THE VARIETIES OF CONDITIONAL PROBABILITY

Kenny Easwaran

Alan Hájek discusses many interesting features and applications of the notion of conditional probability. In addition to this discussion, he also gives arguments for some specific philosophical views on the nature of conditional probability. In general I agree with most of his points, and in particular with his arguments that Kolmogorov's "ratio analysis" of conditional probability is not correct. However, he suggests three more points that I would like to contest — that there is a single correct analysis of conditional probability across interpretations, that conditional probability is in general a *more* fundamental notion than unconditional probability, and that Popper's account of conditional probability is the correct one.. I discuss all of these issues (and especially the third) in my dissertation [Easwaran, 2008], but here I will focus more on the first two. I will argue in section 1 that although the different interpretations of probability have many similarities, this similarity is not exact, even at the purely formal level. In particular, I will argue that the bearers of probability in the propensity, subjective, and logical interpretations of probability are distinct classes of objects, and that this changes the mathematical relations that must hold between conditional and unconditional probabilities. As for Hájek's argument that conditional probabilities are fundamental, in section 2 I will demonstrate a distinction that I think his argument collapses.

1 PLURALISM ABOUT CONDITIONAL PROBABILITY

The fact that a single word, "probability", is applied to many diverse phenomena, and the existence of a branch of mathematics called "probability theory" suggests that these diverse phenomena can all be understood as applications of the same formal system. However, there are many different formal systems that all claim to describe probability — for instance, both [Kolmogorov, 1950] and [Rényi, 1970] take the objects of the probability function to be sets, although only the latter allows for conditioning on events of probability zero; the system described in [Hailperin, 1984; Hailperin, 1997; Hailperin, 2000; Hailperin, 2006] and that of [Popper, 1959] both allow for the objects not to be sets, but only the latter takes conditional probability as fundamental. Because of this variety of available systems, it seems that an argument is needed to show that different interpretations of probability should use the same one.[1]

[1]There are some terminological issues here — some suggest that what it *means* to be an interpretation of probability requires satisfying some specific set of mathematical axioms. [Eells,

Handbook of the Philosophy of Science. Volume 7: Philosophy of Statistics.
Volume editors: Prasanta S. Bandyopadhyay and Malcolm R. Forster. General Editors: Dov M. Gabbay, Paul Thagard and John Woods.

At first it might seem that principles like the "Principal Principle" [Lewis, 1980] give the required link to show that the different interpretations should use the same formalism. However, such arguments are not definitive. A standard, simplified version of the Principal Principle states that $CR(A|CH(A) = x) = x$, where CR is an agent's credence function, and CH is the objective chance function. As stated, this principle requires that the objects of the two functions are the same sort of thing (since A is an object of both). It also gives conditions (when the agent knows all the chances) under which the two functions are apparently identical, so there should be no general axioms for subjective probability that rule out any function that satisfies the axioms for objective chance. However, phrasing the principle more carefully, we will be able to see that these two points are not actually correct.

First, Paul Humphreys has argued that the objects of the chance function are events, rather than sentences or sets of worlds. [Humphreys, 2004, p. 669] Even if this is right, a version of the Principal Principle can be stated that is compatible with the arguments of the credence function being sets, as Kolmogorov suggested. Consider the set of all epistemic possibilities for an agent.[2] Let $[A]$ be the set of epistemic possibilities according to which the physical event A occurs. Then the Principal Principle can be reformulated by saying that $CR([A]|CH(A) = x) = x$, so that the different functions can still be defined over different objects. The principle only requires that there be some way of associating some objects of the chance function and some objects of the credence function, not that the objects actually be the same things.

Further, although the Principal Principle may entail that there be no constraints on *unconditional* credences that don't also apply to chances, it doesn't directly relate *conditional* chances to credences, so there is still room for the conditional notions to satisfy different mathematical constraints.[3] Although in many cases the values of conditional probabilities can be derived from the values of unconditional probabilities, this is not true when unconditional probabilities take the value 0, which is a case that Hájek and I agree is often interesting. And as I will suggest later, conditional chances may not be directly constrained by the unconditional ones even in cases where they are all non-zero. Similar moves can be made for other principles that link different interpretations of probability together — these principles don't guarantee that there is a single unified mathematical theory that all interpretations must follow.

1983] I just take the term "interpretation of probability" to apply to the notions that have historically been considered paradigmatic cases. Since my arguments suggest that there is no unified formalism that applies to all of them, this will mean that most of them won't really count as interpretations of probability in this other sense.

[2]These may be possible worlds, impossible worlds, maximal consistent sets of sentences, or some other sort of entities.

[3][Lewis, 1994] gives a "New Principal Principle" that suggests equating conditional probabilities in some circumstances as well. But on both principles it is only under special conditions that the functions must agree. This suggests that the functions may disagree and have different formalisms in other conditions, though I will not investigate this possibility here.

Thus, the primary objection to the possibility of different formal accounts of different interpretations of probability is incorrect — the principles connecting different interpretations of probability have natural modifications that allow the objects of the functions to be different, and allow different relations between conditional and unconditional probability to hold in each. In the rest of this section I will consider several different interpretations of probability and argue for each that it differs from the standard mathematical account of probability given in the first few chapters of [Kolmogorov, 1950], on which probability is a countably-additive, non-negative, normalized function on an algebra of sets, and the conditional probability $P(A|B)$ is given by $P(A \cap B)/P(B)$, whenever $P(B)$ has a precise non-zero value. Since they differ from this account in different ways, there can be no unified account of all of them that is complete for any one.

1.1 Propensity

The first interpretation of probability I will consider is Hájek's "propensity interpretation". (I will use the terms "objective chance" and "propensity" interchangeably.) I will argue that the standard account of conditional probability is incorrect here. Hájek gives some brief arguments in Section 2.2 that various interpretations of probability must satisfy $P(A|B) = P(A \cap B)/P(B)$, and then raises problems in Section 4 in which $P(B)$ is zero, imprecise, vague, or undefined. However, none of the arguments he presents in favor of the ratio account applies specifically to the propensity interpretation. I will suggest that there may be reason to think that there are further failures of the ratio account for the propensity interpretation, even when $P(B)$ is well-defined and non-zero.

The beginnings of this criticism come from [Humphreys, 1985], which argues that conditional chances don't obey Bayes' Theorem. This claim is known as "Humphreys' Paradox." The idea is that if B comes causally before A, then $P(A|B)$ makes sense in a standard way, while $P(B|A)$ should either be 0, $P(B)$, undefined, or otherwise different from the value $P(A|B)P(B)/P(A)$. Several objections to Humphreys' argument are raised in [McCurdy, 1996] and [Gillies, 2000], but there are responses in [Humphreys, 2004].

Regardless of the status of that argument however, I think it brings up an interesting point about the interpretation of *conditional* probability, when probability is interpreted as objective chance, or propensity. Under most accounts of this interpretation, this means that the probability function measures some sort of disposition. If this is right, then we might think of the distinction between conditional and unconditional chances as being parallel to a distinction between dispositions that are triggered by some external conditions, and those that have some tendency to manifest regardless of the conditions. However, some responses to Humphreys' Paradox rely on interpreting the conditional probability $P(A|B)$ differently, as a measure of the system's unconditional disposition to produce both A and B simultaneously, out of the cases where it happens to produce B. This interpretation (which Humphreys calls the "co-production" interpretation) effectively ends up *defining* conditional chance by the ratio formula, rather than taking conditional chance as a notion in need of its own analysis.

For a potential counterexample to the ratio formula, where $P(B)$ is well-defined and positive, but $P(A|B) \neq P(A \wedge B)/P(B)$, consider a fair coin-flipping machine whose trigger is inside a locked room. The easiest way to open the door is to turn on a huge magnet, which unlocks the door, but also prevents the coin from rotating in the air, so that while the magnet is activated, the coin almost always lands heads. Then on the co-production account, the probability of the coin coming up heads given that it is flipped is close to 1, because the propensity of the situation to give rise to a coin landing heads is almost as great as the propensity of the situation to give rise to a coin being flipped at all. However, while the magnet is off and the room is locked, it seems that we may want to say that the propensity for the device to produce a coin landing heads, given that it produces a flip at all, is close to 1/2 — this is exactly what we mean to say that it is a fair coin-flipping machine. This example suggests that the co-production account of conditional propensities gives rise to a notion that is not actually dispositional — it is a parallel to examples of finks that are standard in the literature on dispositions. [Fara, 2006] I suspect that similar examples can be devised to correspond to other problematic cases in the literature on dispositions, like so-called "masks" and "mimics". To get the behavior I have suggested may be intuitive here, we may need to observe the distinction between intervening and conditioning given in [Meek and Glymour, 1994].

It might be objected to this apparent counterexample that often this sort of non-dispositional behavior is exactly what we want for conditional propensities. For example, [Gillies, 2000, p. 828] includes an example of a barometer, in which it seems that we want a notion of conditional propensity for which the probability of a storm, given that the barometer level drops, is fairly high, even though there is no causal influence of the barometer on the storm. The co-production account gives a high conditional probability in this situation (since most cases in which the barometer level drops are cases in which there is a storm), but the more directly dispositional account Humphreys and I suggest gives a fairly low value (since lowering the barometer level has no tendency to cause storms).

But I think it may be that we want such a notion of conditional *probability*, but it's not clear that a notion of conditional *propensity* is needed here. I think all the work of this conditional probability can be done with a notion like conditional *degree of belief*, even though an important part of the work here will be done by the notion of *unconditional* propensity (or objective chance).

To see how this works, we can observe how the desired conditional probability comes out of the unconditional chances by means of Lewis' Principal Principle. A simplified version of this principle states that if an agent knows at some time t the chance at that time that some event A will occur, then her degree of belief in that event should equal the chance of that event. In a toy example, we might consider an agent who knows that the chance of a storm is .25, the chance of a drop in the barometer is .25 and the chance of both occurring is .2. Because she knows these chances, these will also be her credences. Because conditional credences are generally related to unconditional credences by the ratio formula $P(A|B) = P(A \wedge B)/P(B)$, she will have credence .8 in a storm conditional on a drop in the barometer. The conditional chances never enter into her credence function here, so her reasoning can work just as Gillies wants, without having to accept his argument as saying anything about conditional chances. Now, most actual agents don't know the precise chances of any of these events, but they do have some ideas about them. All that is needed to make a modified version of this account work is that the agent be fairly confident that the ratio of chance of storm together with drop in barometer to the chance of a drop in the barometer is higher than the chance of a storm.

Given that an account of conditional chance doesn't need to accommodate these evidential correlations, it seems that we can get a more useful and interesting theory by avoiding the co-production interpretation — just as chances are dispositions on the propensity account, conditional chances are conditional dispositions. As Humphreys puts it in his reply to Gillies and McCurdy, "the conditional propensity constitutes an objective relationship between two events and any increase in our information about one when we learn of the other is a completely separate matter." [Humphreys, 1985, p. 563] But this ends up leading to the conclusion I mentioned above — conditional propensity doesn't obey the mathematical relationships traditionally used for conditional probability.

To make the minimal modification to the standard framework, one natural suggestion (embraced by Fetzer and others) is that in cases like the ones Humphreys describes, conditional propensities are defined in one direction, but undefined in the other. That is, $P(A|B)$ makes sense when A is the event of a coin flip coming up heads, and B is the event of it being flipped fairly, but not when the two are reversed. (This is a difference from Humphreys' picture, on which the inverse conditional propensity has a value of 0, 1, or the value of a related unconditional probability, rather than being undefined.) Then, we just require that conditional propensities obey the standard probability axioms whenever they are defined.

However, the example of the coin flipping device in a magnetic room seems to suggest otherwise - in that case, the only way to get the intuitively correct conditional probabilities involves violating the ratio account of conditional probability, even though there is no division by zero or anything else mathematically strange going on. But I think that making this much of a departure from the standard Kolmogorov axioms for probability makes sense in this case, at least in part because of the metaphysics of the objects of the probability functions. Humphreys suggests that "the arguments of the propensity functions are names designating specific physical events. They do not pick out subsets of an outcome space as in the measure-theoretic approach." [Humphreys, 2004, p. 669] This, if correct, would be a further departure from the standard mathematical theory. Since much of the motivation for the ratio formula comes from taking the objects of the probability function to be sets or sentences, this helps further undermine the argument that chance must satisfy this formula.

Other considerations on conditional chance come up when considering its uses. Humphreys' paradox (which motivates moving away from the standard mathematical formalism) proceeds from an intuition about what conditional chance means. Hájek mentions that evidential decision theory uses conditional probability — it may be that evidential decision theory makes use of conditional credences, while *causal* decision theory [Joyce, 1999] makes use of the agent's expectation of the conditional chances. If so, then the differences between the decision theories motivates a difference between the notions of conditional probability.

In [Lewis, 1980], Lewis suggests another use for conditional chance, which is to describe how *unconditional* chances change over time. The idea is that chances are explicitly indexed to times. Thus, at two different times t_1 and t_2, there are generally two distinct probability functions P_1 and P_2 giving the chances for various events. Lewis argues that $P_2(A) = P_1(A|B)$, where B is "the complete history of the interval between" the two times. [Lewis, 1980, p. 280] In [Lange, 2006], there are some arguments against this claim — however, these depend on whether objective chances can change without any particular non-dispositional fact being the basis of such a change. At any rate, the modified claim is that instead of B being the history of all events occurring in the interval, B is specified to be the conjunction of outcomes of all chance events in the interval.

But in either case, only very specific conditional probabilities are needed — these are always probabilities of a later event conditional on an earlier one (so none of Humphreys' inverse conditional probabilities need to be used), and these are in fact always probabilities conditional on complete sets of occurrences between two times. It is plausible that such cases are never like the case I described with the coin-flipping device in a magnetized room, so that the question of whether those sorts of cases alo violate the ratio account of conditional probability doesn't come up. Thus, this use of conditional chance doesn't seem to cut one way or the other in a dispute about whether conditional chances should behave like other conditional probabilities. If anything, this use suggests that Humphreys may be right about the order of the events being relevant to whether the conditional chance is even defined.

In conclusion, I have argued that chances are properties of events, rather than propositions, sets of worlds, or sentences. Because of the connection between chances and dispositions, and the role of chance in causal decision theory, it seems that the ratio analysis of conditional probability must fail, even in cases where the probabilities are non-zero. I don't have a substantive positive theory to offer in place of this, but this is a question for future research.

1.2 Subjective Probability

The second interpretation of probability is the "subjective" or "Bayesian" interpretation. The probability function is said to give an agent's "degrees of belief" or "credences".[4] In this case, I think the standard account is correct as far as it goes — the difference only arises in the case where $P(B)$ doesn't have a precise non-zero value. In these cases, I argue that conditional probability is best analyzed by the account described by Hájek in Section 5, as "Kolmogorov's refinement". The primary reason is that such an account is entailed by the principle of "conglomerability," (which Hájek mentions in Section 8.4) which states that there cannot be a partition $\mathbf{G}$ (a set of propositions such that the agent is certain that exactly one proposition in the set is true) such that for every $G \in \mathbf{G}$, $P(A|G) > P(A)$. Basically, this means that no experiment can be such that every single conceivable outcome of the experiment would confirm A. (Since most interpretations of probability lack a notion corresponding to confirmation, this motivation for the principle only extends to the subjective interpretation.) A slightly weaker principle that also pushes in this direction is an identification of two notions of independence — conglomerability entails that if A is independent of partition $\mathbf{G}$ in the sense that $P(A|G)$ is constant for all $G \in \mathbf{G}$, then A is independent of $\mathbf{G}$ in the sense that $P(A|G) = P(A)$ for all $G \in \mathbf{G}$.

I also suggest that the examples of "impropriety" that Hájek attributes in Section 5.2 to Seidenfeld, Schervish, and Kadane, are irrelevant for probability as degree of belief. The particular probability spaces and partitions they use are so complicated that they can only be grasped by minds that are far more complicated than the ones that we normally attribute subjective probabilities to. (In particular, to actually construct a specific example of such a space, an agent would need to independently consider uncountably many distinct propositions.) However, other interpretations of probability don't obviously require that every partition involved be graspable by a finite mind. For those interpretations, conglomerability may well fail, and thus a different mathematical account of conditional probability will be relevant.[5]

A further apparent problem for Kolmogorov's refinement is that a single event G may be an element of two different partitions $\mathbf{G}$ and $\mathbf{G}'$, and $P(A|G)$ may be constrained to have two distinct values by these different partitions. ([Kadane *et al.*, 1986] shows that there are certain types of probability space in which this is unavoidable.) However, I suggest that this means we must view conditional probability as (in general) a *three*-place function, depending not only on A and G, but also the partition $\mathbf{G}$ defining the set of "relevant alternatives" to G. In particular cases, this partition will be specified by the experiment an agent is considering G as an outcome to, or the set of alternative hypotheses under consideration, or some other contextual factor. Thus, we must think of conditional degree of belief

[4]For much more on this topic, including more detailed versions of these arguments, and responses to some objections, see my dissertation. [Easwaran, 2008] I also explain in more detail what role I think conditional and unconditional probability play on this interpretation, which is an important part of arguing for any particular formalism.

[5]The failure of conglomerability Hájek mentions in Section 8.5, in the two-envelope paradox, depends essentially on the possibility of unboundedly large payoffs. I endorse only conglomerability for probabilities and not for expectations, and since probabilities are never greater than 1, these sorts of cases can't arise.

as a function $P(A|G, \mathbf{G})$ rather than just $P(A|G)$.[6] In my dissertation I argue that this relativization of the notion to a further parameter doesn't cause any problems for the applications of conditional degree of belief in its usual settings.

I do, however, think that the standard account is right to treat the objects of the probability function as something like sets of possible outcomes (in this case, sets of epistemic possibilities, whatever those are). Hájek claims in Section 4.4 that Kolmogorov's account conflates probability zero events with impossibilities — however, I think Kolmogorov's account actually preserves this distinction better than some of the alternatives. On his account, events are represented by sets, and their probability is assigned by a function. It's true that this function assigns zero both to impossible and non-impossible events, but this distinction shows up in the sets themselves — impossible events are represented by the empty set, while non-impossible events are represented by non-empty sets. In my dissertation, I use this very argument to motivate the claim that at least subjective probabilities are best understood by taking the bearers of probability values to be sets of some sort. This contrasts with accounts like Hailperin's, on which the events themselves have no internal structure to distinguish impossible from non-impossible ones.[7]

In summary, subjective probabilities are relations between an agent and a set of epistemic possibilities, rather than a sentence or an event. Although Kolmogorov's axioms are all appropriate constraints for the relation between conditional and unconditional degrees of belief, the need for conglomerability (to preserve some basic ideas about evidence and independence) means that conditional degrees of belief must actually be relativized to partitions, and must obey Kolmogorov's more sophisticated account of conditional probability, rather than the simple ratio analysis.

1.3 Logical probability

This interpretation is intended to give some sort of generalization of deductive logic's notion of entailment to a notion of "partial" or "inductive" entailment. When B entails A deductively, $P(A|B) = 1$; when B entails $\neg A$, $P(A|B) = 0$; in other cases, intermediate values are possible. In this particular case, I agree with Hájek's claim that conditional probability must be the fundamental notion — just as a notion of logical entailment must underlie any notion of logical truth, it looks like a notion of logical conditional probability must underlie any notion of logical probability. Thus, the proper account must again differ from the standard

[6] One might worry that this would cause further problems with our ability to state the Principal Principle. However, the Principal Principle only considers probabilities conditional on specific values of the chance function, which suggests that we should use the partition by values of chances to fill in the third spot in the conditional probability function. Although this might seem not to put many constraints on credences with respect to other partitions, it does (together with conglomerability) entail that an agent's unconditional credence in an event match her expected value of the unconditional chance, which is most of the work the principle tends to be used for.

[7] Popper's account does better here — although it, like Hailperin's, has no internal structure to the events, the impossible events can be distinguished as those events B such that for every A, $P(A|B) = 1$.

account mentioned above, in a way different from the previous two interpretations. Although this might suggest that Popper's account is the right one for this interpretation, I think this is not totally clear.

For one thing, the standard notion of logical consequence is not fundamentally a relation $A \vdash B$ between two sentences, but rather a relation $\Gamma \vdash B$ between a *set* of sentences Γ and a sentence. To be sure, there is a special case of this relation where Γ contains exactly one sentence. Additionally, when Γ is a finite set, the set-based relation holds between Γ and B iff the sentence-based relation holds between the conjunction of the elements of Γ and B. Thus, if the set is finite, one might try to reduce a function taking a set and a sentence as arguments to a function taking two sentences as arguments.[8]

In the case of complete entailment (that is, when the conditional probability is 1 or 0), classical first-order logic allows for the infinite case to be reduced to the finite case. The compactness theorem states that if Γ is an infinite set and $\Gamma \vdash B$, then there is some finite $\Gamma_0 \subset \Gamma$ such that $\Gamma_0 \vdash B$. However, although complete entailment by an infinite set can be reduced to complete entailment by a finite subset, there is no guarantee that this will be the case for partial entailment. The logic of complete entailment is monotonic — if a subset of Γ entails B, then Γ entails B. But if a set of sentences only *partially* entails a sentence, then surely adding more sentences to the collection of premises can make this partial entailment either better or worse. Thus, there seems to be no hope of reducing the infinite-premise case to the single-premise case in general. If the underlying notion of entailment to be generalized is not classical first-order logic, then the reduction may even be impossible in the case of full entailment.

Thus, if the logical interpretation of probability makes sense (in spite of the many objections that have been raised), it may need to be given by a sort of conditional probability function that allows for conditionalizing on a *set* of sentences, rather than just a single sentence, as all current formalisms for conditional probability do.

2 CONDITIONAL PROBABILITY VS. PROBABILITY GIVEN THE BACKGROUND

Hájek gives several different arguments for the claim that conditional probability ought to be taken as the basic notion of probability theory, rather than unconditional probability, as is standard. Given what I have said above, these arguments will need to be made separately for each interpretation of probability, rather than for all of them simultaneously. But at any rate, I think an important distinction to be made in these arguments is the distinction in degree of belief terms between probability *conditional* on some evidence E, and probability on a corpus of background knowledge B. This distinction may be more clear for objective

[8] Even in the finite case though this seem to blur some distinctions. After all, the fact that $\{A, B\} \vdash A \wedge B$ surely tells us something different than the fact that $A \wedge B \vdash A \wedge B$.

chances — if P is the probability function giving the chances at time t, then this is the distinction between $P(A|E)$ where E is itself some possible event *later in time* than t, and $P(A|B)$, where B is the complete description of the history of the universe *up to* time t. Similar distinctions may apply for some of the other interpretations of probability. To distinguish these, I will call the latter notion "background probability" and reserve the term "conditional probability" for the former. The distinction is approximately that the "background" consists of a collection of information that is necessary for the assignment of probability to even make sense, while a "condition" is some further information that could itself have had a probability given the background. If there really are such sets of background information that are necessary for the assignment of probabilities, then background probabilities seem to play the role that we originally assigned for *unconditional* probability, and not for conditional probability. There would be no room for a truly "unconditional" probability independent even of a background.

Hájek's argument at the end of Section 6 that the fundamental notion of probability theory is conditional probability seems to be based on this notion of background probability, rather than the more specific notion of conditional probability that I distinguish from it. Thus, if these two notions require different formalisms, then he has only established that background probabilities are fundamental, and not conditional probabilities.

To see this distinction in the case of the propensity interpretation, consider Hájek's example of a use for conditional probability: "Probability is a measure of the tendency for a certain kind of experimental set-up to produce a particular outcome . . . [I]t is a conditional probability, the condition being a specification of the experimental set up." (Section 3.1) This contrasts with the account suggested in [Lewis, 1980], on which the propensity function is indexed to a world and time. With the former point of view, unconditional probabilities don't even make sense (what's the probability that a coin lands heads, if nothing about the coin's history or composition or manner of flipping is definite?) while on the latter point of view, both conditional and unconditional probabilities can be understood. If t, w is the time/world-pair at which a coin is about to be flipped, then we can make sense of both $P_{t,w}(H)$ and $P_{t,w}(H|A)$, where H is the event of the coin coming up heads, and A is (for example) the event of a substantial amount of matter quantum tunneling away from the coin in mid-air, thus disrupting the rotation of the coin as it flips. Indexing to the time obviates the need to conditionalize on the set-up, or the history, or anything else in determining a probability value. Thus, there is one picture of the propensity interpretation on which all propensities are conditional on a background (but this is not real conditional probability, in which we might conditionalize on more information than what is necessary to set up the experiment), and another on which there are truly unconditional propensities (though they are indexed to worlds and times, which suffice to determine the background conditions). In neither case do the truly conditional propensities play the fundamental role.

In the case of degree of belief, there are again two options for describing things.

We can index the probability function to an agent at a time, who happens to already have certain knowledge and belief. In that case, there are unconditional probabilities (her degrees of belief in particular propositions that she doesn't already know or fully believe) and conditional probabilities (her degrees of belief in particular propositions when she temporarily supposes further propositions). The other option is to always explicitly include all the agent's knowledge and information and beliefs in the proposition that is being conditioned on. This is the picture that seems to fit Hájek's account (citing de Finetti) of all probabilities really being conditional probabilities. But this conflates the things the agent *actually* knows or believes with things that she is merely supposing for the sake of a conditional probability. Additionally, it invites consideration of probabilities conditional on sets of information that *don't* include her full set of knowledge and beliefs. In some cases we can perhaps make sense of these alternate conditional probabilities as expressing counterfactual possibilities of what she imagines she would believe if she didn't have some of the information she does now. (This would make the conditional probability function conflate *three* types of attitudes an agent might have to a proposition.) But in cases where we consider a radically impoverished set of information in the condition, it seems that the conditional probability doesn't really tell us anything meaningful about the agent. At least, the only way around this seems to be to take seriously the notion of a hypothetical prior probability function that the agent *would* have had in the absence of any information whatsoever.

In the case of logical probability, the distinction I am making between background and truly conditional probabilities doesn't seem to hold up. There is supposed to be one correct logical probability function, and this function measures degree of entailment between sentences. No background is necessary for assigning these probabilities — all probabilities are really conditional. This is different from the other cases, where it seems that some sort of background is necessary to even begin to assign probabilities (there are no chances without a world and time, there are no degrees of belief absent some agent and time), but that other information might be hypothetically added in a way that gives an interesting (but distinct) notion of conditional probability. For logical probabilities it seems that only the latter exists, but that it really is the basic notion.

This distinction can be challenged in both the propensity and degree of belief cases. But in the propensity case this would involve arguing that there really is some objective chance function that specified the chances of various events even in the absence of an experimental set-up, including any of the specifications of how the universe works. And in the degree of belief case it would involve arguing that there really are hypothetical priors that have some real meaning for a particular agent. The difficulty of arguing for such unified functions that can assign probabilities in the absence of information is one of the serious challenges that has been raised against the logical interpretation of probability. Instead, it seems more promising to me to just index chances to worlds and times, and index degrees of belief to agents and times, so that the background doesn't have to be explicitly mentioned in every probability statement. But this just means that there is sense in these cases to be made of a notion of unconditional probability that is not merely derivative from conditional probability, and that some arguments suggesting a priority for conditional probability identify the background information of the situation with the actual objects of the conditional probability function.

Thus, although it may be the case that all non-logical interpretations of probability depend on some set of information in order to assign probabilities in particular cases, the role this information plays can be very different from the role that the antecedent of a conditional probability plays in that interpretation. Thus, this background information is not a good argument for the claim that conditional probabilities are more fundamental than unconditional.

BIBLIOGRAPHY

[Easwaran, 2008] K. Easwaran. *The Foundations of Conditional Probability.* PhD thesis, University of California, Berkeley, 2008.

[Eells, 1983] E. Eells. Objective probability theory theory. *Synthese*, 57:387–442, 1983.

[Fara, 2006] M. Fara. Dispositions. *Stanford Encyclopedia of Philosophy*, 1006.

[Gillies, 2000] D. Gillies. Varieties of propensity. *British Journal for the Philosophy of Science*, 51:807–835, 2000.

[Goel and Zellner, 1986] P. K. Goel and A. Zellner, eds. *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*. North-Holland, 1986.

[Hailperin, 1984] T. Hailperin. Probability logic. *Notre Dame Journal of Formal Logic*, 25(3):198–212, 1984.

[Hailperin, 1997] T. Hailperin. Ontologically neutral logic. *History and Philosophy of Logic*, 18:185–200, 1997.

[Hailperin, 2000] T. Hailperin. Probability semantics for quantifier logic. *Journal of Philosophical Logic*, 29:207–239, 2000.

[Hailperin, 2006] T. Hailperin. Probability logic and combining evidence. *History and Philosophy of Logic*, 27:249–269, 2006.

[Humphreys, 1985] P. Humphreys. Why propensities cannot be probabilities. *The Philosophical Review*, 94(4):557–570, 1985.

[Humphreys, 2004] P. Humphreys. Some considerations on conditional chances. *British Journal for the Philosophy of Science*, 55:667–680, 2004.

[Joyce, 1999] J. Joyce. *The Foundations of Causal Decision Theory.* Cambridge University Press, 1999.

[Kadane *et al.*, 1986] J. B. Kadane, M. J. Schervish, and T. Seidenfeld. Statistical implications of finitely additive probability. In *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*, P. K. Goel and A. Zellner, eds. North-Holland, 1986

[Kolmogorov, 1950] A. N. Kolmogorov. *Foundations of the Theory of Probability.* Chelsea, 1950.

[Lange, 2006] M. Lange. Do chances receive equal treatment under the laws? or: Must chances be probabilities? *British Journal for the Philosophy of Science*, 57:383–403, 2006.

[Lewis, 1980] Lewis, D. (1980). A subjectivist's guide to objective chance. In Jeffrey, R., editor, *Studies in Inductive Logic and Probability*, volume II. University of California Press.

[Lewis, 1994] D. Lewis. Humean supervenience debugged. *Mind*, 109(412):473–490, 1994.

[McCurdy, 1996] C. McCurdy. Humphreys's paradox and the interpretation of inverse conditional probabilities. *Synthese*, 108:105–125, 1996.

[Meek and Glymour, 1994] C. Meek and C. Glymour. Conditioning and intervening. *British Journal for the Philosophy of Science*, 45:1001–1021, 1994.

[Popper, 1959] K. Popper. *The Logic of Scientific Discovery*, chapter iv*, pages 326–348. Harper & Row, 959.

[Rényi, 1970] A. Rényi. *Foundations of Probability.* Holden-Day, 1970.