr_{[\neg C]}(A \to C)$? Rothschild insists: no. After all, given Strong Centering, the probability of $A \to C$ given $\neg C$ and A is just the probability of C and A given $\neg C$ and A, i.e., obligatorily 0. Thus, whenever $Pr_{[\neg C]}(A \to C) > 0$, the Independence condition *cannot* be satisfied: of necessity, $Pr_{[\neg C]}(A \to C|A) \neq Pr_{[\neg C]}(A \to C)$.

This strategy cannot, however, be extended to the analysis we are considering here. First, is it *ever* possible that $Pr_{[\neg C]}(A \to C) > 0$ (when $Pr_{[\neg C]}(A) > 0$)? I would say not (see §3.3).[21] More important, however, is that, given the Reflection Lemma, this is simply not possible *on a Kratzerian treatment* of $A \to C$ (given §4.3). I am not, to be clear, suggesting that ST should be assumed to hold in full generality for either $Pr_{[C]}$ or $Pr_{[\neg C]}$. I am, however, suggesting that the relevant *instances* of ST do hold for $Pr_{[C]}(A \to C)$ and $Pr_{[\neg C]}(A \to C)$, on the Kratzerian analysis—and must hold, more generally, on any plausible analysis.

4.5. Avoiding LT by Challenging CT

A better strategy for resisting LT is to challenge CT, the relevant version of which is reproduced here:

K-CT For $\sigma = \langle s, Pr \rangle : Pr_{[A]}(C) = Pr(C|A)$

The idea is that the proof we have constructed of LT is *unsound*, since CT cannot be retained in the probabilistic framework of §4.2.[22] Here, to get the feel of things, is a simple counterexample. Consider the probability, in a fair three-ticket lottery, for the absolute epistemic necessity: *ticket x must have won*. Intuitively it is 0; hence the probability of this proposition, relative to the probability function that results from conditionalizing on the proposition that ticket x won, is also, of necessity, 0.[23] Nevertheless, the probability of this claim, given *update* (in the sense of Definition 2) with the proposition that ticket x won seems to be greater than 0. Thus, K-CT fails.

To fill the point in, while it is ordinarily fine to understand $Pr(\cdot|X)$ suppositionally—to evaluate the probability of Y given X by supposing X and evaluating Y's probability—the present suggestion is that this is unsustainable in the system sketched here, hence that this understanding of conditional probability must be relinquished. The point is worth underscoring: if correct, it would seem that Kratzer's analysis of conditionals blocks a plausible formulation of the assumptions needed to run Lewis' proof of LT. Still, it bears noting that this way of avoiding LT is *substantive*: the proof of LT is formulated in a way that is prima facie plausible. It is just that one of its assumptions seems difficult to retain in the probabilistic framework of §4.2. This should worry the Kratzerian—it raises the specter of Triviality, even without yielding an immediate threat—particularly if the substantive claims involved can be called into question (as I will suggest they can be).

There are two different ways I can see this challenge being developed, one of which highlights the supposed context-sensitivity of the conditional, the other of which challenges the apparatus we have developed on grounds of inconsistency. I will take the latter challenge first.

4.5.1. Inconsistency? CT assumes that probabilities are updated by standard conditionalization. CT was crucial in our proof of LT: it allows us to go from K-ACCEPT* to K-ACCEPT, as well as from K-REJECT* to K-REJECT. Standard conditionalization, however, might seem inadequate for determining the probability of (e.g.) modals relative to an updated information state. As noted above, even when $Pr(must(\top)(P)) = 0$, it is a *consequence of Reflection* that $Pr_{[P]}(must(\top)(P)) = 1$, even though $Pr(must(\top)(P)|P)$, as standardly defined in terms of the ratio of the likelihood of $must(\top)(P) \wedge P$ to the likelihood of P, must equal 0. This cannot hold, on pain of inconsistency, unless CT is relinquished.[24]

4.5.2. Handling Inconsistency This challenge begins by *conceding* that standard conditionalization fails for certain cases—e.g. cases where a prior probability is extremal (1 or 0). So standard conditionalization must be replaced. Suppose that this pushes us—as I am inclined to admit—to a probability calculus in which the notion of conditional probability of a proposition P given Q is treated as *basic* or *primitive*, rather than in terms of ratios between unconditional probabilities. Such a shift is independently called-for—not at all ad hoc—if we think, as seems plausible, that the probability of, e.g., P given $must(\top)(P)$ is well-defined, even where $must(\top)(P)$ itself has probability 0. We might imagine a context in which we are quite sure that it might be sunny, but in which we suppose that it must be raining. What is the probability, at this context, given this supposition, that it is raining? Intuitively, it is 1; $must(\top)(P)$, after all, entails that P (cf. von Fintel and Gillies, 2010). It is, at the very least, not *undefined*.[25]

How best to develop a formal Primitive Conditional Probability (PCP) system is an open question, but one that we need not settle here. That is because it seems simply to be a desideratum on any reasonable PCP system that it secure a principle like the following:

CT (PCP) For $\sigma = \langle s, Pr \rangle : Pr_{[X]}(w)(Y|\top) = Pr(w)(Y|X)$

This principle says that the "unconditional" probability of Y, given an update with X—that is to say, the probability of Y conditional on a tautology, given an update with X—should equal the conditional probability of Y, given X. This seems very plausible. What, after all, *is* conditional probability? More precisely, what is the notion that a system in which conditional probability is treated as primitive is intended to model? Fundamentally, the conditional probability of Y, given X, seems as if it should track the probability of Y under indicative supposition that X. This is just what the principle in question seems to be enforcing.[26]

While it is a non-trivial challenge to actually state a consistent PCP system in which this principle is secured—see technical Appendix B—so long as it *is* secured, K-ACCEPT* will entail K-ACCEPT, K-REJECT* will entail K-REJECT, and the proof of Triviality given above will go through.

4.5.3. Context-Sensitivity? The prior section proposes to secure K-ACCEPT and K-REJECT by showing that the probability of the Kratzer conditional on *update* with C or $\neg C$ is, respectively, 1 or 0. Of course, however, K-ACCEPT and K-REJECT, reproduced here, are claims about the probability of $must(A)(C)$ *at the original state*, given C or $\neg C$, not claims about the the probability of $must(A)(C)$ relative to a new state—one updated with C or $\neg C$.

K-ACCEPT For $\sigma = \langle s, Pr \rangle : Pr(w)\,(must(A)(C)|C) = 1$

K-REJECT For $\sigma = \langle s, Pr \rangle : Pr(w)\,(must(A)(C)|\neg C) = 0$

It is natural to read this objection as an appeal to the *context-sensitivity* of the conditional: even if it is unobjectionable to endorse K-ACCEPT* and K-REJECT*, this is no case for endorsing K-ACCEPT and K-REJECT, if the "proposition expressed" (to a rough approximation) by $must(A)(C)$ relative to $\sigma[C]$ or $\sigma[\neg C]$ is distinct from the "proposition expressed" relative to σ.

The formulation of CT required for the proof of LT to go through is, in other words, this:

CT*(PCP) For $\sigma = \langle s, Pr \rangle : Pr_{[X]}(w)(\llbracket Y \rrbracket^{\sigma[X]}|\top) = Pr(w)(\llbracket Y \rrbracket^{\sigma}|\llbracket X \rrbracket^{\sigma})$

The present charge, however, is that we have said nothing to motivate *this* sort of principle. There is no general expectation that the probability of Y on update with X should equal the probability of Y conditional on X, if update on X would *affect the proposition* expressed by Y. The probability of an epistemic necessity at σ, *conditional on* its scope, may be 0, while its probability *on update* with its scope must, given Reflection, be 1. The suggestion is that there is just no contradiction here—even in a system that retains CT and the ordinary, ratio understanding of conditional probability—if the epistemic necessity expresses one proposition relative to σ, a different proposition relative to σ updated with the epistemic necessity's scope.[27]

4.5.4. Dealing with Context-Sensitivity I know no definitive response to the charge of context-sensitivity. It is, of course, well-known that appeals to context-sensitivity offer a way around purported Triviality results.[28] The most definitive thing I have to say in reply is that appeal to context-sensitivity would in fact bolster the main point of this paper: that Triviality results present the Kratzer semantics the same choices it presents any other propositional semantics for the conditional. This said, I think there is something constructive to say in reply to the charge of context-sensitivity. I will try to say some of it here.

The first thing to note is that information states—which I take to represent *subjective states* of qualitative and probabilistic information—bear no obvious relationship to *discourse contexts*, as standardly conceived. Contexts in semantics are theoretical entities that fix the content of context-sensitive expressions (e.g. indexicals); they are public, in the sense that their semantically relevant characteristics are generally taken to be part of the Common Ground.

If the content of the conditional is, as claimed, sensitive to the relevant information state, then the semantic function of the conditional is to describe features of the relevant information state. If the information state is interpreted as a state representing a subject's qualitative and probabilistic information, then the semantic function of the conditional is to describe features of the relevant subject's qualitative and probabilistic information. The first reaction to this state of affairs is liable to be discomfort: "This suggests that our semantic theory will not be able to assign a general meaning to $A \rightarrow C$ which applies across different credal states . . . This is not a happy situation, since $A \rightarrow C$, intuitively, has some sort of uniform meaning" (Rothschild, 2013: 53).[29]

Haven't I, however, on Kratzer's behalf, *relativized the truth condition of the conditional to an information state*, i.e., a representation of a subjective state of qualitative and probabilistic information.[30] Mustn't I allow that such states can go proxy for contexts in determining the content of a conditional on an occasion of use?

True, I have relativized the conditional's intension to a subjective state of qualitative and probabilistic information. But—while a reader would be forgiven for misconstruing this—in the case of a context-sensitive sentence P, a metalinguistic expression of the form $[\![P]\!]^{\sigma,w} = 1$ does not generally represent that P is *true in the actual context* of the agent whose information σ represents. Instead, the truth of

P at $\langle\sigma, w\rangle$ serves roughly to represent that agent's *estimate* of the truth-value of the proposition expressed in that context by P at w, as evaluated from the standpoint of *that agent's information*. Agents can estimate truth-values in this way because the truth of the proposition expressed by P is in fact information-sensitive—sensitive to relevant information—and the agent generally takes her subjective state of qualitative and probabilistic information to *accurately represent* the information relevant for determining the truth of P. Similarly:

- $[\![P]\!]^{\sigma}$ represents the set of indices of evaluation w such that the agent estimates the truth-value of the proposition expressed in the relevant context by P at w as true
- $Pr(w)([\![P]\!]^{\sigma})$ represents the agent's estimate of the objective likelihood (given the relevant information at w) of the proposition expressed in the relevant context by P

I do not deny, of course, that when $[\![P]\!]^{\sigma,w} = 1$, P is true at $\langle\sigma, w\rangle$. What I reject is the move from this to a claim like *the proposition expressed by* P *in the relevant context for the agent* σ *represents is true*. While there is a tight connection between an agent's estimates of truth and likelihood for an information-sensitive proposition and the actual truth and probability of that proposition—namely, both may be governed by semantic rules like Strong Centering, Closure, and Inheritance, as well as probabilistic rules like Reflection—I claim they are to be distinguished.

What is the case for this way of interpreting the apparatus developed in this paper? I take the position I am stating here to be based on a fairly minimal metasemantic commitment: namely, *Fallibilism* about information-sensitive expressions like modals and conditionals. Even when $[\![P]\!]^{\sigma,w} = 1$—i.e. the agent estimates that P is true at w from the standpoint of information σ—it may turn out that P is actually false. Similarly, an agent's estimate of the conditional or unconditional likelihood of an information-sensitive proposition may turn out to be incorrect. Obviously, if σ *determined* the proposition expressed by P in the agent's context, Fallibilism would not be a sensible position. But Fallibilism is a sensible (and, plausibly, correct) position.[31] So σ cannot determine the proposition expressed by P in the agent's context. A Fallibilist metasemantics is surely the most natural metasemantics for modals and conditionals, if one is committed to the notion that modals and conditionals are in the business of expressing propositions.[32]

Let us now return to the version of CT that context-sensitivity was claimed to undermine.

CT*(PCP) For $\sigma = \langle s, Pr\rangle : Pr_{[X]}(w)([\![Y]\!]^{\sigma[X]} | \top) = Pr(w)([\![Y]\!]^{\sigma} | [\![X]\!]^{\sigma})$

On the suggested interpretation, this version of CT says an agent's estimate of the likelihood of the truth of the proposition expressed by Y given the truth of the proposition X should track the agent's estimate of the unconditional likelihood of the truth of the proposition expressed by Y, when she updates with X. Phrased thus, I see nothing objectionable about this principle. I also see no incompatibility with the idea that a semantic theory of the conditional should "assign a general meaning to $A \rightarrow C$ which applies across different credal states."

There are possible metasemantic views concerning modals and conditionals on which they do express propositions and on which what I have called Fallibilism fails. I do not take such views to be very plausible, but then I might be wrong.[33] Here, then, is a possible loophole that advocates of the Kratzer conditional might exploit in trying to avoid LT. Of course I do not claim to have here closed off such loopholes to advocates of Kratzer conditionals, or to have shown definitively that Kratzer conditionals are as susceptible to problems of Triviality as any other propositional semantics of the conditional. However I do think I have shown that the standard story about Kratzer conditionals—on which they can avoid facing up to Triviality on what amounts to a technicality—must be revised.

5. Conclusion

A basic constraint of adequacy on a propositional semantics for the conditional is that it not trivialize a probability distribution defined for the conditional. It has generally been assumed that Kratzer's analysis of the conditional automatically meets this desideratum. This paper has tried to portray this consensus as overly simple. This paper has also suggested, somewhat more tentatively, that problems of Triviality—more precisely, problems of trivializing a probability distribution over a range of cases—*do arise* for Kratzer's analysis of the conditional, even without assuming that Stalnaker's Thesis holds in any kind of generality.

Where might an advocate of Kratzer conditionals—that is to say, one dissatisfied with the suggestion that the Kratzer analysis is committed to LT—lay the fault? Some obvious candidates for blame are: Inheritance (§2.2); Euclideanness (§2.2); or the notion that the information that is relevant for evaluating certain kinds of claims in natural language is not just qualitative, but also probabilistic, in nature (§4.2). Denying Inheritance rather clearly means giving up the entrenched analysis of modalities as quantifiers. I cannot see how to make such a denial plausible. Denying Euclideanness means facing up the possibility of the truth of claims that are, on the face of it (cf. Yalcin, 2007), modal contradictions, e.g., *it might be raining but maybe it must not be*. Again, I find it hard to see how to make such a denial plausible. Perhaps it is my suggested reconceiving of "relevant information" as both qualitative and probabilistic in nature that will be challenged? However, my initial assumptions here are so weak that I find it hard to see how to get along without them, once one allows that the information relevant for evaluating certain claims in natural language has both qualitative and probabilistic dimensions.

The best hope for the Kratzer analysis seems to me to be a challenge to CT—specifically a challenge to the notion that we should expect the probability of a Kratzer conditional, conditional on its consequent (or the negation of its consequent), to track the probability of a Kratzer conditional, on update with its consequent (or the negation of its consequent). The best way of filling out this challenge seems to me to appeal to context-sensitivity—specifically, the notion that assessing the probability of the conditional, on update with its consequent, should not be regarded as a guide to the probability of the conditional, conditional on its consequent. I have suggested that this response plausibly rests on a misunderstanding of the semantic and probabilistic apparatus this paper has developed.

Here, perhaps, I may be wrong. Triviality for Kratzer conditionals can conceivably be sidestepped. This, however, holds for basically any plausible propositional semantics of the conditional. Triviality is universally regarded as a reductio of *something*. The onus is on advocates of a propositional semantics of the conditional—a group that should now be understood to include advocates of the Kratzer semantics—to say what, if not their semantics, it might be a reductio of.

Appendix A: Bradley-Triviality

We understood Triviality to be the property an analysis of the conditional has when, roughly, it trivializes a probability distribution (and Restricted Triviality to be, roughly, the property of trivializing a probability distribution over a range of cases). But there are other ways for a semantics for the conditional to be trivial. If, for instance, the semantics for $A \to C$ requires that, when A does not entail C, that $A \to C$ entail C, the conditional is seemingly semantically trivial (see, e.g., Bradley, 2000). (I will expand on this below.) In this appendix, I argue that Triviality in this sense can also be shown for the Kratzer conditional.

A.1 Proof Something close to Bradley's Preservation Condition—according to which, roughly, the probability of a conditional must be 0 when the probability of its consequent is 0—is a corollary of Reflection.

K-Pres For $\sigma = \langle s, Pr\rangle$: if $Pr(w)([\![A]\!]^{\sigma}) > 0$ and $\sigma(w) \models \neg C$,

then $Pr(w)([\![must(A)(C)]\!]^{\sigma}) = 0$

From here it is not far to:

Bradley-Triviality If $must(A)(C)$ is persistent, and $[\![A]\!]$ and $[\![C]\!]$ are state-invariant, then, for any σ, $A \models_{\sigma} C$ or $must(A)(C) \models_{\sigma} C$

We will prove this using (i) the below definitions of the consequence relation $\models$ and Persistence, (ii) the Preservation condition, (iii) a basic Existence Lemma.

Definition 4. $X \models_{\sigma} Y$ *iff* $\forall w : \sigma(w)[X] \models Y$

Definition 5. *Given* $\sigma = \langle s, Pr\rangle$, $\sigma' = \langle s', Pr'\rangle$ *is an* ***extension*** *of* σ *iff (i)* σ' *is an information state, (ii) for any* w, $s'(w) \subseteq s(w)$ *and* $\exists Q : Pr'(w)(\cdot) = Pr(w)(\cdot|Q)$

Definition 6. P *is* ***persistent*** *iff, for any* $\sigma = \langle s, Pr\rangle$ *and extension* $\sigma' = \langle s', Pr'\rangle$ *of* σ *and* v, *if* $[\![P]\!]^{\sigma,v} = 1$, *then* $[\![P]\!]^{\sigma',v} = 1$.

Persistence amounts to the requirement that truth at w—more precisely, an estimate of a sentence as true at w—is preserved on addition of information to an information state.[34]

The Existence Lemma says simply that when $X \not\models_\sigma Z$ and $Y \not\models_\sigma Z$, there is some "accessible" Pr that treats X and Y as possible but Z as impossible (on the assumption that X is persistent, while Y and Z have state-invariant intensions):

Existence Suppose that X is persistent, $[\![Y]\!]$ and $[\![Z]\!]$ are state-invariant. Then if $X \not\models_\sigma Z$ and $Y \not\models_\sigma Z$, for some $\sigma' = \langle s', Pr' \rangle$ and v:

i. $Pr'(v)([\![X]\!]^{\sigma'}) > 0$
ii. $Pr'(v)([\![Y]\!]^{\sigma'}) > 0$
iii. $\sigma'(v) \models \neg Z$

Proof. Suppose, for some $\sigma = \langle s, Pr \rangle : X \not\models_\sigma Z$ and $Y \not\models_\sigma Z$. Then there is some $v, u \in s(w) : [\![X]\!]^{\sigma,v} = 1$; $[\![Z]\!]^{\sigma,v} = 0$; $[\![Y]\!]^{\sigma,u} = 1$; $[\![Z]\!]^{\sigma,u} = 0$. Let $s'(v) = \{v, u\}$, and $Pr'(v)(\cdot)$ be a probability measure for $s'(v)$ such that $Pr'(v)(\{v\}) > 0$ and $Pr'(v)(\{u\}) > 0$. Suppose $\sigma' = \langle s', Pr' \rangle$ is an information state. Then $\sigma' = \langle s', Pr' \rangle$ extends σ. If X is persistent (and $[\![Y]\!]$, $[\![Z]\!]$ state-invariant), $Pr'(v)([\![X]\!]^{\sigma'}) > 0$ and $Pr'(v)([\![Y]\!]^{\sigma'}) > 0$, while $\sigma'(v) \models \neg Z$. □

This follows from the general fact that when neither X nor Y entail Z, there is some estimate of the probabilities of X and Y according to which each exceeds that of Z.[35]

Given the Existence Lemma and the Preservation Condition, Bradley-Triviality follows directly:

Proof. Suppose $must(A)(C)$ is persistent, and $[\![A]\!]$ and $[\![C]\!]$ are state-invariant. Suppose for reductio: $A \not\models_\sigma C$ and $must(A)(C) \not\models_\sigma C$. Then, by Existence, for some $\sigma' = \langle s', Pr' \rangle$, v:

i. $Pr'(v)([\![must(A)(C)]\!]^{\sigma'}) > 0$
ii. $Pr'(v)([\![A]\!]^{\sigma'}) > 0$
iii. $\sigma'(v) \models \neg C$

But, by K-Pres, $Pr'(v)([\![must(A)(C)]\!]^{\sigma'}) = 0$. Contradiction. □

A.2 Significance According to Bradley-Triviality, on the assumption that $must(A)(C)$ is persistent (and $[\![A]\!]$ and $[\![C]\!]$ are state-invariant), whenever $A \not\models_\sigma C$, $must(A)(C) \models_\sigma C$. The Kratzer conditional, then, is said (very loosely) to entail its "consequent" whenever its "antecedent" does not logically entail its "consequent." This might seem a terrible result—contexts in which a possible antecedent fails to logically entail a possible consequent (hence in which the conditional has a trivial semantics) seemingly being preponderant—but this impression is in fact somewhat misleading.

Suppose, first, it were the case that, whenever $A \not\models_\sigma C$, then $[\![must(A)(C)]\!]^\sigma = \emptyset$. Then, of course, trivially $must(A)(C) \models_\sigma C$. Note, next, that the truth condition for the *pure strict conditional* $\Box(A \supset C)$ relative to σ is simply that A entail C:

$A \models_\sigma C$. Thus:

$$[\![\Box(A \supset C)]\!]^\sigma \; = \; \{w : A \models_\sigma C\} \; = \; \begin{cases} W, \text{ if } A \models_\sigma C \\ \emptyset, \text{ otherwise} \end{cases}$$

Thus, if a *Kratzer conditional were to logically entail the corresponding strict conditional...*

$$\forall\sigma : must(Q)(R) \models_\sigma \Box(Q \supset R)$$

... we would have a more or less direct explanation for why $must(A)(C) \models_\sigma C$ when $A \not\models_\sigma C$: $A \not\models_\sigma C$ would imply that $[\![\Box(A \supset C)]\!]^\sigma = \emptyset$, hence that $[\![must(A)(C)]\!]^\sigma = \emptyset$).

More generally, Bradley-Triviality can be independently derived for any analysis of the conditional on which a conditional entails the corresponding pure strict implication (so long as the pure strict implication receives an interpretation along the lines suggested here). Thus, Bradley-Triviality is hardly the end of the world for, e.g., advocates of strict conditional analyses of the indicative conditional.[36] It is, in fact, an *independent* consequence of their view.

For the Kratzerian, this eases the pain of Bradley-Triviality a bit. But not, I think, by very much. So long as the Kratzer conditional is (as it must be) bounded "from above" by strict implication (and so long as we assume that it does not entail its "consequent"), the Kratzer conditional *must be* truth-conditionally equivalent to the pure strict conditional.

$$must(Q)(R) \; \equiv \; \Box(Q \supset R)$$

Taken at face value, Bradley-Triviality reveals a basic underlying incoherence in the Kratzer semantics for conditionals. The Kratzer semantics makes liberal use of devices (e.g. the *ordering source*, which is roughly analogous to the similarity relation deployed by Stalnaker and Lewis; see Kratzer 1979, 1981; Stalnaker 1968; Lewis 1973) to *block* its equivalence with an ordinary pure strict conditional semantics, and to account for facts for which a strict conditional semantics cannot, on the face of it, account (e.g., the felt invalidity of Strengthening of the Antecedent, Transitivity, and Contraposition). But—given the friendliest explanation of why $must(A)(C) \models_\sigma C$ when $A \not\models_\sigma C$—the Kratzer conditional must be equivalent to a pure strict conditional. In any case, then, where the Kratzer semantics delivers a verdict of truth for a conditional and the pure strict conditional semantics does not, then, we will—given very natural assumptions—be able to use Bradley-Triviality to *derive a contradiction*. On pain of contradiction, the ordering source can *never* play an active role in the semantics of a Kratzer conditional. It is much as if we had shown that the semantics of the Lewis-Stalnaker conditional $A \,\Box\!\!\rightarrow C$ required that it entail $\Box(A \supset C)$. Lewis-Stalnaker conditionals would be forced to validate Strengthening of the Antecedent, Transitivity, and Contraposition. Since the Kratzer semantics, like the Lewis-Stalnaker semantics, is designed, in part, to invalidate exactly these inference patterns (Kratzer, 1979, 1981), the parallel here is in fact fairly direct.[37]

Insofar as problems of Triviality—namely, the problems of *theoretical incoherence* to which problems of Triviality give rise—do not arise for a different (e.g. non-propositional) analysis of the conditional, there is a case for thinking (contra linguistic orthodoxy) we should prefer it to the Kratzerian alternative.[38]

A.3 Persistence and Context-Sensitivity We have proved that if $must(A)(C)$ is persistent (and $[\![A]\!]$ and $[\![C]\!]$ are state-invariant), then $A \models_\sigma C$ or $must(A)(C) \models_\sigma C$. The stipulation that $[\![A]\!]$ and $[\![C]\!]$ are state-invariant is not generally up for challenge. But what about the stipulation that $must(A)(C)$ is persistent?

Given the interpretation of the apparatus sketched in §4.5.4, I understand the assumption that the conditional is persistent as the assumption that, if an agent with information σ estimates the conditional's truth-value at v as true, and the agent gains some information (shifts to some extension of σ), the agent should continue to estimate the conditional's truth-value at v as true.

However, this might seem straightforwardly contradicted by the following sort of case. I think the marble you have concealed in your hand might be red. I also think that if it's red, it's large; I estimate the truth-value of this conditional at v as true, for every v compatible with my information. I then come to believe that the marble couldn't possibly be large. I still think it might be red. I must, then, deny that it could be large if it's red. But, unless the conditional entails that the marble is large—which surely it does not—v may still be compatible with my information. But, of course, I no longer estimate the truth-value of the conditional *if the marble is red, then it is large* at v as true.

It should be clear that there is a kind of appeal to context-sensitivity involved here: if I have gained information, the only way I can change my mind about the truth of the conditional at v is if the conditional, roughly, expresses something different relative to my new information.

We now find ourselves again in the dialectic of §4.5.4. If the suggestion is that acquiring information changes the proposition expressed by the conditional, we have subjectivism (and infallibilism) about the conditional. Suppose we reject subjectivism. Then we are left with the suggestion that it is possible for an agent to estimate the truth-value of a proposition that is *independent of her information* at v as true, for her information to strictly *grow*, and subsequently for her to estimate the truth-value of that same proposition at v as false. This would seem obviously incoherent.[39]

As before, a metasemantics on which conditionals express *information-sensitive propositions*, in the special sense of Rothschild and Yalcin (2012), may offer a way around this. I do not want to remark on this idea's prospects here. My goal here is to clearly describe the options for the advocate of Kratzer conditionals interested in avoiding Bradley-Triviality.

Appendix B: PCP

This appendix briefly describes a PCP system in which the following principle is secured.

CT(PCP) For $\sigma = \langle s, Pr \rangle : Pr_{[X]}(w)(Y|\top) = Pr(w)(Y|X)$

The presentation here is due largely to Weisberg (2011). The system is a version of a Popper-Rényi system. In such a system, the value of the unconditional probability function for a proposition P may be understood in terms of the probability of P given a tautology $\top$: $Pr(w)(P) := Pr(w)(P|\top)$. Update and Support are understood as in Definitions 2–3. Information states and their PCP functions are axiomatized as follows:

Definition 7. *$\sigma = \langle s, Pr \rangle$ is a* ***PCP information state*** *iff for all w:*

i. $\emptyset \subset s(w) \subseteq W$
ii. $Pr(w)(\cdot|X)$ *is a function from subsets of W into* $[0, 1]$ *such that for any $X, Y, Z \subseteq W$:*
  a. If $\sigma(w)[X] \models Y$, $Pr(w)(Y|X) = 1$ *(Normalization)*
  b. If $P \wedge Q = \bot$, $Pr(w)(P \vee Q|X) = Pr(w)(P|X) + Pr(w)(Q|X)$ *(Additivity)*
  c. $Pr(w)(Y \wedge Z|X) = Pr(w)(Y|X) \cdot Pr(w)(Z|X \wedge Y)$

In such a system, Reflection and the Law of Total Probability are retained. CT is a trivial theorem:

Proof. By Definition 2, $Pr_{[X]}(w)(Y) = Pr(w)(Y|X)$. But $Pr_{[X]}(w)(Y) = Pr_{[X]}(w)(Y|\top)$. □

The system seems to avoid the problem of inconsistency described in §4.5.1. The ratio is relinquished, so it seems there is no independent pressure to hold that $Pr(must(\top)(P)|P) = 0$.

This system may have its own problems of (in)consistency. But little rides on this: *any* consistent system securing an intuitively correct notion of conditional probability—that Y's probability given X tracks Y's probability under indicative supposition that X—will secure LT, given the Law of Total Probability and Reflection. Any such system will also run into the issues of context-sensitivity discussed above. Thus none of the issues I explore in this paper depend on the particular system sketched here.[40] This said, I will review the putative inconsistency and try to make it palatable.

Notice: $Pr(P \wedge must(\top)(P)) = Pr(P) \cdot Pr(must(\top)(P)|P) = Pr(P) \cdot 1 = Pr(P)$. Thus, when $Pr(P) > 0$, $Pr(must(\top)(P)|P) = \frac{Pr(P \wedge must(\top)(P))}{Pr(P)}$. On one way of looking at things, then, the way this system secures the desired result—that $Pr(must(\top)(P)|P) = 1$—is, in fact, *not* by relinquishing the ratio definition of conditional probability in such cases (though the ratio does not, of course, hold generally). Rather, it is by predicting that $Pr(P \wedge must(\top)(P)) = Pr(P)$.

But now notice that the probability of the conjunction of P and $must(\top)(P)$ can have different values depending on which conjunct is treated as "primary." As an illustration, assuming $Pr(X \wedge Y) = Pr(X) \cdot Pr(Y|X)$, whenever $Pr(P) > 0$ and $Pr(must(\top)(P)) = 0$, $Pr(P \wedge must(\top)(P)) = Pr(P)$, while $Pr(must(\top)(P) \wedge P) = 0$.

Does this make the resulting system is inconsistent? Not per se. It means that conjunctions *do not commute* in the probability calculus. This, however, seems to me to be desirable, given the fact that conjunction itself fails to be commutative once we introduce expressions that are broadly "sensitive" to the information that has been introduced into a discourse, like modals.[41]

If, however, the reader is uncomfortable with this form of non-commutativity, an alternative system in which $Pr(must(\top)(P)|P) = 1$—and, more generally, Y's conditional probability given X tracks the probability of X under indicative supposition that Y—may be entertained.

Notes

[1] I use the phrase 'Kratzer conditional,' loosely, to refer to Kratzer's restrictor analysis of the simple, unembedded indicative conditional. Thus, when I say something like "problems of Triviality arise for Kratzer conditionals," I will mean something like "Triviality results can be constructed for the analysis of the simple, unembedded indicative conditional that Kratzer advocates." I will further refine the use of this phrase below.

[2] Stalnaker's Thesis—roughly, the thesis that a conditional's probability equals the corresponding conditional probability—is variously known as *The Equation* and *Adams' Thesis*. See Hájek (2012) for why the latter should be dispreferred.

[3] This presentation is largely due to Bennett (2003). Note: when there is no problem doing so, I will be sloppy about whether Pr is defined over elements of an implicitly given Boolean algebra, or over sentences that characterize elements of that algebra. Note also that obvious definedness conditions of probability claims—e.g., that ST, the Conditionalization Thesis (and so the If-And Lemma and LT) are meant to hold only when $Pr(A) \neq 0$; or that Lewis-Triviality is meant to hold only when $Pr(A \wedge C) > 0$ and $Pr(A \wedge \neg C) > 0$—are generally taken as understood.

[4] Its level of acceptance is such that Kaufmann (2012) is able to argue, not too outlandishly, for a truth-conditional analysis *of imperatives* partly on the grounds that (i) imperatives occur in genuine indicative conditionals, (ii) genuine indicative conditionals have the meaning that Kratzer proposed for them (hence indicative conditional imperatives involve the restriction of the quantificational operator expressed by the imperative). Similarly, the fact that the default Kratzer analysis for so-called "anankastic" conditionals assigns them manifestly incorrect truth conditions is occasionally taken as evidence, not against the Kratzer semantics, rather for logical forms for anankastics in which overt matrix clause deontic modals scope (quite exceptionally) under covert epistemic modals (see the discussion and references in von Fintel and Iatridou, 2005). For a somewhat different view of these dialectics, see Charlow (2013, 2014). (See also Isaacs and Rawlins, 2008; Charlow, 2010 for attempts to make the idea of "domain restriction" do work for operators that do not have quantificational force.)

[5] Kratzer seems to confirm this interpretation in (pc): "It makes no sense to look at an English conditional . . . and start wondering about the probability of the proposition expressed by that sentence . . . That [is] precisely the kind of mistaken conception I targeted with my work on conditionals."

[6] Égré and Cozic (2011, 25) mention a subtler version of this position, according to which an expression like $Pr((A \to C)|C)$ may be undefined on the Kratzer analysis, because (i) it is equivalent to $\frac{Pr((A\to C)\wedge C)}{Pr(C)}$; (ii) the probability of a Boolean compound $(A \to C) \wedge C$ is hard to make sense of for Kratzer: "If the role of the if-clause A is to restrict the domain of an operator, then here it appears that A cannot restrict the domain of the probability operator... since the conditional clause in this case is syntactically dominated by the conjunction." But this can be sidestepped by understanding $Pr((A \to C) \wedge C)$ to estimate the probability of the conjunction of the *proposition* that is (according to Kratzer) expressed by the bare conditional $A \to C$ in some standard context of utterance with the

proposition expressed by C. If the conditional expresses a proposition, we may (indeed, should) flatly insist that the probability of that proposition must, e.g., obey the Law of Total Probability.

[7] I will not cite further linguistic data to support this. This is not fundamentally a linguistic claim, and one worries that advocates of Kratzer conditionals will treat the presence of any linguistic material outrunning the conditional as signaling the presence of a potentially restrictable quantifier.

[8] Douven (2007) (addressing, actually, Richard Bradley's Preservation Condition, on which more in Appendix A) suggests we may be misled into thinking this by the fact that an indicative conditional $A \to C$ (whose logical form is, suppose, simply $\neg A \vee C$) *conventionally implicates* that the probability of C given A is high. This might be able to explain the linguistic data mentioned in the prior paragraph. It does not, however, make palatable the suggestion that $Pr((A \to C)|\neg C) = Pr(\neg A|\neg C)$. (Douven, to be fair, does not claim otherwise.) Moreover, the explanation's traction depends on the dubious assumption that the material conditional analysis is a viable account of the meaning of the conditional.

[9] For a broadly similar claim, see Égré and Cozic (2011). (But see Rothschild, 2013) for a suggestion that a good theory of the conditional should indeed make room for such probabilities.) Égré and Cozic suggest that Kratzer's analysis of sentences like (9)—on which it has exactly the truth condition we would expect given ST (see below), but bare conditionals like *if it doesn't snow it will rain* have trivial probabilities—is a point in favor of the notion that indicative antecedents are domain-restrictors. That is consistent with the argument of this paper: I will not be taking issue with the idea that indicative antecedents are restrictors. My topic is the *specific truth condition* that Kratzer assigns basic indicative conditionals, not the manner in which that truth condition is compositionally derived. (Of course, as a proponent of a non-restrictor analysis of indicative antecedents, I would take issue with Égré and Cozic. But that goes beyond the scope of this paper.)

[10] The position under consideration in this section—embrace LT, reject ST—might look at least potentially incoherent, since it might seem that ST is supported by the same sorts of considerations that would lead us to accept that $Pr((A \to C)|C) = 1$ and $Pr((A \to C)|\neg C) = 0$. We might seem to assess these claims by asking whether $A \to C$ is acceptable *if*, on the one hand, C and, on the other hand, $\neg C$. In other words, we might seem to determine that $Pr((A \to C)|C) = 1$ and $Pr((A \to C)|\neg C) = 0$ by implicitly deploying two instances of ST. The suggestion is, however, incorrect. Given a commitment to a propositional semantics for the conditional, it may well be $A \to C$ must receive probability 1, conditional on C, and probability 0, conditional on $\neg C$—things that do follow from ST—without ST holding generally. Conditionals might, e.g., always receive extremal probabilities (1 or 0), but, conditional on C, the extremal probability is forced to be 1, and, conditional on $\neg C$, the extremal probability is forced to be 0. (I will address this below.) This avoids the most unpalatable consequence of LT, i.e., the trivialization of the probability distribution. But of course this is not a way of avoiding LT altogether. The notion that simply adopting Kratzer's analysis would allow one to avoid LT is, again, hard to credit.

[11] See Rothschild (forthcomingb) for an explanation of why the Kratzer analysis—if not the semantics she in fact defends (Kratzer, 1979, 1981), which is equivalent to Lewis' ordering semantics (see Lewis, 1981)—is compatible with all of these.

[12] A common objection to this entailment involves so called "causal" or "nomological" readings of the indicative conditional. My view is that such readings bear *strictly more content* than the basic indicative conditional. (A suggestive data point: it is generally not possible to criticize a conditional claim, on grounds of *falsity*, in a situation where its antecedent and consequent are each true, but the truth of the antecedent is not, e.g., causally responsible for the truth of the consequent.) Note that I am not here presupposing the incorrectness of, e.g., the Relevance Theoretic treatment of causal or nomological conditionals, if these indeed exist. I am only suggesting that a preponderance of conditionals in natural language probably obey a principle of Strong Centering, while directing our focus to these (and away from those conditionals that do not).

[13] For instance, learning that Alex has taken the serum may raise one's confidence in the conditional *if Alex takes the serum, he will live*, if Alex would not take the serum unless it were effective. In this, conditionalizing on the proposition that Alex has taken the serum puts one in a position to be confident that if Alex takes the serum, he will live. The argument here rests only on the claim that there are many cases in which Independence does seem to hold.

[14] I am *not* here making the point that, if $Pr(A \to C) = Pr(C|A)$, one can derive LT for $A \to C$ (by the proof in §2.1), and, subsequently, show $Pr(C|A) = Pr(C)$. From the fact that ST holds for $Pr(A \to C)$, it does not necessarily follow that it holds for $Pr(A \to C|C)$ and $Pr(A \to C|\neg C)$. For instance, Rothschild (2013: 62) argues that *Independence* may fail for $Pr(A \to C|C)$ and $Pr(A \to C|\neg C)$, hence ST may fail, hence the relevant applications of the If-And Lemma may fail. (More on this below.) Thanks to Jonathan Weisberg for discussion here.

[15] To simplify, I avoid mention of the role played by the Ordering Source. (This affects none of the points I will make here, which depend only on understanding the modalities as quantifiers.) The reader should recognize, however, that the Ordering Source is playing a critical (if implicit) role: as the Similarity Relation of Lewis (1973); Stalnaker (1968) distinguishes the truth condition of a Variably Strict conditional from that of an ordinary strict conditional, it is the Ordering Source that ultimately differentiates the truth condition of a Kratzer conditional from that of an ordinary strict conditional (see Kratzer, 1979, 1981). This will be important in our later discussion of Bradley (2000) (Appendix A).

[16] Here I wish to avoid the assumption that there is a domain of quantification in the sense required by the Limit Assumption (Lewis, 1973) (which would, inter alia, require the domain to be finite). I will only make assumptions about the domain of quantification to which Kratzer helps herself, or which it is hard to see getting along without.

[17] The reader should note: when I use such examples in discussing constraints like Euclideanness, *might* and *must* are intended to express absolute epistemic possibility and necessity—the truth of the prejacent at some possibility in $s(w)$ and at each possibility in $s(w)$, respectively. Thus they are importantly different from the restricted epistemic modals considered in the main text—modals of the form $must(X)(Y)$—which do not, according to Kratzer, express absolute epistemic possibility and necessity, but rather—due to the role of the ordering source (cf. note [15]) in refining the domain of quantification—restricted versions of these absolute notions.

[18] Closure follows from Euclideanness together with any constraint sufficient to guarantee that $s(w)$ is Transitive. This is important since some—including Kratzer (1991b)!—are inclined to deny Reflexivity (though see von Fintel and Gillies, 2010 for a compelling defense of it). Assuming anything that entails Closure might appear quite strong, since, for any information state $\sigma = \langle s, Pr \rangle$ and possibility w, this will commit one to assigning claims expressing bare characteristics of $s(w)$ (e.g., a claim of absolute epistemic necessity) probability 1 or 0. We will see this below, but here is a brief sketch of why. Closure implies that if X is absolutely epistemically necessary at $s(w)$, it is also absolutely epistemically necessary at $s(v)$, for any $v \in s(w)$—and, conversely, if X is not absolutely epistemically necessary at $s(w)$, it is also not absolutely epistemically necessary at $s(v)$, for any $v \in s(w)$. (The former fact is something analogous to positive introspection for absolute epistemic necessity, the latter analogous to negative introspection for absolute epistemic necessity.) This is enough to guarantee that a claim of absolute epistemic necessity or possibility must have probability 1 or 0. Perhaps where these assumptions lead is surprising, but it is hard to see how to resist it.

[19] There is room for disagreement about how to account for the interaction of epistemic and probabilistic operators in natural language. I will try not to assume anything contentious here.

[20] For the sake of understanding update on modal sentences, I am inclined to allow that, in such cases, update at $\sigma(w)$ does not necessarily proceed by adding the set of indices expressed by the modal sentence relative to σ to the relevant information. Rather, update consists in moving to the nearest (according to some measure of nearness) information state in which the modal sentence is supported. This complication is largely irrelevant, for our purposes. While it does bear on issues like what probability to assign P conditional on $must(\top)(P)$ (see §4.5.1), I will discuss it no further here.

[21] Voices like Weatherson (ms); Douven (2007) notwithstanding. Weatherson presciently connects the issues here to those surrounding Euclideanness, remarking that "As a general rule, the principle $\Diamond q \to \Box\Diamond q$ is wildly implausible when $\Diamond$ is epistemic *might*." But I take Yalcin (2007) to have shown that Weatherson was wrong about this principle.

[22] This objection is originally due to conversation with Daniel Rothschild.

[23] This remark assumes that B's conditional probability given A is defined in terms of the ratio of (i) the unconditional probability of $A \wedge B$, (ii) the unconditional probability of A. I will, to be clear, suggest below that we reject this assumption.

[24] Another way of putting the point is that our definition of update (Definition 2) does not guarantee that $\sigma(w)[P]$ is an information state (in the sense of Definition 1). Definition 2 defines $\sigma(w)[P]$ to be $\langle s(w) \cap P, Pr(w)(\cdot|P)\rangle$. Notice, however, that if $Pr(w)(must(\top)(P)|P) = 0$, it follows that $Pr(w)(\cdot|P)$ is not in fact Normalized to $s(w) \cap P$.

[25] Considerations in favor of primitive conditional probability are reviewed in Hájek (2003); Weisberg (2011).

[26] Alan Hájek (pc) raises a worry about this: is the probability of $must(\top)(P)$ under indicative supposition that P really equal to 1? Suppose (indicatively) that your partner is cheating on you. Does it follow that it is known to you that your partner is cheating on you? Perhaps not, but all this shows is that epistemic modals are not inter-paraphrasable with claims about one's knowledge. Attempting to entertain $\neg must(\top)(P)$ under the indicative supposition that P seems to be inconsistent: you cannot consistently suppose that your partner is cheating on you, but that s/he might not be (cf. Yalcin, 2007).

Brian Leahy and Timothy Williamson (pc) raise a different, but related, worry: if, as I argue, $Pr(must(P)(Q)|Q) = 1$, then $Pr(must(P)(Q) \wedge Q) = Pr(Q)$, in which case Q entails (very counterintuitively) $must(P)(Q)$. Similarly, if, as I argue, $Pr(must(P)(Q)|\neg Q) = 0$, then $Pr(must(P)(Q) \wedge \neg Q) = 0$, in which case $must(P)(Q)$ entails (also very counterintuitively) Q. While I agree that these things are counterintuitive, their conjunction implies (and is implied by) Lewis' Triviality Result for the Kratzer conditional (i.e. K-LT). It is not an objection to a paper charging the Kratzer conditional with Triviality that it succeeds in its aim.

[27] This response's plausibility is perhaps enhanced by the fact that Kratzer herself is generally taken as suggesting that the truth condition of a modal (like that covertly expressed by the conditional) is context-dependent. For Kratzer (1981) (see also the updated discussion in Kratzer, 2012), the "discourse context" supplies a modal base and ordering source—actually, she says, given the vagueness of the average discourse context, a range of admissible modal bases and ordering sources—which settle both the flavor of the modal (epistemic, deontic, bouletic, teleological, etc.) and its specific truth-condition in context.

[28] For a brief overview, see Korzukhin (2014). Korzukhin shows that context-sensitive accounts of the conditional—given plausible constraints on the shape they can take—have difficulty holding onto ST. This is not directly relevant to this paper's argument, since we have declined to treat ST as a desideratum for a theory of the conditional.

[29] Complicating this argument is the fact that Kratzer herself famously endorses—in attempting to deal with the Sly Pete conditionals of Gibbard (1981)—a kind of subjectivism (namely, Speaker Subjectivism) for the conditional. She writes:

> Zack's and Jack's utterances are both true. But again: If Zack had uttered If Pete called, he lost (given his evidence) and if Jack had uttered If Pete called, he won (given the situation he was in) the two men's claims would have been both false. Like the epistemic modals above, these particular indicative conditionals are interpreted with respect to the evidence available to their utterers. (1991a: 655)

Few, however, have elected to follow her here. Speaker Subjectivism is generally recognized as a dead letter in the semantics of both modals and conditionals (see Yalcin, 2007; MacFarlane, 2011; Rothschild, forthcominga). The prospects are certainly better for *Contextualism* for modals and conditionals, according to which the content of a modal or conditional is context-of-utterance-sensitive, rather than speaker-sensitive. But, again, if information states serve to represent an agent's qualitative and probabilistic information, they will seem ill-suited to playing a content-determining role in the semantics of conditionals. Thanks here to Una Stojnic.

[30] More precisely, I have relativized the truth condition of the conditional to the qualitative dimension of the information; the probabilistic dimension is semantically idle.

[31] In claiming this I do not mean to beg questions against the Contextualist about, e.g., epistemic modality. It is true that retraction data—wherein an agent rejects an epistemic modal she once accepted as incorrect—is often cited by Relativists as motivation for their view against Contextualism (see esp. MacFarlane, 2011). Nevertheless, Contextualists do not, so far as I know, react to this data by dismissing it. Rather, they attempt to account for it within their theory.

[32] How to make sense of Fallibilist intuitions about a kind of claim K on a non-propositional account of the semantics of K-claims is an interesting project about which much has been written. I will not address this here.

[33] It might, e.g., turn out that conditionals express *information-sensitive propositions*, in the special sense of Rothschild and Yalcin (2012). Indeed, this might be an appealing way of developing Kratzer's own subjectivist-sounding remarks on Gibbardian standoffs. Kratzer's recent emphasis on the ubiquity of "unresolved indeterminacy" and "unspecified or underspecified situations of evaluation" in resolving the relevant dimensions of context (Kratzer, 2012, 44–5, 103–4) seems potentially relevant here. One natural reading of these remarks is, perhaps, Fallibilist—highlighting an agent's *uncertainty* about the relevant features of the context—but I can imagine alternative readings. For what it is worth, I see the discussion of Rothschild and Yalcin as drawing attention to a formal possibility for the interpretation of theories, like that of Veltman (1996), that are normally understood in dynamic terms. It is a possibility of which I have, however, found it hard to make theoretical sense.

[34] Persistence tends strongly to be a feature of proposition-expressing sentences. On, e.g., the classic treatment of Veltman (1996), only epistemic possibilities are potentially non-persistent in this sense. According to the most common interpretation of Veltman, this makes his treatment of epistemic possibility modals *non-propositional* (if not non-truth-conditional).

[35] This justification may seem unintuitive, and the discussion of it at Bradley (2000, 220), while certainly correct, will unfortunately not help anyone who remains puzzled. To better explain, suppose X and Y jointly entail Z though neither individually entails Z. Then according to Existence, there is a coherent probability function that assigns X and Y each a nonzero probability, while assigning their logical consequence probability 0. But this might seem probabilistically incoherent: if X and Y entail something one regards as utterly ruled out, one knows that either X or Y must be false, hence cannot assign both X and Y nonzero probability. (Indeed this is precisely the confusion Bennett (2003) self-reports (at 105).) While there is something appealing about this reasoning, its mistake is obvious enough. It is quite right that if X and Y entail something one regards as utterly ruled out, one must assign $(X \wedge Y)$ probability 0. But it remains at least possibly probabilistically coherent for one to assign both $(X \wedge \neg Y \wedge \neg Z)$ and $(\neg X \wedge Y \wedge \neg Z)$ nonzero probability. Hence it is at least possibly probabilistically coherent to assign both X and Y nonzero probability.

[36] It might be alleged that the type of strict conditional analysis under consideration, on which indicative conditionals express global properties of σ—i.e. $[\![\Box(Q \supset R)]\!]^{\sigma} = W$ (if $[\![Q]\!]^{\sigma} \subseteq [\![R]\!]^{\sigma}$) or $[\![\Box(Q \supset R)]\!]^{\sigma} = \emptyset$ (if $[\![Q]\!]^{\sigma} \nsubseteq [\![R]\!]^{\sigma}$)—is in fact not a truth-conditional analysis at all, since the indicative conditional does not, on this view, have a robust possible worlds truth condition. Possible worlds truth conditions are hardly, however, the only kind of truth condition (cf. Yalcin, 2012, 1009–10).

[37] Relevant here is Lewis (1981)'s proof of an equivalence between his "ordering semantics" and Kratzer's "premise semantics."

[38] As noted above, it may be that the Kratzerian analysis—more precisely, the notion that *if*-clauses are domain restrictors—looks better when we consider embedded conditionals (under, e.g., adverbs of quantification, as suggested in Khoo, 2011; see also Égré and Cozic, 2011). I doubt that is right, but I cannot do justice to this issue here. Note, however, that, whatever the outcome of this issue, there will remain an evident problem of theoretical incoherence that the Kratzerian analysis needs to address.

[39] Unless the agent has revised one or more priors. This possibility, however, is supposed to be ruled out by the assumption that the agent's information has strictly grown.

[40] I am not certain that there *is* a consistent system in which Y's probability given X can generally track probability of Y under indicative supposition that Y (if X and Y are allowed to range freely over non-modal and modal propositions). If it turns out that there is no such system, I think the best explanation is that we must restrict Pr's application to non-modal propositions—that is to say, granting Kratzer's assumption that conditionals are equivalent to modal quantifiers, to non-conditional propositions—which is, I have noted, where I see Triviality pushing us to begin with.

[41] See Veltman (1996); von Fintel and Gillies (2007); Klinedinst and Rothschild (2012). Rothschild and Yalcin (2012) sketch an alternative explanation of non-commutativity that connects to their own sophisticated understanding of information-sensitivity (see §4.5.4). My present suggestion is that any disagreement here will track the disagreement discussed in §4.5.4.

References

Austin, J. L. 1961. The meaning of a word. In J. O. Urmson and G. J. Warnock (eds.) *Philosophical Papers by J. L. Austin*. Oxford: Clarendon Press.

Barwise, Jon and Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4: 159–219. doi:10.1007/BF00350139.

Bennett, Jonathan. 2003. *A Philosophical Guide to Conditionals*. New York: Oxford.

Bradley, Richard. 2000. A preservation condition for indicative conditionals. *Analysis* 60: 219–222. doi:10.1093/analys/60.3.219.

———. 2007. A defence of the Ramsey test. *Mind* 116: 1–21. doi:10.1093/mind/fzm001.

Charlow, Nate. 2010. Restricting and embedding imperatives. In M. Aloni, H. Bastiaanse, T. de Jager, and K. Schulz (eds.) *Logic, Language, and Meaning: Selected Papers from the 17th Amsterdam Colloquium*. Amsterdam: ILLC. doi:10.1007/978-3-642-14287-1_23.

———. 2013. Conditional preferences and practical conditionals. *Linguistics and Philosophy* 36: 463–511. doi:10.1007/s10988-013-9143-3.

———. 2014. Logic and semantics for imperatives. *Journal of Philosophical Logic* 43: 617–664. doi:10.1007/s10992-013-9284-4.

Douven, Igor. 2007. On Bradley's preservation condition for conditionals. *Erkenntnis* 67: 111–118. doi:10.1007/s10670-007-9043-4.

Égré, Paul and Mikaël Cozic. 2011. If-clauses and probability operators. *Topoi* 30: 17–29. doi:10.1007/s11245-010-9087-y.

Ellis, Brian. 1978. A unified theory of conditionals. *Journal of Philosophical Logic* 7: 107–124. doi:10.1007/BF00245924.

Gibbard, Allan. 1981. Two recent theories of conditionals. In W. Harper, R. Stalnaker, and G. Pearce (eds.) *Ifs*, 211–247. Dordrecht: Reidel.

Gillies, Anthony S. 2010. Iffiness. *Semantics and Pragmatics* 3: 1–42. doi:10.3765/sp.3.4.

———. 2012. Indicative conditionals. In D. G. Fara and G. Russell (eds.) *Routledge Companion to the Philosophy of Language*, 449–65. New York: Routledge.

Hájek, Alan. 2003. What conditional probability could not be. *Synthese* 137: 273–323. doi:10.1023/B:SYNT.0000004904.91112.16.

———. 2012. The fall of 'Adams' thesis'? *Journal of Logic, Language, and Information* 21: 145–61. doi:10.1007/s10849-012-9157-1.

Isaacs, James and Kyle Rawlins. 2008. Conditional questions. *Journal of Semantics* 25: 269–319. doi:10.1093/jos/ffn003.

Kaufmann, Magdalena. 2012. *Interpreting Imperatives*. Dordrecht: Springer.

Khoo, Justin. 2011. Operators or restrictors? a reply to Gillies. *Semantics and Pragmatics* 4: 1–25. doi:10.3765/sp.4.4.

Klinedinst, Nathan and Daniel Rothschild. 2012. Connectives without truth tables. *Natural Language Semantics* 20: 137–175. doi:10.1007/s11050-011-9079-5.

Korzukhin, Theodore. 2014. Contextualist theories of the indicative conditional and Stalnaker's thesis. *Thought* 3: 177–83. doi:10.1002/tht3.126.

Kratzer, Angelika. 1979. Conditional necessity and possibility. In A. von Stechow, R. Bäuerle, and U. Egli (ed.) *Semantics from different points of view*, 117–147. Berlin: Springer.

———. 1981. The notional category of modality. In H. Eikmeyer and H. Rieser (eds.) *Words, Worlds, and Contexts*, 38–74. Berlin: De Gruyter.

———. 1986. Conditionals. *Chicago Linguistics Society* 22: 1–15.

———. 1991a. Conditionals. In A. von Stechow and D. Wunderlich (eds.) *Semantics: An International Handbook of Contemporary Research*, 651–56. Berlin: De Gruyter.

———. 1991b. Modality. In A. von Stechow and D. Wunderlich (eds.) *Semantics: An International Handbook of Contemporary Research*, 639–51. Berlin: De Gruyter.

———. 2012. *Modals and Conditionals*. Oxford: Oxford UP.

Lewis, David. 1973. *Counterfactuals*. Malden: Basil Blackwell.

———. 1976. Probabilities of conditionals and conditional probabilities. *Philosophical Review* 85: 297–315. http://www.jstor.org/pss/2184045.

———. 1981. Ordering semantics and premise semantics for counterfactuals. *Journal of Philosophical Logic* 10: 217–34. doi:10.1007/BF00248850.

MacFarlane, John. 2011. Epistemic modals are assessment-sensitive. In B. Weatherson and A. Egan (eds.) *Epistemic Modality*. Oxford: Oxford UP.

Rothschild, Daniel. 2013. Do indicative conditionals express propositions? *Noûs* 47: 49–68. doi:10.1111/j.1468-0068.2010.00825.x.

———. forthcominga. Conditionals and propositions in semantics. *Journal of Philosophical Logic*.

———. forthcomingb. A note on conditionals and restrictors. In John Hawthorne and Lee Walters (eds.) *Conditionals, Probability, and Paradox: Themes from the Philosophy of Dorothy Edgington*. Oxford: Oxford UP.

Rothschild, Daniel and Seth Yalcin. forthcoming. On the dynamics of conversation. Noûs.

Stalnaker, Robert. 1968. A theory of conditionals. In N. Rescher (ed.) *Studies in logical theory (American Philosophical Quarterly Monograph Series 2)*, 98–112. Oxford: Blackwell.

———. 1970. Probability and conditionals. *Philosophy of Science* 37: 64–80.

Swanson, Eric. 2008. Modality in language. *Philosophy Compass* 3: 1193–1207. doi:10.1111/j.1747-9991.2008.00177.x.

Veltman, Frank. 1996. Defaults in update semantics. *Journal of Philosophical Logic* 25: 221–61. doi:10.1007/BF00248150.

von Fintel, Kai. 2011. Conditionals. In K. von Heusinger, C. Maienborn, and P. Portner (eds.) *Semantics: An International Handbook of Meaning*, 1515–38. Berlin: DeGruyter. doi:10.1515/9783110255072.1515.

von Fintel, Kai and Anthony S. Gillies. 2007. An opinionated guide to epistemic modality. In T. Gendler and J. Hawthorne (eds.) *Oxford Studies in Epistemology: Volume 2*. Oxford: Oxford UP. http://mit.edu/fintel/fintel-gillies-2007-ose2.pdf.

———. 2010. *Must* . . . stay . . . strong! *Natural Language Semantics* 18: 351–383. doi:10.1007/s11050-010-9058-2.

von Fintel, Kai and Sabine Iatridou. 2005. What to do if you want to go to Harlem. http://mit.edu/fintel/www/harlem-rutgers.pdf. Ms., MIT.

Weatherson, Brian. ms. Moore, Bradley, and indicative conditionals. http://brian.weatherson.org/mbaic.pdf.

Weisberg, Jonathan. 2011. Varieties of bayesianism. In D. Gabbay, S. Hartmann, and J. Woods (eds.) *Handbook of the History of Logic*, vol. 10, 477–552. Amsterdam: Elsevier.

Yalcin, Seth. 2007. Epistemic modals. *Mind* 116: 983–1026. doi:10.1093/mind/fzm983.

———. 2010. Probability operators. *Philosophy Compass* 5: 916–37. doi:10.1111/j.1747-9991.2010.00360.x.

———. 2012. A counterexample to modus tollens. *Journal of Philosophical Logic* 41: 1001–1024. doi:10.1007/s10992-012-9228-4.