A Counterexample to Modus Tollens

Seth Yalcin

Abstract This paper defends a counterexample to Modus Tollens, and uses it to draw some conclusions about the logic and semantics of indicative conditionals and probability operators in natural language. Along the way we investigate some of the interactions of these expressions with *knows*, and we call into question the thesis that all knowledge ascriptions have truth-conditions. A probabilistic dynamic semantics for probability operators, conditionals, and acceptance attitudes is developed around the idea of representing the common ground of a conversation as a set of probability spaces.

1 A Counterexample to Modus Tollens

An urn contains 100 marbles: a mix of blue and red, big and small. The breakdown:

	blue	red
big	10	30
small	50	10

S. Yalcin (✉)
University of California, Berkeley, CA, USA
e-mail: yalcin@berkeley.edu

A marble is selected at random and placed under a cup. This is all the information given about the situation. Against this background, the following claims about the marble under the cup are licensed:

(P1)   If the marble is big, then it's likely red.
(P2)   The marble is not likely red.

However, from these, the following conclusion does not intuitively follow:

(C1)   The marble is not big.

But this conclusion would follow, were Modus Tollens (MT) valid. So MT is not generally valid.

We do well to put the point in various ways. For example: it is entirely possible to believe (P1) and (P2), but fail to believe (C1), even after full and complete rational reflection. Contrariwise, if one believed only (P1) and (P2), and on the basis of these concluded (C1), one would be making a mistake. Further, it is possible to suppose (P1) and (P2), without (C1) in any way following from one's suppositions. Further, one can be given the information that (P1) and (P2) express, without the information given by (C1)'s being part of one's information in any way.

I take it that the probability operators *likely* and *probably* are synonyms, so I will use them interchangeably. Schematically, our case is of this form:

$$\phi \rightarrow \text{ PROBABLY } \psi$$

$$\neg (\text{PROBABLY } \psi)$$

$$\therefore \neg\phi$$

where ϕ, ψ are themselves assumed to be probability operator-free.[1] This argument form is invalid. Since it is just a special case of MT, it is a counterexample to the claim that MT is a generally valid pattern.

Let me stress before continuing that I am by no means the first to advance a counterexample to MT. Lewis Carroll's barbershop paradox [5] effectively supplies a counterexample to MT, one employing a right-nested indicative conditional.[2] Frank Veltman developed a similar counterexample in his pioneering dissertation, citing Carroll for inspiration [30]. On a natural reading, Forrester's gentle murder paradox [7] calls attention to what looks like a

---

[1]And epistemic modal-free. (Probability operators just are epistemic modals, in the relevant semantic respects; see [18, 37].) Or at least, where whatever probability operators/epistemic modals there be in ϕ, ψ are safely nestled in the appropriate embedded context—under 'John believes', for instance. Unless otherwise noted, all of my schematic letters range over 'unhedged' sentences of this sort.

[2]Cleaned up, here are Carroll's premises: *If Carr is out, then if Allen is out, Brown is in*; and *it's not the case that if Allen is out, Brown is in*. The first we know because we know one of the three is in; the second we know because Allen never leaves without Brown. But we cannot conclude from our premises (by MT) that *Carr is in*. After all, if Allen is in, Carr might well be out.

counterexample to MT involving deontic modals in the consequent.[3] Building on Forrester, [4] suggests that MT admits of failures in cases wherein the consequent is deontically or epistemically modalized. Kolodny and MacFarlane [15] make the same claim.[4]

Aside from examples involving modalized or conditional consequents, we might also reach for examples involving consequents which superficially appear to contain adverbs of quantification. For instance:

> It is not the case that the alarm always sounds.
>
> If there is a break-in, the alarm always sounds.
>
> ⊭ There is no break-in.

If one is worried about a break-in, the premises here are less than comforting.[5]

My objective in this note is not to defend all of these counterexamples, though I am sympathetic with most of them. Rather, my aim to to defend just one—the one we began with—in a sustained way, and in a way which meets some replies left unmet in existing discussions. I focus on the interaction of the conditional with probability operators for three reasons: first, because the intuition that the relevant patterns are invalid is relatively clear; second, because probability operators don't admit of the diversity of possible readings that many other ordinary modal auxiliaries do; and third, because the role of probabilities in the interpretation of conditionals is of considerable independent interest.

A further advantage of focusing on probability operators is that they embed comparatively easily under *knows*. Once we motivate our counterexample to MT, we will show how it can be leveraged to learn something about the logic of knowledge operators. Specifically, we will show how it can be leveraged to argue that at least some knowledge ascriptions do not have truth-conditions.[6]

2 Objections and Replies

Here are three possible replies to our counterexample. The first reply is that I have misrepresented the logical form of (P1). The probability operator in the sentence is really taking scope over, not under, the conditional operator; and as a result the pattern is a non-instance of MT. The second reply is that, thanks to

---

[3]A variation of Forrester's theme: *If you will kill him, then you ought to kill him gently*; and *it's not the case that you ought to kill him gently*. It seems we should conclude from these by MT *you will not kill him*. And yet our premises don't intuitively entail anything at all about what you will do next.

[4]We might add that MT fails already on the dynamic conditional semantics developed by [9], and on the static conditional semantics given by [35].

[5]One can construct similar examples with *usually*, *often*, *typically* and the like.

[6]Understanding 'truth-condition' here to correspond to a way the world might have been, the sort of thing model-able as a function from possibilities to truth-values.

the context-sensitivity of *likely*, its semantic contribution is not constant across the premises, and so what is negated in (P2) is not the consequent of (P1). The third reply is indirect: it merely stresses the abundance of cases wherein MT is obviously valid, and says that it will always be more reasonable to suppose that something is suspect with my example than to give up MT. Let me take these objections in reverse order.

We can be brief with the third response. There are indeed many instances of MT which are semantically valid.[7] Notable are those cases wherein the conditional is free of explicit modals ('bare conditionals'). If we snip the probability operators from our premises, for example:

(P1′)   If the marble is big, then it is red.
(P2′)   The marble is not red.

The conclusion (C1) certainly does seem to follow. Does this counteract the force of our counterexample to MT? It is hard to see how. What we should want is simply a semantics and definition of consequence which match the data: an account which will validate the inference from (P1′) and (P2′) to (C1), but which will invalidate the inference from (P1) and (P2) to (C1). (And do the same for all structurally parallel examples.) Now if there were some principled difficulty in giving an account that could do this, that might be a theoretical motivation to reconsider our judgments about our counterexample. But there is no such difficulty. There are existing accounts of indicative conditionals and of probability operators that will do the job.[8] Meanwhile, let us agree wholeheartedly that MT is valid, except for when it is not.[9]

Turn now to the second reply, that the context-sensitivity of *likely* is what generates the illusion of a counterexample. Now there is no question that probability operators exhibit context-sensitivity. For instance, when one says a certain outcome is likely, what is said depends on what the alternatives to that outcome are taken to be in context (see [37]). The question is, what positive reason is there to think that there is some illicit context-shifting going on between our two premises (P1) and (P2)? There appears to be little independent motivation for the idea. In conversation I have heard some cite the feeling that the thing being said to be likely in (P1) is a *conditional probability*, namely, the conditional probability that the marble is red if big, whereas the thing said to be likely in (P2) is not a conditional probability. Or to put essentially the same thought in less loaded terms, the idea is that the *likely* in (P1) is in some sense semantically evaluated with respect to more

---

[7]One valid instance occurred in the first paragraph of this paper, when I concluded that MT is not generally valid.

[8]See, *inter alia*, [9, 10, 15, 30, 35, 37] for relevant work and the appendix for explicit examples.

[9]Could one retreat to the claim that MT should be construed as a generalization only about bare conditionals? This would leave out valid patterns that seem to be instances of MT: for example (where '◇' corresponds to epistemic possibility): 'If ϕ, then $\Diamond\psi$; $\neg\Diamond\psi$; therefore $\neg\phi$.' We could deny that this is an instance of MT if we like, but it is difficult to see what advantage there is to talking this way.

information than the *likely* in (P2). This is taken to be evidence in support of the idea of illicit context-shifting.

It seems to me the intuition that the *likely* in (P1) is in some sense semantically evaluated with respect to more information than the *likely* in (P2) is surely correct. But the inference from this intuition to illicit context-shifting is a non-sequitur. There is no tension between (i) the intuition that the thing said to be likely in (P1), but not (P2), is a conditional probability, and (ii) the idea that the semantic contribution of the probability operator in (P1) and (P2) is the same. To square these, we merely require the assumption that probability operators are sensitive in their interpretation to a parameter which is semantically shiftable by other linguistic expressions—and in particular, by conditional environments. And there is already solid evidence that modal operators are in general sensitive to some such semantically shiftable parameter; indeed, that has claim to being the standard view of the matter.[10] So, bracketing some independent evidence, appealing to context-sensitivity and context-shifting here is quite unmotivated.

We come finally to the third objection. Again, this objection says that at the relevant level of logical form, (P1) is not well-schematized by

$$\phi \rightarrow \text{ PROBABLY } \psi$$

Rather, the compositionally revealing logical form is something else, namely:

$$\text{PROBABLY } (\phi \rightarrow \psi)$$

Hence our first premise is not really a conditional; and hence our argument is not really an instance of MT.

The first point to make here is a clarificatory one. The worry I press about MT depends only on the availability of the narrow scope reading of PROBABLY, not on the unavailability of the wide scope reading. So the theorist who insists that PROBABLY takes wide scope in (P1) carries a strange burden: he must explain what *obligatorily rules out* the possibility of the scopal order $\rightarrow$, PROBABLY—the superficial order.

We should be clear about the nature of this burden. It would be one thing if we could detect some semantic difference between these two allegedly logically possible scopal orders, and then simply declare that intuitively, the scopal order in (P1) is obligatorily PROBABLY, $\rightarrow$. This style of argument certainly works in principle.[11] But the trouble is, if there are two scopal orders logically possible for sentences like (P1), it seems that they yield semantically

---

[10]See [18]. For further discussion see [35].

[11]To give an example of a case where this style of reasoning works, consider: "Everyone probably lost the lottery". In this sentence there are two logically possible scopal orders for the interpretation of PROBABLY and the quantifier EVERYONE, each of which would yield a (truth-conditionally) different reading of the sentence. But intuitively, only the reading with *probably* taking wide scope is actually available. See [32].

equivalent readings.[12] And if they yield equivalent readings, what motivation is there for maintaining that the probability operator takes, and must take, wide scope? For this assumption would not be required to explain how the sentence is interpreted. Rather, what would need explaining from a semantic point of view is just why the two logically possible scopal orders yield equivalent readings.

So the burden is on the wide-scoper to motivate her view. (I take it "Otherwise, MT would fail" is not adequate motivation.) Moreover the burden appears to be a very difficult one to carry, for at least the following two reasons.

First, from (P1) and

(1)   The marble is big.

we can apparently conclude

(2)   The marble is probably red.

by Modus Ponens (MP). But this could only be an instance of MP if the probability operator in (P1) does not have wide scope—hence only if the problematic scopal order is in fact available. So the wide-scoper is obliged to deny a whole class of what seem to be routine applications of MP.

Second and more damaging, we have problems with conjunctive consequents whose conjuncts are of variable modality. Plainly

(3)   If Sally is at the party, then Isaac is at the party and Steve is probably at the party.

is not equivalent to:

(4)   Probably, if Sally is at the party, then Isaac is at the party and Steve is at the party.

For instance (3), together with

(5)   Sally is at the party.

entails

(6)   Isaac is at the party.

whereas (4) and (5) together intuitively entail, not (6), but

(7)   Isaac is probably at the party.

So it is clear that we should not represent the probability operator in (3) as taking widest scope in the sentence (let alone mandatorily taking wide scope). If we are working under the hypothesis that the sentence contains a conditional

---

[12]The point that sentences superficially of the form IF ϕ, PROBABLY ϕ and sentences of the form PROBABLY, IF ϕ, ψ generally strike us as semantically equivalent has been made by many others; for instance [29]. See [25] for the view that there are some subtle cases in which these readings can in fact be teased apart.
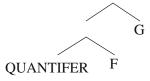
operator, the natural thing to say is that it takes scope over the probability operator. But if the conditional operator may take scope over the probability operator in (3), we should of course expect that it may in (P1).

Pleaing wide scope is thus unmotivated. Other things being equal, we may that assume scope-taking expressions can assume the relative scopes that they superficially occupy. If things are not equal here, we are owed some account why.
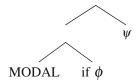
3 Flank Attack from the Restrictor Analysis

There is another reason to worry about the general validity of MT. This is a very different worry. It is the worry that the compositional semantics of conditional constructions is such that required notions of *antecedent* and *consequent* do not really make sense. Let me explain.

In accord with most philosophical discussions of conditionals, we have so far assumed that conditional meanings are the result of the semantic contribution of a single dyadic conditional operator. But many find a second idea about the semantics of conditionals much more plausible. On this idea, we assume that modals are much like quantifiers, and we suppose that IF-clauses are devices for restricting the quantificational force of modals. (The idea is most famously developed by [16, 17], building on [20]; see also [12].) Just as quantificational determiners combine with a (perhaps tacit) domain restriction and a nuclear scope, so modals combine with a (perhaps tacit) modal restriction and a matrix clause. Supposing that quantificational sentences have this kind of structure:

$$
\begin{array}{c}
\diagup\diagdown \quad G \\
\diagup\diagdown \quad F \\
\text{QUANTIFER}
\end{array}
$$

The idea would be that conditional constructions are analogous:

$$
\begin{array}{c}
\diagup\diagdown \quad \psi \\
\diagup\diagdown \\
\text{MODAL} \quad \text{if } \phi
\end{array}
$$

As quantifiers express quantification over individuals, so modals express quantification over possibilities. As quantifiers express relations between properties, so modals express relations between propositions. On this view, it isn't that IF adds to natural language a new dyadic modal operator. Rather, various modal operators of natural language, operators traditionally supposed to be monadic, are now hypothesized to be fundamentally dyadic in nature. Moreover, IF-clauses mark restrictions and express no modality on their own. Multiple IF-clauses may correspond to multiple restrictions, but needn't entail the presence of multiple modals. Where no modal operator is superficially

apparent in a conditional construction, the presence of a tacit epistemic necessity modal is typically assumed.[13] As for our (P1) above, the relevant modal operator would be PROBABLY.

Now it is clear that if this kind analysis is correct, the inference we began with is *not* plausibly an instance of MT. For on such an analysis, what is negated in (P2) is *not even a constituent* in (P1); *a fortiori*, what is negated is not the consequent of (P1). (I take it the consequent of a conditional, whatever else it is supposed to be, is at least a constituent of the conditional.)

This may appear to be good news for the Modus Tollens Lobby, for it eliminates our counterexample. On the contrary: it is bad news. For the larger upshot of this analysis is that MT is based on a mistake. MT is most naturally taken as a generalization concerning a certain dyadic modal operator, the conditional operator. But there is no such operator according to the restrictor analysis. There is no single sort of dyadic modal operator figuring in every conditional, or even in every indicative conditional. Rather, there are just various dyadic modal operators, corresponding to the various modals of natural language. And MT certainly does not characterize *all* of these operators.

One way to bring this out is to see that there is no stable notion of 'antecedent' and 'consequent' in the setting of the restrictor analysis which could vindicate MT. Suppose we try to construct an instance of MT from (P1). How to proceed? We need only add a premise which negates its consequent.—But what is its consequent? We cannot identify its consequent with its matrix clause, for that would, absurdly, make the following an instance of MT:

 P1. If the marble is big, then it's likely red.
 P2″. The marble is not red.
 C1. The marble is not big.

It would also, absurdly, make the following an instance of MP:

 P1. If the marble is big, then it's likely red.
 P2″. The marble is big.
 C1. The marble is red.

The larger upshot is that if modals are really just quantifiers over worlds and IF-clauses just ways of restricting these quantifiers explicitly, it is not obvious that we should expect there to be any distinctive logic of conditionals going beyond the logic of generalized quantifiers.

Note further that the restrictor analysis problematizes MP just as it does MT, for just the same reasons. It seems to me that if there is a good reason to doubt MP over and above MT, it is this abstract one stemming from the

---

[13]At least in the context of 'epistemic' indicative conditionals, our focus here. In other cases, the tacit necessity modal needs to be a frequency adverb, as in the sort of cases originally discussed by [20] (e.g., *If you pet him, he (always) bites*).

consequences of the restrictor analysis.[14] For MP, unlike MT, does not give rise to the same abundance of apparent counterexamples.[15]

4 Truth and Consequence

My brief against MT is essentially complete. Let me close it by anticipating a line of concern I expect some readers will have.

Consider the following argument:

(i)   The semantic value of an indicative conditional relative to context is, or has, a possible worlds truth-condition.
(ii)   If the semantic value of an indicative conditional relative to context is, or has, a possible worlds truth-condition, then the only plausible truth-condition for it will be a truth-condition making MT valid.
(iii) So MT is valid.

The argument above is of course intuitively valid,[16] so if our preceding conclusions about MT are correct, one or both of the premises must be rejected. Now, it is not our burden to explain where every argument in support of MT goes wrong, and we are not particularly obliged to respond to the above argument until its premises are defended. But as many readers will want to embrace both of the above premises, it is worth pausing to note some respects in which these premises are both highly nontrivial—indeed, contentious. This will help clarify the burden on those wishing to defend the premises, and it will help to clarify what rejecting MT does and does not entail.

Some would embrace (i) for the following reason: they believe (plausibly) that indicative conditionals have a compositional semantics, and they also believe:

(iv)   If indicative conditionals have a compositional semantics, then (i) must be true.

---

[14]Kolodny and MacFarlane [15] reject MP in addition to MT. They do so, not in the face of any worry stemming from the restrictor analysis, but rather *en route* to solving a certain puzzle about sentences expressing conditional obligation (the miners puzzle). Here I wish to note two points: (i) In connection with solving the miners puzzle, the particular way of defining of consequence Kolodny and MacFarlane adopt is not superior to rival definitions which would validate MP, such as the notion called 'informational consequence' in [35]. (ii) Kolodny and MacFarlane's preferred formalization of consequence has a problematic feature: the semantic analogue of the deduction principle (i.e., the principle that if $\Gamma, \phi \vDash \psi$, then $\Gamma \vDash \phi \rightarrow \psi$) fails. That suggests an unexpected disconnect between conditionals and consequence. (The rival notion of consequence recently mentioned, by contrast, does vindicate this principle.)

[15]The best known putative counterexamples to MP are due to [23]. In his discussion, McGee crucially assumes a certain strong connection obtains between the notion of *having good reason to believe* and the notion of *consequence* MP (putatively) characterizes. Had I the space I would question this connection, and I would worry about whether McGee's examples really have the status of linguistic explananda.

[16]Not to say that it is valid *because* MP is; as recently noted, that is controversial.

The thesis (iv) seems to be presupposed with little or no argument in many discussions of conditionals.[17] But it is a quite nontrivial thesis. Indeed, in light of the existence of well-motivated compositional semantic systems for conditionals which effect no straightforward semantic association between indicative conditionals and possible worlds truth-conditions,[18] it appears to be false. The larger point, in any case, is that there is no in-principle tension between accepting compositionality and rejecting (i).

Others would embrace (i) because they believe (plausibly) that indicative conditionals participate nontrivially in valid arguments, and they also believe the following two claims:

(v)   If indicative conditionals participate nontrivially in valid arguments, then they have truth values.

(vi) If indicative conditionals have truth values, then (i) must be true.

One will believe (v) if one assumes that consequence is to be modeled in terms of truth-preservation, roughly along Tarksian lines: a valid argument is one such that if the premises are true, the conclusion must be true. Many discussions of consequence proceed under the assumption that adopting this way of characterizing consequence is a theoretically innocent move. But let us be explicit that this is not so. Whether consequence should be modeled in terms of truth-preservation is a substantive and debated question. Indeed, in the context of natural language semantics, the view that consequence is to be modeled in terms of truth-preservation is really a (high-level) empirical thesis. (One with a number of competitors.[19]) For a view about the compositional semantics of indicative conditionals in natural language only makes robust predictions when paired with some characterization of consequence. It is really the package of the two that makes predictions. In response to data, either part of the package may in principle be revised. And amongst the data that must be factored into the cost-benefit analysis for any given formal characterization of consequence are our judgments about the argument discussed at the opening of this paper. Thus it would make no sense—indeed, it would get things quite backward—to reject the counterexample merely on the grounds that it sits uneasily with a Tarksian definition of consequence.

This is why, incidentally, our counterexample is not an attempt to *describe a world* with respect to which the premises are true, but the conclusion false. To assume that this is what would be required to refute MT is to illicitly assume something like (i), and a Tarksian account of consequence, in advance. But

---

[17]See, e.g., [3, 6, 8, 21, 23].

[18]See, among others, [9, 12, 13, 15, 30, 35].

[19]Examples abound, but the dynamic tradition provides an especially clear and relevant alternative: see [2, 11, 12, 30], among many others. The static account of consequence defended in [35] is another example.

one of the issues the counterexample raises just is the question whether such a notion of consequence could be adequate for modeling natural language.[20]

The thesis (vi) is also substantive, although here the reason is more technical. A sentence may have a compositional semantic value which determines a truth-value with respect to a point of evaluation (in a model), without the relevant points of evaluation needing themselves to be possible worlds (or context-world pairs). In such a setting, indicative conditionals might have truth-values relative to points of evaluation, but fail to have truth-values with respect to worlds in any interesting sense—fail to correspond to a way the world might be.[21] So (vi), too, is nontrivial.

Finally, consider (ii). If validity is understood in terms of truth-preservation, and if (i) is true, then it is indeed difficult to see what the possible worlds truth-conditions of indicative conditionals could be if MT is not valid for indicatives. MT is of course valid on the leading possible world semantics for conditionals, namely the Stalnaker–Lewis analysis [19, 27, 28]; and it is valid on the two other most widely-discussed truth-conditional accounts, namely the strict conditional analysis and the material conditional analysis. But the difficulty here is mainly for the case in which we assume, additionally, that the conditional involves a dyadic sentential operator. If instead we adopt the Kratzer-Lewis style syntax for conditionals described in the previous section, it is perfectly clear how (i), but not (ii), could hold; just see [17]. Suffice to say there is nothing trivial about (ii).

5 Knowledge Operators in MT Inferences

The evidence suggests that MT is either not generally valid, or based upon a mistake. We should take this result as a constraint on semantic theory. We should like a semantic theory for indicative conditionals and for probability operators which, together with an appropriate formal characterization of consequence, shows MT to be invalid in the kind of case we have discussed.

In this section I want to consider a question about the scope of the failure of MT. Setting aside the syntactic worry for the moment, let us pretend the notions of 'antecedent' and 'consequent' do make sense. Above we noted that MT inferences wherein the consequent lacks an overt modal generally strike us as valid. We can add that when modals appearing in the consequent are nestled

---

[20]A relevant comparison for the notion of consequence at work here might be with the theoretical notion of *grammaticality* in natural language syntax. The syntactician will often characterize subjects as judging that some sentences are grammatical and that some are not. When she does this, she is employing a theoretical notion of grammaticality in a empirically-driven enterprise, one which may be (usually is) alien to the subjects being described. Her use of this notion, and her characterization of subjects, is justified insofar as it plays a role in a theory which best explains a target range of facts—in this case, certain linguistic capacities of the relevant subjects. I am taking it that the notion of consequence employed in natural language semantics has a parallel theoretical status.

[21]See the works cited in footnote 19 above.

in some *appropriate embedded context*, the result is also valid. For example the following MT inference, with a probability modal appearing in the consequent under a belief operator, is valid:

(P3)   If the marble is big, then John believes it is probably red.
(P4)   John does not believe that the marble is probably red.
(C1)   The marble is not big.

What constitutes an "appropriate embedded context"? I will not attempt to answer that question in full generality here, but let us consider the special case of the factive attitude verb *knows*.

Initial appearances suggest nothing different from *believes*. Consider:

(P5)   If the marble is big, then John knows it is probably red.
(P6)   John does not know that it is probably red.
(C1)   The marble is not big.

The argument is intuitively valid.

Trouble is not far off, however. To see the difficulty, consider first:

(8)   # If the marble is not big, then it is probably big.

I take it this conditional is incoherent. I take it also that if we add a knowledge operator to the consequent, the result is still marked and uninterpretable:

(9)   # If the marble is not big, then John knows it is probably big.

That (9) is defective is unsurprising, given that (8) is defective. For (9) seems obviously to entail (8); and sentences that obviously entail defective sentences are generally defective themselves.[22]

Predicably, if we negate the consequent of (9), the result is acceptable:

(10)   If the marble is not big, then John does not know it is probably big.

Now to come to the difficulty, fix your attention upon (10). Let us consider an MT inference involving it. We add the relevant minor premise:

(11)   John knows the marble is probably big.

Question: do (10) and (11) together entail (1) (repeated)?

(1)   The marble is big.

---

[22]The entailment from (9) to (8) would follow from the factivity of knowledge operators ($K\phi \vDash \phi$) and from transitivity for indicatives. Transitivity for indicatives is controversial in some quarters; see [3]. The fact that transitivity would help explain the defect in (9) (by reducing it to the defect in (8)) is a reason to favor it.

To clarify, we are asking about arguments fitting this schema:

$$\neg\phi \rightarrow \neg K(\text{PROBABLY } \phi)$$

$$K(\text{PROBABLY } \phi)$$

$$\therefore \phi$$

The validity of arguments fitting this schema would follow from MT and Double Negation. We have seen no independent reason to question Double Negation, so I will just assume without question that if the argument here is invalid, it is another case of the failure of MT.

Intuitive judgments about this example may be less clear than they were with our original counterexample to MT. But consider the following line of thought:

> If the marble is not big, then (*a fortiori*) it is not the case that it is probably big. Hence it is not the case that anyone, for instance John, knows that the marble is probably big. So the conditional (10) is really quite trivial. Now to this triviality, let us add an assumption about John's knowledge state, namely the assumption that John knows that the marble is probably big. This assumption obviously does not, by itself, entail that the marble is big. So why think that it would together with a triviality? Thus the inference from (10) and (11) to (1) is invalid; and similarly for any argument of the same schematic form.

If this line of reasoning is not sound, then presumably what must be rejected is the idea that (10) is in some sense trivial. But it is difficult to see why we should not regard it is as trivial. Doesn't the marble's not being big preclude it's likely being big, and hence preclude anyone's knowing that it is likely big?

In support of the thought, we can note that there is undoubtedly some logical tension between (12) and (13):

(12)   The marble is not big.
(13)   John knows the marble is probably big.

The conditional (9) clearly illustrates this. We observe a tension also when we simply attempt to hypothetically entertain their conjunction:

(14)   # Suppose the marble is not big and John knows the marble is probably big.

As the trouble with (9) is rooted in the factivity of *knows* together with the badness of (8), so the trouble with (14) is plausibly rooted in the factivity of *knows* together with the badness of (15):

(15)   # Suppose the marble is not big and the marble is probably big. (cf. [35])

These considerations favor the idea that the inference from (10) and (11) to (1) is indeed invalid.

Is there a way of pushing back against this conclusion? Consider the following rejoinder:

Granted (12) and (13) seem incompatible. *If they are incompatible, however, then the truth of the latter would entail the negation of the former.* But that would be absurd: If we are given merely that John knows the marble is probably big, it does not follow that the marble is big. So we should resist the superficially compelling idea that (12) and (13) are incompatible.

We should agree that it would be absurd to hold that (13) entails the negation of (12). But why can't we reject that entailment, and *also* maintain that the two sentences are incompatible? To assume that we cannot—as the italicized remark above does—is to beg the question in favor of a classical account of consequence. And a natural thought is that these data call classicality into question if anything does. On balance, what these data *prima facie* suggests is simply that we need a semantics and an account of consequence according to which (12) and (13) incompatible, despite the latter's not entailing the negation of the former. That is, generally speaking, we should like a semantic theory equipped with an account of consequence according to which:

(I) $(\neg\phi \wedge K(\textsc{probably } \phi)) \vDash \bot$
(II) $\neg\phi \vDash \neg K(\textsc{probably } \phi)$
(III) $K(\textsc{probably } \phi) \nvDash \phi.$

(I) is supported by data like (9) and (14); (II) is supported by the factivity of knowledge plus the idea, suggested by (8) and (15), that $\neg\phi \vDash \neg\textsc{probably}\phi$; and (III) is obvious. Or to put it differently, the properties are desiderata for our semantic theory because the following are desiderata:

(I′) $(\neg\phi \wedge \textsc{probably } \phi) \vDash \bot$
(II′) $\neg\phi \vDash \neg\textsc{probably } \phi$
(III′) $\textsc{probably } \phi \nvDash \phi$

—and because *knows* is factive. Obviously a classical account of consequence could not satisfy (I)–(III) together. (Or (I′)–(III′) together.)

Yalcin [35] notes that it is difficult to see how to reconcile (I′)–(III′) with possible worlds truth-conditions for $\textsc{probably}\phi$-sentences. Given the way consequence is usually defined in the setting of possible worlds semantics,[23] no truth-conditions could satisfy these demands. To this we can add that the same is true for $K(\textsc{probably}\phi)$-sentences, with respect to (I)–(III). Given the way consequence is usually defined in the setting of possible worlds semantics, there is no way of associating $K(\textsc{probably}\phi)$ with possible worlds truth-conditions that would satisfy these constraints.

---

[23]I have in mind a formalization of consequence along the following lines: $\phi_1, ..., \phi_n \vDash \psi$ just in case for any world w, if $[\![\phi_1]\!]^w = 1, ..., [\![\phi_n]\!]^w = 1$, then $[\![\psi]\!]^w = 1$, for all sentences $\phi_1, ..., \phi_n, \psi$.

Where does this leave us? First, it leaves us tentatively of the view that the inference from (10) and (11) to (1) is indeed invalid, as is anything of the same schematic form. A probability operator in the consequent position of a conditional will invalidate MT reasoning on that conditional even when the probability operator is embedded under a knowledge operator, which is in turn embedded under a negation operator. Second and more interestingly, it leaves us with some evidence against the idea that all knowledge ascriptions can be associated with possible worlds truth-conditions. Thanks to the factivity of *knows*, knowledge ascriptions inherit the unusual semantic features of the epistemic modal sentences they embed.

6 Closing

How to understand the failure of MT, and this surprising conclusion about the truth-conditions for knowledge ascriptions embedding epistemic modals? Our objective here has been entirely negative: the aim was just to shift the burden to those who would take the general validity of MT as a desideratum for a theory of conditionals. We have already noted a number of contemporary frameworks in which MT is rejected. Let me close by sketching the direction of explanation I favor, gestured at in various places above, without pretending to fully motivate it here (see [35, 38], and the Appendix below for further discussion).

We observed that it is entirely possible to believe (P1) and (P2) while failing to believe (C1), even after full and complete rational reflection.[24] Let us step back and ask: what exactly is it to believe (P1)? Suggestion: it is, ideally, to be in a credal state giving the outcome that the marble is red is better-than-even odds, conditional on the marble's being big. What is it to believe (P2)? Suggestion: roughly, it is to be in a credal state that gives the outcome that the marble is red even-or-lower odds. What is it to believe (C1)? Suggestion: it is to be in a state whose content rules out that possibility that the marble is big.

If these suggestions are on track, then it is clear enough why it can be rational to believe (P1) and (P2), but fail to believe (C1). We can make the matter clearer if we model idealized credal states, and information states generally, as probability spaces. Then trivially there will be credal states which satisfy the requirement that

(a) $Pr(\texttt{red}|\texttt{big}) > .5$

and the requirement that

(b) $Pr(\texttt{red}) \leq .5$

---

[24]Not to suggest that this is somehow the *sine qua non* of an invalid argument, but merely that it is a characteristic mark of invalid arguments.

but which nevertheless fail to rule out big. (Indeed there will be rational credal states satisfying (a) and (b), and yet also such that $Pr(\texttt{big}) > .5$).

To get from these observations back to semantics, note that on this way of thinking about what it is to believe sentences such as (P1) and (P2), believing what they say is not tantamount to ruling some possibilities in or out. (At least not without nontrivial further assumptions.) It is instead a matter of one's doxastic state's satisfying certain global features, features that do not simply reduce to the way the state represents the world to be. This works nicely with the idea already motivated, namely that the semantic values of PROBABLYϕ-sentences are not given by possible worlds truth-conditions (functions from possibilities to truth-values, sets of possibilities, ways of dividing logical space, ways the world might have been, etc.). It points to a different idea about the semantic values of our target sentences, namely the idea that they correspond to constraints (not on the way the world is, but) on states of information. What we should want out of the compositional semantic values of sentences like (P1) and (P2) are just constraints on states of information, of the sort delivered by (a) and (b). Formally such constraints correspond to sets of information states—or sets of probability spaces, if that is how we elect to model states of information like credal states. This simple thought here is that with such sentences we can (*inter alia*) express aspects of our credal states, aspects which do not correspond in any straightforward way to a view about the way the world is.[25]

This gives us a new kind of object to design a compositional semantics around. Instead of possible worlds truth-conditions, we have constraints on states of information—or 'probability conditions', if states of information are probabilistically modelled. (Cf. the notion of a "probasition" in [14].) Now suppose we could compositionally associate sentences in general with something like probability conditions. Then consequence could be modeled as a relation between a set of sentences (premises) and a sentence (conclusion) which holds when the satisfaction of the constraints expressed by the premises suffices for the satisfaction of the constraint expressed by the conclusion, for any given information state. It would be a relation that preserves probability conditions. Such a model of consequence would predict the failure of MT we have pointed to. And it may even allow that although (C1) does not follow from (P1) and (P2), the three statements together are not jointly consistent (i.e., not jointly satisfied by any information state).

The central burden of this approach, of course, is to compositionally associate sentences in general with probability conditions, and to motivate that semantics over rival accounts on a broad array of data. See [35, 37, 38] and the appendix for a start at this burden.

---

[25]The *inter alia* here is important. I don't mean to suggest the *only* attitude states we can express with words like *probably* is (doxastic) credence. That is not the case. Various attitudes of acceptance may be probabilistically articulated—that is, be representable in terms of probability conditions. Credence is just a convenient, familiar, and important example.

Finally, what of the apparently non-truth-conditional character of knowledge ascriptions embedding epistemic modals like *probably*? We have just suggested that having a view about what is likely is not, or not merely, to have a view about which possibilities are still open; it is also a matter of how one distributes the probabilities over the possibilities one takes to be open. Given this, the same holds about having a view concerning whether some other agent *knows*, where what is at issue is whether he knows that something is likely. To agree that John knows that the marble is probably big is partly to have the view that the marble is probably big, which is itself not a view about a purely factual matter—not purely a matter of what kind of possibilities one rules out. And this yields a restricted form of nonfactualism about the state of knowing itself.[26]

Appendix

We describe two compositional semantic accounts for conditionals and probability operators, each of which will invalidate MT in the relevant cases when paired with the appropriate definition of consequence. Both semantic theories are defined over the same formal language $\mathcal{L}$:

Syntax of $\mathcal{L}$.   The primitive expressions include sentence letters p_1, p_2, ...; the one-place operators $\neg$, $\Diamond$ (*might*), $\triangle$ (*probably*); the two-place operators $\wedge$, $\rightarrow$; and parentheses. Sentence letters are well-formed. If ϕ and ψ are well-formed, so too are $\neg\phi$, $\Diamond\phi$, $\triangle\phi$, $(\phi \wedge \psi)$, and $(\phi \rightarrow \psi)$.

We take $\mathcal{L}$ to reflect the logical forms of the relevant fragment of English at the suitable level of abstraction. Both semantic theories are defined relative to models of the same kind:

Defs.   A **model** $\mathcal{M}$ for $\mathcal{L}$ is a pair $\langle \mathcal{W}, \mathcal{I} \rangle$ where $\mathcal{W}$ is a set of possible worlds, $\mathcal{I}$ is an **interpretation function** mapping the propositional letters of $\mathcal{L}$ to subsets of $\mathcal{W}$.

---

[26]Here I am indebted to Moss [24].

And both employ the notion of an *information state*:

Def. An **information state** i in $\mathcal{M}$ is a pair $\langle s, Pr \rangle$ of a set of worlds s (some subset of $\mathcal{W}_{\mathcal{M}}$; call it the **domain** of i) and a function Pr assigning the elements of some Boolean algebra of subsets of $\mathcal{W}$ a number in $[0, 1]$ satisfying the following:

1. Normalization. $Pr(s) = 1$
2. Additivity. $Pr(p \cup q) = Pr(p) + Pr(q)$, where p, q disjoint.

An information state is a probability space conditionalized on a primitively given set of possibilities. Semantically these will play the role of epistemic possibilities in each system below.

Probabilistic Static Semantics

Our first formal semantics builds upon [22, 35, 37]. It is a static account assigning truth values to sentences relative to *indices*:

Def. An **index** in $\mathcal{M}$ is any world-information state pair $w_{\mathcal{M}}, i_{\mathcal{M}}$.

The semantics takes the form of a recursive characterization of truth at an index for sentences of $\mathcal{L}$, as follows:

Def. For any $\mathcal{M}$, a **valuation** $[\![\ ]\!]^{w,i}$ for $\mathcal{M}$ is a function assigning either 0 or 1 to each wff relative to each index w, i in $\mathcal{M}$ subject to the following constraints, where α is any propositional letter, ϕ and ψ are any wffs:

$$
\begin{aligned}
[\![\alpha]\!]^{w,i} = 1 \quad &\text{iff} \quad w \in \mathcal{I}(\alpha) \\
[\![\neg\phi]\!]^{w,i} = 1 \quad &\text{iff} \quad [\![\phi]\!]^{w,i} \neq 1 \\
[\![\phi \wedge \psi]\!]^{w,i} = 1 \quad &\text{iff} \quad [\![\phi]\!]^{w,i} = 1 \text{ and } [\![\psi]\!]^{w,i} = 1 \\
[\![\Diamond\phi]\!]^{w,i} = 1 \quad &\text{iff} \quad s_i \cap [\![\phi]\!]^{i} \neq \emptyset \\
[\![\triangle\,\phi]\!]^{w,i} = 1 \quad &\text{iff} \quad Pr_i([\![\phi]\!]^{i}) > .5
\end{aligned}
$$

(Where $[\![\phi]\!]^{i}$ is an abbreviation for $\{w : [\![\phi]\!]^{w,i} = 1\}$.) Observe negation and conjunction have their classical semantics. To add the indicative conditional operator $\rightarrow$, two further definitions will be helpful:

Def. An information state i **accepts** ϕ iff $\forall w \in s_i : [\![\phi]\!]^{w,i} = 1$.

Acceptance is meant to track the intuitive idea of an information state incorporating the information associated with a sentence. We also want the idea of the *nearest* information state to a given information state accepting a sentence.

Def. The **nearest information state to i accepting ϕ**, $i + \phi$, is defined as follows:

$$
i + \phi =_{\text{DEF}} \langle s_i \cap [\![\phi]\!]^{i}, Pr_i^{\phi} \rangle, \text{ where}
$$

$$
Pr_i^{\phi}(x) =_{\text{DEF}} Pr_i(x | [\![\phi]\!]^{i})
$$

I assume the usual definition of conditional probability in terms of unconditional probability. With these definitions in place, we add the following clause to our recursive characterization of truth with respect to an index:

$$[\![\phi \to \psi]\!]^{w,i} = 1 \text{ iff } i + \phi \text{ accepts } \psi$$

Observe the connection between indicative conditionals and conditional probabilities this semantics supplies.

Finally, we define consequence over the semantics. As in [35], we understand the consequence relation $\vDash$ to preserve, not truth with respect to an index, but rather acceptance, as follows:

Def. $\phi_1, ..., \phi_n \vDash \psi$ iff no information state which accepts $\phi_1, ..., \phi_n$ fails to accept ψ.

Now it is easy to see why MT fails on this account. Considering our initial counterexample, our semantics associates the premises with the following acceptance conditions:

(P1) If the marble is big, then it's likely red.
$Pr_i(\texttt{red}|\texttt{big}) > .5$
(P2) The marble is not likely red.
$Pr_i(\texttt{red}) \leq .5$

These conditions can be satisfied simultaneous with the constraint corresponding to the negation of the MT conclusion:

(C1) The marble is not big.
$s_i \cap \texttt{big} \neq \emptyset$

Examples of information states with these properties are trivial to construct.[27]

For independent motivation of various aspects of the above semantics and further discussion, see [10, 15, 22, 25, 26, 35, 37, 38], .[28] For extensions to attitude verbs, see [26, 35, 38]. Anand and Hacquard [1] contains additional relevant discussion.

---

[27]For the sake of explicitness: consider i where $s_i = \{w_1, w_2, w_3\}$, $Pr_i(\{w_1\}) = .4$, $Pr_i(\{w_2\}) = .3$, $Pr_i(\{w_3\}) = .3$, only $w_1, w_3 \in \texttt{big}$, and only $w_1 \in \texttt{red}$.

[28]Note that on the above semantics, expressions with stacked epistemic modals (e.g., $\Diamond\Diamond\phi$) will generally be semantically equivalent to the corresponding expression with the most narrow modal ($\Diamond\phi$). Stacked epistemic modals often are interpreted as vacuous in this way, especially when the modals stacked are the same (a phenomenon called *modal concord*); other times, such stacking is just anomalous; still other times, such stacking does allow for coherent interpretations not equivalent to corresponding expression with the most narrow modal. The latter case is not provided for by the above semantics. In such cases I would be inclined to appeal to tacit shifting of the information state parameter, akin to free indirect discourse. See [35, 36] for some discussion. The issue is beyond the scope of this appendix.

Probabilistic Dynamic Semantics

Our second formal semantics essentially extends the semantics of [31] (see [2] for a nice overview), incorporating ideas described in [34, 35, 37]; see [40] for detailed discussion. First, rather than taking sentential semantic values to be functions from points of evaluations to truth values, we take them to be functions from information states to information states (*update functions*, or *context change potentials*), joining the dynamic tradition going back to [12]. Second, rather than modeling information states as sets of possible worlds (as in Veltman's update semantics) or as files (as in Heim's file change semantics), we continue to take them to be probability spaces as defined above. Sentences are recursively associated with update functions as follows:[29]

Def. An **dynamic valuation** $[\cdot]$ is a function from wffs of $\mathcal{L}$ to functions from information states to information states subject to the following constraints, where α is any propositional letter, ϕ and ψ are any wffs:

$$i[\alpha] = \langle s_i \cap \mathcal{I}(\alpha),\ Pr_i(x|s_i \cap \mathcal{I}(\alpha)) \rangle$$

$$i[\neg\phi] = \langle s_i - s_{i[\phi]},\ Pr_i(x|s_i - s_{i[\phi]}) \rangle$$

$$i[\phi \wedge \psi] = i[\phi][\psi]$$

$$i[\Diamond\phi] = i \text{ iff } s_{i[\phi]} \neq 0,\ \text{else } \langle \emptyset,\ Pr(x|\emptyset) \rangle$$

$$i[\triangle\phi] = i \text{ iff } Pr(s_{i[\phi]}) > .5,\ \text{else } \langle \emptyset,\ Pr(x|\emptyset) \rangle$$

$$i[\phi \rightarrow \psi] = i \text{ iff } i[\phi] = i[\phi][\psi],\ \text{else } \langle \emptyset,\ Pr(x|\emptyset) \rangle$$

The semantics for negation and conjunction here in essence go back to [12, 13]. The idea for the treatment of epistemic possibility modals here goes back at least to [31]. The conditional semantics here is a probabilistic analogue of the semantics of [9].[30] Note the tight connection between indicative conditionals and the corresponding conditional probabilities built into this account.

Consequence may be defined in various ways over this semantics (see [2, 31]). A relatively conservative choice, and one adequate for our purposes, would be the following:

Def. $\phi_1, ..., \phi_n \vDash_d \psi$ iff for all i, if i is a fixed point of each of $[\phi_1], ..., [\phi_n]$, it is also a fixed point of $[\psi]$.

---

[29]I an indebted to Justin Bledin here for suggestions which lead to a considerable simplification of the following semantics.

[30]An alternative clause for the conditional, in the spirit of [13], would be the following: $i[\phi \rightarrow \psi] = \langle s_i - (s_{i[\phi]} - s_{i[\phi][\psi]}),\ Pr_i(x|s_i - (s_{i[\phi]} - s_{i[\phi][\psi]})) \rangle$. This entry has attractions in the case of conditionals not containing modals. In cases of modalized consequents, however, the semantics would lead to anomalous results. For instance, for any i such that $i[\phi] \neq \langle \emptyset, Pr(x|\emptyset) \rangle$ but $i[\phi][\triangle\psi] = \langle \emptyset, Pr(x|\emptyset) \rangle$, the update of $i[\phi \rightarrow \psi]$ would equal that of $i[\neg\phi]$.

Now it is again routine to observe the failure of MT in the context of this semantics. Consider again any information state i satisfying the following properties:[31]

$$Pr_i(\texttt{red}|\texttt{big}) > .5$$

$$Pr_i(\texttt{red}) \leq .5$$

$$s_i \cap \texttt{big} \neq \emptyset$$

Thanks to the first property, it follows that $i[\texttt{big}] = i[\texttt{big}][\texttt{red}]$; so i is a fixed point of $[\texttt{big} \rightarrow \triangle\texttt{red}]$ (i.e., (P1)). Thanks to the second property, i is a fixed point of $[\neg\triangle\texttt{red}]$ (i.e., (P2)). But thanks to the third property, $i[\neg\texttt{big}] \neq i$, so i is not a fixed point of $[\neg\texttt{big}]$ (i.e., (C1)). So (P1), (P2) $\nvDash_d$ (C2).

Ultimately, the failure of MT here is explained in the same fashion as in the static semantics. An information state which accepts the premises need not accept the conclusion.

The dynamic semantics presented here presupposes a thesis I have elsewhere [34, 35] called *context probabilism*:

Context probabilism: the common ground of a conversation is characterizable as a probability space, or as a set of such spaces.

(The static semantics above also naturally lends itself to this thesis, given the definition of consequence assumed there.) Now the question arises, should a state of presupposition be modeled as a probability space (an information state), or a set of such spaces? If we stick to a single probability measure, then if states of presupposition are assumed to be perfectly coordinated in the ideal case, the problem immediately arises how agents can be expected to coordinate on the probabilities of myriad propositions that go undiscussed in context. It seems implausible that propositions which are completely open in conversation must be assigned precise probabilities; among other things, these probabilities would reflect nothing about how we are conversationally coordinated.

We can escape this problem by supposing that states of presupposition, hence informational contexts, are representable by sets of probability spaces (sets of information states). A conversation which presupposes nothing about a proposition p is one which leaves open every possible way of associating a probability with it. That is to say, for every probability n, there will be an information state i in the common ground such that $Pr_i(p) = n$. On this approach, the common ground is representable as a probability condition. Then we can define on update on sets of information states I in a manner parasitic on our recursive characterization of updates for single information states, as follows:[32]

For any set of information states I, $I[\phi] = \{i \in I : i[\phi] = i\}$.

---

[31]I will be harmlessly loose about use and mention for expressions in teletype.
[32]I am inspired by a structurally analogous idea developed by [33].

This has the additional advantage of supplying a more substantive account of the conversational dynamic change initiated by the epistemic language for which we have supplied a 'test'-like semantics (i.e., epistemic modals, probability operators, and indicative conditionals). Appeal to sets of probability spaces may also be ultimately needed to handle disjunction; see [26].

As goes the state of presupposition, so too, I think, should go other acceptance-like attitudes—for example, supposition, belief, and knowledge. We can assume that these states too determine sets of probability spaces. For example, suppose to any agent x and possible world w there corresponds the set of probability spaces $\mathcal{B}_x^w$ which reflect that agent's state of belief—an idea familiar from the formal epistemology literature. Then we can extend our semantics to belief ascriptions as follows:

$$i\,[B_x\phi] = \left\langle \left\{ w \in s_i : \mathcal{B}_x^w[\phi] = \mathcal{B}_x^w \right\}, Pr_i \left(x | \left\{ w \in s_i : \mathcal{B}_x^w[\phi] = \mathcal{B}_x^w \right\} \right) \right\rangle$$

(Compare [26], who develops a similar idea in different semantic setting.)

It is easy to imagine an analogous semantics for knowledge ascriptions. There we mainly need to incorporate the fact that knowledge ascriptions generally presuppose their complements. Suppose, in the spirit of [2], we introduce an artificial presupposition operator ∂:

$$i[\partial\phi] = i \text{ iff } i[\phi] = i; \text{ otherwise undefined.}$$

Letting an agent x's state of knowledge in w be the set of information states $\mathcal{K}_x^w$, we can give dynamic semantics for knowledge ascriptions as follows:[33]

$$i\,[K_x\phi] = \left\langle \left\{ w \in s_{i[\partial\phi]} : \mathcal{K}_x^w[\phi] = \mathcal{K}_x^w \right\}, Pr_i \left(x | \left\{ w \in s_i : \mathcal{K}_x^w[\phi] = \mathcal{K}_x^w \right\} \right) \right\rangle$$

This aims to capture presupposition projection via the percolation of undefinedness in the calculation of a sentence's context change potential, as in [12]. Knowledge ascriptions will go undefined with respect to contexts which do not already accept ϕ, the desired result.

On what seems to me a perfectly natural interpretation, these formal proposals about what it is to accept and assert probabilistic sentences, whether developed statically or dynamically, yield a kind of expressivism about this fragment of language. If we restrict attention to those cases where the language is used to express or describe doxastic states, we could call the view *credal expressivism*. Again, to be in a state of mind which accepts a sentence of the form $\triangle\phi$ is generally not, on this picture, merely a matter of representing the world as being some particular way—not merely a matter of ruling some possibilities in and others out. It is a matter also of how one distributes the probabilities over the possibilities.

Once we see this way of formalizing the idea of expressing credence, another possibility—the idea of expressing preference, or utility—comes into view. Like states of credence, states of preference are not states that merely rule some possibilities in and others out. Just as the preceding story shows how we

---

[33]I am indebted here to conversations with Julien Dutant.

might use language to coordinate on something other than a way of representing how the world is, so we can imagine an analogous story about the language by which we express preference and utility. As epistemically modalized and probabilistic sentences correspond to conditions on information states, so too might we describe deontically modalized sentences as expressing conditions on the allocations of utility one's state of mind leaves open. As we added probabilistic structure to the common ground, so too might we add utility-theoretic structure. If that worked out, epistemic modals and deontic modals would be, in a sense, the Bayesian modalities, and we would have what we could call *Bayesian expressivism*. I investigate the possibility of this view elsewhere [39].

References

1. Anand, P., & Hacquard, V. (2008). Attitudes and epistemics. In T. Friedman, & S. Ito (Eds.), *Proceedings of SALT XVIII* (pp. 37–54). Eau Claire: Cornell.
2. Beaver, D. (2001). *Presupposition and assertion in dynamic semantics. Studies in logic, language and information*. Stanford: CSLI Publications.
3. Bennett, J. (2003). *A philosophical guide to conditionals*. Oxford: Oxford University Press.
4. Cantwell, J. (2008). Changing the modal context. *Theoria, 74*(4), 331–351.
5. Carroll, L. (1894). A logical paradox. *Mind, 3*(11), 436–440.
6. Edgington, D. (1995). On conditionals. *Mind, 104*(414), 235–329.
7. Forrester, J. W. (1984). Gentle murder, or the adverbial samaritan. *The Journal of Philosophy, 81*(4), 193–197.
8. Gibbard, A. (1981). Two recent theories of conditionals. In W. Harper, R. Stalnaker, & G. Pearce (Eds.), *Ifs* (pp. 211–47). Dordrecht: D. Reidel.
9. Gillies, A. (2004). Epistemic conditionals and conditional epistemics. *Nous, 38*, 585–616.
10. Gillies, A. (2010). Iffiness. *Semantics and Pragmatics, 3*, 1–42.
11. Groenendijk, J., & Stokhof, M. (1991). Dynamic predicate logic. *Linguistics and Philosophy, 14*, 39–100.
12. Heim, I. (1982). *The semantics of definite and indefinite noun phrases*. PhD thesis, University of Massachusetts.
13. Heim, I. (1983). On the projection problem for presuppositions. In M. Barlow, D. Flickinger, & M. Wescoat (Eds.), *WCCFL 2: Second annual west coat conference on formal linguistics* (pp. 114–125). Stanford: Stanford University.
14. Jeffrey, R. (1983). Bayesianism with a human face. In J. Earman (Ed.), *Testing scientific theories. Minnesota studies in the philosophy of science* (Vol. 10, pp. 133–156).
15. Kolodny, N., & MacFarlane, J. (2010). Ifs and oughts. *Journal of Philosophy, CVII*(3), 115–143.
16. Kratzer, A. (1981). The notional category of modality. In H.-J. Eikmeyer, & H. Rieser (Eds.), *Words, worlds, and contexts, new approaches to word semantics* (pp. 38-74). Berlin: Walter de Gruyter.
17. Kratzer, A. (1983). Conditionals. In A. Farley, P. Farley, & K.-E. McCullough (Eds.), *Papers from the parasession on pragmatics and grammatical theory*. Chicago: Chicago Linguistics Society.
18. Kratzer, A. (1991). Modality. In A. von Stechow, & D. Wunderlich (Eds.), *Semantics: An international handbook of contemporary research* (pp. 639–650). Berlin: Walter de Gruyter.
19. Lewis, D. K. (1973). *Counterfactuals*. Malden: Blackwell.
20. Lewis, D. K. (1975). Adverbs of quantification. In E. L. Keenan (Ed.), *Formal semantics of natural language* (pp. 3–15). Cambridge: Cambridge University Press.
21. Lewis, D. K. (1976). Probabilities of conditionals and conditional probabilities. *Philosophical Review, 85*, 297–315.

22. MacFarlane, J. (2011). Epistemic modals are assessment-sensitive. In A. Egan, & B. Weatherson (Eds.), *Epistemic modality* (pp. 144–178). Oxford: Oxford University Press.
23. McGee, V. (1985). A counterexample to Modus Ponens. *Journal of Philosophy, 82*(9), 462–471.
24. Moss, S. (2012). Epistemology formalized. *Philosophical Review* (forthcoming).
25. Rothschild, D. (2011). Do indicative conditionals express propositions? *Nous*. doi:10.1111/j.1468-0068.2010.00825.x
26. Rothschild, D. (2012). Expressing credences. In *Proceedings of the Aristotelian Society* (forthcoming).
27. Stalnaker, R. (1968). A theory of conditionals. In N. Rescher (Ed.), *Studies in logical theory* (pp. 98–112). Oxford: Basil Blackwell.
28. Stalnaker, R. (1973). Indicative conditionals. *Philosophia, 3*, 269–286.
29. Van Fraassen, B. (1976). Probabilities of conditionals. *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science, 1*, 261–308.
30. Veltman, F. (1985). *Logics for conditionals*. PhD thesis, University of Amsterdam.
31. Veltman, F. (1996). Defaults in update semantics. *Journal of Philosophical Logic, 25*(3), 221–261.
32. von Fintel, K., & Iatridou, S. (2003). Epistemic containment. *Linguistic Inquiry, 34*(2), 173–198.
33. Willer, M. (2010). *Modality in flux*. PhD thesis, University of Texas, Austin.
34. Yalcin, S. (2005). Epistemic modals. In J. Gajewski, V. Haquard, B. Nickel, & S. Yalcin (Eds.), *New work on modality* (pp. 231–272). MIT Working Papers in Linguistics (Vol. 51).
35. Yalcin, S. (2007). Epistemic modals. *Mind, 116*(464), 983–1026.
36. Yalcin, S. (2009). More on epistemic modals. *Mind, 118*(471), 785.
37. Yalcin, S. (2010). Probability operators. *Philosophy Compass, 5*(11), 916–937.
38. Yalcin, S. (2011). Nonfactualism about epistemic modality. In A. Egan, & B. Weatherson (Eds.), *Epistemic modality* (pp. 295–332). Oxford: Oxford University Press.
39. Yalcin, S. (2012a). Bayesian Expressivism. In *Proceedings of the Aristotelian Society* (forthcoming).
40. Yalcin, S. (2012b). Context probabilism. In M. Aloni, V. Kimmelman, F. Roelofsen, G. W. Sassoon, K. Schulz, & M. Westera (Eds.), *Proceedings of the 18th Amsterdam Colloquium, FoLLI, lecture notes in computer science*. Springer. (forthcoming).