Connecting Conditionals and Conditional Probability: A Sketchy Primer

Daniel Rothschild

March 24, 2009

There are a lot of things to read on conditionals and conditional probability (Lewis, 1976; Adams, 1975; Edgington, 1995; Bennett, 2003). However, very few pieces explore the relationship between Kratzer's theory of conditionals and the standard discussions of conditionals and conditional probabilities. Nor, for obvious reasons, are there any discussions of the relationship between Kratzer's account of conditionals and the recent work of Swanson and Yalcin on epistemic modals and conditionals. Unfortunately this primer is very sketchy at this stage.

1 Conditional Probability

You probably remember that the conditional probability of A given B, written $P(A|B)$, is defined, when $P(B) > 0$ as $\frac{P(A\&B)}{P(B)}$. If this doesn't make sense to you, then you probably need to read a textbook on probability. But, briefly, the idea is that the credence one should put in A given B is just the ratio of the credence you put on A and B both happening to the credence you put on B happening. Making this a *definition* doesn't mean you can only *know* the conditional probability if you know these two unconditional probabilities, so using this definition doesn't assert a epistemic or conceptual priority of unconditional probability over conditional probability, but rather just states, in the form of a definition, the relationship between them (some might disagree with me here, but whatever). Alan Hajek argues we should also discuss the conditional probability of A given B when $P(B) = 0$. I think doing so would only complicate matters so I'll just assume, as most non-specialists do, that conditional probability is undefined in these circumstances.

2 Conditionals and their Relation to Conditional Probability

We'll just discuss what people call indicative conditionals which basically are conditionals that don't have any strange tense or mood in them. So, as should be familiar, (1-a) is an indicative conditional and (1-b) is not.[1]

(1) a. If Oswald didn't shoot Kennedy, someone else did.

b. If Oswald hadn't shot Kennedy, someone else would have.

Let me put a seemingly harmless thesis on the table: indicative conditionals, such as (1-a), like most declarative sentences, are capable of expressing claims that we can have more or less credence in. Let's unpack this by discussing credence in non-conditional statements.

(2) The cat is on the table.

An statement such as (2), uttered in a context which resolves its ambiguities, expresses a statement which we might judge to be more or less likely. Such judgements, we will say, are ways of assigning credence to the statement. For instance, you might be 90% sure that (2) is true. Or you might have some less determinate credence. But let's just simplify and assume that your credences come in the form of real numbers between 1 and 0, i.e. they are probabilities. Now, it seems like your degree of credence in (2) *is* just your degree of credence in the fact that the cat is on the table.

Let's discuss your degree of credence in complex sentences:

(3) John is a communist and Michael is an anarchist.

Obviously this is a conjunction of two other sentences:

(4) a. John is a communist.

b. Michael is an anarchist

[1]Note that (1-b) seems to have a sort of uninterpreted extra past tense in the antecedent as well as a modal in the consequent. For discussion of why this might be so see Iatridou (2000).

Now, one's credence in (3) depends, it seems, in some way, on one's credences in (4-a) and in (4-b). On way of looking at it, it seems that one's credence in (3) is just one's credence in (4-a) times one's conditional credence in (4-b) given (4-a). This is of course, just an application of the definition of conditional probability above. If (4-a) and (4-b) are probabilistically independent, then this is just the product of the probability of (4-a) and the probability of (4-b).

Now, what about conditionals like this:

(5) If John is a communist, then Michael is an anarchist.

What is one's credence in (5) in terms of one's credences in the parts? Well... what's standardly said is that it's simply one's conditional probability of (4-b) given (4-a). Why? Well, it seems like you believe (5) to the extent that you think that it likely that Mike is an anarchist given that John is a communist. In the general case, expressing the the conditional by means of the operator $\rightarrow$ we say that:

$$P(A \rightarrow B) = P(A|B)$$

This is what is often called "the equation."

There are a variety of reasons to think the equation is basically correct. One is that using conditionals is often how we talk about conditional probabilities: "The probability of B given A"can be easily rephrased as "the probability if A, of B".

More motivation can be found than just the language we use to discuss conditional probabilities. Consider a sentence like this:

(6) It's likely that if John is a communist then Michael is an anarchist.

This seems to be true iff the conditional probability of (4-b) given (4-a) is greater than .5. The obvious explanation of this is that the probability of the embedded sentence in (6), which is just (5), is given by the equation.

3 Triviality

People get very excited about the equation. The main reason is that if you accept the equation, there are variety of arguments that conditionals do not, properly speaking, express propositions.

The start of these arguments is a famous paper by David Lewis (1976), which contains a couple versions of his so-called triviality theorems. All the proofs of triviality theorems that I know of share the same basic features:

- A language L formed from a stock of atomic propositions, a conditional operator, $\rightarrow$, and the usual Boolean operators.

- A set of probability function, P over all sentences in L that obeys the usual axioms of probability theory.

- The equation: for any given sentences θ and ϕ in L,

$$P(\theta \rightarrow \phi) = P(\phi|\theta)$$

Lewis originally proved the following: Suppose a probability function, P is such that it satisfies all the conditions above and is non-trivial (which means that there are at least three statements with distinct probabilities). Then consider the class of probability functions related to P by conditionalization on arbitrary formulas in L. It cannot be the case that all the members of this class also satisfy the conditions above.

Let me just sample one of the other results in this vein: Hajek proved that any non-trivial probability function that has a maximum number d such that for any sentence, ϕ in L, $P(\phi) > d$ cannot satisfy all the equation.

There are other theorems showing that various probability functions that have certain features cannot satisfy the equation.

4 Assigning Probabilities to Conditionals

Maybe we should be a bit more explicit about how we typically understand probability theory. On the standard, measure-theoretic way of thinking about probability there are a number (perhaps an infinite number) of different possibilities (let's call them worlds) and to each possibility we assign a number–or a function (satisfying certain conditions–that assigns positive numbers to sets of worlds, namely a probability *density* function).[2] This number corresponds to how likely that individual world (or set of worlds) is: in other words probabilities are a kind of measure. The likelihood of the set of all worlds is, of course, 1 (this is one of the conditions one puts on the measure). The probability of any given statement is just the total likelihood of all the worlds that it is true in, and we thus assume that we can correspond to each "statement" a set of possible worlds.[3]

Think about the conditional statement **if** A **then** B. I guess I have some theoretical intuitions here that maybe are a bit too crazy to rely on, but I take it that when we talk about the likelihood of **if** A **then** B (i.e. how likely is it that if A is true B is true) we're simply *talking* about the ratio of the measure of the worlds where A and B is true to the worlds where A is true, and hence we're not talking about the $\neg A$ worlds at all. If someone denied that, I guess I'd think they didn't really understand conditional probability or probability generally. (I'm a little hesitant to say this, since Stalnaker perhaps doesn't share this intuition, and I think Stalnaker must understand (much better than I) probabilities and conditional probabilities.)

The dialectic of the triviality theorems mentioned above is that you assume that **if** A **then** B expresses a statement which is true or false in every possibility (world). That's just what it means for **if** A **then** B to express a proposition in the language of probability theory thought of in the measure-theoretic way. It's natural to think that **if** A **then** B is going to be true in all the A and B worlds and false in all the A and $\neg B$ worlds. Now the problem is that the probability of the statement **if** A **then** B is meant to be the total measure of *all* the worlds in which it's true. So what we now need to ensure is that somehow that's equal to the ratio of $A \wedge B$ worlds to A worlds. How do we do that? Well, we just need to somehow assign **if** A **then** B truth values in

[2] Actually we don't need to assign probabilities to all sets of possible worlds but just to a sigma algebra of them.

[3] Some discussions of probability that they do not view probability this way. Instead, they think of the Kolmorogov axioms as being about a logical language (which doesn't necessarily have a set-theoretic interpretation). However, I think that the measure-theoretic way of thinking about probability is more natural and intuitive than the logical way.

$\neg A$ worlds in a way that *maintains* the ratio of worlds where it's true to worlds where it's false. Now what all the triviality theorems are doing is proving (one way or another) that doing this has a lot of fucked-up consequences.

But this whole way of thinking about conditionals and probability departs from the central intuition I just mentioned, i.e. that the probability of **if** A **then** B simply doesn't haven anything to do with the not-A worlds. We can, modulo the triviality results, of course, *pretend* that when we evaluate the probability of **if** A **then** B we are looking at all the worlds in the normal sense and not just the A worlds, and rig the system so that we always get the same results as just looking at the A worlds would do, but this seems to go against what I take to be the basic intuition propelling the equation anyway. So, if you think that the only thing determining the probability of **if** A **then** B is the conditional probability of B given A then you probably think that **if** A **then** B doesn't have truth values at every possible world.

I realize, however, these remarks might seem unconvincing if you don't share these intuitions, but I'll just put them out for those who do. Let me clarify one point: I'm not claiming that we think that **if** A **then** B is only true or false when A is true. There are plenty of decent examples that suggest this is not right, so certainly we can't rely on it as an intuition. What I am claiming, though, is that when we evaluate the probability of **if** A **then** B we are simply not interested in the not-A worlds. I guess you could say that's because we know that whatever the probability of **if** A **then** B is we know we can calculate it just by finding $P(B|A)$. This seems like a cop-out, though: the default view should be that we can calculate it this way because it only *depends* on what's happening in the the A worlds: not because there's some cosmic law relating the A worlds and the $\neg A$ worlds with respect to the truth of **if** A **then** B. This isn't a knock-down argument, but it's a way of pointing out just how *perverse* the attempt to treat **if** A **then** B as a sentence in the language of probability is. The triviality theorems are ways of drawing consequences from this perversity that make manifest its theoretical costs.

Another way of thinking about it is that the "formula" (or rather string of three letters) "$A|B$" which appears in "$P(A|B)$" has a fundamentally different status in probability theory then, say, the formula "$A \wedge B$" as it appears in "$P(A \wedge B)$". $A|B$ isn't a sentence that's true or not in every world. If we think that $P(A \rightarrow B) = P(B|A) = \frac{P(A \wedge B)}{P(A)}$ (when defined) then the most natural thing to think is that we are in some way treating **if** A **then** B as something akin to $A|B$ which is, as I've said, something that *doesn't have an interpretation by itself in the probability calculus*. So, from the perspective of probability theory it doesn't express a proposition. As I said this is just the most *natural* way of thinking about conditionals from the perspective of probability theory. It's not the only way, but the triviality theorems suggest that if you take seriously the equation you're going to be pushed in this direction. I want to reserve judgment, for the moment, on how much this amounts to, in general, treating conditionals as not expressing propositions: the degree to which it does really depends on difficult issues about how you understand probability and its relation to natural language semantics generally.

5 Explaining Conditional Probability/Conditional Connection

As I've suggested, I think asking whether conditionals express propositions is perhaps not the most productive line of inquiry. In the system of probability theory, they don't seem to, but maybe that's just an illusion of one sort or another. We can still asks some interesting questions though: for instance, how is it that in natural language a probability operator can connect with a sentence and not directly assess the probability of the sentence in the normal way:

(7) It's likely that if A then B.

This *looks* to be of the form "it's likely that p" where $p =$ **if** A **then** B. But our usual way of understand things of *that* form is that we take them to be talking about the probability simpliciter of p, but we've just said that a sentence like **if** A **then** B don't have probabilities in the usual sense since they're not statements over which probabilities are defined. So what the fuck?

This is just a way of asking how a semantics for the natural language words **if** and **likely** can be such that we treat conditionals underneath "likely" as behaving differently from normal statements underneath likely.

I guess before we get into the details of this problem, we should say something about what **it's likely that ...** means. As it happens, addressing this problem raises some of the same worries with respect to propositionally that the question of what **if** A **then** B raises! For a statement that the probability of A is .6 is not itself a statement that we can easily think of as being in the probability calculus, so it's not propositional from the perspective of probability theory. Of course, if we took **it's likely that** P just to report the epistemic state of some person (or other source of credences/objective probabilities) then we could understand it as being a statement about the world. I'll call such approaches *ascriptional* takes on probability modals. As we'll see there's another view, what I'll call the MIT-view, which is more radical. But the point for now is that the same issues about propositionally arise in understanding the semantics of probability modals as those that arise from the equation.

Let's put all that stuff aside. **It's likely that . . .** , at the least, expresses some sort of probability of what happens in the dots. The question is why, when we have a conditional, it doesn't do it in the *usual* way.

I'll discuss a few different sorts of views here, especially Kratzer's view, but for now let me discuss a rather intuitive answer.

Let's suppose that **if** A **then** B really means something like every A world is a B world. This is at least plausible for non-habitual conditionals of various sorts. The reason this is plausible is that if you accept **if** A **then** B then you've got to accept (about all the worlds you think are possible) that any world in which A is true is a one in which B is true (and nothing more). Of course this is a statement about the whole set of worlds, not about individual worlds. But nonetheless we can think of it as having a *derivative* truth value in certain worlds, in particular it's clearly false (in a certain sense) in $A \wedge \neg B$ worlds and true in $A \wedge B$ worlds. Well, anyway, that's one way of thinking about it. So what happens when we put it under a **it's likely that . . .** operator? What we're coming close to is the view on which **if** A **then** B is only *defined* in worlds where A is true. In this case calculating probability in the normal way is not really possible, because we don't know what to do about worlds in which the statement is undefined. A natural fix would be to suppose that we just re-calibrate our probability function to only look at the worlds where the statement in question *is* defined. In that case the probability of **if** A **then** B will equal $P(B|A)$.[4] If we want to go this way, it might be nice to find some independent motivation for this treatment

[4]I believe this suggestion is mentioned by Lewis (1975), in a slightly different context, as well as Kai von Fintel—for exactly this purpose—in some unpublished lecture notes here.

of the interaction between probability and undefinedness. I won't consider this view any further, though, as I'm not really sure what to make of it.

6 Kratzer

In this context, the importance of Kratzer's work on conditionals shows itself particularly clearly (Kratzer, 1981, 1986). Kratzer essentially extends David Lewis's idea that the function of conditionals is to restrict quantifiers (Lewis, 1975). Her critical claim, in which she differs from Lewis, is that this is what *all* instances of if-clauses serve to do, not just the special instances where there are overt (or implicit) adverbs of quantification.

How does Kratzer's work apply to the relationship between conditionals and conditional probability? Well, recall the question we were left with in the last section: why does the interpretation of **it's likely that if** A **then** B not just involve finding a set of worlds where **if** A **then** B is true and then trying to determine how likely those worlds are. The answer Kratzer can give is that we have misunderstood the *syntax* of the sentence. What **if** A serves to do is restrict some quantifier. In this case, it is the probabilistic epistemic quantifier **it's likely**. So we can treat the sentence as having a logical form a bit like this:

(8) It's likely $[A][B]$

Here we understand **it's likely** to be a generalized quantifier expressing a relation between two sets (with measures on them). **It's likely** $[X][Y]$, generally speaking is true iff the measure of the worlds in $X \wedge Y$ divided by the measure of all the Y worlds is greater than 1/2. Generally speaking, the restrictor of **it's likely** is taken to be something like all the worlds in whatever probability space one is talking about (maybe the default is our knowledge), but in this case we restrict that further to all those worlds in the probability space that are X worlds. Note, too, that this account is independently motivated, since viewing conditionals as restrictors on quantifiers is useful for other purposes, such as the analysis of adverbs of quantification.

There are several issues, however, that this analysis raises. The most pressing question (to my mind) is how we treat instances where there is no probability operator in the sentence with the conditional. Consider just the plain conditional **if** A **then** B. We convinced ourselves before that we often think of the probability of this as being the conditional probability, $P(B|A)$. However, when we merely judge a sentence to have a certain probability we do not need to entertain another *sentence* that has an actual syntactically real probability operator that the if-clause in the conditional restricts. So there is no trivial way of extending Kratzer's analysis of conditionals to cover judgments of the probability of sentences with normal indicative conditionals. In other words, even if Kratzer's account yields a neat account of **it's likely that if** A **then** B it doesn't immediately yield an explanation of why the plain judgments of probabilities of conditionals that we normally make follow the equation.

What sort of further work will do? There are really two different approaches: one is to argue that conditionals serve to restrict operators in judgments of probability, as well as in sentences involving probability operators. This is a radical, extra-syntactic construal of Kratzer's hypothesis that conditionals act as restrictors. Not only does **if** A serve to restrict higher-up operators, such as **it's likely**, in a sentence , but, in fact, when we *think* about the probability of a statement like **if** A **then** B, then **if** A serves to restrict the domain which we think about the probability in!

Another, less radical approach is to bite the bullet and admit that Kratzer's approach only explains the relationship between conditionals and conditional probability in certain, special cases. This, then, requires us to have other stories to explain the extra cases. In many senses, I think this is a deeply unsatisfying way of going, but it might at least be willing to pursue what these supplementary stories are. What does Kratzer take an unembedded indicative conditional sentence to express? Her idea is that there are silent epistemic modal operators in conditionals and *these* are what the conditional restricts. So a sentence of the form **if** A **then** B should really be understood as **must**$[A][B]$ where **must** is the epistemic necessity modal. In particular she thinks **must**$[A][B]$ is true generally speaking iff all the A worlds compatible with one's knowledge are also B worlds. Where there is no explicit restrictor, of course, **must** P is true iff P is true in all worlds compatible with one's knowledge. I say "one's knowledge" but for Kratzer it is the speaker's knowledge. However, the literature since Kratzer has provided many seeming counterexamples to this claim (e.g. von Fintel and Gillies, 2005), so I won't burden her account with it.

The problem left then is to explain how a sentence that means *that* (i.e. what we've just said about epistemic modals) could possibly be such that its probability is equal to the conditional probability of the consequent given the antecedent. This task may seem impossible, but actually I think it's not that hard: Let's suppose you don't know whether or not A is true, but you think there's some source of knowledge that can tell you whether or not if A happens B will happen also. You could then take the knowledge basis of that source of knowledge, call it s, to be what the epistemic modal is "about". If that's what's going on then your credence in the statement $\mathbf{must}_s[A][B]$ should be the same in your conditional probability, $P(B|A)$.[5] So it turns out that Kratzer, if she is allowed to do a little fiddling with the knowledge source for epistemic modals, might have something to say about the judgments of the probabilities of conditionals without invoking her syntactic mechanism.

7 Variations: If-Clauses as Parameter Shifters

What does the relation between conditionals and conditional probability tell us? Let's take a sentence like this:

(9) It's likely that if A then B.

Let's suppose, that the real structure of the sentence is rather this:

(10) If A, then it's likely that B.

Unlike Kratzer then we no longer understand if A as specifically restricting the modal but rather as being an operator on the entire sentence. What sort of operator could it be? Well, what sort of operator it is, is going to need to depend on what we understand **it's likely** to do, since if A is going to have to do something to effect the interpretation of the embedded **it's likely**.

Let's suppose, following Yalcin (2007), that probability modals are given a meaning only relative to an information parameter i, where the information parameter specifies, at the minimum, a set

[5] If you can't see why this is so, you should probably just think about it a bit harder, it took me a bit of time to convince myself. Note that nothing guarantees that such knowledge sources really exist! If the world is indeterministic, in fact, such sources shouldn't always exist, but it's not clear to me that that's any argument that we don't implicitly posit them in making certain judgments.

of credences over possible worlds, i.e. a probability function, p_i. This parameter can be thought of as playing a semantic role akin to the standard world or time parameters in intensional semantics: it is set by default by context (in one way or another) but can be "shifted" by operators.[6]

So, let's suppose any given sentence S has a truth value relative to an information parameter i as well as a world w. In other words, $[\![\mathbf{S}]\!]^{i,w}$ specifies a truth value. Now we can define the meaning of "it's likely that S": $[\![\textbf{it's likely that } S]\!]^{i,w}$ is true iff the probability function determined by i, p, satisfies the condition that $p(w' : [\![\mathbf{S}]\!]^{i,w'}\text{is true}) > 1/2$.

Now we can understand **if** A as an operator that shifts the information parameter, i. Specifically, $[\![\textbf{if } A\textbf{, } S]\!]^{i,t}$ is true iff $[\![\mathbf{S}]\!]^{i',w}$ is true where i' is the information parameter obtained from i by probabilistically conditionalizing on A. This will then give us the requisite reading of (9).

Of course, most *normal* sentences without probability operators or epistemic modals won't be sensitive to the information parameter (e.g. $[\![\textbf{it's raining}]\!]^{i,w}$ is true iff it's raining in w). So what do we do with garden variety indicative conditionals like this:

(11) If it's cold then it's raining.

If the consequent is just **it's raining** then shifting the information parameter will have no effect on the truth of (11), so it will just be true or false when **it's raining** is true or false. To rectify this situation we can just take a leaf from Kratzer and assume that this sentence contains a silent epistemic modal, so it's real structure is as follows:

(12) If it's cold, it must be raining.

A simple semantics for **must** will be as follows (following Yalcin again) $[\![\textbf{must } S]\!]^{i,w}$ is true iff for every w' s.t. $p(w') > 0, [\![\mathbf{S}]\!]^{i,w'}$ is true.[7] Adding this silent **must** makes the conditional operator's shifting of the information parameter have an effect, which is what we need. Indeed, the effect is essentially equivalent to restricting the epistemic modal, which is just what happens on Kratzer's

[6]Briefly we assume a "double-indexing" system where there are both contextual parameters that cannot be shifted by operators (what, for example generally determines the meaning of "now") as well as indices that can be shifted by operators–such as a world index that's shifted by modal operators. See, e.g., Lewis (1980).

[7]Yalcin's actual treatment of conditionals doesn't involve a silent epistemic necessity modal, and thus is less Kratzerian, but I think it is equivalent.

analysis.[8]

So this analysis shares the same basic strengths and weaknesses of Kratzer's analysis. Again, we have the problem of understanding why our judgments of the probability of sentences like (11) seem to obey the equation. Again, there are two basic routes to explaining this: on the one hand we can say that when we judge the probability of sentences with if-clauses we allow the if-clause to restrict our probability judgment. We will also have to say that, somehow, the judgment of probability "erases" the epistemic necessity modal that we put in to get (11) to make sense. This is not an impossible story to tell, but it will certainly take some work to motivate it. The other way of going, again, is to say that when we judge the probability of (11) we take the knowledge source controlling the epistemic modal to be one other than our own. We can then again (by the reasoning in the previous section) try to explain why we would then judge the probability of this sentence to be just the conditional probability of it raining given that it's cold.[9]

8 Rethinking Semantic Values

Another way of looking at the lesson about conditionals and conditional probability is to think that what it shows is that sentences do not convey first-order information about the language that a probability calculus is defined over, but rather information about the probabilities themselves. Let's suppose we have a language L with a probability function, p defined over it. Let's further assume that sentences in L pick out sets of possible worlds, and probabilities are determined by a density function over the space of possible worlds. In standard semantics, we think of sentences in natural language as being akin to sentences in L. However, what the connection between conditionals and conditional probability might be thought to show is that, whatever **if** A **then** B expresses, it's *not* equivalent to a sentence in L, since no sentence in L can reasonably be thought of as having as its probability the conditional probability of B given A.

We might, then, go for a radical rethinking of the relationship between the language of probability and the semantic values of natural language sentences. Perhaps natural language sentences

[8]Kratzer's analysis however *might* give a more unified account of how if-clauses work in indicative conditionals and adverbs of quantification.

[9]Much of Yalcin's paper is concerned with how exactly we should understand the information parameter and its relationship to context. I've been deliberately ignoring this, but it is likely to be relevant here in figuring out how desirable the different options are.

don't correspond to sets of possible worlds (as sentences in L do), but rather correspond, generally, to sets of probability functions. In a certain sense, this is exactly the idea that Yalcin's semantics (sketched above) is doing. For him, one of the parameters a sentence is evaluated with respect to is the information parameter, which determines a probability function. Certain sentences such as **it's likely that it's raining**, **if it's cold then it's raining** and **it must be raining** are in fact such that they do not depend on w at all for their truth but just on the information parameter i. Such sentences can then be thought of as functions from probability functions to truth values, and thus as sets of probability functions compatible with them. In particular, **if it's raining then it's cold** picks out the set of probability functions where the conditional probability that it's cold given that it's raining is 1. **It's likely that it's raining**, on the other hand, picks out the set of probability functions that give a probability of greater than .5 to the proposition that its raining. Thus what we effectively have in Yalcin's semantics is a (partial) movement of semantic values from being about statements in L to being about probabilities directly.

Eric Swanson (2006) provides a very similar treatment of such modal and conditional expressions. Instead of having probability functions be a parameter that sentences are evaluated with respect to, he treats the semantic values of sentences as just being constraints on credal states. Thus, for him, **must** P simply denotes a certain kind of constraint on one's credal state. If take semantic values to express constraints on credal states, then we can obviously find a natural treatment of probability operators, too. I won't go through his rich semantic system in detail, here, but it is clear working out of the idea that the semantic values of all sentences are higher order, not just constraining the objects of belief states but the states themselves.[10]

Note that if we take the value of a sentence like **if it's cold, then it's raining** to be a constraint on probability functions then it's very hard to see directly how we should evaluate its probability. This is a problem and a virtue for those who wish to enrich semantic values. It's a problem because we *do* make judgments of probability about **if it's cold, then it's raining**. But it's a virtue to since we don't have as a default that **if it's cold, then it's raining** expresses some standard sentence in L and thus don't burden ourselves with the trouble of explaining how any

[10]The main difference between Yalcin and Swanson that I can detect, in this connection, is that Swanson effectively generalizes to the worst case and makes even simple, non-epistemic, non-conditional, sentences express constraints on credal states, whereas Yalcin gives a more traditional semantics for such sentences—there are doubtless further, perhaps deep, differences but I haven't thought them through.

sentence in L should always have as its probability a conditional probability of two other sentences. What we need then is some non-standard treatment of **it's likely** that applies not just to sets of worlds but also to constraints on credence functions. It should be possible to find one that does the right thing and thus preserves the equation.

9 Summing Up

We haven't here discussed the sorts of positions held by Lewis and Jackson, who think that semantically-speaking conditionals are material, but that they have special rules which makes us, in practice uphold the equation. I take it such positions are pretty difficult to maintain in that the judgments supporting the equation are judgements about the probability of the *truth* of various sentences. It's unappealing to allow one's semantics to run almost entirely free of judgments about *truth*. Anyway, people like Lewis and Jackson, I take it, just don't really care that much about natural language semantics, hence their willingness to give entirely different treatments of counterfactual and indicative conditionals.

I also haven't discussed in detail the school of thought that Bennett (2003) and Edgington (1995) represent. I guess this is because the main thing that they say is that conditionals don't express propositions and I don't take this to be the key question in the area. However, I should say I'm just no the familiar with exactly what Adams (1975) has to say about conditionals and probability and perhaps I'm missing some crucial insight.

So, what, fundamentally, does the relationship between conditionals and conditional probability show us? I think the simplest answer is that it's very hard to think of the semantic value of conditionals as being a set of truth-conditions over which we defined a probability measure. What I'll call the MIT-school (Yalcin's and Swanson's) response is to *radically* alter what counts as a semantic value. Both essentially make statements with conditionals and modals tests on probability functions (or characteristic functions over a set of probability functions). It's then up to them to rewrite the pragmatics of communication to work with these new objects which are playing the role of semantic values instead of the traditional worldly truth-conditions. In this respect they're in a similarly revisionary semantic position to that of the recent relativism advocated by John MacFarlane and others, which also postulates a different format for semantic values (MacFarlane, 2005, forthcoming; Egan et al., 2005).

The other way to go is the one suggest by Kratzer: make conditional statements about some entity's state of knowledge (the ascriptional view). This allows us to have them be about probabilities while preserving the idea that conditionals have real world truth-conditions. As I've suggested, Kratzer's semantics and syntax of the conditional allows one to do quite a bit, then, to explain away certain pieces of evidence for the equation. Not all of it, though. Some will need to be explained directly: that is, we will need to say that there is really some proposition that $A \rightarrow B$ expresses which we routinely assess to have as its probability $P(B|A)$. Even this, however, does not seem an impossible task, and I pointed out how we might start to make that argument. Nonetheless the MIT-school has a series of arguments against the ascriptional view of epistemic modals (and, implicitly, probability operators and conditionals) which I haven't reviewed here. I recommend reading their work (Swanson, 2006, forthcoming; Yalcin, 2007, forthcoming).

References

Adams, Ernest W. (1975). *The Logic of Conditionals: An application of probability to deductive logic*. Dordrecht.

Bennett, Jonathan (2003). *A Philosophical Guide to Conditionals*. Oxford.

Edgington, Dorothy (1995). On conditionals. *Mind*, 104:235–329.

Egan, Andy, Hawthorne, John, and Weatherson, Brian (2005). Epistemic modals in context. In G. Preyer and G. Peter (eds.), *Contextualism in Philosophy: Knowledge, Meaning and Truth*. OUP.

von Fintel, Kai and Gillies, Anthony (2005). *Might* made right. Ms, MIT and University of Michigan.

Iatridou, Sabine (2000). The grammatical ingredients of counterfactuality. *Linguistic Inquiry*, 31:231–270.

Kratzer, Angelica (1981). The notional category of modality. In H.-J Eikmeyer and H. Reiser (eds.), *Words, Worlds, and Contexts*, pages 38–74. Walter de Gruyter.

— (1986). Conditionals. *Chicago Linguistics Society*, 22(2):1–15.

Lewis, David (1975). Adverbs of quantification. In Edward L. Keenan (ed.), *Formal Semantics of Natural Language*. Cambridge University Press.

— (1976). Probabilities of conditional and conditional probabilities. *Philosophical Review*, 8:297–315.

— (1980). Index, context, and content. In S. Kranger and S. Ohman (eds.), *Philosophy and Grammar*, pages 79–100. Reidel.

MacFarlane, John (2005). Making sense of relative turth. *Proceedings of the Aristotelian Society*, 105:321–39.

— (forthcoming). Epistemic modals are assessment-sensitive. In B. Weatherson and A. Egan (eds.), *Epistemic Modals*. Oxford.

Swanson, Eric (2006). *Interactions with Context*. Ph.D. thesis, MIT.

— (forthcoming). How not to theorize about the language of subjective uncertainty. In Andy Egan and Brian Weatherson (eds.), *Epistemic Modality*. OUP.

Yalcin, Seth (2007). Epistemic modals. *Mind*, 116:983–1026.

— (forthcoming). Nonfactualism about epistemic modality. In Andy Egan and Brian Weatherson (eds.), *Epistemic Modality*. OUP.