Synthese (2022) 200:282
https://doi.org/10.1007/s11229-022-03730-x

ORIGINAL RESEARCH

Triviality and the logic of restricted quantification

Nate Charlow[1]

Received: 16 February 2020 / Accepted: 2 May 2022

Abstract
This paper clarifies the relationship between the Triviality Results for the conditional and the Restrictor Theory of the conditional. On the understanding of Triviality proposed here, it is implausible—pace many proponents of the Restrictor Theory—that Triviality rests on a syntactic error. As argued here, Triviality arises from simply mistaking the feature a (e.g. conditional-involving) claim has when that claim is logically unacceptable for the feature a claim has when that claim is unsatisfiable (in a model). Triviality rests on a semantic confusion—one which some semantic theories, but not others, are prone to making. On the interpretation proposed here, Triviality Results thus play a theoretically constructive role in the project of natural language semantics.

Keywords Triviality · Restricted quantification · Indicative Conditionals · Probabilities of conditionals · Epistemic Modality · Dynamic Semantics · Higher-order uncertainty

1 Introduction

David Lewis (1976) proved there is no such thing as the probability of a conditional, if the probability of a conditional is required to satisfy Stalnaker's Thesis, according to which, whenever $\Pr(C \mid A)$ is defined, $\Pr(A \Rightarrow C) = \Pr(C \mid A)$ (Stalnaker, 1970). Angelika Kratzer has claimed that an appropriate understanding of the logical form of natural language conditionals, along the lines of her Restrictor Theory (Kratzer, 1986, 1991, 2012), "divert[s]" the "threat" of Triviality:

> The reason why the probabilities of conditional propositions can't seem to match the corresponding conditional probabilities is that there *are* no conditional propositions that are embedded under probability operators... [Q]uantificational operators, including adverbs of quantification, probability operators, and other modal operators, do not operate over 'conditional propositions'. The persistent

✉ Nate Charlow
nate.charlow@utoronto.ca

1 Department of Philosophy, University of Toronto, Toronto, ON M5R 2M8, Canada

Published online: 30 June 2022

> 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. (2012, pp. 93, 107)

The vague notion that Kratzer's theory of restricted quantification diverts Triviality's threat is common.[1] The general idea here seems to be that Kratzer's theory in some fashion undermines Stalnaker's Thesis. Kratzer herself rejects Stalnaker's Thesis by rejecting the idea that there is any such thing as the probability of a conditional claim or proposition. Belief in such a thing "is based on a simple syntactic mistake." (A fortiori, then, there is no such thing as the probability of a conditional proposition, *if* it is required to equal the corresponding conditional probability, which is what Lewis proved.) Kratzer is skeptical of the very idea of conditional propositions *qua* contents of assertion and attitudes like (degreed) belief: whenever we appear to evaluate $A \Rightarrow C$ as likely (to degree d), Kratzer thinks we are *in fact* evaluating C as likely (to degree d) given A; whenever we appear to assert that $A \Rightarrow C$ is likely (to degree d), Kratzer thinks we are *in fact* asserting that C is likely (to degree d) given A. Of course, it will not be possible to trivialize a probability measure over conditional claims, if probability measures are not even defined over conditional claims. Nevertheless the idea that (the content of) a conditional claim cannot be *assessed for likelihood* (nor *asserted*, nor *believed*) is itself a Triviality "result" for conditionals. If Triviality represents a "threat", this seems an odd way to divert it.

For this reason this paper will focus on a "moderate" version of Kratzer's Restrictor Theory. According to the **Moderate Restrictor Theory** (MRT):

- If-clauses are analyzed syntactically as restrictors.
- The content of a conditional is the type of thing that is assessible for likelihood.

According to the MRT (and contra Kratzer's stated view), *Triviality proofs are syntactically well-formed* (and of course valid). They fail for the mundane reason that they are unsound, i.e., they have a false premise. On this line of thought, though Lewis' original Triviality Result is a straightforward reductio of Stalnaker's Thesis in its full generality, it is no reductio of the idea that for a wide (even preponderant) range of probabilistic contexts Pr and conditionals $A \Rightarrow C$, $\Pr(A \Rightarrow C) = \Pr(C \mid A)$.[2] Section 2 will introduce the particulars of the Restrictor Theory and will describe this way of theorizing the connection between the Restrictor Theory and Triviality.

Section 3 will present a simple objection to this understanding of the Restrictor Theory vis-à-vis Triviality: facially sound Triviality results remain provable in the MRT (as I show by adapting an argument from Charlow (2016, 2019)). For example: in a classical semantic setting, there is provably no proposition $A \Rightarrow C$ such that it is inconsistent to accept that $A \Rightarrow C$ while also accepting that there is a chance that A and no chance that C. But the interpretation of $\ulcorner A \Rightarrow C \urcorner$ *is* ordinarily such that it is inconsistent to accept that $A \Rightarrow C$ while also accepting that there is a chance that A and no chance that C. In such contexts the interpretation of $\ulcorner A \Rightarrow C \urcorner$, $A \Rightarrow C$, cannot be a conditional proposition.

[1] Work in this vein includes Égré and Cozic (2011); Rothschild (2013); Rothschild (2020); Vandenburgh (2020).

[2] This is in essence the line on Triviality developed by Rothschild (2013, 2020).

The Triviality results I describe here are generally well-known in the literature. The overarching goals of this paper are (i) to leverage these results to develop a careful reading of the theoretical impact of Triviality on a moderate version of the Restrictor Theory, (ii) to pitch a particular version of the MRT that is attentive to this reading. Triviality, on the reading I present and argue for here, is suggestive of a *gap in the theory of content* implied by/accompanying the MRT: what is the content of '$A \Rightarrow C$', if not a conditional proposition (on the ordinary understanding of that notion)? Section 4 addresses this question in detail. I begin by sketching an answer to this question that is inspired by Dynamic Semantics, according to which $A \Rightarrow C$ is an information-sensitive proposition (or "test"). Although this strategy has its attractions—which I try to describe in some detail—its has its pitfalls too. I end the paper by pitching a different answer to this question, drawing on the semantic theory of Charlow (2020). According to this theory, semantic content is polymorphic in type; in particular, if A and C are semantic type p (the type of propositions), $A \Rightarrow C$ is of a higher, non-propositional semantic type $\langle pv, t\rangle$ (a set of real-valued measures defined over propositions). Probability operators, too, are polymorphic in type—they take contents of any type as their arguments; probability measures are devices for measuring the degree of acceptance/belief an agent has in a representation. On this theory, Triviality arises from a logical/semantic (not syntactic) confusion. The assumptions that induce Triviality are logically valid, in a sense, but not the exact sense that is required to induce Triviality. On the theory I prefer, and in agreement with Kratzer, the content of a conditional claim is not generally a conditional proposition. But, contra Kratzer, measures of comparative likelihood are *defined over* such claims and their contents.

2 Triviality and restricted quantification

2.1 Triviality as inexpressibility

Égré and Cozic (2011) drew the literature's attention to an illuminating parallel between certain Triviality results for the conditional and Inexpressibility results from generalized quantifier theory (Barwise & Cooper, 1981). Consider a language with a two-place conditional probability operator $[\alpha]$ ($\alpha \in [0, 1]$), where $\ulcorner[\alpha](\phi, \psi)\urcorner$ says that the probability of ψ conditional on ϕ equals α. Such a language is interpreted relative to a structure we will call a **Probability Model**, which is a triple $\mathcal{M} = \langle W, \Pr, V\rangle$ (with W a universe of worlds, Pr a probability measure over subsets of W that is normalized to W, and V a valuation mapping atoms to subsets of W). The semantics for the Boolean fragment (as well as the consequence relation appropriate to such a semantics) is standard:

$$\mathcal{M}, w \vDash A \text{ iff } w \in V(A)$$
$$\mathcal{M}, w \vDash \neg\phi \text{ iff } \mathcal{M}, w \nvDash \phi$$
$$\mathcal{M}, w \vDash (\phi \wedge \psi) \text{ iff } \mathcal{M}, w \vDash \phi \text{ and } \mathcal{M}, w \vDash \psi$$
$$\phi \vDash \psi \text{ iff } \forall \mathcal{M}, w \vDash \phi : \mathcal{M}, w \vDash \psi$$
$$\mathcal{M} \vDash \phi \text{ iff } \forall w \in W : \mathcal{M}, w \vDash \phi$$

Relative to $\mathcal{M} = \langle W, \mathrm{Pr}, V \rangle$, we let $[\![\phi]\!]^{\mathcal{M}}$ denote the set of points in W such that $\mathcal{M}, w \vDash \phi$. The semantics for the probabilistic fragment is as expected.

$$\mathcal{M}, w \vDash [\alpha](\phi, \psi) \text{ iff } \mathrm{Pr}([\![\psi]\!]^{\mathcal{M}} \mid [\![\phi]\!]^{\mathcal{M}}) = \alpha$$

Using an adaptation of the Wallflower argument (Hájek, 1989, 2012), Égré & Cozic prove a Triviality theorem for this language. The theorem they demonstrate is that there is no way of extending this language with a conditional connective $\rightarrow$ so that for all α, ϕ, and ψ:[3]

$$\vDash [\alpha](\phi, \psi) \equiv [\alpha](\phi \rightarrow \psi) \text{ [Stalnaker's Thesis]}$$

Proof Consider a probability model $\mathcal{M} = \langle W, \mathrm{Pr}, V \rangle$. Let $W = \{a, b, c\}$; $a, b \in V(A)$; $a \in V(C)$; and $\mathrm{Pr}(\{a\}) = \mathrm{Pr}(\{b\}) = \mathrm{Pr}(\{c\}) = 1/3$. Clearly $\mathcal{M}, w \vDash [.5](A, C)$. But $\forall p \subseteq W$, $\mathrm{Pr}(p) \in \{0, 1/3, 2/3, 1\}$. So $\mathcal{M}, w \nvDash [.5](A \rightarrow C)$. $\square$

As Égré & Cozic emphasize, such results reflect the fact that there is no meaning-preserving translation from a language of binary conditional probability operators into a language of unary probability operators—a fact that ought to be, in light of the inexpressibility results on which the theory of binary quantification was founded (Barwise & Cooper, 1981), completely unsurprising.

> From a semantic point of view, the result can be interpreted as showing that the binary probability operator $[\alpha](\cdot, \cdot)$ is more expressive than the unary probability operator $\alpha(\cdot)$... The result thus shows a positive side to Lewis's original Triviality result, since it establishes that restriction of unary probability operators by if-clauses adds expressiveness to a language with unrestricted probability operators taking scope over more complex formulae. (Égré and Cozic 2011, p. 22)

2.2 The restrictor theory

These two theses comprise the core of Kratzer's Restrictor Theory:

i. If-clauses are restrictors. When $\ulcorner$*if* $A, C\urcorner$ occurs with/under Q (as $\ulcorner Q$ *if* $A, C\urcorner$ or $\ulcorner$*if* $A, QC\urcorner$), the logical form of the sentence is (ordinarily) $Q(A, C)$.[4]
ii. If-clauses are *always* restrictors. When no appropriate operator is syntactically explicit (e.g., a speaker utters $\ulcorner$*if* $A, C\urcorner$ alone/bare), to interpret what the speaker says, it is necessary to posit an unpronounced binary quantifier in the logical form of their utterance.

To the Kratzerian, the semantic unsuitability of the binary conditional connective comes as no surprise, given its syntactic unsuitability. It is a corollary of thesis (i) that

[3] I will generally leave the assumption that the antecedent has non-zero probability implicit.

[4] The 'ordinarily' qualification is inserted to account for the possibility that $\ulcorner$ *if* $A, QC\urcorner$ can in certain contexts be interpreted as 'doubly' quantified, i.e., $Q'(A, Q(\top, C))$. For discussion of the 'nested' quantifier interpretation of such conditionals, see e.g. von Fintel and Iatridou (2005).

when $\ulcorner$*if* $A, C\urcorner$ occurs with/under a binary quantifier Q, $\ulcorner$*if* $A, C\urcorner$ is not interpreted as a syntactic constituent, and that it does not contribute to semantic computation the semantic value it would have if it were to occur alone/bare (i.e., a "conditional proposition"). Instead, according to Kratzer, in such cases the subordinate clause A provides the restriction argument for Q, while the matrix clause C provides its nuclear scope (Kratzer, 1986; Lewis, 1975). If $\ulcorner$*if* $A, C\urcorner$ were uniformly—regardless of syntactic context—translated with a two-place, proposition-forming conditional connective $>$, there would be no space for this syntactic possibility: $A > C$ would provide Q's nuclear scope. The result of the last section demonstrates that there are semantic consequences to these syntactic facts: distorting the syntax of natural language conditionals means distorting their truth conditions.

This point bears emphasis: *Kratzer's theory is devastating to Stalnaker's Thesis* (although that will turn out to be less relevant to the question of Triviality than one might have been led to expect, given the literature's focus on Stalnaker's Thesis). Given thesis (i), Stalnaker's Thesis is syntactically ambiguous. Supposing that $\ulcorner\phi \Rightarrow \psi\urcorner$ is a direct translation of the natural language $\ulcorner$*if* $\phi, \psi\urcorner$, Stalnaker's Thesis could be parsed as (ST1) or (ST2), which say that, for all α, ϕ, and ψ:

$$\text{(ST1)} \quad \vDash [\alpha](\phi, \psi) \equiv [\alpha](\top, \phi \Rightarrow \psi)$$
$$\text{(ST2)} \quad \vDash [\alpha](\phi, \psi) \equiv [\alpha](\phi, \psi)$$

In (ST1), $\phi \Rightarrow \psi$ is the result of composing $\phi \Rightarrow \psi$ with an unpronounced two-place quantifier (e.g., an epistemic necessity modal, à la Kratzer, or perhaps a two-place conditional connective, à la Stalnaker-Lewis). (ST2) is a triviality, and (ST1) would appear (in light of the inexpressibility result above) to be an absurdity. Conditional on Stalnaker's Thesis saying anything, it says something absurd.

What's more, the intuitive equivalences that Stalnaker's Thesis is posited to explain *are explained* by a Kratzerian syntactic mapping from natural language sentences, like (1) and (2), into a *single* sentence of our logical metalanguage $\ulcorner[.3](L, R)\urcorner$ (see esp. Rothschild 2020, Sect. 3.13).

(1) The probability of rain, given low pressure, is 30%

(2) ≈It is 30% likely that it will rain if there is low pressure

Such intuitive equivalences are *not* to be explained by attempting (per impossibile) (i) to provide a syntactic mapping from (1) into $\ulcorner[.3](L, R)\urcorner$ and from (2) into $\ulcorner[.3](L \Rightarrow R)\urcorner$ and (ii) to guarantee that, relative to an arbitrary probability model $\mathcal{M}$, $\ulcorner[.3](L, R)\urcorner$ is satisfied just when $\ulcorner[.3](L \Rightarrow R)\urcorner$ is satisfied. It would be as if we attempted to explain a felt equivalence like the following...

(3) Half the red blocks are square.

(4) ≈Half the blocks are square if they're red.

...by providing a mapping from (3) into a sentence of the form $\ulcorner Qx(Rx, Sx)\urcorner$ and from (4) into a sentence of the form $\ulcorner Qx(\phi x)\urcorner$, and stipulating that a first-order model satisfies $\ulcorner Qx(Rx, Sx)\urcorner$ iff it satisfies $\ulcorner Qx(\phi x)\urcorner$. The maneuver makes no sense in either case.

2.3 The moderate restrictor theory

All that said, there is nothing to stop a proponent of the Restrictor Theory from adopting the (extremely natural!) view that, when no appropriate binary quantifier is syntactically explicit, the posited binary operator is generally a *conditional operator*, by which I mean a binary operator (like the Stalnaker-Lewis conditional) that composes with an antecedent (restrictor) and a consequent (scope), yielding a *conditional proposition*: a proposition with the logic of the indicative conditional.[5]

On this line of thought, bare conditionals express conditional propositions, which can be the objects of assertion and (degreed) belief (and which may often, if not always, have probabilities that are equal to the corresponding conditional probabilities). Still, in line with thesis (i), conditionals do not contribute conditional propositions when they semantically compose with binary quantifiers. This section will develop and motivate this "moderate" version of the Kratzerian theory.

Begin with the notion of a probability model being proper for a conditional.

Definition 1 A probability model $\mathcal{M}$ is **proper** for $\ulcorner A \Rightarrow C \urcorner$ iff

$$[\![A \wedge C]\!]^{\mathcal{M}} = [\![(A \Rightarrow C) \wedge A)]\!]^{\mathcal{M}} \quad (A \Rightarrow C \text{ Strongly Centered in } \mathcal{M})$$

$$\Pr([\![A \Rightarrow C]\!]^{\mathcal{M}} \mid [\![A]\!]^{\mathcal{M}}) = \Pr([\![A \Rightarrow C]\!]^{\mathcal{M}}) \quad (A \Rightarrow C\, A\text{-independent in} \mathcal{M})$$

Restricting our attention to models that are proper for conditionals, we will find no counterexamples to Stalnaker's Thesis (Ellis 1978; Rothschild 2013). (Corollary: the Wallflower model is improper.)

Proof Suppose $\mathcal{M}$ is proper for $\ulcorner A \Rightarrow C \urcorner$. Then:

$$\Pr([\![A \Rightarrow C]\!]^{\mathcal{M}}) = \frac{\Pr([\![(A \Rightarrow C) \wedge A]\!]^{\mathcal{M}})}{\Pr([\![A]\!]^{\mathcal{M}})} = \frac{\Pr([\![A \wedge C]\!]^{\mathcal{M}})}{\Pr([\![A]\!]^{\mathcal{M}})} = \Pr([\![C]\!]^{\mathcal{M}} \mid [\![A]\!]^{\mathcal{M}})$$

□

Stalnaker's Thesis can thus be retained in *some* degree of generality. The degree of generality will depend on, very roughly, the generality of the "requirement" that a probability model be proper for a conditional. If one wants Stalnaker's Thesis to

[5] Kratzer (1991) argues that the silent operator should be modeled—at least in the case of indicative conditionals—as an operator expressing epistemic necessity (from the point of view of the speaker). For some simple considerations against this analysis, see Rothschild (2015) and Charlow (2016).

Stalnaker (2014) writes, about his own proposed semantics for the conditional: "The conditional was represented as a binary connective, not because the word 'if' is a conjunction with a syntax just like 'and' and 'or', but because what a conditional [which has been asserted] is used to say is a function of what is expressed by two sentential clauses" (183). This passage—and the overall focus of Stalnaker's corpus on conditionals—makes it clear that Stalnaker's focus is on modeling something we might dub the *propositional content of a conditional sentence*. There is no conflict here with Kratzer's observation that conditional sentences, *when embedded*, lack propositional content. On a charitable reading, Stalnaker is simply not trying to offer any theory of how those constructions work: he is offering a substantive theory which tells us what propositional content a conditional sentence has—*when it has a propositional content*, but not otherwise. Here and in the main text I am generally echoing Rothschild's understanding of the dialectic (see esp. Rothschild 2013; 2020).

hold in a maximal degree of generality one could take it to be a semantic or logical presupposition of the conditional that it is Strongly Centered and Independent of its antecedent (and hence that conditionals cannot strictly be semantically evaluated relative to probability models like the Wallflower model). Probably that is too strong: even if Strong Centering is a logical truth, Independence is not (Rothschild, 2013). Still, for a wide range (possibly even preponderance) of ordinary cases, these assumptions seem to hold.

I'll talk through this dialectic with a mundane case. An urn contains 10 marbles; 9 are big, and of those 9, 8 are red; so the chance of drawing a big marble is 8/10, while the chance of drawing a red marble, conditional on drawing a big one, is 8/9. An agent is to draw a marble from the urn, and a speaker delivers a prediction in which they are fairly confident:

(5) If you draw a big marble, it will be red.

Claim: a semantic theory should deliver a true/acceptable reading of (6):[6]

(6) The likelihood of that/what they said is 8/9.

In the likely ($p = .8$) event that $B \wedge R$, there is a pull to judge the prediction as true (since, if we had bet on the truth of the prediction, we would win the bet). In the unlikely ($p = .1$) event that $B \wedge \neg R$, there is a pull to judge the prediction as false (since we would lose that bet). In the event that $\neg B$, it seems that we don't know what to think about the prediction or regard its truth value as indeterminate or undecided (although I believe we can say that the likelihood that it *would have been true* if a big marble had been drawn was 8/9).[7]

	$B \wedge R$ ($p = .8$)	$B \wedge \neg R$ ($p = .1$)	$\neg B$ ($p = .1$)
$B \Rightarrow R$	true	false	??

Since the ratio of the likelihood of a scenario in which what a speaker predicts with (5) is assessed as *true* to the likelihood of a scenario in which his prediction is assessed

[6] I am aware of a judgment that (6) is false: if we are asked to decide whether what the speaker said is *true or false*, there is a some pull (which I share) to judge it false; *given this judgment*, the likelihood that what the speaker said is true is 0 (see e.g. Égré and Cozic 2011) It is plausible to take this as evidence for the *existence of* a Kratzerian reading of (5), according to which it expresses, roughly, the (false) claim that all relevant epistemic possibilities where a big marble is drawn are possibilities in which that marble is red.

[7] One way of semantically encoding these observations is to adopt a truth-functional, but trivalent, semantics for conditionals that are interpreted as expressing conditional propositions, according to which $\ulcorner A \Rightarrow C \urcorner$ lacks a truth value relative to any point w such that $w \notin V(A)$. For discussions of this possibility, see Reichenbach (1949); de Finetti (1936); McDermott (1996); Bradley (2002); Rothschild (2014); Rothschild (2020). McDermott makes an especially plausible argument, which I have drawn on here, for a trivalent semantics that connects the likelihood that $A \Rightarrow C$ is true to its "fair betting quotient", which he understands as "the probability that the bet [on $A \Rightarrow C$] will be won, given that it will be won or lost (=not called off)," i.e., given A (4). But the Trivalent semantics has some curious features; most obviously, it predicts that any conditional with a false antecedent is truth-value-less (which seems to be in tension with the observation that the likelihood that what the speaker said would have been true, had a big marble been drawn, is 8/9). It would be better to capture the observations that the trivalent semantics attempts to capture in a different semantic framework. I will gesture at one such framework later on.

as having a clear *truth value* is $\frac{.8}{.8+.1} = 8/9$, the likelihood that the prediction is true is naturally assessed as 8/9. The likelihood that the prediction is true is also assessed as being independent of the marble's size, while the truth of $B \wedge R$ is also assessed as sufficient for the truth of $B \Rightarrow R$.

We controlled for syntactic ambiguity by using expressions anaphoric to the content of what the speaker said ('that'/'what they said').[8] We will therefore represent (6)—the sentence whose truth we are trying to guarantee—with $\ulcorner[8/9](\top, B \Rightarrow R)\urcorner$. Given the linguistic context—in particular, the fact that $B \wedge (B \Rightarrow R)$ is contextually equivalent to $B \wedge R$—for any probability model for the case $\mathcal{M} = \langle W, \Pr, V\rangle$, we expect the probability of $B \wedge (B \Rightarrow R)$ (in $\mathcal{M}$) to equal the probability of $B \wedge R$ (in $\mathcal{M}$). That guarantees (F1) and (F2). Given the linguistic context—in particular, the fact that the conditional appears to be logically and probabilistically independent of its antecedent in that context—we expect the probability that $B \Rightarrow R$ is true (in $\mathcal{M}$) to equal the probability that $B \Rightarrow R$ is true, conditional on B (in $\mathcal{M}$). That guarantees (F1), which with (F2) and (F3) guarantees (F4).

$$\exists\alpha : \mathcal{M}, w \vDash [\alpha](\top, B \Rightarrow R) \wedge [\alpha](B, B \Rightarrow R) \tag{F1}$$

$$\mathcal{M}, w \vDash [.8](\top, B \wedge R) \tag{F2}$$

$$\mathcal{M}, w \vDash [.8](\top, B \wedge (B \Rightarrow R)) \tag{F3}$$

$$\Pr([\![B \Rightarrow R]\!]^{\mathcal{M}}) = \Pr([\![R]\!]^{\mathcal{M}} \mid [\![B]\!]^{\mathcal{M}}) = 8/9 \tag{F4}$$

So: there appears to be *strong reason for a moderate Kratzerian to endorse some (appropriately restricted) version of Stalnaker's Thesis*. The reason is simple: in a range of linguistic contexts for $\ulcorner A \Rightarrow C\urcorner$, the class of probability models that is *appropriate to* that context is a class of models that are proper for $\ulcorner A \Rightarrow C\urcorner$ (and which therefore validate the relevant instance of Stalnaker's Thesis). Often—as in the case above—the linguistic context c is such that:

i. It is common ground in c that $\ulcorner(A \Rightarrow C) \wedge A\urcorner$ and $\ulcorner A \wedge C\urcorner$ are equally likely.
ii. It is common ground in c that $\ulcorner A \Rightarrow C\urcorner$ is independent of $\ulcorner A\urcorner$.

The linguistic context for $\ulcorner A \Rightarrow C\urcorner$ sometimes guarantees that the probability models *compatible with* that linguistic context satisfy the appropriate instance of Stalnaker's Thesis.[9] If our interest is in representing the probabilistic information that is common

[8] For an earlier argument that puts pressure on Kratzer's views about conditional propositions using propositional anaphors, see Huitink (2008, §5.4.1), who credits von Fintel (2007). There is an interesting "Kratzerian" possibility for understanding this propositional anaphor that is worth outlining here (though it is not clear to me that it can actually be developed in a satisfactory way). (I note here that Kratzer (2012, 108) claims that the anaphor here resolves to C, which is not plausible, for reasons covered in Von Fintel & gillies (2015, §5).) The possibility is that a "propositional" anaphor to $A \Rightarrow C$ may resolve, *not* (as would be expected) to the claim expressed by a speaker of $\ulcorner A \rightarrow C\urcorner$ (i.e., $A \Rightarrow C$), but instead to its *semantic value*. For a Kratzerian, this semantic value can be modeled, *not* as a proposition, instead as a sequence $\langle A, C\rangle$, which delivers values for the restrictor and nuclear scope arguments of the quantifier under which the anaphor occurs. On this analysis of 'what they said', in the case of (6), the probability quantifier 'the likelihood of...is 90%' receives restriction B and nuclear scope R, and may thus be represented with a sentence $\ulcorner[8/9](B, R)\urcorner$. In this paper, however, I will be working with what I take to be the null hypothesis about propositional anaphors: that $\ulcorner[8/9](\top, B \Rightarrow R)\urcorner$ is a faithful representation of (6).

[9] This way of talking about the relationship between linguistic context and probability models seems to be implicit in theorists who regard Stalnaker's Thesis (or nearby corollaries) as holding by default, or in

ground in such a linguistic context, we have reason to restrict our focus to probability models that are proper for $\ulcorner A \Rightarrow C \urcorner$. (To the extent that such linguistic contexts predominate among possible contexts of interpretation for natural language conditionals, there will be reason to retain Stalnaker's Thesis in a high degree of generality.)

At first glance, the MRT seems well-positioned theoretically. Since "propriety" is surely in some sense a contextual norm, MRT will retain most of the explanatory power of Stalnaker's Thesis: in any probability model proper for $\ulcorner A \Rightarrow C \urcorner$, $\ulcorner [\alpha](\top, A \Rightarrow C) \equiv [\alpha](A, C) \urcorner$ is *valid*. But, in restricting the Thesis, we will avoid Wallflower-type worries, and more besides. For example, Lewis' original Triviality result (Lewis, 1976) assumes that Stalnaker's Thesis holds for all Pr and $\ulcorner A \Rightarrow C \urcorner$, and thus holds for both $\Pr(\cdot \mid C)$ and $\Pr(\cdot \mid \neg C)$. Given these assumptions, Lewis derives an absurdity—that according to Pr C and A are independent—using a single instance of the Law of Total Probability:

$$\begin{aligned}\Pr(A \Rightarrow C) &= \Pr(A \Rightarrow C \mid C)\Pr(C) + \Pr(A \Rightarrow C \mid \neg C)\Pr(\neg C)\\ &= \Pr(C \mid A \wedge C)\Pr(C) + \Pr(C \mid A \wedge \neg C)\Pr(\neg C)\\ &= \Pr(C) = \Pr(C \mid A)\end{aligned}$$

Of course, though, the moderate Kratzerian has every reason to reject the extension of Stalnaker's Thesis to $\Pr(\cdot \mid \neg C)$. That is because $\Pr(\cdot \mid \neg C)$ isn't proper for $\ulcorner A \Rightarrow C \urcorner$: $A \Rightarrow C$ is not A-independent in $\Pr(\cdot \mid \neg C)$, since $\Pr(A \Rightarrow C \mid \neg C) > 0$ although $\Pr(A \Rightarrow C \mid A \wedge \neg C) = 0$. Triviality isn't a "threat" for the moderate Kratzerian theory; it just reveals that we can't expect Stalnaker's Thesis to hold in cases where the probabilistic context is improper.[10]

3 Revenge

The MRT attempts to offer a well-motivated synthesis of the Restrictor Theory with philosophical commonplaces about the content of the indicative conditional. MRT's synthetic ambitions represent a large part of its theoretical appeal. But the real appeal of MRT, I will now argue, is unclear. The basic difficulty is that there appear to be (many) probabilistic contexts for which there is some pressure to say $\Pr(A \Rightarrow C \mid \neg C) = \Pr(C \mid A \wedge \neg C) = 0$. This constraint induces triviality *wherever it holds* (cf. Charlow, 2016; 2019). Suppose that $\Pr(A \Rightarrow C) > 0$. Then:[11]

$$\text{(T1)}\quad [0](\neg C, A \Rightarrow C) \vDash [1](A \Rightarrow C, C)$$

(T1) is a Triviality theorem. Consider the following class of probability models:

Footnote 9 continued
a normal linguistic context or case (see, e.g., Ellis, 1978; Rothschild, 2013; Khoo and Mandelkern, 2018a; 2018b).

[10] This is in essence the line on Stalnaker's Thesis that is proposed by Rothschild (2013), Sect. 6.

[11] Proof. Suppose that $\Pr(A \Rightarrow C \mid \neg C) = 0$ and $\Pr(A \Rightarrow C) > 0$. Note that $\Pr(A \Rightarrow C \mid \neg C) = \frac{\Pr(\neg C \mid A \Rightarrow C) \cdot \Pr(A \Rightarrow C)}{\Pr(\neg C)} = 0$. Since $\Pr(A \Rightarrow C) > 0$ and $\Pr(\neg C) > 0$, $\Pr(\neg C \mid A \Rightarrow C) = 0$. Since $\Pr(\neg C \mid A \Rightarrow C) = 1 - \Pr(C \mid A \Rightarrow C) = 0$, $\Pr(C \mid A \Rightarrow C) = 1$.

$$\mathfrak{M} = \{\mathcal{M} : \mathcal{M} \vDash [1](A \Rightarrow C, C) \text{ or } \forall\alpha > 0 : \mathcal{M} \nvDash [\alpha](\top, A \Rightarrow C)\}$$

$\mathfrak{M}$ is a class of *trivial probability models* for $A \Rightarrow C$: it is the class of models according to which either C is certain given $A \Rightarrow C$ or $A \Rightarrow C$ lacks a (positive) probability. But we have shown that for any $\mathcal{M}$ such that $\mathcal{M} \vDash [0](\neg C, A \Rightarrow C)$, $\mathcal{M} \in \mathfrak{M}$. That is to say, the class of probability models that witness condition (C1) is a class of trivial probability models for $A \Rightarrow C$.

$$\text{(C1)} \quad \Pr(A \Rightarrow C \mid \neg C) = 0$$

That should be regarded as surprising. Indeed, at least one author (Bradley, 2000, 2007) has contemplated the stipulation that for *all* Pr such that $\Pr(A) > 0$ and $\Pr(C) = 0$ (and so, therefore, for all $\Pr(\cdot \mid \neg C)$ such that $\Pr(A \mid \neg C) > 0$): $\Pr(A \Rightarrow C \mid \neg C) = 0$. Bradley's argument for this stipulation, which he calls the *Preservation Condition*, is brief but compelling:

> You cannot... hold that we might go to the beach, but that we certainly won't go swimming and at the same time consider it possible that if we go to the beach we will go swimming! To do so would reveal a misunderstanding of the indicative conditional (or just plain inconsistency). (2000, p. 220)

Now the Preservation Condition represents a *very* strong constraint on probability models, which I take to be reason enough for being wary of it.[12] However, in comparison with the Preservation Condition, the assumption that (C1) *is witnessed* is quite modest: like an instance of Stalnaker's Thesis, it can be justified as a constraint on contextually "proper" probability models. To illustrate, recall the marbles case. In this case, the conditional proposition $B \Rightarrow R$ is assessed as having a positive probability (8/9). But how likely is $B \Rightarrow R$ to be true, conditional on $\neg R$ being true? The natural answer is obtained by asking how likely it is that a bet on $B \Rightarrow R$ will pay off conditional on the bet being resolved (i.e., conditional on B) and on $\neg R$; the natural answer, that is to say, is 0. But respecting this answer renders a probability model trivial.

An advocate of MRT might seem well-positioned to explain a case like (7) without appeal to anything resembling (T1) or the Preservation Condition.

(7) #If the marble isn't red, there's a chance that if it is big it will be red.

For there are two options for translating this sentence into the language of binary quantifiers (assuming the antecedent of (7) is parsed as a restrictor for the binary quantifier 'there's some chance that'):

$$\text{(Tr1)} \quad [\alpha](B \wedge \neg R, R)$$
$$\text{(Tr2)} \quad [\alpha](\neg R, B \Rightarrow R)$$

[12] In the background here is an implicit modus tollens. Bradley shows that the Preservation Condition implies a Triviality result: whenever $A \nvDash C$, $A \Rightarrow C \vDash C$. Since that result is an absurdity, we should conclude that the Preservation Condition is false (see, e.g., Weatherson, 2004; Khoo and Mandelkern, 2018a).

Clearly (Tr1) $\vDash \bot$ (for any $\alpha > 0$). Notice, too, that for any $\mathcal{M} \vDash [0](\neg R, B \Rightarrow R)$ (i.e., any model that falsifies (Tr2) for all $\alpha > 0$), $\mathcal{M}$ witnesses (C1) for $B \Rightarrow R$, in which case $\mathcal{M}$ is a trivial probability model for $B \Rightarrow R$. We could, it seems, scarcely hope to have better evidence in favor of the Kratzerian approach to representing (7)'s logical form: not only does (Tr1) explain why (7) is degraded; the alternative (Tr2) induces triviality *whenever it is falsified* by a probability model (for all $\alpha > 0$).

As before, though, progress is fleeting: the same phenomenon is observed, controlling for syntactic ambiguity. Compare the replies to the claim 'if you draw a big marble, it will be red' in (8) and (9).

(8) #If the marble isn't red, there's (still) a chance that will be true.

(9) If the marble isn't red, there's no chance that will be true.

Although it would seem reasonable to represent (9) with a sentence of the form $\ulcorner[0](\neg R, B \Rightarrow R)\urcorner$, any probability model that satisfies this sentence or falsifies (8) is a trivial probability model.[13] Triviality results arise from seemingly plausible claims that are considerably weaker than the type of thesis to which moderate Kratzerianism was devastating (e.g., Stalnaker's Thesis in full generality).

4 Informational theories of content

How then *do* we divert the threat (assuming that we are, as I think we ought to be, proponents of MRT)? That is my topic for the remainder. This section will begin by considering how appeal to ideas familiar in Dynamic Semantic theorizing generally (and in a particular form of Dynamic Semantic theorizing known as Context Probabilism) divert Triviality. Although going Dynamic has many desirable features, there are pitfalls too—in particular, the phenomenon of higher-order uncertainty, or uncertainty that is directed toward chance hypotheses, is rendered mysterious. I conclude this section, and the paper, by outlining a different spin on an informational theory of content, which is prima facie well-suited to accommodating the phenomenon of higher-order uncertainty.

4.1 Information-Sensitivity

It is possible for a broadly "Dynamic Semantic" theory to avoid the Revenge worry just posed.[14] The key idea is to draw a distinction between:

- The "true" conditional probability of $A \Rightarrow C$ on $\neg C$, $\Pr(A \Rightarrow C \mid \neg C) = \frac{\Pr((A \Rightarrow C) \wedge \neg C)}{\Pr(\neg C)}$.

[13] As noted above, there are possible strategies for handling "propositional anaphors" that make (Tr1) an available parse for (8), none of which I discuss here. Given the null hypothesis that propositional anaphors resolve to contents, there is ample reason to seek a different explanation of the judgment in (8).

[14] This strategy was first suggested by Daniel Rothschild (pc). It receives a brief (and skeptical) discussion in Charlow (2016).

- The "suppositional" probability of $A \Rightarrow C$ given $\neg C$, $\mathrm{Pr}_{|\neg C}(A \Rightarrow C)$: the probability of $A \Rightarrow C$, relative to the probability model $\mathcal{M}_{|\neg C} = \langle W - C, \mathrm{Pr}(\cdot \mid \neg C), V \rangle$.

The thought behind the distinction is this: while it is natural to require that the suppositional probability of $A \Rightarrow C$ relative to such a probability model is 0, it makes no sense to do so for the true conditional probability. That is because $\frac{\mathrm{Pr}((A \Rightarrow C) \wedge \neg C)}{\mathrm{Pr}(\neg C)} = 0$ implies that $\mathrm{Pr}((A \Rightarrow C) \wedge \neg C) = 0$. But, ordinarily anyway, $\mathrm{Pr}((A \Rightarrow C) \wedge \neg C) > 0$; ordinarily $A \Rightarrow C$ and $\neg C$ aren't incompatible!

Given that $\mathrm{Pr}_{|\neg C}$ was defined as $\mathrm{Pr}(\cdot \mid \neg C)$, how is it possible to distinguish $\mathrm{Pr}(A \Rightarrow C \mid \neg C)$ from $\mathrm{Pr}_{|\neg C}(A \Rightarrow C)$? There is room to do so if the "proposition" $A \Rightarrow C$ *shifts* when we consider the context that results from conditionalizing the model on $\neg C$. The idea here is that $A \Rightarrow C$ is to be modeled, *not* as a proposition *sans phrase*, but instead as a function from probability models into propositions—an *information-sensitive proposition*. So, the notion of "the proposition that" $A \Rightarrow C$ is only well-defined relative to a specification of a probability model. We do not need to say much more by way of developing the basic idea: all that is needed is a story on which $A \Rightarrow C$ is a function mapping any probability model in which $\mathrm{Pr}(C) = 0$ and $\mathrm{Pr}(A) > 0$ into the absurd proposition $\perp$.

Your first impression may be that this is a radical or arbitrary kind of context-sensitivity—what reason could there be for introducing it? Although that was once my reaction, it was wrong: it is the kind of context-sensitivity already exemplified by a family of familiar Dynamic Semantic theories of epistemic modality. Dynamic theories represent the content of an utterance as a *diachronic informational constraint* (i.e., a "context change potential" or update instruction, modeled as a function from states of information into states of information). On the classic theory of Veltman (1996), the update expressed by an epistemic possibility claim $\ulcorner \Diamond \phi \urcorner$ is a peculiar kind of update, which he calls a *test*. Writing $s[\phi]$ for the result of updating a set of worlds s with ϕ, Veltman proposes:

$$s[\Diamond \phi] = \begin{cases} s, & \text{if } s[\phi] \neq \emptyset \\ \emptyset, & \text{otherwise} \end{cases}$$

We obtain an equivalent[15] "static" understanding of propositional content by modeling the content of $\ulcorner \Diamond \phi \urcorner$ as a function from states of information into propositions—an information-sensitive proposition:

$$[\![\Diamond \phi]\!] = \lambda s. \begin{cases} \top, & \text{if } s \cap [\![\phi]\!] \neq \emptyset \\ \perp, & \text{otherwise} \end{cases}$$

There is no such thing as "the" proposition $\Diamond \phi$, on this staticization of Veltman's test semantics: $\Diamond \phi$ is either $\top$ or $\perp$, depending on the characteristics of the information state against which $\Diamond \phi$ is evaluated.

[15] Equivalence here is understood as the definability of an update function f such that (i) $\langle s, s' \rangle \in f(\phi)$ iff $s' = s \cap [\![\phi]\!]$, (ii) $\langle s, s' \rangle \in f(\phi)$ iff $s[\phi] = s'$. Compare Rothschild and Yalcin (2017), who also contemplate a reading of Veltman along these lines.

Conditionals, like epistemics, are fruitfully modeled as tests on states of information (equivalently, information-sensitive propositions) (see a.o. Gillies, 2004; 2010; Willer, 2010). Let $\mathcal{M} = \langle W, \text{Pr}, V\rangle$ and suppose that $\text{Pr}(C) = 0$ and $\text{Pr}(A) > 0$. Then:

$$\text{(PCI)} \;\; \forall\phi : \text{Pr}_{|A\Rightarrow C}(\phi) \text{ is undefined}$$

PCI represents the claim that $A \Rightarrow C$ is evaluated as $\perp$ relative to any such $\mathcal{M}$: $\text{Pr}_{|A\Rightarrow C}(\phi) = \text{Pr}(\phi \mid A \Rightarrow C)$ is undefined for arbitrary ϕ iff $A \Rightarrow C$ is evaluated as $\perp$ relative to any such $\mathcal{M}$. Notice that PCI is in no tension with the probabilistic compatibility of $A \Rightarrow C$ and $\neg C$: generally, $\text{Pr}((A \Rightarrow C) \wedge \neg C) > 0$, and hence $\text{Pr}(A \Rightarrow C \mid \neg C) > 0$. And so this appeal to information-sensitive propositions does avoid allow the moderate Kratzerian to divert the threat of Triviality, as posed just above.

What is more, though PCI does not imply the Preservation Condition, it predicts the linguistic judgments that were cited in its favor. Given PCI, in any probability model in which we certainly won't go swimming, but there's some chance we go to the beach, the proposition expressed by the conditional 'if we go to the beach we will go swimming' is $\perp$; and there would be no way to accept or update on the conditional (without revising one of these priors).

4.2 Context probabilism

Although this rough sketch of a Dynamic theory is enough to disrupt the Revenge argument of Section 3, it doesn't yet address the following tension. In any non-trivial probability model $\mathcal{M} = \langle W, \text{Pr}, V\rangle$ such that $\mathcal{M} \vDash [\alpha]B$, $\text{Pr}((B \Rightarrow R) \wedge \neg R) > 0$. And yet there is evidently something semantically defective[16] with the following set of claims (for $\alpha > 0$):

$$\Gamma_1 = \{\ [\alpha]B\ ,\ \ B \Rightarrow R\ ,\ \ \neg R\ \}$$

"Given" $[\alpha]B$, $\ulcorner(B \Rightarrow R) \wedge \neg R\urcorner$ is judged defective/inconsistent. How is it possible to account for this judgment (without inducing Triviality)?

This section will describe and evaluate the "Bayesian"/"Context Probabilist" theory of Yalcin (2012a, b), which is a concrete version of the Information-Sensitive strategy described just above, and explains how it threads this needle (and another).[17] Here are the basic semantic clauses for that theory:

[16] One might object that these claims are jointly Moore Paradoxical, and so pragmatically, not semantically, defective. But Moore Paradoxicality dissolves under supposition (Yalcin, 2007). One cannot suppose that the claims in Γ_1 are all jointly so.

[17] For a broadly similar theory see Rothschild (2012).

$$\begin{aligned}
\mathcal{M}_{|A} &= \langle W \cap V(A), \Pr(\cdot \mid V(A)), V\rangle \\
\mathcal{M}_{|\neg\phi} &= \langle W - W_{|\phi}, \Pr(\cdot \mid W - W_{|\phi}), V\rangle \\
\mathcal{M}_{|\phi\wedge\psi} &= \mathcal{M}_{|\phi|\psi} \\
\mathcal{M}_{|[\alpha](\phi,\psi)} &= \mathcal{M} \text{ iff } \Pr(W_{|\psi} \mid W_{|\phi}) = \alpha, \text{ otherwise } \emptyset \\
\mathcal{M}_{|\phi\Rightarrow\psi} &= \mathcal{M} \text{ iff } \mathcal{M}_{|\phi|\psi} = \mathcal{M}_{|\phi}, \text{ otherwise } \emptyset \\
\mathcal{M} \Vdash \phi &\text{ iff } \mathcal{M}_{|\phi} = \mathcal{M} \\
\phi \Vdash \psi &\text{ iff } \forall \mathcal{M} \Vdash \phi : \mathcal{M} \Vdash \psi
\end{aligned}$$

Note that the fundamental semantic relation $\Vdash$ is of an "informational", rather than "alethic" character: a model satisfies ϕ when, we might say, the model contains the information that ϕ. Notice that, according to CP, $\Gamma_1 \Vdash \bot$; that is to say, it is impossible for a model to simultaneously contain all of the information in Γ_1. Proof: suppose $\mathcal{M} = \langle W, \Pr, V\rangle$ satisfies Γ_1. Since $\mathcal{M} \Vdash \neg R$, $\Pr(R) = 0$. Since $\mathcal{M} \Vdash B \Rightarrow R$, $\mathcal{M}_{|B|R} = \mathcal{M}_{|B}$. So, $\Pr(B) = 0$. So, $\mathcal{M} \nVdash [\alpha]B$, for any $\alpha > 0$, and we have our contradiction. However, although Γ_1 isn't simultaneously satisfiable on CP, the Triviality argument above is blocked. In particular, on the assumption that $\mathcal{M} \Vdash [\alpha]B$, although it will follow that $\mathcal{M} \nVdash (B \Rightarrow R) \wedge \neg R$, it will not follow that $\Pr((B \Rightarrow R) \wedge \neg R) = 0$. CP appears to thread the needle: it explains why Γ_1 is semantically defective. But the explanation appeals to an informational, rather than alethic, understanding of logical consequence, and so Triviality seems to be avoided.

In a recent paper Santorio (forthcoming) has proved a Triviality result that is prima facie worrisome for theories, like CP, that rely on a relation of informational consequence ($\Vdash$) that is extensionally distinct from the relation of classical logical consequence ($\vDash$). Santorio's result is this:

$$\text{(TIC)} \;\; \text{If } \phi \Vdash \psi \text{ then } \phi \vDash \psi.$$

Santorio's proof of (TIC) uses the supposition that probability is classical.[18]

Classical Probability. If $\phi \nvDash \psi$ then $\exists \mathcal{M} : \mathcal{M} \vDash [1]\phi, \mathcal{M} \vDash [\beta]\psi \; (\beta < 1)$.

Proof Suppose that $\phi \Vdash \psi$. Then for any $\mathcal{M}$ such that $\mathcal{M}_{|\phi} = \mathcal{M}$, $\mathcal{M}_{|\psi} = \mathcal{M}$. Notice that according to CP, $\phi \Vdash \psi$ implies $[1]\phi \Vdash [1]\psi$.[19] Suppose for reductio: $\phi \nvDash \psi$. By Classical Probability, $\exists \mathcal{M} : \mathcal{M} \vDash [1]\phi, \mathcal{M} \vDash [\beta]\psi \; (\beta < 1)$. So $\phi \nVdash \psi$. Contradiction.

□

According to CP, an informational consequence preserves certainty. But if there are more informational consequences than classical consequences, then there are ϕ and

[18] Santorio derives Classical Probability from two more basic assumptions in classical probability theory (his "Plenitude" and "Bound" assumptions). The idea is straightforward: a suitable witness for Classical Probability can be constructed by considering a probability model $\mathcal{M}$ according to which ϕ is likelier than ψ (such a model is guaranteed to exist, according to Plenitude). Then $\mathcal{M}_{|\phi}$ witnesses Classical Probability. See santorio (forthcoming: Sect. 5.1).

[19] Proof: suppose $\phi \Vdash \psi$ and $\mathcal{M} \nVdash [1]\phi$. Then $\Pr(W_{|\phi}) = \Pr(W_{|\phi|\psi}) = \Pr(W_{|\psi}) = 1$, and $\mathcal{M} \nVdash [1]\psi$.

ψ such that $\phi \Vdash \psi$ and $\phi \nvDash \psi$. CP has it that $\phi \Vdash \psi$ implies $[1]\phi \Vdash [1]\psi$. Classical probability theory has it that $\phi \nvDash \psi$ implies that ϕ's probability can exceed ψ's (in particular: it is possible for ϕ to have probability 1 while ψ's probability is less than 1). Contradiction. So there are no more informational consequences than classical consequences. But then Γ_1, for example, must be classically inconsistent, and we have our Triviality result, by the argument that began this section.

CP's response to this challenge is direct: if CP is true, Classical Probability is (demonstrably) false. Here is a counterexample. Note that although $A \Vdash [1]A$, $A \nvDash [1]A$. Consider any $\mathcal{M} = \langle W, \mathrm{Pr}, V\rangle$ such that $\mathcal{M} \vDash [1]A$, and note that $\mathrm{Pr}(W_{|A}) = 1$. Since $A \Vdash [1]A$ and $[1]A \Vdash [1][1]A$, $\mathrm{Pr}(W_{|A}) = \mathrm{Pr}(W_{|[1]A}) = \mathrm{Pr}(W_{|[1][1]A}) = 1$. So $\mathcal{M} \vDash [1][1]A$. So $\mathcal{M} \nvDash [\beta][1]A$, for any $\beta < 1$.

I have been explaining why Dynamic theories like Context Probabilism could be regarded as having desirable properties vis-à-vis the Triviality results. But in the present context, these properties aren't really to-the-point. In fact theories like CP do *not* ultimately sidestep the theoretical threat that is associated with Triviality (no more, anyway, than Kratzer's version of the Restrictor Theory or the Dynamic theory of the prior section). The basic difficulty is that, like those other theories, dynamic theories like CP run roughshod over "higher-order" uncertainty: uncertainty targeting, or directed at, claims with probabilistic (or otherwise informational) content. According to CP:

$$\forall\alpha : [\alpha](\phi \Rightarrow \psi) \vDash [0](\phi \Rightarrow \psi) \vee [1](\phi \Rightarrow \psi)$$

Proof Suppose $\mathcal{M} \vDash [\alpha](\phi \Rightarrow \psi)$. Then $\mathrm{Pr}(W_{|\phi\Rightarrow\psi}) = \alpha$. By the semantic clause for $\Rightarrow$: either $W_{|\phi\Rightarrow\psi} = W$ or $W_{|\phi\Rightarrow\psi} = \emptyset$. So, either $\mathrm{Pr}(W_{|\phi\Rightarrow\psi}) = \mathrm{Pr}(W) = 1$ or $\mathrm{Pr}(W_{|\phi\Rightarrow\psi}) = \mathrm{Pr}(\emptyset) = 0$. □

According to CP, if $\phi \Rightarrow \psi$ has a probability, it is extremal.[20] For the advocate of MRT hoping to uphold Stalnaker's Thesis in *any* degree of generality, that is a disappointing result.

4.3 Content polymorphism

In the remainder I want to explain why the semantic theory of Charlow (2020) is well-positioned to explain the relevant facts, and in particular to deliver on the theoretical desideratum that has been operating in the background of this paper: that chance hypotheses—including the contents of conditional claims—are often *themselves the objects of non-extremal estimates of likelihood*.[21]

20 More generally if ϕ is a test (a function from a state of information into $\top$ or $\bot$) then $[1]\phi \vee [0]\phi$ is a logical truth (Willer, 2013 raises a related point for Veltman's theory). Thanks to a referee for pressing me to' make the point here. The proper representation of higher-order uncertainty—uncertainty targeting chance hypotheses (a semantic category that includes indicative conditionals)—is a central theme of Moss (2015, 2018) and Charlow (2020).

21 Other semantic work in this vein includes Moss (2015); Goldstein (2021); Goldstein and Santorio (2021). Though a comparison is beyond the scope of the present paper, see Charlow (2020) for discussion of Moss' theory.

I'll begin with a short statement of the formal theory of content, then spend the rest of this section talking through its core features. We begin by generalizing the notion of a probability model, which we now take to be a triple $\mathcal{M} = \langle W, \Pr, V\rangle$, with W a universe of *possibilities* (of the same logical or semantic type, but not necessarily possible worlds), Pr a probability measure over subsets of W that is normalized to W, and V a valuation mapping atoms to sets of possible worlds.

$$\begin{aligned}
[\![A]\!]^{\mathcal{M}} &= \{w \in W^{\mathcal{M}} : w \in V(A)\} \\
\llbracket [\alpha](\phi, \psi)\rrbracket^{\mathcal{M}} &= \{\Pr : \Pr([\![\psi]\!]^{\mathcal{M}} \mid [\![\phi]\!]^{\mathcal{M}}) = \alpha\} \\
[\![\neg\phi]\!]^{\mathcal{M}} &= W^{\mathcal{M}} - [\![\phi]\!]^{\mathcal{M}} \\
[\![\phi \wedge \psi]\!]^{\mathcal{M}} &= [\![\phi]\!]^{\mathcal{M}} \cap [\![\psi]\!]^{\mathcal{M}} \\
\mathcal{M}_{|\phi} &= \langle W \cap [\![\phi]\!]^{\mathcal{M}}, \Pr(\cdot \mid [\![\phi]\!]^{\mathcal{M}}), V\rangle
\end{aligned}$$

The basic idea is familiar: the content of a claim is a set of points at which it is evaluated as true. Probabilistic claims are, however, true at *different types of points* than ordinary atomic claims; the possibilities that witness $[\alpha]p$ are ways (for one) *to distribute* (their) credence, while the possibilities that witness p are ways *the world* could be.

On this theory, a probability operator is a kind of *type-raiser*: picturesquely applying a probability operator to a proposition transforms a claim about the world into a *way of representing with respect to that claim*. Given an ordinary propositional content $p_{\langle s,t\rangle}$, the semantic type of a probabilistic claim $[\alpha]p$ is $\langle\langle st, v\rangle, t\rangle$ (a function from real-valued measures defined over propositions into truth values). The semantic type of an "iterated" probabilistic claim $[\beta][\alpha]p$ is higher still: $\langle\langle\langle stv, t\rangle, v\rangle, t\rangle$—a set of probability measures Pr defined over a domain consisting of sets of (lower type) probability measures, such that the chance hypothesis $[\alpha]p$ is β-likely according to Pr.

Note that in this "polymorphic" semantic theory, there is no question about the status of Stalnaker's Thesis, $\Pr(A \Rightarrow C) = \Pr(C|A)$. On the present theory, there is *no* Pr that is defined over both $A \Rightarrow C$ and C, which is a consequence of the fact that $A \Rightarrow C$ and C are elements of algebras of *different semantic types*. Stalnaker's Thesis—indeed any instance of it—is semantically malformed.

Like Context Probabilism, this theory explains the defectiveness of Γ_1, without inducing Triviality. I'll approach this point indirectly. On the present theory $\neg p$ and $[\alpha]p$ ($\alpha > 0$) are of distinct semantic types. In a compositional semantic setting this means that $\llbracket [\alpha]p \wedge \neg p\rrbracket$ will be undefined, absent application of some way of coordinating the semantic types of the claims on either side of $\wedge$ (for the classic discussion of this phenomenon, see Partee and Rooth 1983).

The natural way to obtain semantic coordination is to raise $\neg p$'s type to $[\alpha]p$'s, by application of the type-shifting operator $*$:[22]

[22] This is close to the type-shifter $\mathscr{C}$ proposed by Moss (2015). The general idea is that $*\phi$ is a property that a representation—a set of possibilities—satisfies when according to it ϕ is certain. The use of type-shifting does invite broadly logical worries. Mandelkern (2019, Sect. 7.4) notes that a type-shifting theory yields bizarre predictions for cases like $\neg([1]p \vee p)$, assuming that it is parsed as $\neg([1]p \vee [1]p) = \neg[1]p$. Note

$$*\phi = [1]\phi$$

On this parse, $\ulcorner[\alpha]p \wedge \neg p\urcorner$ is semantically well-formed. But it is inconsistent, since no probability measure can assign $\neg p$ probability 1 while assigning p probability $\alpha > 0$: $[\![[\alpha]p \wedge *\neg p]\!] = \emptyset$.

Similarly evaluating Γ_1 for *joint* acceptability requires evaluating each sentence in Γ_1 as bearing on the same "subject matter" or "representational question", which in turn requires interpreting each sentence in Γ_1 as a chance hypothesis, type $\langle stv, t\rangle$. These chance hypotheses are evaluated for joint acceptability relative to *a way of assigning probabilities to chance hypotheses*—relative, that is to say, to a probability model "living at" the appropriate level of the type hierarchy. Concretely, suppose that $[\alpha]B$ and $B \Rightarrow R$ are each of chance hypothesis type, $\langle stv, t\rangle$. Then $[\alpha]B$ characterizes the set of measures according to which B is α-likely. We'll assume that $B \Rightarrow R := [1](B, R)$, so that $B \Rightarrow R$ characterizes the set of measures according to which R is certain given B. Finally to achieve semantic coordination, we raise $\neg R$ to $[1]\neg R$. On this "parse", Γ_1 is *unsatisfiable*, since for any $\mathcal{M}$ and $\alpha > 0$:

$$[\![[\alpha]B]\!]^{\mathcal{M}} \cap [\![B \Rightarrow R]\!]^{\mathcal{M}} \cap [\![[1]\neg R]\!]^{\mathcal{M}} = \emptyset$$

This theory is designed to accommodate the truism that it is possible to have non-extremal confidence in chance hypotheses. And certainly it does that. As a moderate Kratzerian, however, I take it that this isn't sufficient: a semantic theory should deliver an acceptable reading of (6), the claim that $B \Rightarrow R$ is 8/9 likely. On the present theory, (6) will be satisfied by a probability measure Pr iff $\Pr(B \Rightarrow R) = 8/9$. But, according to the present theory, a Pr such that $\Pr(B \Rightarrow R) = 8/9$ is to be understood as a distribution of confidence *over chance hypotheses*—typed as sets of probability measures—so that $\Pr(B \Rightarrow R) = 8/9$ iff Pr maps $\mathcal{C}_1$ to 8/9:

$$\mathcal{C}_1 = \{\Pr{}' : \Pr{}'(R|B) = 1\} = \{\Pr{}' : \Pr{}'(B \supset R) = 1\}$$

Does the context for (6) supply the information that $[8/9](R \Rightarrow B)$, on this understanding of the nature of that information? Yes it does. The context makes salient two "ways of representing" (i.e., chance hypotheses) on the relevant question: what to conclude about the marble's color, given that the marble is big. They are $\mathcal{C}_1 = B \Rightarrow R$ (roughly: given it's big, conclude that it's red) and $\mathcal{C}_2 = B \Rightarrow \neg R$ (roughly: given it's big, conclude that it's not red).

$$\mathcal{C}_2 = \{\Pr{}' : \Pr{}'(R|B) = 0\} = \{\Pr{}' : \Pr{}'(B \supset \neg R) = 1\}$$

Footnote 22 continued
however that there are two ways that $[1]p$ and p can be coordinated on the present theory; in particular, $[1]p$ might be lowered to $\langle s, t\rangle$—e.g., the proposition that is true at w iff p is certain according to the w-relevant measure. (I don't want to commit to any more specific account of type-lowering here.) Since the type-shifting theory can assign a reasonable meaning to a claim like $\neg([1]p \vee p)$, it is not clear what the objection to the type-shifting theory is supposed to be. Hawke and Steinert-Threlkeld (2021) raise a worry for the theory of Moss (2015), which treats the logical connectives as context-sensitive. But Charlow (2020) does not treat the connectives as context-sensitive.

The information in the case tells me to distribute my confidence over $\mathcal{C}_1$ and $\mathcal{C}_2$ as follows: be 8 times more confident in the way of representing $\mathcal{C}_1$ than you are in the way of representing $\mathcal{C}_2$. (6) "expresses" this state of comparative confidence.[23]

I conclude by scrutinizing Bradley's case for Preservation. Bradley says correctly that it is impossible to "hold that we might go to the beach, but that we certainly won't go swimming and at the same time consider it possible that if we go to the beach we will go swimming". We explain this judgment as follows. Preservation is *Acceptance-Valid*, in the following sense: conditioning a probability measure over chance hypotheses Pr on $[\alpha]B$ and $[0]S$ means Pr must assign the chance hypothesis associated with $B \Rightarrow S$ probability 0 (and the alternative chance hypothesis associated with $B \Rightarrow \neg S$ probability 1):

$$\mathrm{Pr}_{|[\alpha]B\wedge[0]S}\ (\ \{\mathrm{Pr}' : \mathrm{Pr}'(S|B) = 1\}\ ) = 0$$
$$\mathrm{Pr}_{|[\alpha]B\wedge[0]S}\ (\ \{\mathrm{Pr}' : \mathrm{Pr}'(S|B) = 0\}\ ) = 1$$

So, just as Bradley says, one cannot *hold that* there's a chance that B, *and that* there's no chance that S, and at the same time retain any confidence in $B \Rightarrow S$.[24] Contra Bradley, though, the Preservation condition is false; indeed, on the present theory it is, like Stalnaker's Thesis, a *type-theoretically nonsensical* addition to the theory I have proposed in this section. The Preservation Condition is valid, in a sense; just not the sense required to run the Triviality argument of Sect. 3.

Obviously I don't aim to have convinced you of my theory. My purpose has been more limited. I've provided some desiderata on a logic of chance hypotheses (the accommodation of which leads to Triviality in standard logical frameworks, even within the context of a moderate Kratzerian theory of conditionals). And I have tried to explain why conjoining the moderate Kratzerian theory to the semantic theory of Charlow (2020) would yield a theory with some prima facie desirable features.

5 Conclusion

The questions about the conditional that Triviality results are used to pose—in particular, questions about the relationship between the probability of a conditional and the corresponding conditional probability—are often supposed to have been rendered moot (even nonsensical) by Kratzer's restrictor theory. I hope to have convinced you that this piece of theoretical folklore rests on a misunderstanding of the dialectic. Kratzer's skepticism about "conditional propositions" is not a logical consequence of the Restrictor Theory (indeed, in adopting it, Kratzer embraces a form of Triviality: she denies that what is said with a conditional is ever the object of degreed belief or epistemic necessity). And although the moderate Kratzerian theory—i.e., a Kratzerian theory that countenances conditional claims as the objects of degreed belief or

[23] A similar story will apply more generally—wherever conditional confidence of d in C given A implies that for any contextually appropriate confidence measure Pr over chance hypotheses, $\frac{\mathrm{Pr}(A\Rightarrow C)}{\mathrm{Pr}(A\Rightarrow C)+\mathrm{Pr}(A\Rightarrow\neg C)} = d$.

[24] For the same reason, recalling cases (8) and (9), one cannot hold that there is a chance the marble is big and that there's no chance it's red, while holding that there's a chance that if it's big, it will be red.

epistemic necessity—undermines Stalnaker's Thesis, in full generality, the threat still looms.

This paper also explored how appeal to an *information-sensitive* logical and semantic theory—for conditionals, epistemic modals, and probabilistic operators—might allow us to constructively theorize the kinds of linguistic judgments that induce Triviality when they are accommodated within broadly classical logical and semantic frameworks. According to the interpretation developed here, Triviality reveals a logical tension between basic facts about the logic of acceptance, presupposition, and similar states, and classical understandings of logical inconsistency. The choice is simple, and confronts any theorist working on the semantics and logic of conditionals, epistemic modals, and probability operators, Kratzerian or otherwise: reject the data, or amend the logic.

References

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

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

Bradley, Richard. (2002). Indicative conditionals. *Erkenntnis, 56*, 345–378. https://doi.org/10.1023/A:1016331903927

Bradley, Richard. (2007). A defence of the Ramsey Test. *Mind, 116*, 1–21. https://doi.org/10.1093/mind/fzm001

Charlow, Nate. (2016). Triviality for restrictor conditionals. *Noûs, 50*, 533–564. https://doi.org/10.1111/nous.12111

Charlow, Nate. (2019). The spectre of triviality. *Analysis, 79*, 595–605. https://doi.org/10.1093/analys/anz011

Charlow, Nate. (2020). Grading modal judgement. *Mind, 129*, 769–807. https://doi.org/10.1093/mind/fzz028

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

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

de Finetti, Bruno. (1936). La logique de la probabilité. In Hermann & Cie (eds.) Actes du Congrès International de Philosophie Scientifique, Vol. 4, 565–573. http://www.brunodefinetti.it/Opere/logiquedelaProbabilite.pdf.

von Fintel, Kai. (2007). *If*: The biggest little word. http://web.mit.edu/fintel/gurt-slides.pdf. Talk delivered at Georgetown University Roundtable.

von Fintel, Kai & Anthony S. Gillies. (2015). Hedging your *if*s and vice versa. http://mit.edu/fintel/fintel-gillies-2015-hedging.pdf. Ms., MIT/Rutgers.

von Fintel, Kai and Sabine Iatridou. (2005). What to Do If You Want to Go to Harlem. https://web.mit.edu/fintel/fintel-iatridou-2005-harlem.pdf.

Gillies, Anthony S. (2004). Epistemic conditionals and conditional epistemics. *Noûs, 38*, 585–616. https://doi.org/10.1111/j.0029-4624.2004.00485.x

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

Goldstein, Simon. (2021). Epistemic modal credence. Philosophers' Imprint 21: 1–24. https://philpapers.org/rec/GOLEMC.

Goldstein, Simon & Paolo Santorio. (2021). Probability for epistemic modalities. Philosophers' Imprint 21: 1–37. https://philpapers.org/rec/GOLPFE-2.

Hájek, Alan. (1989). Probabilities of conditionals–revisited. *Journal of Philosophical Logic, 18*, 423–428. https://doi.org/10.1007/BF00262944

Hájek, Alan. (2012). The fall of 'Adams' Thesis'? *Journal of Logic, Language, and Information, 21*, 145–161. https://doi.org/10.1007/s10849-012-9157-1

Hawke, Peter, & Steinert-Threlkeld, Shane. (2021). Semantic expressivism for epistemic modals. *Linguistics and philosophy, 44*, 475–511. https://doi.org/10.1007/s10988-020-09295-7
Huitink, Janneke. (2008). Modals, Conditionals, and Compositionality. Ph.D. thesis, Radboud Universiteit Nijmegen. Ph.D. Dissertation.
Khoo, Justin & Matthew Mandelkern. (2018). Against preservation. Analysis. Forthcoming.
Khoo, Justin, & Mandelkern, Matthew. (2018). Triviality results and the relationship between logical and natural languages. Mind https://doi.org/10.1093/mind/fzy006. Forthcoming.
Kratzer, Angelika. (1986). Conditionals. *Chicago Linguistics Society, 22*, 1–15.
Kratzer, Angelika. (1991). Conditionals. In A. von Stechow & D. Wunderlich (Eds.), *Semantics: An International Handbook of Contemporary Research* (pp. 651–656). Berlin: De Gruyter.
Kratzer, Angelika. (2012). *Modals and Conditionals*. Oxford: Oxford University Press.
Lewis, David. (1975). Adverbs of quantification. In E. L. Keenan (Ed.), *Formal semantics of natural language.* Cambridge: Cambridge University Press.
Lewis, David. (1976). Probabilities of conditionals and conditional probabilities. *The Philosophical Review, 85*, 297–315. https://doi.org/10.2307/2184045
Mandelkern, Matthew. (2019). Bounded modality. *Philosophical Review, 128*(1), 1–61. https://doi.org/10.1215/00318108-7213001
McDermott, Michael. (1996). On the truth conditions of certain 'If'-Sentences. *The Philosophical review, 105*, 1–37. https://doi.org/10.2307/2185762
Moss, Sarah. (2015). On the semantics and pragmatics of epistemic vocabulary. *Semantics and Pragmatics, 8*, 1–81. https://doi.org/10.3765/sp.8.5
Moss, Sarah. (2018). *Probabilistic Knowledge*. Oxford: Oxford University Press.
Partee, Barbara H., & Rooth, Mats. (1983). Generalized conjunction and type ambiguity. In R. Bäuerle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, Use, and Interpretation of Language* (pp. 361–383). Berlin: Walter de Gruyter.
Reichenbach, Hans. (1949). The Theory of Probability: An Inquiry Into the Logical and Mathematical Foundations of the Calculus of Probability. University of California Press.
Rothschild, Daniel. (2012). Expressing credences. *Proceedings of the Aristotelian Society CXII, Part, 1*, 99–114. https://doi.org/10.1111/j.1467-9264.2012.00327.x
Rothschild, Daniel. (2013). Do indicative conditionals express propositions? *Noûs, 47*, 49–68. https://doi.org/10.1111/j.1468-0068.2010.00825.x
Rothschild, Daniel. (2014). Capturing the relationship between conditionals and conditional probability with a trivalent semantics. *Journal of Applied Non-Classical Logics, 24*, 144–152. https://doi.org/10.1080/11663081.2014.911535
Rothschild, Daniel. (2015). Conditionals and propositions in semantics. *Journal of Philosophical Logic, 44,* 781–791. https://doi.org/10.1007/s10992-015-9359-5.
Rothschild, Daniel. 2020. A note on conditionals and restrictors. In John Hawthorne & Lee Walters (eds.) Conditionals, Probability, and Paradox: Themes from the Philosophy of Dorothy Edgington, 19–39. Oxford University Press. https://doi.org/10.1093/oso/9780198712732.003.0003.
Rothschild, Daniel, & Yalcin, Seth. (2017). On the dynamics of conversation. *Noûs, 51*, 24–48. https://doi.org/10.1111/nous.12121
Santorio, Paolo. forthcoming. Trivializing informational consequence. Philosophy and phenomenological research https://doi.org/10.1111/phpr.12745.
Stalnaker, Robert. (1970). Probability and conditionals. *Philosophy of Science, 37*, 64–80.
Stalnaker, Robert. (2014). *Context*. Oxford: Oxford University Press.
Vandenburgh, Jonathan. (2020). Triviality results, conditional probability, and restrictor conditionals. https://philpapers.org/rec/VANTRC-4.
Veltman, Frank. (1996). Defaults in update semantics. *Journal of Philosophical Logic, 25*, 221–261. https://doi.org/10.1007/BF00248150
Weatherson, Brian. (2004). Moore, Bradley, and indicative conditionals. Ms.
Willer, Malte. (2010). New surprises for the ramsey test. *Synthese, 176*, 291–309. https://doi.org/10.1007/s11229-009-9494-z
Willer, Malte. (2013). Dynamics of epistemic modality. *The Philosophical Review, 122*, 45–92. https://doi.org/10.1215/00318108-1728714
Yalcin, Seth. (2007). Epistemic modals. *Mind, 116*, 983–1026. https://doi.org/10.1093/mind/fzm983
Yalcin, Seth. (2012). Bayesian expressivism. *Proceedings of the Aristotelian Society CXII, Part, 2*, 123–160. https://doi.org/10.1111/j.1467-9264.2012.00329.x

Yalcin, Seth. 2012b. Context probabilism. In M. Aloni, V. Kimmelman, F. Roelofsen, G. Sassoon, K. Schulz & M. Westera (eds.) Proceedings of the 18th Amsterdam Colloquium, 12–21. ILLC. https://doi.org/10.1007/978-3-642-31482-7_2.