Version of 27 March 2019. Forthcoming in *Mind*.

Grading Modal Judgment

Nate Charlow*

Abstract

This paper proposes a new model of graded modal judgment. It begins by problematizing the phenomenon: given plausible constraints on the logic of epistemic modality, it is impossible to model graded attitudes toward modal claims as judgments of probability targeting epistemically modal propositions. This paper considers two alternative models, on which modal operators are non-proposition-forming: (1) Moss (2015), in which graded attitudes toward modal claims are represented as judgments of probability targeting a "proxy" proposition, belief in which would underwrite belief in the modal claim. (2) A model on which graded attitudes toward modal claims are represented as judgments of credence taking as their objects (non-propositional) modal representations (rather than proxy propositions). The second model, like Moss' model, is shown to be semantically and mathematically tractable. The second model, however, can be straightforwardly integrated into a plausible model of the role of graded attitudes toward modal claims in cognition and normative epistemology.

1 Introduction

Agents can bear graded attitudes (e.g., intermediate or high credence) towards epistemic modalities.[1] Sentences expressing such graded attitudes are commonplace; consider the following triad (adapted from Moss 2015: 4):

(1) It is probably the case that Trump might be impeached.

(2) It is probably the case that Trump will be impeached.

(3) Trump might be impeached.

Moss remarks that "our judgments suggest that [(1)] is weaker than either [(2)] or [(3)]." Believing (2) "is intuitively sufficient reason to bet at even odds" that Trump will be impeached, "whereas merely believing [(1)] is not" (Moss 2015: 4). Meanwhile, asserting

*Special thanks to Simon Charlow for generous and extended feedback on this paper (and for introducing me to this paper's core formal idea, type polymorphism), without which I really don't think it would or could have worked out. The same is true of the feedback provided by this journal's referees. Simon Goldstein, Sarah Moss, Paolo Santorio, and Jonathan Weisberg also offered feedback or encouragement at crucial junctures for this project. My thanks to all.

[1] By 'epistemic modality', I mean a sentence (or sometimes the content of a sentence) of the form $O\phi$, where O is an autocentrically interpreted epistemic operator and ϕ is its sentential prejacent. An autocentric interpretation of a sentence of the form $O\phi$ is an interpretation according to which the speaker is interpreted as making an epistemic claim, "based on", or from the "vantage" of, their own information/evidence (cf. Lasersohn 2005). Epistemic operators are here understood to encompass genuinely modal operators ('must', 'might'), epistemic or probabilistic adverbs ('probably', 'certainly', 'possibly'), numerical probability operators ('it is n-likely that'), and more.

(3) represents the speaker as believing that Trump might be impeached; (1) does not.[2]

The basic data point can be established in various ways. Consideration of (1)–(3) suggests the existence of sentences of natural language serving to express graded attitudes towards epistemic modalities. Graded attitudes towards epistemic modalities also appear to be presupposed by platitudes about the conversational role of epistemic modalities. Willer (2013), for instance, observes that assertions of epistemic modalities are understood as non-trivial proposals to add information to—that is, address a question within—a discourse. Assertion of a sentence like (3) addresses a question about whether Trump might be impeached:

(4) Might Trump be impeached?

But the notion of such a question seems to *presuppose* the possibility of a graded attitude (i.e., a degree of confidence greater than 0 and less than 1) toward a sentence like (3). Such an attitude typically forms at least part of the cognitive basis for entertaining (or explicitly posing) such a question; the question is generally occasioned by the questioner's bearing a graded attitude toward an epistemically modal representation.

This paper proposes a new model of graded modal judgment. It begins (§2) by problematizing the phenomenon for classical, truth-conditional accounts of the semantic content of epistemic modalities: given plausible constraints on the logic of epistemic modality, it is actually *impossible* to model graded attitudes toward modal claims as judgments of *probability* taking modalized propositions as their objects (Charlow 2016b, forthcomingb). In response to this problem, this paper considers two alternative models, on which modal operators are *non-proposition-forming* operators:

- §3: Moss (2015), on which a graded attitude toward a modal claim is represented as a degreed belief taking a "proxy" proposition, belief in which would underwrite belief in the modal claim, as its object.
- §4: A model on which a graded attitude toward a modal claim is represented as a degreed belief taking as its object a (non-propositional) modal representation (rather than a proxy proposition).

[2] Yalcin (2009) argues (in service of a more general skepticism about the semantic productivity of iterating epistemic vocabulary) that a speaker who asserts that it might be the case that Trump might be impeached is committed to allowing that Trump might be impeached. I deny this (but also its relevance to the target phenomenon for this paper). Claim: the truth (or assertability) in a context c of a sentence of the form $\Diamond\phi$ implies that ϕ is a relevant epistemic possibility in c (and vice versa). Therefore, a speaker who asserts this claim at c has *not* made a mistake, if Trump's being impeached is not a relevant epistemic possibility at c (i.e., it would be improper to assert (3) in c). This suggests the content of the speaker's assertion is weaker than with (3). In any case, it is hard to see how to extend Yalcin's argument—as it must be extended, if the aim is to deny the semantic productivity of iterating epistemic vocabulary—to possibility modals scoping over 'probably'. Someone who says that it's possible that Trump will probably be impeached is not committed to allowing that Trump will probably be impeached (cf. Moss 2018: 46). Finally, even if successful, arguing that a speaker who asserts that it might be the case that Trump might be impeached is *committed to allowing that* Trump might be impeached is insufficient to establish that the content of the speaker's assertion entails that Trump might be impeached—particularly given a plausible argument that the content of the speaker's assertion is weaker than the content of (3).

The second model is shown to be theoretically tractable—a feature that does *not* ultimately distinguish it from Moss' model. Since, however, Moss argues against accounts of the second type, such a model is worth developing, even if only as a proof of concept. In §5, I argue that such a model deserves attention, not only as a proof of concept, but also because it is straightforwardly integrated into a plausible understanding of the functional role of graded attitudes toward modal claims in both cognition and normative epistemology. A view of this shape has some claim to being regarded as the null hypothesis about the target phenomenon.

2 No Uncertainty?

Most everyone would agree that the base semantic clause for the epistemic possibility modal $\Diamond$ (and its dual operator $\Box$) is information-sensitive—i.e., involves reference within the semantic metalanguage to a state of information—and that, relative to a "base" state of information—for present purposes, this is modeled as a (possibly constant and/or partial) function from worlds of evaluation into sets of possible worlds—epistemic possibility modals quantify existentially over possibilities compatible with that state. Relative to a choice of information state σ and a choice of index of evaluation w, the appropriate semantic clause for $\Diamond$ is as follows:

$$[\![\Diamond\phi]\!]^{\sigma,w} = \mathsf{T} \Leftrightarrow \exists v \in \sigma_w : [\![\phi]\!]^{\sigma,v} = \mathsf{T}$$

A sentence of the form $\Diamond\phi$ thus expresses a possible worlds proposition, namely:

$$[\![\Diamond\phi]\!]^{\sigma} = \{w : \exists v \in \sigma_w : [\![\phi]\!]^{\sigma,v} = \mathsf{T}\}$$

Such a proposition is the sort of thing to which a probability function can, in principle, assign a probability, and is the sort of thing toward which agents can, in principle, bear graded attitudes (e.g., being 10% confident in this proposition).

On the other hand, there is apparently strong evidence that sentences of the form $\Diamond\phi$ *cannot* generally express possible worlds propositions with these sorts of characteristics. First, assume that, for any w and σ, $[\![\Box\phi \wedge \neg\phi]\!]^{\sigma,w} = \mathsf{F}$ and $[\![\Diamond\phi \wedge \neg\Box\Diamond\phi]\!]^{\sigma,w} = \mathsf{F}$.

(5) #It must be raining, but it isn't.

(6) #It may be raining, but maybe it can't be.

In the present setting, this is equivalent to assuming that σ is governed by constrains of Reflexivity and Euclideanness.[3]

Reflexivity: $w \in \sigma_w$	$\forall w, \sigma : [\![\Box\phi \supset \phi]\!]^{\sigma,w} = \mathsf{T}$
Euclideanness: $v \in \sigma_w \Rightarrow \sigma_w \subseteq \sigma_v$	$\forall w, \sigma : [\![\Diamond\phi \supset \Box\Diamond\phi]\!]^{\sigma,w} = \mathsf{T}$

[3] On Euclideanness, see Appendix B.1 and Charlow (2016b). These are standard assumptions in the semantics/logic of epistemic modalities (see, e.g., Holliday & Icard III 2010; Gillies 2010; von Fintel & Gillies 2010, 2011, 2018). The phenomena of interest in this paper will also arise for modalities of belief (axiomatized by KD45, rather than S5).

These constraints imply that information states are epistemically transparent:

Transparency: $v \in \sigma_w \Rightarrow \sigma_w = \sigma_v$

$$\forall w, v, \sigma : v \in \sigma_w \Rightarrow [\![\Diamond\phi]\!]^{\sigma,w} = [\![\Diamond\phi]\!]^{\sigma,v}$$

Given Transparency, epistemic modalities are "rigid" relative to a choice of σ and w: if ϕ is a sentence of the form $\Diamond\psi$ or $\Box\psi$ and $[\![\phi]\!]^{\sigma,w} = \mathsf{T}$, then, for any $v \in \sigma_w$, $[\![\phi]\!]^{\sigma,v} = \mathsf{T}$. Hence, whenever $[\![\phi]\!]^{\sigma,w} = \mathsf{T}$:

$$\sigma_w \subseteq [\![\phi]\!]^{\sigma}$$

The basic difficulty this gives rise to is this: like an agent's degree of belief in any proposition, an agent's degree of belief in an epistemically modal proposition *ought to be* a probability. And yet it provably *cannot* be. (Conclusion: a graded modal judgment is not to be understood in terms of a degree of belief in an epistemically modal proposition.) To illustrate, consider a probabilistically coherent agent[4] A bearing a graded attitude—confidence $\in$ (0,1)—towards $\Diamond p$ (alternatively, $\Box p$). Although this may already be apparent to some readers, it is worth underlining that (as well as why) such an agent is impossible to represent within the present framework.

Let σ^A_w represent A's information at w. If p is possible (alternatively, necessary) for A at w, then $[\![\Diamond p]\!]^{\sigma^A,w} = \mathsf{T}$ (alternatively, $[\![\Box p]\!]^{\sigma^A,w} = \mathsf{T}$). But then, in view of Transparency, $[\![\Box\Diamond p]\!]^{\sigma^A,w} = \mathsf{T}$ (alternatively, $[\![\Box\Box p]\!]^{\sigma^A,w} = \mathsf{T}$). The difficulty is this: if A is probabilistically coherent, and A's information entails $\Diamond p$ (alternatively, $\Box p$), then A must assign $\Diamond p$ (alternatively, $\Box p$) probability 1. It follows that A's confidence in $\Diamond p$ or $\Box p$ must be extremal (0 or 1).[5]

Proof. Consider any probabilistically coherent agent A; let σ^A_w be A's information at w and Pr^A_w be a probability measure for A at w. Either $\exists v \in \sigma_w : [\![p]\!]^{\sigma,v} = \mathsf{T}$ or $\forall v \in \sigma_w : [\![p]\!]^{\sigma,v} = \mathsf{F}$. If $\exists v \in \sigma_w : [\![p]\!]^{\sigma,v} = \mathsf{T}$, then $[\![\Diamond p]\!]^{\sigma,w} = \mathsf{T}$, in which case $\sigma_w \subseteq [\![\Diamond p]\!]^{\sigma}$ and $Pr^A_w([\![\Diamond p]\!]^{\sigma}) = 1$. If $\forall v \in \sigma_w : [\![p]\!]^{\sigma,v} = \mathsf{F}$, then $[\![\Diamond p]\!]^{\sigma,w} = \mathsf{F}$, in

[4] A probabilistically coherent agent is one whose degrees of belief in some σ-algebra of some subset of W are representable with a probability function.

[5] Worth underlining: this is *not* an artifact of the use of sets of possibilia to model states of information. So long as the class of models for an epistemically modal language is required to satisfy object language analogues of Reflexivity ($\Box\phi \supset \phi$) and Euclideanness ($\Diamond\phi \supset \Box\Diamond\phi$)—and logical consequence is closed under modus ponens, i.e., $\Gamma \vdash \phi \supset \psi$ implies that $\Gamma \cup \{\phi\} \vdash \psi$—the logic of epistemic modality will be constrained by the following entailments:

$$\Box\phi \dashv\vdash \Box\Box\phi \qquad \Diamond\phi \dashv\vdash \Box\Diamond\phi$$

(To preview, on the account defended here, these entailments will fail in the left-to-right direction, even though for any context c, $[\![\Box\phi \wedge \neg\phi]\!]^c = \emptyset$ and $[\![\Diamond\phi \wedge \neg\Box\Diamond\phi]\!]^c = \emptyset$; on this point, see Appendix B.1.) Let I^A_w designate A's information at w; we will *not* assume that I^A_w is a set of possible worlds. Now either $[\![\Diamond p]\!]^{I^A,w} = \mathsf{T}$ (if I^A_w is compatible with p) or $[\![\Diamond p]\!]^{I^A,w} = \mathsf{F}$ (otherwise). If $[\![\Diamond p]\!]^{I^A,w} = \mathsf{T}$, then $[\![\Box\Diamond p]\!]^{I^A,w} = \mathsf{T}$, in which case I^A_w is incompatible with $\neg\Diamond p$ (i.e., I^A_w entails $\Diamond p$). Since A is probabilistically coherent, $Pr^A_w(\Diamond p) = 1$. If, on the other hand, $[\![\Diamond p]\!]^{I^A,w} = \mathsf{F}$, then $[\![\Box\neg\Diamond p]\!]^{I^A,w} = \mathsf{T}$, in which case I^A_w is incompatible with $\Diamond p$ (i.e., I^A_w entails $\neg\Diamond p$). Since A is probabilistically coherent, $Pr^A_w(\Diamond p) = 0$. Thus, again, either $Pr^A_w([\![\Diamond p]\!]^{\sigma}) = 1$ or $Pr^A_w([\![\Diamond p]\!]^{\sigma}) = 0$.

which case $\sigma_w \cap [\![\Diamond p]\!]^\sigma = \emptyset$ and $Pr^A_w([\![\Diamond p]\!]^\sigma) = 0$. Thus, either $Pr^A_w([\![\Diamond p]\!]^\sigma) = 1$ or $Pr^A_w([\![\Diamond p]\!]^\sigma) = 0$. □

Within the "classical" semantic setting presupposed in this section, it seems that we confront a hard choice: between a revisionary logic of epistemic modality, and a revisionary understanding of the attitudes it is possible to bear toward sentences expressing subjective uncertainty. Given this sort of "classical" setting, if we maintain the assumption that (5) and (6) are inconsistent, the phenomenon of degreed belief in epistemic modalities is mystified: such a degree of belief cannot be represented as a judgment of probability targeting an epistemically modal proposition.

Nor, therefore, can we represent degrees of belief toward epistemic modalities using *sets* of probability measures taking epistemically modal propositions as their objects. According to the "Bayesian" proposal for representing such attitudes (Yalcin 2012; Rothschild 2012), "Where an agent assigns a determinate probability to a proposition, every measure in their credal set [i.e., the set of probability measures compatible with their information] assigns that probability to it. A probabilistic claim is true of a credal set just in case it is true on every probability measure in the set" (Rothschild 2012: 110). The difficulty is that, given the arguments of this section, a set of probability measures S is constrained so that, if ϕ is epistemically modal:

$$\forall Pr \in S : Pr(\phi) = 0 \text{ or } Pr(\phi) = 1$$

Attitudes of intermediate confidence (e.g., confidence n) toward a sentence ϕ are represented, according to the Bayesian proposal, with sets of probability measures, all of which assign probability $\geq n$ to ϕ. No probability measure assigns intermediate confidence to ϕ if ϕ is epistemically modal. And so, given the Bayesian proposal, no set of probability measures can represent attitudes of intermediate confidence toward epistemic modalities.

3 Linguistic Gradability and Gradability of Content

Moss (2015) provides a semantics for epistemic operators, including a sentential operator $\triangle$ expressing high confidence in its complement. On Moss' semantics, a sentence of the form $\triangle\phi$ (read: it is probable that ϕ) expresses a constraint on probability measures, namely, the constraint that a probability measure satisfies just if it is in the following set:

$$[\![\triangle_1 \phi]\!]^c = [\![\triangle_1]\!]^c([\![\phi]\!]^c) = \{m : m(\bigcup\{p \in g_c(1) : m|_p \in [\![\phi]\!]^c\}) > .5\}$$

Epistemic operators are, in general, interpreted relative to contextually salient partitions (i.e., contextually relevant questions); numerical indices (e.g., subscripted '1') are mapped to contextually salient partitions by a contextual variable assignment g_c. Thus, $\triangle\phi$ expresses a constraint on probability measures that m satisfies iff m assigns this proposition a value exceeding .5:

$$\bigcup\{p \in g_c(1) : m|_p \in [\![\phi]\!]^c\}$$

The object that receives a probability value (according to a probability measure m

satisfying the constraint semantically expressed by the sentence) is the disjunction of those propositions in the salient partition that confirm ϕ—i.e., the disjunction of those propositions p such that, if m were conditionalized on p, m would satisfy the constraint expressed by ϕ.

Moss' account handles iterated epistemic operators with ease. To illustrate, suppose $\phi = \Box_2 p$. Then $\triangle_1 \phi$ expresses a constraint on probability measures that m satisfies iff m assigns this proposition a value exceeding .5:

$$\bigcup \{q \in g_c(1) : m|_q \in [\![\Box_2 p]\!]^c\}$$

This is, roughly, the constraint that m satisfies just if m regards a salient propositional disjunction as probable: in particular, the disjunction of those answers q to the contextually salient question $g_c(1)$ such that conditioning m on q would make it the case that m regards p as necessary. In shorthand, it is the constraint that m satisfies just if m regards as likely some disjunction such that any way of m's coming to accept that disjunction would amount to m's regarding p as necessary.[6] It is obvious why, on this model, it is sensible for a probabilistically coherent agent to think (as well as to express the thought) that p is probably necessary, without at the same time thinking (or being committed to expressing the thought) that p is necessary: $[\![\Box_2 p]\!]^c$ encodes a stronger constraint on measures—in Moss' system, it is the constraint that m satisfies just if, for any way r of answering $g_c(2)$, $m|_r([\![p]\!]^c) = 1$ (Moss 2015: 27). Of course, this is a constraint that a measure satisfying the constraint encoded in $[\![\triangle_1 \Box_2 p]\!]^c$ will not generally satisfy.

3.1 Worldly Representation

A noteworthy feature of Moss' account is that the semantic content of an epistemically modalized sentence—an epistemically modal representation—is not *itself* assessable within the semantic metalanguage as (e.g.) probable. For Moss, $\triangle \phi$ semantically rules out probability measures that do not regard as likely the disjunction of the propositions conditionalization on which is sufficient for believing ϕ. Strictly speaking, $[\![\triangle \phi]\!]^c$ does *not* rule out measures according to which $[\![\phi]\!]^c$ is not likely; when ϕ is epistemically

[6] An informative comparison is with Gaifman (1988)'s classic approach to higher-order uncertainty. Gaifman represents higher-order uncertainty as uncertainty about the "true" probabilities: the claim "there is a 70% chance that later this evening the chance of rain tomorrow will be 80%" is represented as an assignment of probability to a proposition about some future event: in this case, the event in which the "true" probability of rain tomorrow (according to some contextually salient estimate of the probability of rain tomorrow) is 80%. Like Moss, Gaifman treats higher-order uncertainty (to degree d) about epistemically modal ϕ as *factual uncertainty* (to degree d) about the truth of some proposition p such that, if the agent conditionalizes on p, the agent thinks that ϕ (here, see Gaifman's discussion of his Axiom VI at 281–2.). Unlike Moss, Gaifman takes the content of 'the chance of rain tomorrow will be 80%' to be a possible worlds proposition that is true at w iff the "true" probability of rain tomorrow at w (according to the contextually salient estimate at w) is 80%. Apart from the metasemantic quandaries associated with this proposal (in virtue of what is the "true" probability of some event at w determined relative to a context?), Gaifman's account (unlike Moss') is caught up in the dialectic of §2: if $\Diamond$ is a propositional operator, the logic of $\Diamond$ is S5, which trivializes degreed belief in sentences of the form $\Diamond \phi$. (Thanks to a referee for this journal for drawing my attention to Gaifman's paper.)

modal, Moss denies that $[\![\phi]\!]^c$ is the sort of thing that over which a measure is defined (since it is not, on Moss' semantics, a proposition).

To be clear, a view like this certainly makes sense—indeed, given §2, would be *forced*—if probability measures (qua devices for representing agential degrees of belief) may be defined only over σ-algebras of sets of possible worlds. And—given that the theoretical purpose of invoking probability measures (for this application) is to model an agent's degree of belief *in some representation*—assuming that probability measures are defined only over σ-algebras of sets of possible worlds makes sense, too, if the cognitive state of representation is to be understood in terms of representing some way the actual world could (not) be (as, e.g., in Stalnaker 1984).

3.2 Generalizing Representation

This section will try to motivate an alternative picture. It identifies some reasons in favor of adopting a "flexible" or "generalized" notion of representation, according to which (1) expresses, and (7) ascribes, a high degree of belief in the representation encoded in (3).

(7) Marcy thinks it is probably the case that Trump might be impeached.

This section suggests, more generally, that it is theoretically natural to treat the semantic value of an epistemic modality—an epistemically modal representation—as a kind of *content*: an entity toward which we may represent agents as bearing representational attitudes (like degreed belief).[7]

Degreed attitudes towards representations of any type are rationally subject to the same laws of probability. Why is this? Why, e.g., is it *irrational* to believe either of (8) or (9)?

(8) It is probable that Trump might be impeached, and it's probable that he can't be.

(9) Trump will probably be impeached, and he probably won't be impeached.

Generally, why is it that, for any ϕ whatever, it is irrational to believe $\triangle\phi$ while believing $\triangle\neg\phi$, i.e., to think that both ϕ and its negation are probable? A natural, if incomplete, explanation: for *any* representation R, an agent's degree of belief in R summed with her degree of belief in $\neg R$ should never exceed a designated value representing full/outright acceptance. Notice that developing this sort of off-the-shelf explanation means countenancing degrees of belief as properties of representations of any type (in order to state, and ultimately account for, rational norms that govern the suite of graded attitudes agents

[7]This is already part and parcel of the "classical" propositional account of epistemic modalities discussed in §2. Moss relinquishes this bit of the "classical" account to provide a treatment of graded modal judgment; this section offers some pushback. To be sure, the considerations raised in this section are tentative, and are not intended as a serious brief against Moss' theory. The relevant context for this discussion is Moss' remark that "It is difficult to independently motivate such an arcane model of our mental life" as the sort of model that this paper goes on to propose (Moss 2015: 30). This section (and §5) argues that the model of mental life that this alternative picture evokes is relatively commonsensical, once we decide to make room for higher-order uncertainty in language and thought.

can bear toward representations of any type).[8]

To similar effect, consider a decision situation in which the desirability of an action *depends* on an epistemically modal representation. Agents who are not risk-neutral confront such decision situations routinely. On a simple model of attitudes toward risk (cf. Buchak 2013), risk-inclined agents assign additional weight to low probability payoffs, while risk-averse agents assign a negative weight to low probability payoffs; rational agents that are not risk-neutral seek to maximize risk-weighted expected value. For any decision situation in which a risk-inclined agent believes to degree n that there is a low probability that p—for simplicity, imagine the agent treats p as low probability just when the agent accepts $\neg\triangle p$—an agent who seeks to *maximize* risk-weighted expected value may compute (when her information allows her to do so[9]) an *expected* risk-weighted expected value.[10]

---

[8]Schroeder (2011); Staffel (forthcoming) argue in a similar fashion. There is certainly work to do to build a "natural" account of this phenomenon that is both functional and theoretically appealing—but that will be the work of this paper. To preview, on the account defended here (§5), an agent who believes $\triangle\phi \wedge \triangle\neg\phi$ makes a specific kind of rational mistake: they are subject to (Generalized) Dutch Books. Interestingly, such an agent, unlike an agent who believes (9), will *not* be subject to having Accuracy-Dominated credences—something that should be seen as congenial to a generally "non-factualist" or "Bayesian" perspective on probabilistic thought and talk (cf. Yalcin 2011).

[9]This qualification is necessary since I do not want to assume an agent is *always* in a position to compute such a value. That said, there are certainly decision situations in which an agent is in a position to compute such a value. Imagine α is entertaining two probability functions Cr_1 and Cr_2 such that (i) $Cr_1(p) = .7$ and $Cr_2(p) = .3$; (ii) if α were to adopt either Cr_1 or Cr_2, $REU_(X)$ would be defined for α. In such a decision situation, evidently, $REU(X|\neg\triangle p)$ is just the value of $REU(X)$ computed using a credence function Cr $(=Cr_2) \in [\![\neg\triangle p]\!]^c$, whereas $REU(X|\triangle p)$ is just the value of $REU(X)$ computed using a credence function Cr $(=Cr_1) \in [\![\triangle p]\!]^c$.

This example provides a good opportunity to note that there is an epistemological debate (from which I will prescind in this paper) about how probabilistically unspecific evidence (evidence that fails to settle a precise probability for a relevant possibility) constrains an agent's credences (and the choices an agent might base on her credences). (On certain models of "imprecise" credence, entertaining a set of possibilities such as $\{Cr_1, Cr_2\}$ would be disallowed. See the discussion of the controversial "convexity" requirement in Moss (2019) as well the overview in Joyce (2005).)

This paper will also prescind from the debate over general decision rules for imprecise decision problems (for several criteria of adequacy for such a decision rule, see Joyce 2010: 313ff). This paper does attempt to state a natural decision rule for *one type* of imprecise decision problem in which an agent seeks to maximize some value that is dependent on an epistemically modal representation. Although I believe the ability to formulate such a decision rule represents an advantage of this account over the relatively coarse-grained decision rules put forward in alternative frameworks for representing higher-order uncertainty (e.g., Gärdenfors & Sahlin 1982), I defer comparison to another time.

[10]Risk-sensitivity is just one illustration of how epistemically modal representations can impinge on practical reasoning: decision situations in which an agent confronts "higher-order uncertainty" (uncertainty about the probabilities of the relevant states in their "base" decision situation) serve the purpose too. In decision situations characterized by higher-order uncertainty, a maximizing agent—one who seeks to maximize any quantity that depends on an epistemically modal representation (including ordinary expected value)—will tend to prefer an action they regard as likely to maximize that quantity to actions that they regard as unlikely to maximize it. Such preferences can be characterized straightforwardly in the framework this paper proposes.

$$X \quad \overbrace{n \cdot REU(X|\neg\triangle p)}^{\overbrace{\neg\triangle p}^{n}} \quad \overbrace{(1-n) \cdot REU(X|\triangle p)}^{\overbrace{\triangle p}^{1-n}}$$

Suppose I strive to maximize risk-weighted expected value, and suppose I only care about the risk of D (=drinking from the lake) when it's probable that A (=there's an algae bloom). Suppose that my evidence is unspecific as to the "base" probabilities in my decision situation—for example, it does not settle whether or not A is probable—although I do have evidence about the higher-order likelihoods, according to which it is .6 likely that A is not probable. Here is a decision table for this decision situation:[11]

	$\overbrace{\neg\triangle A}^{.6}$	$\overbrace{\triangle A}^{.4}$
D	$REU(D\|\neg\triangle A)$	$REU(D\|\triangle A)$
$\neg D$	$REU(\neg D\|\neg\triangle A)$	$REU(\neg D\|\triangle A)$

If I strive to maximize risk-weighted expected value, and my decision situation is accurately represented in the above decision table, it seems to make rational sense for me to drink, if:

$$(.6) \cdot REU(D|\neg\triangle A) + (.4) \cdot REU(D|\triangle A) \;>\; (.6) \cdot REU(\neg D|\neg\triangle A) + (.4) \cdot REU(\neg D|\triangle A)$$

Three (I hope modest) observations. First, decision situations like this should be regarded as fairly commonplace, once we decide to make room for higher-order uncertainty in language and thought—a shared aim of Moss (2015) and this paper. Second, a mathematical model of practical reasoning that could be extended to cover decision making under higher-order uncertainty would be desirable. As Joyce (2010) observes, in frameworks admitting such uncertainty, "in many decisions there are no options that maximize... across [the probability functions compatible with one's evidence], but one must still choose" (314). Third, a natural first step in developing such a model is to adopt a generalized notion of representation, according to which (i) degrees of belief/confidence are treated as values that can attach to representations of any type, including epistemically modal representations; (ii) it is possible to formulate rational norms that govern the relations between the suite of degreed attitudes agents can bear toward representations of any type, and action. §§4–5 will propose just such an account.

[11] An alternative picture: my decision in this situation depends on how likely I regard some ordinary proposition p (e.g., that the pH of the lake water exceeds some value) such that updating on p implies thinking there is probably an algae bloom. But, by stipulation, my decision in this decision situation depends on how inclined I am to think that $\triangle A$. We might additionally suppose that my inclination toward this claim is a *prior*, in the sense that there is no salient proposition p such that (i) I am .4 confident in p and (ii) if I update on p, I'll believe $\triangle A$. Even if my inclination in this regard is a prior—if, for example, I regard $\triangle A$ as low probability, without regarding *any* salient proposition as low probability—it seems that I can confront the decision situation described here.

3.3 Signpost

As seen in §2, on "classical" (i.e., ordinary truth-conditional) accounts of the semantics of epistemic modalities, the content of epistemically modal ϕ, relative to a designated agent A's information, is unfit for being the object of graded attitudes for A (e.g., A being 10% confident that ϕ), according to standard "Bayesian" techniques for modeling those attitudes (and the semantics of sentences ascribing those attitudes). Moss (2015) provides the first analysis of epistemic modality on which it makes sense for speakers to express graded attitudes towards epistemic modalities. But, on that analysis, the semantic value of an epistemic modality—what I've called an epistemically modal representation—is not itself the object of an attitude like degreed belief. This section has identified some tentative motivations for treating such entities as the objects of such attitudes. §4 will develop a semantics for epistemic modality that formalizes and implements this sort of intuitive cognitive model. §5 picks up where §3 leaves off, arguing that there is a sensible, generalized notion of representation undergirding this semantics.

4 Credences in Representations

4.1 Introducing Credence

On the alternative I have in mind, we will define a new quantity, call it *credence*, for probability operators of natural language to uniformly express. To utter a sentence like (1) is simply to express high credence in the content of (i.e., representation semantically encoded in) (3).

$$[\![\triangle\phi]\!] = [\![\triangle]\!]([\![\phi]\!]) = \{Cr : Cr([\![\phi]\!]) > .5\}$$

Here is how I prefer to conceptualize this idea (at a very high level of abstraction). Some credences are *probabilities*: subjective estimates of objective chance of the truth of a worldly representation (alternatively, subjective estimates of actual-worldly truth value). Some credences are *not probabilities* (when a subject's credence cannot be understood as their estimate of objective chance of the truth of a worldly representation, or as a subjective estimate of actual-worldly truth value). There is, nevertheless, no obstacle to defining credences so that they *behave like* probabilities, whether or not the object of credence is a worldly or non-worldly representation.[12]

Begin by assuming that a *representation* is a set of objects of any semantic type—*a set of alternative possibilities* that witness the truth of ("satisfy") sentences of our language.[13] In general, sets of alternative possibilities represent sets of "candidates"

[12]Though the details are very different, this general perspective draws inspiration from Bradley's "Multi-Dimensional" approach towards the probabilities of indicative conditionals (Bradley 2012), as well as remarks in Staffel (forthcoming) discussing how an Expressivist might model the cognitive and normative characteristics of gradable attitudes towards non-factual semantic contents. Jonathan Weisberg (pc) alerts me to an earlier approach to higher-order probability (Hild 1998) that is similar in both spirit and certain modeling choices to the one developed here.

[13]For the purposes of this paper, representations will have a recursive (polymorphic) semantic type,

for different ways of representing; only some sets of alternative possibilities (i.e., sets of possible worlds) represent candidates for actuality (i.e., sets of alternative ways the world could be); other sets of alternative possibilities (e.g., sets of sets of possible worlds) represent candidates for ways of representing a set of candidates for actuality; and so on.[14]

Definition 1. *Let $\mathcal{R}$ be a representation. Then a set of representations $\{\mathcal{R}_1, ..., \mathcal{R}_n\}$* ***partitions*** *$\mathcal{R}$ iff, for all $1 \leq i \neq j \leq n$:*

$$\mathcal{R}_i \cap \mathcal{R}_j = \emptyset \qquad\qquad \bigcup_{i=1}^{n} \mathcal{R}_i = \mathcal{R}$$

Definition 2. *An* ***alternative set*** *for $\mathcal{R}$ is any set $\mathfrak{R}$ that partitions $\mathcal{R}$.*[15]

Definition 3. *If $\mathfrak{R}$ is an alternative set for $\mathcal{R}$, $\mathfrak{R}$'s* ***σ-closure*** *Σ is $\mathfrak{R}$'s closure under $\cup$, $'$.*

Consider any base representation $\mathcal{R}$, alternative set $\mathfrak{R} = \{\mathcal{R}_1, ..., \mathcal{R}_n\}$ for $\mathcal{R}$, and $\mathfrak{R}$'s σ-closure Σ. We will say that a representation $\mathcal{S}$ is **based on** $\langle \mathcal{R}, \mathfrak{R}, \Sigma \rangle$ iff $\mathcal{S} \in \Sigma$. We introduce the notion of a credence function that is based on $\mathcal{R}$, $\mathfrak{R} = \{\mathcal{R}_1, ..., \mathcal{R}_n\}$, and Σ, by requiring that credence functions be normalized to the base representation $\mathcal{R}$, and that it be additive over disjoint elements of Σ.

Definition 4. *A* ***credence function based on*** *$\mathcal{R}$, $\mathfrak{R}$, and Σ is a function $Cr : \Sigma \mapsto [0, 1]$ such that:*

$$Cr(\mathcal{R}) = 1$$

$$Cr(\bigcup_{i=1}^{n} \mathcal{S}_i) = \sum_{i=1}^{n} Cr(\mathcal{S}_i) \quad (i \neq j \Rightarrow \mathcal{S}_i \cap \mathcal{S}_j = \emptyset)$$

constructed from the basic materials of *possible worlds* and *credence functions* (for details, see §4.4). Although nothing here will turn on this, I am pragmatic about the basic materials: agents can represent sets of (and sets of sets of, and sets of sets of sets of) many kind of things—in addition to possible worlds and credence functions, preference orderings or decision rules (Charlow 2015, 2016a, 2018), degree thresholds for gradable adjectives (Charlow forthcominga), features of their subjective perceptual experience (Charlow 2019a), and more besides—as candidate possibilities for some purpose (on this generalized notion of representation, see §5.1). At the most general level, a representation is just a set of *possibilities*, and a possibility is just a semantic object against which a sentence can be *evaluated for satisfaction*, relative to a model.

[14] This is only a rough first pass at stating a functional psychological role for the representation of a set of alternatives of arbitrary semantic type. For an elaboration, see §5.

[15] Present purposes do not require using partitions here, but there are three advantages worth noting: invoking partitions (i) makes the extension of this architecture to decision theoretic applications very smooth (see §5.2); (ii) allows us to avoid certain apparent counterexamples to modus ponens (Charlow 2019b); (iii) will eventually allow us to connect a way of imposing structure on a relevant set of possibilities to a more concrete fact about a discourse (e.g., a set of relevant questions which generates the relevant partition, à la Roberts 1996; Yalcin 2011, 2018).

Possibly, $n = \infty$, in which case Cr is constrained by Normalization and Countable Additivity $[Cr(\bigcup_{i=1}^{\infty} \mathcal{S}_i) = \sum_{i=1}^{\infty} Cr(\mathcal{S}_i)]$. Possibly, $n \in \mathbb{N}$, in which case Cr is constrained by Normalization and Finite Additivity $[Cr(\bigcup_{i=1}^{n} \mathcal{S}_i) = \sum_{i=1}^{n} Cr(\mathcal{S}_i)]$.

Definition 5. *Given a credence function Cr based on $\mathcal{R}$, $\mathfrak{R}$, Σ, and $\mathcal{T} \in \Sigma$ the **conditionalization** of Cr on $\mathcal{T}$ is a function $Cr|_{\mathcal{T}}(\cdot) : \Sigma \mapsto [0,1]$ such that:*

i. *$Cr|_{\mathcal{T}}$ is based on $\langle \mathcal{R} \cap \mathcal{T}, \{\mathcal{R}' \cap \mathcal{T} : \mathcal{R}' \in \mathfrak{R}\}, \{\mathcal{R}' \cap \mathcal{T} : \mathcal{R}' \in \Sigma\}\rangle$*
ii. *If $\mathcal{R} \cap \mathcal{T} \subseteq \mathcal{U}$, then $Cr|_{\mathcal{T}}(\mathcal{U}) = 1$*
iii. *If $\mathcal{R} \cap \mathcal{T} \cap \mathcal{U} = \emptyset$, then $Cr|_{\mathcal{T}}(\mathcal{U}) = 0$*
iv. *Otherwise, $Cr|_{\mathcal{T}}(\mathcal{U}) = \frac{Cr(\mathcal{U} \cap \mathcal{T})}{Cr(\mathcal{T})}$*

Definition 6. *Given Cr based on $\mathcal{R}$, $\mathfrak{R}$, and Σ, a **conditional credence function** based on $\mathcal{R}$, $\mathfrak{R}$, and Σ is a function $Cr(\cdot|\cdot) : \Sigma \mapsto (\Sigma \mapsto [0,1])$ such that $Cr(\mathcal{S}|\mathcal{T}) = Cr|_{\mathcal{T}}(\mathcal{S})$.*

4.2 A Semantics of Representations

The driving semantic ideas are these: sentences of natural language *semantically encode representations*. Representations can be modeled as sets of possibilities. Possibilities come in many different semantic types: possible worlds, sets of possible worlds, sets of sets of possible worlds, credence functions, sets of credence functions, sets of sets of credence functions, and so on.

Consider a language containing a stock of atomic sentences A, Boolean compounds of sentences, the indicative conditional $\rightarrow$, the 'probably' operator $\triangle$, and the epistemic modal $\Diamond$.

$$\phi :: p \in \mathtt{A} \mid \neg\phi \mid \phi \wedge \psi \mid \phi \rightarrow \psi \mid \triangle\phi \mid \Diamond\phi$$

An interpretation function for this language maps sentences into representations. The obvious clauses would be as follows:

$$
\begin{aligned}
[\![p]\!] &= \{w : w(p) = 1\} \ (p \in \mathtt{A})\\
[\![\neg\phi]\!]_\tau &= \mathbb{U}_\tau - [\![\phi]\!]_\tau \ (X_\tau := X \text{ is a set of objects of semantic type } \tau)\\
[\![\phi \wedge \psi]\!] &= [\![\phi]\!] \cap [\![\psi]\!]\\
[\![\phi \rightarrow \psi]\!] &= \{Cr : Cr([\![\psi]\!] | [\![\phi]\!]) = 1\}\\
[\![\triangle\phi]\!] &= \{Cr : Cr([\![\phi]\!]) > .5\}\\
[\![\Diamond\phi]\!] &= \{\mathcal{S} : \mathcal{S} \cap [\![\phi]\!] \neq \emptyset\}
\end{aligned}
$$

These clauses require some elaboration. We have generalized the probability calculus to credence functions, by requiring that credence functions be specified *relative to* a (i) "base" representation $\mathcal{R}$, (ii) an alternative set $\mathfrak{R}$ that partitions $\mathcal{R}$, (iii) $\mathfrak{R}$'s σ-closure Σ. We will therefore say that the representation expressed by such sentences is determined relative to a base representation $\mathcal{R}$, alternative set $\mathfrak{R}$ for $\mathcal{R}$, and $\mathfrak{R}$'s σ-closure Σ. We will call a triple $\langle \mathcal{R}, \mathfrak{R}, \Sigma\rangle$ with these characteristics a **space**, and we will allow a context c to determine (via a contextually determined variable assignment g_c) a space (of the requisite semantic type) for each space-sensitive expression of our language.[16]

$$[\![\phi \rightarrow_1 \psi]\!]^c = \lambda Cr : Cr \text{ is based on } g_c(1) \,.\, Cr([\![\psi]\!]^c | [\![\phi]\!]^c) = 1$$
$$[\![\triangle_1 \phi]\!]^c = \lambda Cr : Cr \text{ is based on } g_c(1) \,.\, Cr([\![\phi]\!]^c) > .5$$
$$[\![\Diamond_1 \phi]\!]^c = \lambda \mathcal{S} : \mathcal{S} \text{ is based on } g_c(1) \,.\, \mathcal{S} \cap [\![\phi]\!]^c \neq \emptyset$$

I shall assume (for technical and empirical reasons) that the operators $\rightarrow_1$, $\triangle_1$, and $\Diamond_1$ presuppose the "visibility" (in the sense of Yalcin 2011, 2018) of their arguments in $g_c(1) = \langle \mathcal{R}_1, \mathfrak{R}_1, \Sigma_1 \rangle$. (In general, a representation $\mathcal{R}$ is visible in a partition $\mathfrak{R}$ iff for each $\mathcal{S} \in \mathfrak{R}$, $\mathcal{S} \cap \mathcal{R} = \mathcal{S}$ or $\mathcal{S} \cap \mathcal{R} = \emptyset$. Applying this notion of visibility, $[\![\triangle_1 \phi]\!]^c$ will be defined only if, for each $\mathcal{S} \in \mathfrak{R}_1$, $\mathcal{S} \cap [\![\phi]\!]^c = \mathcal{S}$ or $\mathcal{S} \cap [\![\phi]\!]^c = \emptyset$.) Taking these clauses in order:

- $[\![\phi \rightarrow_1 \psi]\!]^c$, when defined, is the property a credence function Cr based on $g_c(1)$ has iff Cr treats the representation expressed by $[\![\psi]\!]^c$ as certain conditional on the representation expressed by $[\![\phi]\!]^c$.
- $[\![\triangle_1 \phi]\!]^c$, when defined, is the property that a credence function Cr based on $g_c(1)$ has iff it assigns the representation expressed by $[\![\phi]\!]^c$ a value $> .5$.
- $[\![\Diamond_1 \phi]\!]^c$, when defined, is the property that a representation $\mathcal{S}$ based on $g_c(1)$ has iff $\mathcal{S}$ is compatible with the representation expressed by ϕ.

Equivalently, where the semantics types demand it, such properties (modeled as λ-abstracts) may be understood as sets of possibilities that satisfy the corresponding λ-abstract. For instance, where the semantic types demand it—or, indeed, wherever a set-theoretic understanding is more convenient—we let $[\![\phi \rightarrow_1 \psi]\!]^c = \{Cr : [\![\phi \rightarrow_1 \psi]\!]^c(Cr) = 1\}$.

4.3 Examples

Example 1. Consider the case of $\triangle$ scoping over a propositional atom. Let $g_c(1) = \langle \mathcal{R}, \mathfrak{R}, \Sigma \rangle$.

$$[\![\triangle_1 p]\!]^c = \{Cr \text{ based on } g_c(1) : Cr([\![p]\!]^c) > .5\}$$

Here, semantic types require that $\mathcal{R}$ be a set of worlds, e.g., $\{w_p, v_p, u_{\neg p}\}$; $\mathfrak{R}$ is a partition of $\mathcal{R}$ in which $[\![p]\!]^c$ is visible, e.g., $\{\{w, v\}, \{u\}\}$; Σ is $\mathfrak{R}$'s σ-closure. A sentence of the form $\triangle p$ expresses the property a credence function (based on $\mathcal{R}$, $\mathfrak{R}$, and Σ) has when it assigns the worldly representation encoded in p a value $> .5$.

[16] This notation (and the implementation via contextually determined variable assignments) is from Moss. Note, however, that indices in Moss' semantics uniformly resolve to partitions of W.

Example 2. Next consider an example involving $\triangle$ iterated over $\Diamond$.

$$\begin{aligned}\llbracket\triangle_1\Diamond_2 p\rrbracket^c &= \{Cr \text{ based on } g_c(1) : Cr(\llbracket\Diamond_2 p\rrbracket^c) > .5\}\\ &= \{Cr \text{ based on } g_c(1) : Cr(\{\mathcal{S} \text{ based on } g_c(2) : \mathcal{S}\cap\llbracket p\rrbracket^c \neq \emptyset\}) > .5\}\end{aligned}$$

Let $g_c(1) = \langle\mathcal{R}_1, \Re_1, \Sigma_1\rangle$ and $g_c(2) = \langle\mathcal{R}_2, \Re_2, \Sigma_2\rangle$. Here the semantics requires that:

- As above, $\mathcal{R}_2$ is a set of worlds; $\Re_2$ is a partition of $\mathcal{R}_2$ in which $\llbracket p\rrbracket^c$ is visible; Σ_2 is $\Re_2$'s σ-closure.
- $\mathcal{R}_1$ is a set of sets of worlds, e.g.

$$\{\underbrace{\{u\}_{\neg\Diamond_2 p}}_{A}, \underbrace{\{w, v, u\}_{\Diamond_2 p}}_{B}, \underbrace{\{w\}_{\Diamond_2 p}}_{C}\}$$

- $\Re_1$ is a partition of $\mathcal{R}_1$ in which $\llbracket\Diamond_2 p\rrbracket^c$ is visible, e.g., $\{\{A\}, \{B, C\}\}$

As intended, $\triangle_1\Diamond_2 p$ expresses the property a credence function has when it assigns the (non-worldly) representation encoded in $\llbracket\Diamond_2 p\rrbracket^c$ a value $> .5$. More precisely, it expresses the property a credence function defined over Σ_1 has when it assigns the set of possibilities $\mathcal{S} \in \mathcal{R}_1$ such that $\mathcal{S}$ is based on $g_c(2)$ (i.e., such that $\mathcal{S} \in \Sigma_2$) and $\mathcal{S}$ is compatible with $\llbracket p\rrbracket^c$—a value $> .5$.

Example 3. Next consider the reverse iteration. (As the types increase, I'll compress formal detail for the sake of readability.)

$$\begin{aligned}\llbracket\Diamond_1\triangle_2 p\rrbracket^c &= \{\mathcal{S} : \mathcal{S}\cap\llbracket\triangle_2 p\rrbracket^c \neq \emptyset\}\\ &= \{\mathcal{S} : \mathcal{S}\cap\{Cr \text{ based on } g_c(2) : Cr(\llbracket p\rrbracket^c) > .5\} \neq \emptyset\}\end{aligned}$$

As intended, $\Diamond\triangle p$ expresses the property a set of credence functions has when it contains a credence function that assigns the worldly representation encoded in p a value $> .5$.

Example 4. Finally two examples involving iterated epistemics:

$$\begin{aligned}\llbracket\Diamond_1\Diamond_2 p\rrbracket^c &= \{\mathcal{S} \text{ based on } g_c(1) : \mathcal{S}\cap\llbracket\Diamond_2 p\rrbracket^c \neq \emptyset\}\\ &= \{\mathcal{S} \text{ based on } g_c(1) : \mathcal{S}\cap\{\mathcal{T} \text{ based on } g_c(2) : \mathcal{T}\cap\llbracket p\rrbracket^c \neq \emptyset\} \neq \emptyset\}\end{aligned}$$

As intended, $\Diamond\Diamond p$ expresses the property a set of sets of worlds has when it contains a set of worlds that is compatible with p.

$$\begin{aligned}\llbracket\Diamond_1\Diamond_2\Diamond_3 p\rrbracket^c &= \{\mathcal{S} : \mathcal{S}\cap\llbracket\Diamond_2\Diamond_3 p\rrbracket^c \neq \emptyset\}\\ &= \{\mathcal{S} : \mathcal{S}\cap\{\mathcal{U} : \mathcal{U}\cap\{\mathcal{T} : \mathcal{T}\cap\llbracket p\rrbracket^c \neq \emptyset\} \neq \emptyset\}\}\end{aligned}$$

As intended, $\Diamond\Diamond\Diamond p$ expresses the property a set of sets of worldly propositions has when it contains a set of worldly propositions that is compatible with $\Diamond p$.

4.4 Compositionality and Polymorphic Types

The interesting operators of our language ($\rightarrow$, $\triangle$, $\Diamond$) uniformly take set-type meanings (representations) as arguments. This gives our system the veneer of compositionality, but, for now, only the veneer. Set-type meanings are, strictly speaking, *not* typically the semantic values of these operators' complements; the semantic values of the sentences of our language, in fact, comprise a manifold of *functional* types.[17] Here is an illustration: $\Diamond$ can semantically combine with a worldly representation $[\![p]\!] :: \langle s, t\rangle$,[18] an epistemically modal representation $[\![\Diamond p]\!] :: \langle\langle s, t\rangle, t\rangle$, an epistemically modal representation with epistemically modal content $[\![\Diamond\Diamond p]\!] :: \langle\langle\langle s, t\rangle, t\rangle, t\rangle$, and so on, *ad infinitum*.

Epistemic operators, therefore, have the (perhaps surprising) property of being relatively *unselective* as to the semantic type of their complements—more precisely, of being *type-polymorphic*.[19] The polymorphic type of epistemic operators can be given a recursive characterization:

$$\tau^* ::= \langle \alpha, \langle \alpha, t\rangle\rangle \qquad \alpha ::= \langle s, t\rangle \mid \langle \gamma, t\rangle \mid \langle \alpha, t\rangle$$
$$\gamma ::= \langle \alpha, v_{[0,1]}\rangle$$

One way to think about polymorphic types is this. An expression like $\Diamond$ has a semantic type, in two guises: qua *expression-type* (in which case its type is polymorphic) and qua *expression-token* (in which case its type, as tokened on an occasion of use, is a type drawn from the polymorphic type hierarchy). The semantic type of, e.g., $\Diamond$, as tokened on an occasion of use will "depend" (very loosely speaking[20]) on the semantic type of its complement (but will always be drawn from the hierarchy of types introduced here).[21]

---

[17]I here assume that semantic composition is always via Function-Argument Application.

[18]Notation: s is the type of worlds, t is the type of truth values. A function of type $\langle\tau, \tau'\rangle$ is a function from objects of type τ into objects of type τ'.

[19]On type polymorphism in theoretical computer science, see especially Pierce (2002). For a recent application of polymorphic types to natural language semantics, see Charlow (forthcomingc).

[20]This is no violation of compositionality: the semantic type of $\Diamond$, as tokened on an occasion of use, is not semantically *determined by*, or selected *in virtue of*, the semantic type of its complement. It is simply to say that, if $\Diamond$ occurs in a semantically well-formed expression, its semantic type must be drawn from the hierarchy of types defined above, and must be of the right type to compose, by Function-Argument Application, with the semantic value of its sister.

[21]To be compositional, our system requires a understanding of *semantic coordination*. We currently understand $\wedge$ as expressing $\cap$, but there are two reasons this will not work. First, the semantic values of $\wedge$'s arguments are functions, not sets. (This is trivial to fix, and I will continue to talk as if the difference between a characteristic function and a set is no difference at all.) Second, the semantic values of $\wedge$'s arguments are frequently sentences of different semantic type. This is less trivial to fix: we will require a generalized understanding of conjunction that allows it to coordinate constituents of different semantic type, as in Partee & Rooth (1983). To keep the main discussion maximally simple, I will ignore this sort of complication here (though I will address it in the Appendices).

5 Two Aspects of Mental Life

The last section showed that the notion of credences *in* epistemically modal and probabilistic representations, constrained by the probability axioms, is both mathematically and semantically tractable. But—and I intend this question seriously—does it *make sense*? We have introduced a semantic hierarchy of representations with no upper bound on the complexity of the semantic type of a representation. Is this cognitively realistic? (Here, I will argue: yes.) We have assumed that objects at any level of the type hierarchy can receive credences (where credences are constrained by assumptions of Normalization and Additivity). Is this normatively plausible—do standard justifications for Normalization and Additivity apply, if credences are not assumed to be defined over worldly representations? (Here, I will argue: yes and no.)

5.1 Cognitive

Moss offers an argument against a proposed extension of the Bayesian proposal pursued in Rothschild (2012); Yalcin (2012) to graded modal judgments:

> [I]t is hard to imagine a reason for ruling that embeddings of epistemic vocabulary beyond a certain level of complexity are semantically uninterpretable. In the absence of such a reason, our theory should deliver semantic values for embeddings of arbitrary complexity. Hence in order to repair the [Bayesian] proposal, we would have to model subjects as having not just sets of sets of measures as mental states, but sets of sets of sets of measures, and so on. It is difficult to independently motivate such an arcane model of our mental life. (Moss 2015: 30)

While our proposal isn't quite Bayesian in the sense of Rothschild (2012); Yalcin (2012), Moss' critique clearly applies. The charge that this is an "arcane model" does not, however, really bite. To illustrate, recall that, on the view defended here:

$$[\![\Diamond_1\Diamond_2 p]\!]^c = \{\mathcal{S} : \mathcal{S} \cap \{\mathcal{T} : \mathcal{T} \cap [\![p]\!]^c \neq \emptyset\} \neq \emptyset\}$$

On our view, $\Diamond\Diamond p$ expresses the property a set of sets of worlds (i.e., a set of worldly propositions) has when it contains a set of worlds that is compatible with p. To think or call such a sentence *probable* is to express a property of credences in sets of sets of sets of worlds (i.e., sets of sets of worldly propositions)—namely, the property of assigning a credence $> .5$ to $[\![\Diamond_1\Diamond_2 p]\!]^c$.

An agent can treat *any set of objects* as a set of alternatives **for cognitive purpose** P. An agent can represent sets of possible worlds for the purpose of representing different abstract alternatives (individual possibilia) for accurately representing the world. An agent can represent sets of sets of possible worlds (i.e., sets of propositions) in order to represent different alternatives, *not* for the purpose of accurately representing the world, instead for the purpose of representing alternative ways of representing the world (e.g., alternatives that treat p as possible *versus* those that treat p as impossible). An agent can represent sets of sets of sets of possible worlds (i.e., sets of sets of propositions) in order to represent different alternatives (sets of propositions)—*not* for the purpose of accurately representing the world, *nor* for the purpose of representing alternative ways of representing the world, instead for the purpose of representing alternative ways of representing alternative ways of representing the world.

Representations, as we understand them, have an iterative, or recursive, structure (but is that surprising?). But the cognitive state of *representing* $\mathcal{R}$ *for purpose* P is not arcane: it is the attitude of representing the various alternatives of $\mathcal{R}$ as candidates for fulfilling P. We have understood the attitude of representation more expansively than is traditional[22]—in particular, we have relativized representations to cognitive purposes, and have declined to assume that the functional role of representation is uniformly about representing individual possibilia as candidates for actuality. Generalizing a familiar notion need not, however, render it arcane. Indeed, given this generalized understanding of representation, representing a set of alternatives as candidates for fulfilling P describes a sort of familiar cognitive activity in which agents plausibly can and do engage.

5.2 Normative

Why represent agents as having credences in non-worldly representations? We cited two (related) motivations (§3.2). A theoretical motivation: to vindicate ascriptions of confidence *in* epistemically modal representations within our theoretical metalanguage. And a normative motivation: to describe and justify rational norms governing confidence in epistemically modal representations; and to describe and justify rational norms governing the relationship between confidence in epistemically modal representations, and action. The present account satisfies the theoretical motivation. What about the normative? This section proposes that (i) the rational norms governing the relationship between credence and action are a generalization of the theory of Expected Value; (ii) generalizations of the decision-theoretic notions of a decision problem and of the Expected Value of an action relative to a decision problem can be stated in a basically standard form.

5.2.1 Generalizing Expected Value

We begin by defining the notion of a generalized decision problem.

Definition 7. *A* ***decision problem*** Π *based on* $\langle \mathcal{R}, \mathfrak{R}, \Sigma \rangle$ *is a tuple* $\langle \mathfrak{R}, \mathcal{A}, Cr, Val \rangle$ *where:*

- $\mathfrak{R} = \{C_1, ..., C_n\}$ *is a partition of the possibilities relevant in* Π.
- $\mathcal{A} = \{A_1, ..., A_n\}$ *is a set of actions available in* Π.
- Cr *is a credence function based on* $\langle \mathcal{R}, \mathfrak{R}, \Sigma \rangle$.
- Val *is a conditional value function, such that* $Val(A|C)$ *is a value representing the degree to which* A *is desired conditional on* C *(for each* $C \in \mathfrak{R}$*).*

Decision problems can be presented in a standard tabular format, as follows. As is standard, cells of the table correspond to "payoffs", here understood as degrees of

[22]The traditional view I am attempting to generalize here is, of course, that of Stalnaker (1984).

desirability conditional on the corresponding representation.[23]

Π	$\overbrace{C_1}^{Cr(C_1)}$	...	$\overbrace{C_n}^{Cr(C_n)}$
A_1	$Val(A_1 \mid C_1)$	...	$Val(A_1 \mid C_n)$
...	...	...	...
A_m	$Val(A_m \mid C_1)$	...	$Val(A_m \mid C_n)$

Unlike in standard presentations of decision tables, we do not assume that the contingencies relevant in a decision problem Π based on $\langle \mathcal{R}, \mathfrak{R}, \Sigma \rangle$ form a partition of W (or of a subset of W). $\mathfrak{R}$, rather, partitions a salient representation—picturesquely, the "base" representation against which an agent's deliberation occurs. This base representation is not, however, required to be of any specific semantic type. Given the notion of a generalized decision problem, a corresponding generalization of Expected Value is immediate.

Definition 8. *If $\Pi = \langle \mathfrak{R}, \mathcal{A}, Cr, Val \rangle$ is based on $\langle \mathcal{R}, \mathfrak{R}, \Sigma \rangle$ and $A \in \mathcal{A}$, the* ***expected value*** *of A in Π is a weighted sum of credences multiplied by values:*

$$\sum_{x \in \mathfrak{R}} Cr(x|A) Val(A|x)$$

5.2.2 Justifying Credences

Why should an agent in a generalized decision problem maximize generalized expected value? More specifically (and to bracket certain controversies about formulating a mathematical theory of rational action): why should an agent who wants to maximize expected value in a generalized decision problem compute expected value *using a credence function* (the properties of which are constrained by Definition 4)?

There are two main ways of answering this type of question in the literature. First, Dutch Book Arguments, on which, roughly, agents who have incoherent credences are irrational because subject to sure losses (for an overview, see Hájek 2009). Second, Accuracy Arguments, on which, roughly, agents who want to maximize expected *epistemic* value (roughly, the proximity of one's credences to the truth), but who have incoherent credences, are irrational because coherent credences are always more proximal to the truth (originating with Joyce 1998). Let us see about the prospects of extending these answers to the present account.

[23] Staffel (forthcoming) remarks that, in an Expressivistic system (alike in some, but not all, respects to the one proposed here), "wins and losses can't be determined by checking *what the world is actually like*" (if the relevant contingencies are not worldly propositions that can be "checked" for truth against the actual world). But if, as seems correct, the conditional value $Val(A|C)$ is like the conditional probability $Cr(A|C)$—in that both track degrees of desire or belief, under the *indicative supposition* that C—there is no immediate need for worldly matters to "determine" wins and losses in decision problems based on representations of arbitrary type. The degree to which an agent who indicatively supposes C desires to perform A will determine $Val(A|C)$—nothing worldly required, so long as the degree to which an agent desires to perform A can depend on a non-worldly representation. (Such dependence appears to be commonplace; recall §3.2.) There may yet be a need for worldly matters to determine wins and losses, for a theorist who wants to use the notion of conditional desirability *to run a Dutch Book argument*. More on this just below.

Matters are, not surprisingly, less than straightforward with Accuracy Arguments. *Accuracy* is fundamentally a worldly notion: a representation is said to be accurate when it is satisfied ("true") as evaluated against a possibility taken to represent *actuality* (as also noted by Staffel forthcoming). Accuracy Arguments purport to show that subjective estimates of objective chance that violate the axioms of probability are rationally defective, since, for any such estimate, there is another way of estimating chances that (i) satisfies the axioms of probability and (ii) is guaranteed to be overall more accurate in w, for any possible world w to which the agent assigns some credence (see esp. the accuracy theorem of Joyce 1998). In order to adapt Accuracy Arguments to the framework proposed here, we would require a *non-worldly proxy for the notion of Accuracy* (as well as a non-worldly proxy for the notion of actuality). The prospects here strike me as very dim—particularly given the conceptualization of our theory suggested in §4.

Dutch Books, however, do appear to generalize to this application. The constraints on generalized credence functions we have introduced, therefore, are ultimately motivated by "pragmatic" considerations (although in the case of probability measures over worldly propositions, they may still be motivated by considerations of accuracy).

Nothing in the bare mathematics of the "Dutch Book Theorem" (see, e.g., Hájek 2009) requires that decision-theoretic contingencies ("states") are worldly propositions. To illustrate, here is a Dutch Book for an "overconfident" agent who commits herself to a sentence of the form $\triangle\triangle\phi \wedge \triangle\neg\triangle\phi$ (and thereby commits herself to regarding both $\triangle\phi$ and $\neg\triangle\phi$ as probable). (I assume, just for the sake of illustration, that the operator $\triangle$ expresses a credence $> .6$ in its complement representation.)

	$\triangle\phi$	$\neg\triangle\phi$
Bet 1	-.6	.4
Bet 2	.4	-.6

For concreteness, suppose that ϕ is the proposition that there is an algae bloom in the lake. This table represents Bet 1 (e.g., drinking lake water) as undesirable to degree .6 conditional on $\triangle\phi$, and desirable to degree .4 conditional on $\neg\triangle\phi$; it represents Bet 2 (e.g., purchasing bottled water) as desirable to degree .4 conditional on on $\triangle\phi$, and undesirable to degree .6 conditional on $\neg\triangle\phi$. The agent of this decision problem regards as fair a series of "bets" that, taken together, logically guarantee a "loss" (from the vantage of her own conditional degrees of desirability).[24]

[24] If, for example, the agent is by assumption an REU maximizer, then the conditional degrees of desirability recorded in this decision table (when defined; recall fn9) may be taken to represent REUs computed conditional on the relevant representation. As noted above, once the notion of a decision problem is generalized, words like "bet" and "loss" lose their normal connotations: typically an agent can't bet (in the sense of making a cash wager) that pays out if ϕ is probable. (How, after all, would a winning bet be *determined*—particularly given the broader nonfactualist setting of this paper?) In the present context, when a package of "fair" "bets" is said to guarantee a "sure loss" for an agent, this means that (i) there is a set of actions $\mathcal{A}$ such that for each $A \in \mathcal{A}$, the Expected Value of doing A for the agent is at least as great as the Expected Value of not doing A for the agent (roughly, there is a set of actions all of which the agent regards as "fair" or "worth doing"); (ii) the Expected Value of the complex action *doing everything in* $\mathcal{A}$ is less than the Expected Value of not doing everything in $\mathcal{A}$ (roughly, the agent does not regard the complex action *doing everything in* $\mathcal{A}$ as "fair", even though she regards all of the actions in $\mathcal{A}$ as "fair").

This is, I claim, a rational defect. In general, an agent imposes a partition on a base representation $\mathcal{R}$, thereby generating an alternative set $\mathfrak{R}$ for $\mathcal{R}$, for the sake of representing alternatives whose adoption is relevant for a cognitive purpose P. Whichever alternative in $\mathfrak{R}$ such an agent should accept—however such an agent fulfills P—that agent will be subject to a loss (from the vantage of their own conditional degrees of desirability) in a Dutch Book. Whether or not, that is to say, the agent concludes that there is probably an algae bloom, her incoherent credences make her such that she regards both drinking the lake water and purchasing bottled water as good bets in the Dutch Book presented above. Roughly speaking, in such a Dutch Book, such an agent regards it as okay to spend money on bottled water to avoid drinking lake water that she regards as okay to drink. That is irrational.

5.2.3 Nailing Down Dutch Books

Staffel (forthcoming) develops both Accuracy-style and Dutch Book-style arguments for coherent credences in non-worldly representations (while also registering doubts that such arguments actually meet the theoretical needs that prompt them). In Staffel's Expressivistic Dutch Book—which is in certain respects similar to the one advanced here—an "underconfident" agent (e.g., one who assigns both $\triangle p$ and its negation $\neg\triangle p$ credence .4)...

> can avoid a sure loss by not becoming opinionated. The fact that the underconfident agent *would lose money if she became opinionated* does not point to any obvious rational defect. There are many things I might do that would put me at a great disadvantage in particular circumstances. But *if I have no reason to think I'll find myself in those circumstances, then I have little or no reason to avoid those actions*. (PAGE)

This difficulty certainly does threaten Staffel's Expressivistic Dutch Book (see esp. Staffel forthcoming: PAGE). It might also seem to threaten the version I have pursued here. The "irrationality" that, I claimed, characterizes an incoherent agent is as follows: relative to an alternative set $\mathfrak{R}$ that represents various candidate representations for fulfilling purpose P, the agent would be subject to a loss *if* she selected an alternative from $\mathfrak{R}$ to fulfill P. However, if she is not in a position to select an alternative from $\mathfrak{R}$ to fulfill P—if she is unable to settle on any particular way of resolving the relevant question (e.g., whether there is probably an algae bloom)—the negative conditional desirability (of, e.g., Bets 1 and 2 conditional on $\triangle p$) is never "actualized". The "loss" to which the agent is subject in a Dutch Book is of a *merely hypothetical* character: the agent will be worse off *if* she becomes opinionated, but if she doesn't, she won't. What is irrational about that?

In reply: we said that, in the above Dutch Book:

- An agent entertains a set $\mathcal{R}$ of type $\langle\gamma, t\rangle$ (a set of credence functions).
- She partitions $\mathcal{R}$ into: (i) a cell of credence functions according to which it is probable that there is an algae bloom; (ii) a cell of credence functions according to which it is not probable that there is algae bloom.

The purpose the agent tries to achieve in so-partitioning $\mathcal{R}$ is, we said, to represent alternative ways of representing the world (e.g., alternatives that treat an algae bloom as probable versus alternatives that do not). Conditional on *either* way of representing—i.e., conditional on representing an algae bloom as probable *and* conditional on representing it as not probable—the agent is subject to a loss (from the vantage of her own conditional degrees of desirability) in a Dutch Book.

The irrationality here is, I submit, manifest: the agent is trying to achieve goal g (e.g., figuring out how to represent the likelihood of there being an algae bloom—as probable or not probable), but her credences are such that any way of achieving g presents her with a deficit in desirability in a Dutch Book. That is to say, her credences are structurally such that, conditional on *any way of achieving what she is trying to achieve*, she is subject to a deficit in desirability (from the vantage of her own conditional degrees of desirability) in a Dutch Book. Claim: if your credences in context c are structurally such that they prevent you from doing what you're trying, in c, to do without being subject to a sure loss in a Dutch Book, your credences in c are irrational in c.

6 Conclusion

This paper began by observing that standard models of the semantics of epistemic modals render the phenomenon of graded modal judgment, whether in thought and language, unintelligible. In response, this paper developed a model of graded modal judgment, in both thought and language—one that represented graded modal judgment as a generalization of our cognitive capacity for reasoning with hypotheses about objective chance (i.e., our cognitive capacity for probabilistic reasoning). The generalization was developed as a package of interrelated semantic, cognitive, and epistemological theses:

- **Semantic:** modals compose with representations of arbitrary type. (§4.2)
- **Cognitive:** agents entertain representations of arbitrary type for specific cognitive purposes; the state of *bearing a graded attitude toward a representation of arbitrary type* is a natural cognitive kind (instances of which are, broadly, governed by the purpose for which the agent is entertaining the relevant representation). (§5.1)
- **Epistemological:** part of the functional role of credences in representations of arbitrary type (entertained for cognitive purpose P) is to determine fair "prices" for bets against ways of representing that fulfill P. Agents whose credences violate Normalization or Additivity are thus subject to Dutch Books. (§5.2)

On the model of graded modal judgment developed here, modal sentences are semantically evaluated against complex constructions out of possibilia. But various sentences of our language are not semantically evaluated relative to *individual possibilia*. And so our model exhibits the characteristic insensitivity of logics axiomatized by S5 to a choice of possible world taken to represent "indicative actuality" (as in Kaplan 1989), or to a choice of possible world taken to represent a non-actual circumstance of evaluation. I take this to be one of the main virtues of the present theory: it can accommodate many of the intuitions that motivate axiomatizing the logic of epistemic modality with S5, without rendering the notion of graded modal judgment, whether in thought or in language, unintelligible (for more discussion, see Appendix B.1).

A polemical note to conclude. Notice that "non-factual" theories—theories that do not take modalities of the relevant type to be proposition-forming operators, a description satisfied by both our theory and Moss'—offer the theorist at least two broadly workable models of the cognition, semantics, and epistemology of graded modal judgment. "Factual" accounts of these modalities, so long as they are constrained by S5—and, indeed, even a weaker logic like KD45—are able to offer none of these attractions (see Appendix B.1). It is probably time to move past the philosophical preoccupation with the ability of non-factual theories of operators in natural language to account for environments *embedding* these operators. If anyone has such problems, it seems to be the theorists who have pushed such objections, rather than their targets.

A Indicatives

A.1 Scope-Taking and Type-Raising with Indicatives

As intended, $\triangle(p \to q)$ expresses the property a credence function has when it assigns the non-worldly representation encoded in $p \to q$ a value $> .5$.

$$\begin{aligned} [\![\triangle_1(p \to_2 q)]\!]^c &= \{Cr' : \sum_{Cr \in [\![p \to_2 q]\!]^c} Cr'(\{Cr\}) > .5\} \\ &= \{Cr' : \sum_{Cr([\![q]\!]^c | [\![p]\!]^c) = 1} Cr'(\{Cr\}) > .5\} \end{aligned}$$

Handling the narrow-scope representation $p \to \triangle q$ is trickier. A first attempt:

$$\begin{aligned} [\![p \to_1 \triangle_2 q]\!]^c &= \{Cr : Cr([\![\triangle_2 q]\!]^c | [\![p]\!]^c) = 1\} \\ &= \{Cr : \frac{Cr([\![\triangle_2 q]\!]^c \cap [\![p]\!]^c)}{Cr([\![p]\!]^c)} = 1\} \end{aligned}$$

But this attempt fails, since $Cr([\![\triangle_2 q]\!]^c \cap [\![p]\!]^c)$ is undefined in the present system, as $[\![\triangle_2 q]\!]^c$ and $[\![p]\!]^c$ are of different semantic types. Like Moss (2015: §2.4), and ultimately following Partee & Rooth (1983), we can address this by raising the type of $[\![p]\!]^c$:[25]

$$\begin{aligned} {}_{\texttt{raise}}X_{\langle\tau,t\rangle} &= \lambda Y_{\langle\tau,t\rangle} \,.\, Y \subseteq X && \textbf{(Raise)} \\ {}_{\texttt{p}}X_{\langle\tau,t\rangle} &= \lambda\gamma \,.\, \gamma(\bigcup X) = 1 && \textbf{(Probabilify)} \end{aligned}$$

If $[\![\phi]\!]^c :: \langle\tau, t\rangle$, then ${}_{\texttt{raise}}[\![\phi]\!]^c :: \langle\langle\tau, t\rangle, t\rangle$. That is to say, raising the type of a worldly representation $[\![p]\!]^c$ generates a set of worldly representations (equivalently, again, a characteristic function of worldly representations). In particular, it generates the set of worldly representations that involve representing $[\![p]\!]^c$ as true. Therefore, if $[\![\phi]\!]^c :: \langle\langle\tau, t\rangle, t\rangle$, then ${}_{\texttt{p}}[\![\phi]\!]^c :: \langle\gamma, t\rangle$.[26] Probabilifying a raised worldly representation (`praise`-ing) p yields the set of credence functions that assign probability 1 to some way of representing that p. Type-raising in hand, we have:

$$\begin{aligned} [\![p \rightarrow_1 \triangle_2 q]\!]^c &= \{Cr : Cr([\![\triangle_2 q]\!]^c|_{\texttt{praise}}[\![p]\!]^c) = 1\} \\ &= \{Cr : \frac{Cr([\![\triangle_2 q]\!]^c \cap {}_{\texttt{praise}}[\![p]\!]^c)}{Cr({}_{\texttt{praise}}[\![p]\!]^c)} = 1\} \\ &= \{Cr : \frac{Cr(\{Cr' : \sum_{\mathcal{S}\in[\![q]\!]^c} Cr'(\{\mathcal{S}\}) > .5\} \cap \{Cr' : Cr'(\bigcup {}_{\texttt{raise}}[\![p]\!]^c) = 1\})}{Cr(\{Cr' : Cr'(\bigcup {}_{\texttt{raise}}[\![p]\!]^c) = 1\})} = 1\} \end{aligned}$$

As intended, $p \rightarrow \triangle q$ expresses the property a credence function has when the ratio of the credence it assigns $[\![\triangle_2 q]\!]^c \cap {}_{\texttt{praise}}[\![p]\!]^c$ to the credence it assigns ${}_{\texttt{praise}}[\![p]\!]^c$ is 1.

---

[25] Our Probabilify rule is a close relative of the Type-Shifting rule introduced at Moss (2015: 34) (see also Moss 2018: 234ff). Note, though, that the role that type-shifting plays in Moss' system is quite different from the role it plays in ours. Moss' account, unlike ours, requires that $[\![p]\!]^c$ (understood as a set of worlds) be type-shifted into a probabilistic ("ur," for short) content—the set of probability spaces normalized to $[\![p]\!]^c$—in order to compose with an epistemic operator (and, more generally, to be the sort of object that can play the content-role in Moss' larger theory of probabilistic assertion and belief). This raises the following question: why does $\neg p$ appear to lack a reading on which it means that p is not certain (i.e., means the negation of p's ur-content)? Moss addresses this question for a sentence of the form $\neg p$ with the stipulation that "semantic types of sentences are shifted from sets of worlds to sets of probability spaces if and only if such type shifting is forced" (2018: 238). But how will this work for the following sort of case?

(10) A: It's raining. B: That's not true.

Moss writes that "If a sentence contains no epistemic vocabulary at all, then *assertion itself forces* the semantic type of that sentence to be shifted from a proposition to a set of probability spaces," i.e., an ur-content (2018: 234). Presumably, B's denial can be represented as targeting the content of A's assertion with negation. But that represents B's denial as possibly weaker than it, in fact, could be. Since our account utilizes type-shifting *exclusively* for inducing semantic coordination between syntactically coordinated sentences, this kind of worry does not arise for it.

[26] γ is the type of credence functions (§4.4). So a function of type $\langle\gamma, t\rangle$ is of type $\langle\langle\langle\tau, t\rangle, \nu_{[0,1]}\rangle, t\rangle$. It is natural to assume that whatever credence, if any, someone assigns $[\![p]\!]^c$ will *determine* (or perhaps *rationally constrain*—I do not have a good sense of what is at stake here) their credence in ${}_{\texttt{raise}}[\![p]\!]^c$: if you think of p as i-likely, you are/ought to be i-certain of ${}_{\texttt{raise}}[\![p]\!]^c$ (though often there is no $q \in {}_{\texttt{raise}}[\![p]\!]^c$ such that you are/ought to be i-certain of q).

A.2 Domain Restriction and Triviality

It bears noting that there is another, probably preferable, possibility for representing the scopal interactions of indicatives and modal operators, on which the latter are analyzed as binary (i.e., *restrictable*[27]) operators (Kratzer 1981, 1986):

$$[\![\triangle_1(\phi)(\psi)]\!]^c = \lambda Cr : Cr \text{ is based on } g_c(1) \,.\, Cr([\![\psi]\!]^c | [\![\phi]\!]^c) > .5$$
$$[\![\Diamond_1(\phi)(\psi)]\!]^c = \lambda \mathcal{S} : \mathcal{S} \text{ is based on } g_c(1) \,.\, \mathcal{S} \cap [\![\phi]\!]^c \cap [\![\psi]\!]^c \neq \emptyset$$
$$[\![\Box_1(\phi)(\psi)]\!]^c = \lambda \mathcal{S} : \mathcal{S} \text{ is based on } g_c(1) \,.\, \mathcal{S} \cap [\![\phi]\!]^c \subseteq [\![\psi]\!]^c$$

Kratzer denies that the indicative conditional contributes its own quantificational force; rather, indicative conditionals are syntactic devices for making explicit the restriction argument of a restrictable quantifier.[28] There is no semantic distinction between the "wide scope" $\triangle(p \to q)$ and the "narrow scope" $p \to \triangle q$: both are represented using the restricted modal $\triangle(p)(q)$.

One motivation for adopting Kratzer's analysis of indicative conditionals is explaining the sorts of judgments of equivalence that Stalnaker's Thesis (Stalnaker 1970) attempts to unify—e.g., the judgment that (11) and (12) are equivalent. According to Stalnaker's Thesis, the probability that an indicative $A \to C$ is true equals the conditional probability of C on A. Supposing that probability operators in natural language semantically express degrees of conditional probability, Stalnaker's Thesis predicts, correctly, that (11) and (12) are equivalent.

(11) Rain is likely, conditional on atmospheric pressure being low.

(12) Rain is likely, if atmospheric pressure is low.

In line with Stalnaker's Thesis, the Kratzerian story about probability operators under consideration here renders (11) and (12) equivalent—more precisely, is able to generate equivalent logical forms for these sentences. More generally, and regardless of whether Stalnaker's Thesis holds in its full generality, no version of the Thesis—even massively restricted—can be accommodated without taking probability operators (and, by extension, modal operators) to be *binary* operators. That is because a language with only unary probability operators provably lacks the resources to express a sufficiently wide range of conditional probabilities.[29]

It is difficult to overstate the importance of this idea: it allows the theorist to accommodate (a perhaps appropriately restricted version of) Stalnaker's Thesis, *without* signing

[27] Restrictable quantifiers are Generalized Quantifiers, in the sense of Barwise & Cooper (1981).

[28] Except when no quantifier is provided, in which case a silent restrictable quantifier—which Kratzer (1986), e.g., took to be an epistemic necessity modal—is posited in logical form.

[29] See Égré & Cozic (2011)'s adaption of the theorem of Hájek (1989) to an inexpressibility result for a language with unary probability operators. Sketch of the proof: consider a fair three-ticket lottery, with tickets numbered ONE, TWO, and THREE. According to any probability model for this situation, the conditional likelihood of ONE winning if THREE doesn't is 1/2; but there are probability models for this situation in which no proposition is such that it has probability 1/2.

onto the Thesis in the form in which it is usually presented:

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

That is because the Kratzerian analysis does *not* represent $\triangle(A \to C)$ as expressing a probability judgment whose object is a conditional proposition; rather, it expresses a *restricted probability judgment*: that C is likely (as assessed against the representation expressed by A).

> [I]n saying 'there is one chance in two that if A then C', the conditional 'if A then C' does not express any self-standing proposition... However, this remains compatible with the idea that if-clauses are devices of quantifier restriction. In the scope of an operator, if-clauses do make a systematic truth-conditional contribution to the whole sentence. (Égré & Cozic 2011: 22)

This would seem to be just what is required to avoid challenges to Stalnaker's Thesis on grounds of Triviality results in the mold of Lewis (1976) (see Rothschild 2015; Charlow 2016b).

The dialectic in this neighborhood of issues is, however, a great deal more vexed than this quick summary would suggest. Charlow (2016b) shows that Triviality results in the mold of Lewis (1976) arise for restricted operators (and that such results do not depend on the (mis)understanding of logical form embodied in Stalnaker's Thesis). Indeed, as Charlow (2016b) argues, obstacles of Triviality arise for any treatment of restricted quantification that takes $[\![A]\!]$, $[\![C]\!]$, and $[\![\triangle(A)(C)]\!]$ to be *elements of the same semantic algebra*—i.e., any treatment that takes $[\![A]\!]$, $[\![C]\!]$, and $[\![\triangle(A)(C)]\!]$ to be of the same base semantic type. *This is precisely the assumption that the analysis in this paper discards*. I take this to be another argument in favor of this paper's analysis: unlike any competitor account of which I am aware, it allows the theorist to accommodate the intuitions of equivalence that underlie Stalnaker's Thesis, while also avoiding the specter of Triviality.

B Epistemic Contradiction

B.1 (In)validating S5

If we introduce epistemic modal operators that do not raise the semantic type of their complements, we will observe that, for such operators, the axioms of S5 are validated.

$$[\![\blacklozenge_1\phi]\!]^c = \{\mathcal{S} \in \mathcal{R} : \mathcal{R} \cap [\![\phi]\!]^c \neq \emptyset\} \qquad (g_c(1) = \langle \mathcal{R}, \mathfrak{R}, \Sigma \rangle)$$

Either $[\![\blacklozenge_1\phi]\!]^c = \mathcal{R}$ (if $\mathcal{R} \cap [\![\phi]\!]^c \neq \emptyset$), or else $[\![\blacklozenge_1\phi]\!]^c = \emptyset$ (if $\mathcal{R} \cap [\![\phi]\!]^c = \emptyset$). Given a stipulation that $[\![\blacksquare_1\phi]\!]^c \subseteq [\![\phi]\!]^c$), the logic of $\blacklozenge$ and its dual $\blacksquare$ is the logic of S5:

$$[\![\blacksquare_1\phi]\!]^c = [\![\blacksquare_1\blacksquare_1\phi]\!]^c \qquad [\![\blacklozenge_1\phi]\!]^c = [\![\blacksquare_1\blacklozenge_1\phi]\!]^c$$

But no such consequences[30] hold for the type-raising modal operators of our language:

$$[\![\Box_1\phi]\!]^c \not\subseteq [\![\Box_2\Box_1\phi]\!]^c \qquad\qquad [\![\Diamond_1\phi]\!]^c \not\subseteq [\![\Box_2\Diamond_1\phi]\!]^c$$

To illustrate, let $g_c(1) = \langle \mathcal{R}_1, \mathfrak{R}_1, \Sigma_1\rangle$, $g_c(2) = \langle \mathcal{R}_2, \mathfrak{R}_2, \Sigma_2\rangle$. Then:

$$\begin{aligned} [\![\Diamond_1 p]\!]^c &= \lambda\mathcal{S}_{\langle s,t\rangle} : \mathcal{S} \text{ is based on } g_c(1)\ .\ \mathcal{S} \cap [\![p]\!]^c \neq \emptyset \\ [\![\Box_2\Diamond_1 p]\!]^c &= \lambda\mathcal{S}_{\langle\langle s,t\rangle,t\rangle} : \mathcal{S} \text{ is based on } g_c(2)\ .\ \mathcal{S} \subseteq [\![\Diamond_1 p]\!]^c \\ &= \lambda\mathcal{S}_{\langle\langle s,t\rangle,t\rangle} : \mathcal{S} \text{ is based on } g_c(2)\ .\ \forall\mathcal{S}' \in \mathcal{S} : \mathcal{S}' \cap [\![p]\!]^c \neq \emptyset \end{aligned}$$

$[\![\Diamond_1 p]\!]^c$ is the property a sets of worlds has, when it contains a p-world; $[\![\Box_2\Diamond_1 p]\!]^c$ is the property a set of sets of worlds has when each element in this set is compatible with p.

But if—as our account has it—$\Diamond_1 p$ expresses a property utterly distinct from—indeed, at a different level of the type hierarchy than—the property expressed by $\Box_2\Diamond_1 p$, why does a sentence like (13) of the form $\Diamond_1 p \wedge \Diamond_2\neg\Diamond_1 p$ (and equivalent to $\Diamond_1 p \wedge \neg\Box_2\Diamond_1 p$) sound so terrible?

(13) #It may be raining, but maybe it can't be.

Perhaps because it is Moore-Paradoxical (cf. Weatherson 2004)? Alas, it is not; note that Moore Paradoxicality dissolves in unasserted environments (see Yalcin 2007):

(14) #Suppose it may be raining, but maybe it can't be.

(15) Suppose it is raining, but you don't know it's raining.

Explanation: our account predicts that sentences of the form $\Diamond_1 p \wedge \Diamond_2\neg\Diamond_1 p$ are semantically defective. (Moss' semantics generates the same prediction, in basically the same fashion.) Notice that semantically coordinating $[\![\Diamond_1 p]\!]^c$ and $[\![\neg\Box_2\Diamond_1 p]\!]^c$ requires intersecting $[\![\Diamond_2\neg\Diamond_1 p]\!]^c$ with:

$$_{\texttt{raise}}[\![\Diamond_1 p]\!]^c = \lambda X_{\langle\langle s,t\rangle,t\rangle}\ .\ X \subseteq [\![\Diamond_1 p]\!]^c$$

$_{\texttt{raise}}[\![\Diamond_1 p]\!]^c$ denotes the property a set of propositions $\mathfrak{F}$ has iff each q in that set is compatible with p. As noted above, $[\![\Diamond_2\neg\Diamond_1 p]\!]^c$ denotes the property a set of propositions $\mathfrak{F}$ has iff some q in that set is incompatible with p. Obviously, no $\mathfrak{F}$ satisfies both

[30] I am here thinking of consequence standardly, i.e., in terms of set-theoretic inclusion. Notice that consequence, in this sense, can hold only between sentences of the same semantic type—e.g., $[\![\Box_1 p]\!]^c \not\subseteq [\![p]\!]^c$. This is not difficult, however, to repair—should we decide that it is important to designate the relationship between, say, $\Box_1\phi$ and ϕ as one of logical consequence. (I think this less important than predicting (as this section does) that sentences of the form $\Box_1\phi \wedge \neg\phi$ express defective semantic contents (i.e., $\emptyset$), but your mileage may vary.) Notice that $[\![\Box_1 p]\!]^c$ is a set of p-entailing subsets of W; since each possibility consistent with $[\![\Box_1 p]\!]^c$ is a p-entailing possibility, we say that p is a "birdseye consequence" of $\Box_1 p$. More generally ψ is a birdseye consequence of ϕ in c if $\forall x \in [\![\phi]\!]^c : x \subseteq [\![\psi]\!]^c$. In general, ψ is a birdseye consequence of ϕ_0 iff there exists a sequence $\phi_1, ..., \phi_n$ such that for each $1 \leq i \leq n$, ϕ_i is a birdseye consequence of ϕ_{i-1}, and ψ is a birdseye consequence of ϕ_n. Thus, e.g., ϕ is a birdseye consequence of $\Box_n ... \Box_1\phi$, as desired.

properties:

(16) $[\![\Diamond_1\phi \wedge \Diamond_2\neg\Diamond_1\phi]\!]^c = \emptyset$

And so sentences of the form $\Diamond_1 p \wedge \Diamond_2\neg\Diamond_1 p$ are predicted, on independent grounds, to be semantically anomalous (in spite of the fact that the left conjunct expresses a property utterly distinct from that expressed by the right conjunct).[31]

The data from natural language, therefore, *do support a version of Euclideanness*, namely, the version in (16). This represents an empirical edge over classical truth-conditional accounts of epistemic modality. Accounts of this type cannot, on the face of things, explain why sentences like (13) are semantically anomalous: in such frameworks, regarding sentences like (13) as inconsistent is equivalent to embracing a Euclideanness constraint *on epistemic accessibility* from the "actual" world w (recall §2):

Euclideanness: $v \in \sigma_w \Rightarrow \sigma_w \subseteq \sigma_v$ $\forall w, \sigma : [\![\Diamond\phi \supset \Box\Diamond\phi]\!]^{\sigma,w} = \top$

This makes vivid the dilemma confronting classical truth-conditional accounts of epistemic modals. Such accounts can at most do one of the following:

- Accommodate Euclideanness (while rendering graded modal judgment unintelligible).
- Accommodate graded modal judgment (while rendering (13) semantically impeccable).

The account defended here skirts the dilemma: it accommodates both the clear semantic intuitions motivating Euclideanness, without sacrificing an intelligible model of graded modal judgment.

B.2 Quantification

Yalcin (2015) notes the following data and observes that no standard theory of epistemic modality—including the theory of Yalcin (2007)—is able to account for it:[32]

(17) #Some/#Every person who is not infected might be infected.

---

[31] Inducing semantic coordination via type-raising yields curious results under negation, as observed by Mandelkern (2019: §7.4) and Simon Charlow (pc). For instance, applying the coordination-via-type-raising strategy to $\neg(\Box p \vee p)$ and $\neg(\Diamond p \wedge p)$ wrongly predicts that both are equivalent to $\neg\Box p$. One can avoid this by stipulating that, when a logical operator takes scope over a syntactically coordinated (but semantically uncoordinated) sentence, the sentence is parsed in conjunctive/clausal normal form (cf. Abney & Keshet 2013). Conditional on this stipulation, $\neg(\Box p \vee p)$ is parsed as $(\neg\Box p \wedge \neg p)$, $\neg(\Diamond p \wedge p)$ as $(\neg\Diamond p \vee \neg p)$; on this parsing, both are (correctly) predicted equivalent to $\Box\neg p$ (proofs omitted). (Fans of type-shifting accounts of semantic coordination between epistemic and non-epistemic language should, however, be on the lookout for a less stipulative way of dealing with this kind of issue.)

[32] I ignore the question of the possible order-sensitivity of the phenomenon (i.e., whether swapping the restrictor clause for the nuclear scope affects the sentence's acceptability). Since semantic coordinability is not order-sensitive, the account here predicts that the phenomenon is not order-sensitive—which is, Yalcin (2015) agrees, probably desirable.

Here I will work through how this data is accounted for, more or less automatically, on the present treatment (while also showing how to extend the theory of generalized quantification to the theory under consideration here).

Assume a first-order version of the language defined in §4.2. Here is the natural clause for the two-place existential quantifier; the two-place universal quantifier is its dual.[33]

$$[\![\exists x(\phi(x))(\psi(x))]\!]^{g_c} = \lambda\mathcal{S}.\{d : \mathcal{S} \in [\![\phi(x)]\!]^{g_c[x/d]}\} \cap \{d : \mathcal{S} \in [\![\psi(x)]\!]^{g_c[x/d]}\} \neq \emptyset$$
$$[\![\forall x(\phi(x))(\psi(x))]\!]^{g_c} = \lambda\mathcal{S}.\{d : \mathcal{S} \in [\![\phi(x)]\!]^{g_c[x/d]}\} \subseteq \{d : \mathcal{S} \in [\![\psi(x)]\!]^{g_c[x/d]}\}$$

Roughly: $\exists x(\phi)(\psi)$ expresses the constraint that $\mathcal{S}$ satisfies iff some d of which $\mathcal{S}$ represents ϕ to hold is such that $\mathcal{S}$ represents ψ to hold of d. Picturesquely, it is the constraint of being such that there is some d such that d is represented as satisfying the quantifier's restrictor and scope. $\forall x(\phi)(\psi)$ expresses the constraint that $\mathcal{S}$ satisfies iff every d of which $\mathcal{S}$ represents ϕ to hold is such that $\mathcal{S}$ represents ψ to hold of d. Picturesquely, it is the constraint of being such that any d such that d is represented as satisfying the quantifier's restrictor is such that d is represented as satisfying the quantifier's scope.

This understanding of generalized quantification in hand, we are in an immediate position to explain (17), using a strategy that is effectively the same as our strategy for (13). Notice that, in the case of a sentence of the form $\exists x(\neg Fx)(\Diamond Fx)$, the semantic types demand raising the quantifier's restrictor:

$$[\![\exists x(\neg Fx)(\Diamond Fx)]\!]^{g_c} = \lambda\mathcal{S}.\{d : \mathcal{S} \in {}_{\texttt{raise}}[\![\neg Fx]\!]^{g_c[x/d]}\} \cap \{d : \mathcal{S} \in [\![\Diamond Fx]\!]^{g_c[x/d]}\} \neq \emptyset$$

Consider any $\mathcal{S}$ that satisfies $[\![\exists x(\neg Fx)(\Diamond Fx)]\!]^{g_c}$. By assumption:

$$\{d : \mathcal{S} \in {}_{\texttt{raise}}[\![\neg Fx]\!]^{g_c[x/d]}\} \cap \{d : \mathcal{S} \in [\![\Diamond Fx]\!]^{g_c[x/d]}\} \neq \emptyset$$

In particular, for some $d \in \{d : \mathcal{S} \in {}_{\texttt{raise}}[\![\neg Fx]\!]^{g_c[x/d]}\} \cap \{d : \mathcal{S} \in [\![\Diamond Fx]\!]^{g_c[x/d]}\}$, $\mathcal{S} \subseteq [\![\neg Fx]\!]^{g_c[x/d]}$, but $\mathcal{S} \cap [\![Fx]\!]^{g_c[x/d]} \neq \emptyset$. Clearly there is no such $\mathcal{S}$.

Similarly, consider any $\mathcal{S}$ that satisfies...

$$[\![\forall x(\neg Fx)(\Diamond Fx)]\!]^{g_c} = \lambda\mathcal{S}.\{d : \mathcal{S} \in {}_{\texttt{raise}}[\![\neg Fx]\!]^{g_c[x/d]}\} \subseteq \{d : \mathcal{S} \in [\![\Diamond Fx]\!]^{g_c[x/d]}\}$$

By assumption:

$$\{d : \mathcal{S} \in {}_{\texttt{raise}}[\![\neg Fx]\!]^{g_c[x/d]}\} \subseteq \{d : \mathcal{S} \in [\![\Diamond Fx]\!]^{g_c[x/d]}\}$$

[33] I provide a syncategorematic semantics for quantification in lieu of a compositional version (which would make use of a polymorphic type for generalized quantifiers). In the general case, for any two-place quantifier Qx:

$$[\![Qx(\phi)(\psi)]\!]^{g_c} = \lambda\mathcal{S}.\mathcal{Q}(\{d : \mathcal{S} \in [\![\phi]\!]^{g_c[x/d]}\}, \{d : \mathcal{S} \in [\![\psi]\!]^{g_c[x/d]}\})$$

Here, $\mathcal{Q}$ is the quantificational relationship between sets expressed by Q (as in Barwise & Cooper 1981). Thanks to Simon Charlow for raising the question of generalized quantification (and for suggesting the natural clauses used here).

In particular, for any $d \in \{d : \mathcal{S} \in {}_{\texttt{raise}}[\![\neg Fx]\!]^{g_c[x/d]}\}$, $\mathcal{S} \subseteq [\![\neg Fx]\!]^{g_c[x/d]}$, but $\mathcal{S} \cap [\![Fx]\!]^{g_c[x/d]} \neq \emptyset$. Clearly there is no such $\mathcal{S}$. Thus, for any context c:

$$[\![\exists x(\neg Fx)(\Diamond Fx)]\!]^{g_c} = \emptyset \qquad [\![\forall x(\neg Fx)(\Diamond Fx)]\!]^{g_c} = \emptyset$$

References

Abney, Steven & Ezra Keshet. 2013. Using conjunctive normal form for natural language semantics. http://www.ezrakeshet.com/site/ezrakeshet/papers/Abney-Keshet-2012-Conjunctive-Normal-Form.pdf. Ms.

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

Bradley, Richard. 2012. Multidimensional possible-world semantics for conditionals. *The Philosophical Review* 121: 539–571. doi:10.1215/00318108-1630921.

Buchak, Lara. 2013. *Risk and Rationality*. New York: Oxford University Press.

Charlow, Nate. 2015. Prospects for an expressivist theory of meaning. *Philosophers' Imprint* 15: 1–43.

Charlow, Nate. 2016a. Decision theory: Yes! Truth conditions: No! In N. Charlow & M. Chrisman (eds.) *Deontic Modality*, 47–81. Oxford University Press.

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

Charlow, Nate. 2018. Decision-theoretic relativity in deontic modality. *Linguistics and Philosophy* 41: 251–287. doi:10.1007/s10988-017-9211-1.

Charlow, Nate. 2019a. Gradable predicates of experience. Ms.

Charlow, Nate. 2019b. Modus ponens and the logic of dominance reasoning. Ms.

Charlow, Nate. forthcominga. In B. Dunaway & D. Plunkett (eds.) *Meaning, Decision, and Norms: Themes from the Work of Allan Gibbard*. University of Michigan: Maize Books, Michigan Publishing.

Charlow, Nate. forthcomingb. The spectre of triviality. *Analysis* doi:10.1093/analys/anz011.

Charlow, Simon. forthcomingc. A modular theory of pronouns and binding. In *Proceedings of LENLS 14*.

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

von Fintel, Kai & Anthony S. Gillies. 2010. *Must* . . . stay . . . strong! *Natural Language Semantics* 18: 351–383. doi:10.1007/s11050-010-9058-2.

von Fintel, Kai & Anthony S. Gillies. 2011. *Might* made right. In A. Egan & B. Weatherson (eds.) *Epistemic Modality*, 108–30. Oxford: Oxford University Press.

von Fintel, Kai & Anthony S. Gillies. 2018. Still going strong .

Gaifman, Haim. 1988. A theory of higher order probabilities. In Brian Skyrms & William L Harper (eds.) *Causation, Chance and Credence: Proceedings of the Irvine Conference on Probability and Causation Volume 1*, 191–219. Dordrecht: Springer Netherlands.

Gärdenfors, Peter & Nils-Eric Sahlin. 1982. Unreliable probabilities, risk taking, and

decision making. *Synthese* 53: 361–386. doi:10.1007/BF00486156.
Gillies, Anthony S. 2010. Iffiness. *Semantics and Pragmatics* 3: 1–42. doi:10.3765/sp.3.4.
Hájek, Alan. 1989. Probabilities of conditionals—revisited. *Journal of Philosophical Logic* 18: 423–428. doi:10.1007/BF00262944.
Hájek, Alan. 2009. Dutch book arguments. In P. Anand, P. Pattanaik & C. Puppe (eds.) *Oxford Handbook of Rational and Social Choice*, 173–195. Oxford: Oxford University Press.
Hild, Matthias. 1998. Auto-epistemology and updating. *Philosophical Studies* 92: 321–361.
Holliday, Wesley H. & Thomas F. Icard III. 2010. Moorean phenomena in epistemic logic. In L. Beklemishev, V. Goranko & V. Shehtman (eds.) *Advances in Modal Logic*, vol. 8, 178–199. College Publications.
Joyce, James. 1998. A nonpragmatic vindication of probabilism. *Philosophy of Science* 65: 575–603. doi:10.1086/392661.
Joyce, James. 2005. How probabilities reflect evidence. *Philosophical Perspectives* 19: 153–178. doi:10.1111/j.1520-8583.2005.00058.x.
Joyce, James. 2010. A defense of imprecise credences in inference and decision making. *Philosophical Perspectives* 24(1): 281–323. doi:10.1111/j.1520-8583.2010.00194.x.
Kaplan, David. 1989. Demonstratives. In J. Almog, J. Perry & H. Wettstein (eds.) *Themes from Kaplan*, 481–563. Oxford: Oxford University Press.
Kratzer, Angelika. 1981. The notional category of modality. In H. Eikmeyer & H. Rieser (eds.) *Words, Worlds, and Contexts*, 38–74. Berlin: De Gruyter.
Kratzer, Angelika. 1986. Conditionals. *Chicago Linguistics Society* 22: 1–15.
Lasersohn, Peter. 2005. Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy* 28: 643–686. doi:10.1007/s10988005-0596-x.
Lewis, David. 1976. Probabilities of conditionals and conditional probabilities. *The Philosophical Review* 85: 297–315. doi:10.2307/2184045.
Mandelkern, Matthew. 2019. Bounded modality. *Philosophical Review* 128(1): 1–61. doi:10.1215/00318108-7213001.
Moss, Sarah. 2015. On the semantics and pragmatics of epistemic modals. *Semantics and Pragmatics* 8: 1–81.
Moss, Sarah. 2018. *Probabilistic Knowledge*. Oxford: Oxford University Press.
Moss, Sarah. 2019. Global constraints on imprecise credences. http://www-personal.umich.edu/~ssmoss/GlobalConstraints.pdf. Ms.
Partee, Barbara H. & Mats Rooth. 1983. Generalized conjunction and type ambiguity. In R. Bäuerle, C. Schwarze & A. von Stechow (eds.) *Meaning, Use, and Interpretation of Language*, 361–383. Berlin: Walter de Gruyter.
Pierce, Benjamin C. 2002. *Types and Programming Languages*. MIT Press.
Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. In *OSU Working Papers in Linguistics, Vol 49: Papers in Semantics*. Ohio State University. http://semanticsarchive.net/Archive/WYzOTRkO/.
Rothschild, Daniel. 2012. Expressing credences. *Proceedings of the Aristotelian Society* CXII, Part 1: 99–114. doi:10.1111/j.1467-9264.2012.00327.x.
Rothschild, Daniel. 2015. Conditionals and propositions in semantics. *Journal of*

Philosophical Logic 44: 781–791. doi:10.1007/s10992-015-9359-5.

Schroeder, Mark. 2011. Two roles for propositions: Cause for divorce? *Noûs* 47: 409–430. doi:10.1111/j.1468-0068.2011.00833.x.

Staffel, Julia. forthcoming. Expressivism, normative uncertainty, and arguments for probabilism. In *Oxford Studies in Epistemology*.

Stalnaker, Robert. 1970. Probability and conditionals. *Philosophy of Science* 37: 64–80.

Stalnaker, Robert. 1984. *Inquiry*. Cambridge: MIT Press.

Weatherson, Brian. 2004. Moore, Bradley, and indicative conditionals. Ms.

Willer, Malte. 2013. Dynamics of epistemic modality. *The Philosophical Review* 122: 45–92. doi:10.1215/00318108-1728714.

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

Yalcin, Seth. 2009. More on epistemic modals. *Mind* 118: 785–793. doi:10.1093/mind/fzpl06.

Yalcin, Seth. 2011. Nonfactualism about epistemic modality. In A. Egan & B. Weatherson (eds.) *Epistemic Modality*, 295–332. Oxford: Oxford University Press.

Yalcin, Seth. 2012. Bayesian expressivism. *Proceedings of the Aristotelian Society* CXII, Part 2: 123–160. doi:10.1111/j.1467-9264.2012.00329.x.

Yalcin, Seth. 2015. Epistemic modality *De Re*. *Ergo* 2: 475–527. doi:10.3998/ergo.12405314.0002.019.

Yalcin, Seth. 2018. Belief as question-sensitive. *Philosophy and Phenomenological Research* 97: 23–47. doi:10.1111/phpr.12330.