STEFAN KAUFMANN

CONDITIONING AGAINST THE GRAIN

Abduction and indicative conditionals

Received 4 December 2002; received in revised version 24 October 2003

ABSTRACT. This paper discusses counterexamples to the thesis that the probabilities of conditionals are conditional probabilities. It is argued that the discrepancy is systematic and predictable, and that conditional probabilities are crucially involved in the apparently deviant interpretations. Furthermore, the examples suggest that such conditionals have a less prominent reading on which their probability is in fact the conditional probability, and that the two readings are related by a simple step of abductive inference. Central to the proposal is a distinction between causal and purely stochastic dependence between variables.

KEY WORDS: abduction, conditionals, probability

1. INTRODUCTION

Ramsey (1929) suggested that in arguing about a conditional, the truth value of whose antecedent is unknown, speakers hypothetically add the antecedent to their stock of beliefs and argue on that basis about the consequent. "We can say," he wrote, "that they are fixing their degrees of belief in [the consequent] given [the antecedent]." This last remark can be, and often has been read as making a direct connection to the definition of conditional probability. Many authors since Ramsey (starting with Jeffrey, 1964; Adams, 1965, 1975; Stalnaker, 1970) built their accounts of conditionals on what has come to be called "The Thesis":

(**T**) The probability of a conditional '*if A then C*' is the conditional probability of C, given A.

The Thesis generally accords well with pre-theoretical intuitions. It has also proven valuable in Adams' theory of probabilistic inference, which correctly invalidates certain classically valid but intuitively invalid inference patterns involving conditionals.

Despite these compelling features, however, the Thesis does not always agree with intuitive judgments. In this paper I will discuss examples which illustrate three separate but closely related aspects of this problem:

Journal of Philosophical Logic **33:** 583–606, 2004.

(1) The degree of belief intuitively assigned to a conditional
a. is not always the corresponding conditional probability;
b. is not always the posterior probability of the consequent after observing the antecedent;
c. may be affected by updates which do not affect the conditional probability.

These observations pose no challenge to the assumption that the posterior probability of the consequent upon observing the antecedent should be the conditional probability. What the examples do show is that Ramsey's "hypothetical addition" of the antecedent to the speaker's stock of knowledge does not in general proceed by conditionalization – in other words, *assuming* that the antecedent is true is not the same as *observing* that it is.

The possibility of such a disparity has been noted before; well-known examples are Lewis's (1976) discussion of *imaging vs. conditionalization* and Gibbard's (1981b) distinction between "epistemic" and "nearness" readings of predictive indicative conditionals. These accounts, however, treat the conditionals that obey the Thesis and those that do not as two fundamentally different classes separated by "a profound semantical difference" (Gibbard, 1981b). I will argue instead that (i) both should be analyzed in terms of conditional probabilities, and (ii) they are separated by no more than a simple step of abductive inference.

2. Global and Local Conditional Probabilities

The degree to which a conditional is believed is not always correctly measured by the corresponding conditional probability. Consider the following scenario.[1]

SCENARIO 1. You are about to choose a ball from a bag. It could be one of two bags, X or Y. Bag X contains ten red balls, nine of them with a black spot, and two white balls. Bag Y contains ten red balls, one of them with a black spot, and fifty white balls. By virtue of additional evidence – say, the bag in front of you looks big – you are 75% sure that it is bag Y. For ease of reference, the relevant facts are summarized in the table in (2).

(2)

$P(\text{Bag } X) = 1/4$	$P(\text{Bag } Y) = 3/4$
10 red balls,	10 red balls,
9 of them with a black spot	1 of them with a black spot
2 white balls	50 white balls

In this scenario, is the strength of your belief in the sentence in (3) best characterized as 'high', 'fifty–fifty', or 'low'?

(3) If I pick a red ball, it will have a black spot. □

The reader is invited to stop for a moment and consider how he or she would respond.

The judgment of nine out of ten informants to whom I posed this question in an informal survey, as well as my own intuition, is that the answer should be 'low'. It is easy to trace the reasoning behind this response: In bag X, nine of the ten red balls have a black spot. In bag Y, one of the ten does. We are not sure which bag we are being presented with, but we know that the probability of (3) is either $9/10$ (if it is bag X) or $1/10$ (if is bag Y). Our bias concerning the identity of the bags does the rest: It is likely that it is bag Y, hence the probability of (3) is more likely low than high. Formally, this may be represented as in (4):

$$\begin{aligned} &(4) \quad P(R \to B) &&= P(R \to B|X)P(X) + P(R \to B|Y)P(Y) \\ &(5) &&= P(B|RX)P(X) + P(B|RY)P(Y) \\ & &&= 9/10 \times 1/4 + 1/10 \times 3/4 = 0.3. \end{aligned}$$

The transition from (4) to (5) is not warranted by an "official" rule of the probabilistic calculus. Instead, it is motivated by the intuition that the equality in (6) should hold, at least in this case, if not in general:[2]

$$(6) \quad P(R \to B|X) = P(B|RX).$$

According to (5), the probability of the conditional is the weighted sum of two conditional probabilities, each calculated with respect to one answer to the question of which bag the ball is chosen from.

The problem that this reasoning poses for the Thesis is that in the scenario, the conditional probability of (3) is not "low" but 0.6.[3] Accordingly, the "probabilistically correct" response should be *'high'* or, to allow for some vagueness of judgment, at least *'fifty–fifty'*. Among the various ways of calculating the conditional probability, the following is particularly well-suited to highlight its relationship to (4):

$$\begin{aligned} (7) \quad P(B|R) &= \frac{P(BR)}{P(R)} = \frac{P(BRX) + P(BRY)}{P(R)} \\ &= \frac{P(B|RX)P(X|R)P(R) + P(B|RY)P(Y|R)P(R)}{P(R)} \end{aligned}$$

(8) $$= P(B|RX)P(X|R) + P(B|RY)P(Y|R)$$
$$= 9/10 \times 5/8 + 1/10 \times 3/8 = 0.6.$$

Both (8) and (5) are weighted sums of the very same conditional probabilities, differing only in the weights. The difference lies in whether the antecedent – that is, the hypothesized fact that the ball is red – affects the weights, as it does in (8), or not, as in (5).

A little reflection shows that the probability calculated according to (7) is in fact reasonable under a slightly different interpretation of the conditional. To see this, it is useful to step through the derivation of (7). The unconditional prior probability that the ball will be red is given by (9):

(9) $$P(R) = P(R|X)P(X) + P(R|Y)P(Y) = 1/3.$$

The posterior probability that the bag is X, given that the chosen ball is red, is derived by simple Bayesian reasoning:

(10) $$P(X|R) = \frac{P(R|X)P(X)}{P(R)} = 5/8.$$

Similarly for the posterior probability that the bag is Y, $P(Y|R)$, which is $3/8$. With these posterior probabilities, the weighted sum of the conditional probabilities that the ball has a black spot, given that it is red, for bags X and Y, respectively, is 0.6, as shown in (8) above. In words, the reasoning just illustrated proceeds as follows:

(11) a. I think it is unlikely that I will pick a red ball. (9)
b. But suppose I do.
c. Then this is probably bag X after all. (10)
d. So the ball will probably have a black spot, given that it is red. (8)

In this context, it seems reasonable that the conditional in (11d) should have the high probability it receives according to (7). The step from (11a,b) to (11c) represents an *abductive* inference to the best explanation for the (hypothetical) observation that the ball is red. This step is evidently not performed by those who give the conditional in Scenario 1 a *'low'* rating.

To summarize, both of the following two ways of calculating a probability for (3) correspond to intuitively plausible attitudes towards it, even against the very same factual background:

(5) $$P_\ell(R \to B) = P(B|RX)P(X) + P(B|RY)P(Y),$$

(8) $$P_g(R \to B) = P(P(B|RX)P(X|R) + P(B|RY)P(Y|R).$$

I will refer to these two modes of calculation as "local" and "global," respectively, the motivation being that the latter involves a more sweeping update of the overall probabilities with the antecedent of the conditional. As the example shows, the two are not equivalent.

3. Observing and Assuming

Scenario 1 suggests that the conditional probability is not always the probability of the conditional. This raises the next question: In cases where the two come apart, which of them correctly measures the belief in C that one actually comes to have upon observing A?

Under the assumption that true update (not the hypothetical update involved in conditionals) proceeds by conditionalization, it is predicted that in cases where the two come apart, the posterior probability of the consequent will differ from the prior probability of the conditional. Such a disparity would show that merely hypothesizing that the antecedent is true does not have the same epistemic consequences as the observation that it is.

The following example, adapted from McGee (2000), illustrates that we do indeed, sometimes at least, fail to "come through" on the probabilities we assign conditionals.

SCENARIO 2. Murdoch is dead, possibly killed, possibly by Brown. We have come to believe that his death most likely was an accident, thus that (12) is probably false – until we hear someone who we think is very likely Sherlock Holmes (our champion in a TV show in which the contestants include the real Holmes and each is given the task of convincing the audience that he is the one) declare with great confidence that Murdoch was killed, very likely by Brown, but that in any case, (12) is almost certainly true.

(12) If Brown didn't kill Murdoch, someone else did.

Since we are fairly, though not entirely sure that it was Holmes who asserted (12), we defer to his legendary expertise and believe that the sentence is quite likely true – that is, that the probability that someone else killed Murdoch, given that Brown did not, is high.

Now imagine that it turns out that in fact Brown did not kill Murdoch. Learning this would seriously undermine our belief that the person who asserted (12) was Holmes, who would hardly commit such a blunder. It is much more likely that the speaker was a skilled impostor who really had no clue about the Murdoch case. But then his assertion of (12) is not to be trusted after all, and we revert to our original belief that it is probably false. Thus upon learning that Brown did not kill Murdoch, we conclude that most likely no-one else did either. This conclusion bluntly contradicts the Thesis. □

Or does it? McGee concludes by admitting that he is at a loss for an explanation of the phenomenon, but we find in his text an interesting clue. In light of the newly acquired information that Brown did not kill Murdoch, he writes,

> . . . the most likely explanation for [the contestant's] assertion would be that [the contestant] is not Holmes at all. (p. 109)

Why does McGee speak at this point of the "most likely explanation" for the assertion, rather than simply the conditional probability that the contestant is not Holmes, given that Brown is innocent?

Suppose this conditional probability is high enough to warrant the assertion of (13a). Clearly (13b) is highly probable, since the supposition that the contestant is not Holmes deprives his statement of all its evidential import. Finally, the inference from (13a,b) to (13c) instantiates a probability-preserving variant of Hypothetical Syllogism (Adams, 1975, 1998).

(13) a. If Brown didn't kill Murdoch, the contestant is not Holmes.
b. If Brown didn't kill Murdoch and the contestant is not Holmes, Murdoch's death was an accident.
c. If Brown didn't kill Murdoch, Murdoch's death was an accident.

It seems quite plausible, once the contestant has spoken, that (13a) should have high probability. Nor does there seem to be any flaw in the reasoning in (13a–c). But I also agree with McGee that (12) has high probability in the context of the scenario. Am I being inconsistent?

I think not. I can convince myself that (12) has high probability, and I can convince myself that (13c) has high probability, but I cannot convince myself that both (12) and (13c) have high probability.

Similarly to the earlier example, behind the contrast is the question of whether I let the information that Brown did not kill Murdoch affect my belief that the contestant is Holmes. It is worth while to take a closer look at the structure of McGee's scenario. He does not supply numbers to calculate with. This is reasonable enough, given that the situation cannot be reduced to a simple matter of combinatorics. I will adopt his informal talk of high and low probabilities.

We start out believing that Murdoch's death was likely an accident, even more likely so if Brown did not kill him. We also believe that the contestant is probably Holmes. Initially, these beliefs are independent: Should it turn out, for instance, that Brown did kill Murdoch, we would infer nothing about the identity of the contestant. Likewise, should the contestant turn out not to be Holmes, that would not affect our beliefs about the case. This probabilistic independence between the beliefs we entertain is paralleled by a causal independence: Nothing "out there" in the world is such that the facts of the Murdoch case and the identity of the contestant affect each other. Furthermore, given our beliefs, we think it highly unlikely that the contestant will make a statement of the kind he does, to our surprise, make.

Things change as soon as he has spoken. Most importantly, his statement establishes a probabilistic dependence between our previously independent beliefs: Learning that it is false will now make it unlikely that the contestant is Holmes. This is what eventually happens in McGee's story, and it underlies the inference in (13). But first consider what other changes are effected by the contestant's remarks.

The state of affairs that we considered most likely – that Murdoch's death was an accident and the contestant is Holmes – now turns out to imply that Holmes made a false statement. We are not comfortable with the belief that we probably are in an unlikely situation (where "probably" and "unlikely" refer to posterior and prior probabilities, respectively). Instead, we do what Bayesians are good at, and adjust our beliefs about what we have not observed in such a way as to make our observations unsurprising – in this case, in such a way that it is likely that the contestant would have made the statement. This we do, according to McGee, in deference to Holmes' expertise by concluding that the statement is probably true and we were wrong about the circumstances of Murdoch's death.[4]

The diagram in Figure 1 approximately displays the relative probabilities of the resulting belief state. The event that Brown did not kill Murdoch is represented by the complement of the striped region (the sum of the black and white areas). Within it, the event that someone else killed Murdoch is black, and the event that nobody did is white. Thus informally, the conditional probability that someone killed Murdoch, given that Brown did not, is the probability of *'Black given Black-or-white'*.

We can again distinguish between local and global probabilities for the conditional. The former is the weighted sum of two conditional probabilities that someone killed Murdoch, given that Brown did not, depending on whether the contestant is or is not Holmes, respectively. If the contestant is Holmes (left), that conditional probability is high. If he is not Holmes (right), it is low. Since we consider it likely that the contestant is Holmes,

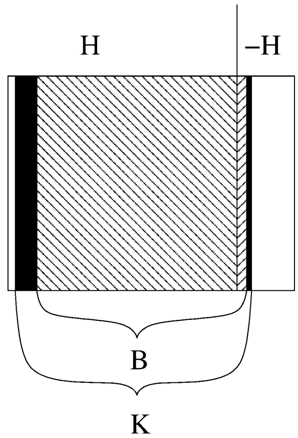

Figure 1. H = *'the contestant is Holmes'* (left); K = *'Murdoch was killed'* (non-white); B = *'Brown killed Murdoch'* (striped).

it is more likely high than low. Hence its expectation is high. This is the prior probability of the conditional according to McGee and most anyone's naïve intuition when first encountering the story.

In contrast, the global conditional probability of (12) is low. To illustrate the difference between it and the local probability, Figure 2 shows the two corresponding ways of dealing with the information that Brown did not kill Holmes. The update itself consists in eliminating the possibility that he did, i.e. the striped region, from Figure 1. Subsequently, the remaining non-zero probabilities are re-normalized. The difference between local and global probabilities corresponds to two ways of carrying out this normalization: Either locally within each cell of the partition *'Holmes/not Holmes'*, leaving the probabilities of these cells unaffected, or globally across the whole belief state, changing the relative probabilities of *'Holmes/not Holmes'*. The respective results are shown in the two lower diagrams in Figure 2.

As is evident from these diagrams, the resulting conditional probability that someone killed Murdoch, given that Brown didn't, is high for the local update (left) and low in the global case (right). The latter accords with the judgment of McGee and his informants that the posterior probability of the consequent is low. It is also what drives the somewhat less salient inference in (13). As the lower left diagram shows, the local update involves a large rise in probability of the unlikely event that Holmes made a false statement. In contrast, the global update on the right avoids this consequence by reducing the probability that the speaker is Holmes.

In comparing the local and global probabilities in (5) and (8) in Section 2 above, I derived both as weighted sums of the same conditional

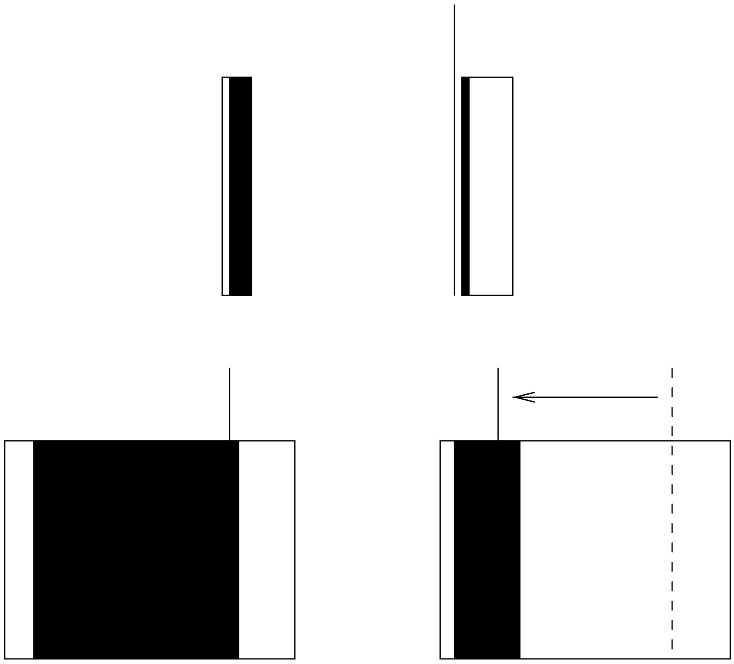

Figure 2. Normalization after removing the possibility that Brown killed Murdoch (top): local (left) vs. global (right).

probabilities, with different weights. The lower diagrams in Figure 2 illustrate the same point: The ratios of the black and white areas within each cell of the *'Holmes/not Holmes'* partition are the same on each side; what differs are the relative sizes of the cells.

4. Dispositions of Subjects and of Things

An example discussed a couple of decades ago by Pollock (1981) and commented on by Gibbard (1981a) turns out to be amenable to the same analysis. The examples shows that our belief in a conditional may be affected by updates which demonstrably do not affect the conditional probability. I quote from Pollock, using his original percentages:

SCENARIO 3. Suppose we know of a vase which was included in a certain shipment of vases. Seventy-five percent . . . of the vases were ceramic and highly fragile, and the other 25% were plastic and virtually unbreakable. We know of this shipment that every ceramic vase which was dropped broke, and none of the plastic vases broke. Furthermore, we know that when the shipment reached its destination, all broken vases and all plastic

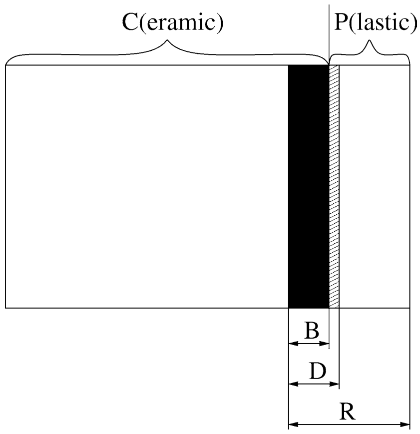

Figure 3. B = *'the vase breaks'*; D = *'it is dropped'*; R = *'it is discarded'*.

vases were discarded, and of the discarded vases, 75% were plastic. This completes our initial background information regarding this shipment. On the basis of this information, we can reasonably believe that

(14) If the vase was dropped, it broke. □

An approximate diagram of the probabilities in this scenario is given in Figure 3. Gibbard (1981a) points out that in order for the example to support Pollock's point, it must in addition be assumed that the probability that a vase is dropped does not depend on whether it is ceramic or plastic. Hence among the vases that were dropped (the non-white region in the figure) the ratios of broken and unbroken vases are just the same as the overall ratios of ceramic and plastic vases (75% and 25%, respectively).[5]

Whether the vase breaks depends not only on whether it is dropped, but in addition on whether it is ceramic or plastic; however, that question is certainly not affected by whether it is dropped or not. Thus we see once again a background factor that is causally independent from the antecedent but determines, jointly with the latter, the probability of the consequent.

Figure 3 also illustrates that the local probability $P_\ell(D \to B)$ of the conditional in (14) is high. In the event that the vase is ceramic (left), the probability that it breaks, given that it is dropped, is 1. In the event that it is plastic (right), the corresponding conditional probability is 0:

(15) $$\begin{aligned} P_\ell(D \to B) &= P(B|DC)P(C) + P(B|DP)P(P) \\ &= 1 \times 0.75 + 0 \times 0.25 = 0.75. \end{aligned}$$

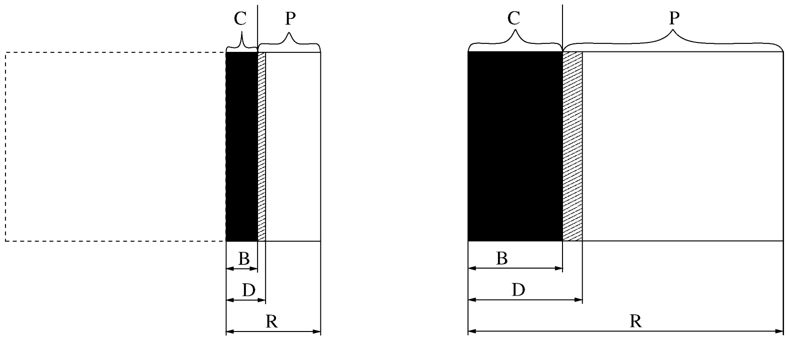

Figure 4. Update with R = *'it was discarded'*: before (left) and after (right) renormalization.

In this example, the global conditional probability $P_g(D \rightarrow B)$ comes out to be the same:

$$\text{(16)} \quad \begin{aligned} P_g(D \rightarrow B) &= P(B|D) \\ &= P(B|DC)P(C|D) + P(B|DP)P(P|D) \\ &= 1 \times 0.75 + 0 \times 0.25 = 0.75. \end{aligned}$$

Gibbard's comment about the independence of the material from the fate of the vases reminds us of why this must be the case: The probability that a vase is ceramic, given that it is discarded, equals the unconditional probability that it is ceramic: $P(C|D) = P(C)$, as seen in both (15) and (16).

Pollock goes on to show that intuitions about the probability of the conditional change in response to the following new information:

> Suppose we are now informed that the vase under consideration was discarded. This is the proposition R. As 75% of the discarded vases were plastic, this makes it unreasonable to believe that if the vase was dropped then it broke. Thus R makes the indicative conditional less reasonable. (p. 251)

This seems plausible, so it may come as a surprise that it is only the local probability that is affected by the update with R, whereas the global conditional probability remains unchanged. Consider Figure 4 for illustration. R is not the antecedent of a conditional, and we can assume that the update proceeds by conditionalization, in two steps: elimination of non-R (left), and renormalization of the probabilities (right). As a result, it is more likely than it was before that the vase is plastic; however, the conditional probability that the vase broke, given that it was dropped (B given D) is 75%, as before.

I will write P' for the new probability distribution which results from updating P with R. For all propositions X, Y, $P'(Y|X) = P(Y|XR)$. Then the posterior probabilities corresponding to (15) and (16) above are (17) and (18), respectively.

$$(17)\quad \begin{aligned} P'_\ell(D \to B) &= P'(B|DC)P'(C) + P'(B|DP)P'(P) \\ &= P(B|DCR)P(C|R) + P(B|DPR)P(P|R) \\ &= 1 \times 0.25 + 0 \times 0.75 = 0.25, \end{aligned}$$

$$(18)\quad \begin{aligned} P'_g(B|D) &= P'(B|DC)P'(C|D) + P'(B|DP)P'(P|D) \\ &= P(B|DCR)P(C|DR) + P(B|DPR)P(P|DR) \\ &= 1 \times 0.75 + 0 \times 0.25 = 0.75. \end{aligned}$$

As Pollock points out (p. 251), the equality between the global probabilities is a consequence of the fact that D entails R: What distinguishes (16) from (18) is the conditionalization on R in the right-hand side of the latter, which due to the entailment has no effect on the values. Not so for the transition from (15) to (17): $P(C|D) = P(C)$, but $P'(C|D) \neq P'(C)$.

To summarize: We started out with a distribution in which the two ways of calculating a probability for the conditional were equivalent. Through an update with the separate but relevant proposition R, we ended up with a new distribution in which they differ, not unlike the situation in Scenario 1. Since the conditional probability is not affected by the update but the probability of the conditional is, the latter must be something else. I suggest that it is the local probability: $P_\ell(D \to B)$ before the update, and $P'_\ell(D \to B)$ thereafter.

As in Scenario 1, here too an argument can be made that besides the more prominent local interpretation, the conditional (14) has a reading under which its probability is in fact high, as predicted by the Thesis. This is what Gibbard (1981a) shows in his response to Pollock's paper, arguing that the following is valid (I use Pollock's percentages again):

(19)
a. The vase was discarded.
b. Of the discarded vases that had been dropped, 75% were ceramic.
c. Therefore if this vase was dropped, it was ceramic.
d. If this vase was dropped and ceramic, then it broke.
e. Therefore if this vase was dropped, it broke.

The inference from (19a) to (19c) rests on the additional premise (19b), which can be inferred from the given facts but is not explicitly mentioned. Gibbard attributes the effect of Pollock's example to speakers' failure to take such additional facts into account. Instead, he suggests, speakers assign the conditional $D \to B$, *'if the vase was dropped then it broke'*, low probability because they consider it equivalent to C, *'the vase was ceramic'*, since all and only the ceramic vases were fragile.

The account I have proposed builds on a similar observation, but the two diverge in important respects. Nothing in my account requires $D \rightarrow B$ and C to be *equivalent* – it would work just as well if the local probability of the conditional, given that the vase is ceramic, were not equal to 1 as in (15) and (17) but only "high" (i.e., higher than the corresponding probability of the conditional, given that the vase is plastic), as in Scenario 1.

Nor do I think that Gibbard is right in dismissing Pollock's judgment as "fallacious." It is wrong as an estimate of the conditional probability, but as evidence of the probability speakers assign to the conditional, it should be taken in its own right. Moreover, I will argue in Section 7 below that it can in fact be rational to interpret a conditional in this way.[6]

Finally, unlike Gibbard, I claim that the two probabilities are closely related and conditional probabilities are the basis of both. The inference from (19a,b) to (19c) is again the by-now familiar abductive step, reflected in (18) in the conditionalization on the hypothetical observation that the vase in question was dropped. Thus Pollock's example is accounted for in just the same way as the others above.[7]

5. SUBJECTIVE AND OBJECTIVE PROBABILITY

All the examples suggest that there is a choice between two non-equivalent ways of bringing conditional probabilities to bear on the probability of conditionals. Of these, the local mode of calculation appears to be preferred in the evaluation of conditionals, while the global mode correctly predicts the posterior belief in the consequent, given the antecedent.

In each case, calculating the global conditional probabilities involves an *abductive* step which, as intuitive judgments suggest, speakers are less prepared to make in *assuming* that the antecedent is true than they are upon *observing* that it is. In essence, the three scenarios share the following characteristics:

1. The probability of the consequent depends, in addition to the antecedent, on a third variable (the identity of the bag in Scenario 1; the identity of the contestant in Scenario 2; the material of the vase in Scenario 3);
2. This third variable is *causally* independent of the antecedent;
3. The third variable is *stochastically* dependent upon the antecedent (i.e., the antecedent provides evidence pertaining to the third variable).

To spell this out in some more detail, in saying that "the probability" of the consequent depends on an additional variable, I have in mind its *objective* probability or *chance*, without loading this use of the term with too much of a metaphysical commitment. Following Skyrms (1988), I am happy to think of chance as belief conditional on a partition "that we have found to be of continuing usefulness." It is helpful, however, to draw a notational distinction in this section between the subjective probabilities 'P' dealt with so far and the corresponding objective chances, for which I will use the symbol '*Ch*'.

Suppose the objective probabilities in a given scenario are known to depend on an unobserved variable **X** whose values are mutually exclusive and jointly exhaust the probability space. In Scenario 1, for instance, the identity of the bag is such an unobserved background variable. I will assume that in general, **X** takes finitely many values $X_1, \ldots, X_n$ and that A, the antecedent of the conditional, is compatible with each X_i (in the sense that it has non-zero probability conditional on each X_i). Neither of these assumptions is essential for the account, but both simplify the exposition.

Each X_i corresponds to the event that the actual objective chances are distributed according to a function $Ch(\cdot|X_i)$. The Bayesian reasoner assigns subjective probabilities P to the cells $X_1, \ldots, X_n$ of the partition induced by **X**. In addition, to draw a straightforward – and idealized – connection between objective and subjective probability, I assume that the agent, though ignorant of the actual value of **X**, nevertheless knows *how* that value determines the probabilities of the sentences in question. This assumption is embodied in (20).[8] For each sentence A and each value X_i of **X**:

(20) $P(A|X_i) = Ch(A|X_i)$

Now the agent's subjective probability of sentence A is the weighted sum of these conditional beliefs, which, given (20), is the speaker's expectation of A's objective chance:

(21)
$$\begin{aligned} P(A) &= \sum_{X_i \in \text{range}\,(\mathbf{X})} P(A|X_i)P(X_i) \\ &= \sum_{X_i \in \text{range}\,(\mathbf{X})} Ch(A|X_i)P(X_i). \end{aligned}$$

Turning to conditionals, their objective probabilities are defined as objective conditional probabilities:

(22) $Ch(A \rightarrow C|X_i) = Ch(C|A, X_i)$.

Assuming that conditional chance, like subjective probability, follows the probabilistic calculus, we can extend (20) to estimates of conditional chances. For all A, C, X_i:

(23) $P(C|A, X_i) = Ch(C|A, X_i)$.

This is all that is required for a general statement of the difference between "local" and "global" probabilities of a conditional $A \rightarrow C$. First, treating it on a par with other sentences of the language, its subjective probability may be defined following (21), as the expectation of its objective chance. This corresponds to the *local* probability:

$$(24) \quad P_\ell(A \rightarrow C) = \sum_{X_i \in \text{range}(\mathbf{X})} P(A \rightarrow C|X_i)P(X_i) = \sum_{X_i \in \text{range}(\mathbf{X})} Ch(C|A, X_i)P(X_i).$$

However, (24) is not equivalent to the conditional probability of C, given A. The latter is given by the *global* probability in (25):

$$(25) \quad P_g(A \rightarrow C) = P(C|A) = \sum_{X_i \in \text{range}(\mathbf{X})} P(C, X_i|A) = \sum_{X_i \in \text{range}(\mathbf{X})} Ch(C|A, X_i)P(X_i|A).$$

Once again, (24) and (25) are weighted sums of the same objective conditional probabilities, the only difference being that the weights are affected by the hypothesized observation of the antecedent in the latter but not the former.

6. TRIVIALITY

In this section and the next, I will address two potential objections to the proposal I have laid out: Its susceptibility to Lewis's (1976) *triviality* results, and the question as to whether it is ever *rational* to give a conditional its local probability.[9] This section is devoted to the first of these, triviality.

Recall that the calculation of the local probability rests on (6) in Section 2, repeated in a more general form in (26), for each cell X_i in the partition induced by a background variable **X**:

$$(26) \quad P_\ell(A \rightarrow C|X_i) = P(C|AX_i).$$

Lewis's argument shows that if the partition is $\{C, \overline{C}\}$, then the weighted sum of these local probabilities invariably equals the probability of the consequent:

$$(27) \quad P_\ell(A \rightarrow C) = P_\ell(A \rightarrow C|C)P(C) + P_\ell(A \rightarrow C|\overline{C})P(\overline{C}) = P(C|AC)P(C) + P(C|A\overline{C})P(\overline{C}) = 1 \times P(C) + 0 \times P(\overline{C}) = P(C).$$

In asking how this result bodes for the present proposal, I should re-emphasize that I do not claim that conditionals have local interpretations with respect to just *any* variable. Such interpretations arise only under certain conditions and with respect to some variables which stand in the right causal relation to the conditional's constituents. The causal structure of a scenario cannot be determined on logical grounds alone or read off the probability distribution,[10] and I did not offer a rule for determining what the causal structure of a given situation is. Still, I take speakers' assumptions about such structure to exist and to be the best explanation for violations of the Thesis.[11]

Now, while it is true in principle that any given conditional *can* receive a local interpretation along the lines of (27), it is an empirical question whether and under what circumstances speakers actually give it that interpretation, and whether the prediction in (27) is counterintuitive in those cases.

I cannot think of a scenario in which the consequent itself is a causal background variable in the relevant sense. There is no sense in which the probability of the consequent could be determined by its truth value (together with that of the antecedent). But one can sometimes find a suitable proposition that is equivalent to the consequent and otherwise eligible for the role of a causal background factor. Consider for instance a variant of Scenario 1, with a different distribution of black spots: Suppose all balls in urn X have a black spot and none of the balls in urn Y does. The question of which urn is being used is still causally independent of R and determines the objective probability of B. In this scenario, X and B are equivalent, hence $P_\ell(R \to B) = P(X) = P(B)$:

$$
\text{(28)} \qquad \begin{aligned} P_\ell(R \to B) &= P(B|RX)P(X) + P(B|RY)P(Y) \\ &= P(B|RB)P(B) + P(B|R\overline{B})P(\overline{B}) \\ &= 1 \times 1/4 + 0 \times 3/4 = 1/4. \end{aligned}
$$

But this prediction is not implausible at all: It amounts to the claim that in the setting described, both of (29a,b) are equivalent to (29c), and all three have probability 1/4, the probability that the urn is X.

(29) a. If the ball is red, it will have a black spot.
b. The ball will have a black spot.
c. The urn is X.

I would consider it incumbent upon the detractor to produce a scenario in which the consequent is a causal background variable in the relevant sense and the resulting equality of the conditional's local probability with the probability of the consequent does not accord with intuition. If and when such a case is found, it will be important to determine how it differs from the ones discussed in this paper.

Meanwhile, notice also that the most disconcerting consequences of (T) do not hold for the local interpretation. Specifically, the equality of $P_\ell(R \to B)$ and $P(B)$ does not imply that R and B must be stochastically independent. Nor does the local interpretation require R and $R \to B$ to be stochastically independent, another consequence of (T) pointed out by van Fraassen (1976). Both of these independencies hold for $P_\ell(B \to R)$ only when it equals $P(B|R)$.

7. Why not Conditionalize?

Assuming that the present proposal is descriptively correct, it raises a deeper question: Is it an account of a fallacy – one that is committed widely and systematically, but fallacious nonetheless – or is the departure from (T), at least in some circumstances, the "correct" interpretation of a conditional?

In order to make sense of a question like this, it is customary to phrase it in terms of rational behavior: Are there situations in which it would be detrimental to base one's actions upon (T) and advantageous to follow the local interpretation? A negative answer would not imply that the local interpretation is not what speakers use, but only that it is not what they *should* use. I will argue in this section, however, that the answer is not negative. I will use a variant of Lewis's "diachronic Dutch book argument" (publicized first by Teller, 1973, and more recently by Lewis, 1999) to show that the local interpretation can indeed be rational in certain cases. However, in those cases it is not rational to conditionalize upon observing the antecedent.

Lewis's argument is not about conditionals, nor does it show that it is always rational to conditionalize. What it does show is that one should hold on to one's beliefs, in the sense that if the strength of your prior belief in C on the supposition that A is x, then the strength of your posterior belief in C upon learning A should also be x. Otherwise you are vulnerable to exploitation, provided that you base your decisions regarding the price at which to sell or buy certain bets on your degrees of belief. (I adopt Lewis' use of the pronouns 'you' and 'I' for the irrational agent and the cunning Dutchman, respectively.)

To apply the argument to conditionals, we should substitute "belief in the conditional 'if A then C' " for what I just referred to as "belief in C on the supposition that A." Lewis shows that it is rational to conditionalize in case your degree of that belief is the conditional probability $P(C|A)$; regarding conditionals, then, it is rational to conditionalize if your degree of belief in the conditional is $P_g(A \to C)$.

I will not reiterate Lewis's argument here, but one important assumption he makes is worth emphasizing: that I possess no more information than you at all times at which the transactions in question are carried out. In fact, Lewis notes, the argument would be pointless otherwise: "It proves nothing derogatory about your rationality that I can exploit you by taking advantage of my greater knowledge."

It is worth asking, however, how exactly I might exploit my greater knowledge. Consider the following circumstances. At time 0, your beliefs are measured by a probability distribution P, and I know no more than you do. Between time 0 and time 1, we will both find out whether A is true or false. If A is true, I will find out in addition "in what way" it is true, i.e., which of $AX_1, AX_2, \ldots, AX_n$ is true, where the X_i are the values of some variable **X** that we both take to be causally relevant, in the sense discussed above: The conditional probability of C is not evenly distributed over all AX_i, and **X** does not causally depend on A. At time 1, if A is true, I will know the actual value X_i and hence the actual, or objective, probability $P(C|AX_i)$. But you won't know it, nor do you know that I will.

Suppose now your probability of the conditional is the conditional probability, and you update your beliefs by conditionalizing. Then I can exploit you with the following strategy.

1. At time 0, sell you the two bets
   [\$1 if AC, \$0 otherwise]
   [\$$P(C|A)$ if $\overline{A}$, \$0 otherwise]
   for the maximum price you will pay, viz. \$$P(C|A)$.
2. Wait to see whether A is true at time 1.
   (We will both have this information.)
3. If A is true, see which AX_i is true.
   (You will not have this information.)
   If $P(C|AX_i) \geq P(C|A)$, buy from you the bet
   [\$1 if C, \$0 otherwise]
   for the minimum price you will accept, viz. \$$P(C|A)$.

If A is false, you get your money back and the game is over. If A is true, and furthermore I learn that my risk of having to pay \$1 to you is higher than $P(C|A)$, I effectively cancel the first bet, without you realizing that this is bad for you because your expected return is in fact positive. If, on the other hand, the actual X_i makes it more likely that I will keep the dollar, I take no further action. In this case, I have the \$$P(C|A)$ from time 0 and the probability that I will pay \$1 is $P(C|AX_i)$. Thus my expected payoff is $\$P(C|A) - \$P(C|AX_i)$, which is positive. Overall, for any possible outcome, you believe that your expected loss is \$0, but in this latter case you are wrong.

But now consider a slightly different version of the same game. Suppose you know at time 0 that I will know at time 1 which X_i is true, so you can see through my strategy; furthermore, I know that you know this. How will you guard against exploitation? First, assuming that I put my own interest first, you know that by accepting the transaction at time 1 you can only decrease your expected payoff.[12] So you will never accept that transaction. Thus I cannot cancel the bet at time 1, and my actual expected payoff can be negative, depending on which X_i is true. My prior expected payoff at time 0 is the weighted sum of these posterior payoffs for each X_i:

$$(30) \qquad \sum_{X_i \in \mathbf{X}} (\$P(C|A) - \$P(C|AX_i))P(X_i).$$

Notice that we can both estimate this at time 0, and that we do so, for we both know that you will not deal with me at time 1. Now while this new expected payoff is a result of your refusal to do business with me at time 1, it affects our opinions about the value of the bet at time 0. If my expected payoff is positive, it is irrational for you to pay $\$P(C|A)$ for the bet. If my expected payoff is negative, it is irrational for me to sell it for $\$P(C|A)$. The only price we can both accept is one which, when substituted for $P(C|A)$ in (30), sets my expected payoff (and thus also yours) to zero:

$$(31) \qquad \begin{aligned} \$0 &= \sum_{X_i \in \mathbf{X}} (\$x - \$P(C|AX_i))P(X_i) \\ &= \$x - \sum_{X_i \in \mathbf{X}} \$P(C|AX_i)P(X_i), \\ \$x &= \$P_\ell(A \to C). \end{aligned}$$

It turns out that the right amount to pay for the bet is just the local probability of the conditional.

One final adjustment is required to make the deal palatable for both of us: Your payoff for the event that A is false (regardless of X_i) should be the price at which you bought the bet, so that cost and return cancel each other out, as they do in Lewis's original example. The resulting strategy – the only one that you and I will both find acceptable – is then the following (no transaction occurs at time 1):

1. At time 0, sell you the two bets
   [\$1 if AC, \$0 otherwise]
   [$\$P_\ell(A \to C)$ if $\overline{A}$, \$0 otherwise]
   for the maximum price you will pay, viz. $\$P_\ell(A \to C)$.

I will have little interest in playing this game with you, however, since I stand to gain nothing from it.

Thus it is indeed sometimes rational to assign to a conditional its local probability. The case I have presented involves knowledge of the fact that the adversary will have the advantage of knowing the actual value of a relevant variable. There may be other situations in which the local interpretation is to be preferred, but this is a particularly illuminating one. The asymmetry of information between you and the Dutchman is the epistemic analog of your uncertainty about objective probabilities. Taking the Dutchman to be Nature Herself, the statement that one variable causally affects another amounts to the claim that Nature knows the value of the former when She estimates the latter.

There is no case, however, in which it is rational for you to assign the local probability at time 0 and then conditionalize upon observing that the antecedent is true at time 1 (unless the two probabilities happen to coincide). This lesson from Lewis's argument remains intact. In this case, the decision to conditionalize after all would be tantamount to admitting that you were wrong in the first place, changed your mind about the value of the bet, and may have to write off some losses as a result. Perhaps this is in fact what happens in cases like Scenario 2.

8. Indicative and Counterfactual

In this last section, I would like briefly to point out a connection between the cases observed above and well-known similar facts about counterfactual conditionals. Consider first the following scenario:

SCENARIO 4. Jim and Jack are not having lunch together as usual. They must have had a quarrel yesterday.

(32) If Jim asks Jack for help this afternoon, Jack will not help him.

But wait: Jim is a prideful fellow. So if he asks Jack for help this afternoon, there has to have been no quarrel yesterday.

(33) So if he asks Jack for help, Jack will help him after all. □

The causal relations in this setup are similar to those in Scenario 1. Whether Jim asks Jack or not depends on whether there was a quarrel, and whether Jack helps Jim depends on both Jim's asking and the occurrence of the quarrel. The quarrel, however, is neither brought about nor prevented by Jim's request for help.

By a similar argument as that for Scenario 1, behind the difference between (32) and (33) is the distinction between the local and the global reading, respectively, and the two are related by the abductive inference from the hypothetical request for help to a revision of the speaker's beliefs about the quarrel.

Scenario 4's better-known original version is due to Downing (1959) and cited here after Lewis (1979):

SCENARIO 5. Jim and Jack quarreled yesterday, and Jack is still hopping mad. We conclude that

(34) If Jim asked Jack for help today, Jack would not help him.

But wait: Jim is a prideful fellow. He never would ask for help after such a quarrel; if Jim were to ask Jack for help today, there would have to have been no quarrel yesterday.[13] In that case Jack would be his usual generous self.

(35) So if Jim asked Jack for help today, Jack would help him after all. □

Here it is an established fact that the quarrel occurred; accordingly, it is reasonable to assume that neither global nor local conditional probabilities are defined. Clearly the simple mechanism outlined in previous sections must be augmented to deal with counterfactuals.

Kaufmann (2001) argued that due consideration for causal independences between variables makes possible a probabilistic treatment of counterfactuals based on the often-made suggestion that their posterior probabilities are closely related to the prior probabilities of the corresponding indicatives. Rephrased in the terms of the present account, the idea is that in examples in which the two seem to come apart, what comes apart are their global probabilities (in terms of the present account), but not their local ones.[14]

Scenario 5 suggests that the ambiguities observed in this paper are mirrored rather closely in counterfactuals. Each of the readings of the indicative conditional in Scenario 4 corresponds to a reading of the counterfactual in Scenario 5, and the abductive inference relating the two readings appears to be the same in both cases.

9. Conclusion

There is no denying that the probabilities of conditionals are not always the corresponding conditional probabilities. The instances of this phenomenon discussed in this paper show that where the two differ, they do so in ways that are predictable, systematic and illuminating. Causal independences between the variables involved in a given scenario are key to predicting and understanding the discrepancy. In addition, the fact that indeterminacies in interpretation that are commonly observed with counterfactuals affect indicative conditionals as well, provides further evidence for the close semantic relationship between these two classes.

"Counterfactuals," Lewis (1979) wrote, "are infected with vagueness." While the present proposal may take us no more than a few steps closer to an account of all such vagueness, it does suggest that at least some of it extends to indicatives and is susceptible to systematic study.

Acknowledgements

I am grateful to an anonymous reviewer for comments which greatly helped improve the paper. Part of this work was supported by the Japan Society for the Promotion of Science (JSPS, Grant #P01063) and carried out at Kyoto University.

Notes

[1] This is a variant of a scenario first suggested to me by Dorothy Edgington (pc).

[2] Lewis (1976) showed that the substitution in (5), combined with the Thesis, leads to triviality if it is to hold as a general principle. See Section 6 below for some discussion.

[3] Using only the conditional probabilities given in the scenario:

$$\begin{aligned} P(B|R) &= \frac{P(B|RX)P(R|X)P(X) + P(B|RY)P(R|Y)P(Y)}{P(R|X)P(X) + P(R|Y)P(Y)} \\ &= \frac{9/10 \times 10/12 \times 1/4 + 1/10 \times 10/60 \times 3/4}{10/12 \times 1/4 + 10/60 \times 3/4} = 0.6. \end{aligned}$$

[4] McGee assumes implicitly that we consider our reasons for believing that the contestant is Holmes more reliable than our reasons for believing that Murdoch's death was an accident. In fact, nothing in the setup suggests that this is more plausible than the opposite conclusion, that the statement is probably false and the contestant not Holmes, but I follow McGee for the sake of argument.

[5] Gibbard considered 75% too low to accept a conditional and changed Pollock's numbers to 95% and 5%. The reader may choose to follow Gibbard; no substantial change results from this move.

[6] Notice that this does not commit me to the claim that *all* beliefs in conditionals are beliefs in dispositions, as Gibbard would have it: “Suppose, for instance, I knew that only plastic vases had been dropped. It seems to me that I could then say, ‘If this vase was dropped, it was plastic’ without supposing that I am talking about a vase that was disposed to being plastic on being dropped. On the dispositional thesis, to do so is incoherent” (p. 255). Not so if abductive inference is granted the important role it deserves.

[7] It may not be a sensible thing to say in this case that the vase’s being ceramic *explains* the (hypothesized) observation that it was dropped. Rather, it is, as it were, “the most likely way for the vase to have been dropped,” given that it was discarded.

[8] Principles of this sort have been postulated by Lewis (1980, 1994), van Fraassen (1981), Skyrms (1990), among others.

[9] I am grateful to an anonymous reviewer for insisting that I address these questions.

[10] For arguments that the causal structure must be encoded separately in the model, see Woodward (2001).

[11] It follows that new information about causal relations can make judgments of conditionals more accurate, even in the absence of any updates about facts or probabilities.

[12] If I expect to repeat the game with you many times over, I might consider displaying some apparently irrational behavior in order to trick you into believing that you can exploit me. I ignore such higher-order considerations here.

[13] From the fact that Lewis discussed the sentence *‘If Jim were to ask Jack for help today . . . ’* in the context of counterfactual dependence, we cannot conclude that he considered it a counterfactual. In his earlier book (Lewis, 1973), he had declared that such “subjunctive conditionals pertaining to the future . . . appear to have the truth conditions of indicative conditionals” (p. 4), which he considered unrelated to those of counterfactuals, for the two are “really two different sorts of conditional” (p. 3). However, regardless of what Lewis thought of this particular sentence, it is easy to check that the phenomenon it illustrates affects sentences that are uncontroversially counterfactual as well: Just shift the hypothetical request for help to an earlier time, such as *‘last night’* or ‘*early this morning*’.

[14] Kaufmann, adapting ideas of Jeffrey (1991) and Stalnaker and Jeffrey (1994), assigns these local conditional probabilities as values point-wise to possible worlds.

References

Adams, E. (1965) The logic of conditionals, *Inquiry* **8**, 166–197.

Adams, E. (1975) *The Logic of Conditionals*, Reidel.

Adams, E. (1998) *A Primer of Probability Logic*, CSLI Publications.

Aristotelian Society (1991) *The Symposia Read at the Joint Session of the Aristotelian Society and the Mind Association at the University of Durham, Proceedings of the Aristotelian Society* **65** (Supplementary Volume).

Downing, P. B. (1959) Subjunctive conditionals, time order, and causation, *Proceedings of the Aristotelian Society* **59**, 125–140.

Eells, E. and B. Skyrms (eds.) (1994) *Probabilities and Conditionals: Belief Revision and Rational Decision*, Cambridge University Press.

Fetzer, J. H. (ed.) (1988) *Probability and Causality*, Studies in Epistemology, Logic, Methodology, and Philosophy of Science, Vol. 192, D. Reidel.

Galavotti, M. C., Suppes, P. and Costantini, D. (eds.) (2001) *Stochastic Causality*, CSLI Publications.

Gibbard, A. (1981a) Indicative conditionals and conditional probability: Reply to Pollock, in (Harper et al., 1981), pp. 253–255.

Gibbard, A. (1981b) Two recent theories of conditionals, in (Harper et al., 1981), pp. 211–247.

Harper, W. L., Stalnaker, R. and Pearce, G. (eds.) (1981) *Ifs: Conditionals, Belief, Decision, Chance, and Time*, Reidel.

Jeffrey, R. C. (1964) If, *Journal of Philosophy* **61**, 702–703.

Jeffrey, R. C. (ed.) (1980) *Studies in Inductive Logic and Probability*, Vol. 2, University of California Press.

Jeffrey, R. C. (1991) Matter-of-fact conditionals, in *The Symposia Read at the Joint Session of the Aristotelian Society and the Mind Association at the University of Durham, Proceedins of the Aristotelian Society* **65** (Supplementary Volume), pp. 161–183.

Kaufmann, S. (2001) Tense probabilism properly conceived, in R. van Rooy and M. Stokhof (eds.), *Proceedings of the Thirteenth Amsterdam Colloquium*, pp. 132–137.

Lewis, D. (1973) *Counterfactuals*, Harvard University Press.

Lewis, D. (1976) Probabilities of conditionals and conditional probabilities, *Philosophical Review* **85**, 297–315.

Lewis, D. (1979) Counterfactual dependence and time's arrow, *Noûs* **13**, 455–476.

Lewis, D. (1980) A subjectivist's guide to objective chance, in (Jeffrey, 1980), pp. 263–293.

Lewis, D. (1994) Humean supervenience debugged, *Mind* **103**, 473–490.

Lewis, D. (1999) Why conditionalize?, in *Papers in Metaphysics and Epistemology*, Cambridge University Press, pp. 403–407.

McGee, V. (2000) To tell the truth about conditionals, *Analysis* **60**, 107–111.

Mellor, D. H. (ed.) (1990) *Philosophical Papers: F. P. Ramsey*, Cambridge University Press.

Pollock, J. L. (1981) Indicative conditionals and conditional probability, in (Harper et al., 1981), pp. 249–252.

Ramsey, F. P. (1929) General propositions and causality, in (Mellor, 1990), pp. 145–163.

Skyrms, B. (1988) Conditional chance, in (Fetzer, 1988), pp. 161–178.

Skyrms, B. (1990) *The Dynamics of Rational Delibaration*, Harvard University Press.

Stalnaker, R. (1970) Probability and conditionals, *Philosophy of Science* **37**, 64–80.

Stalnaker, R. and R. Jeffrey (1994) conditionals as random variables, in (Eells and Skyrms, 1994), pp. 31–46.

Teller, P. (1973) Conditionalization and observation, *Synthese* **26**, 218–258.

van Fraassen, B. C. (1976) Probabilities of conditionals, in W. L. Harper, R. Stalnaker and G. Pearce (eds.), *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, The University of Western Ontario Series in Philosophy of Science, Vol. 1, D. Reidel, pp. 261–308.

van Fraassen, B. C. (1981) A temporal framework for conditionals and chance, in (Harper et al., 1981), pp. 323–340.

Woodward, J. (2001) Probabilistic causality, direct causes and counterfactual dependence, in (Galavotti et al., 2001), pp. 39–63.

Department of Linguistics,
Northwestern University,
2016 Sheridan Road,
Evanston, IL 60208, USA
e-mail: kaufmann@northwestern.edu