Intervention and the Probabilities of Indicative Conditionals[1]

Michael Zhao
(In *Journal of Philosophy* 112 (9):477–503.)
(penultimate draft; please cite published version)

A popular theory of the probabilities of indicative conditionals, the *Adams thesis* (sometimes called the *Stalnaker thesis*), says that the probability of "if p then q" is $P(q|p)$, the probability of q conditional on p.[2] A few anomalies have cropped up in the literature in the last few decades, however: examples in which the intuitively correct probability of a conditional is different from what the Adams thesis predict. If ordinary speakers are not making a mistake in evaluating the conditionals this way, then the existence of these cases implies at least that there sometimes is a second, more salient reading of indicative conditionals whose probability the Adams thesis does not predict.

This paper will discuss and give a theory of this phenomenon. I take as my point of departure a theory recently proposed by Stefan Kaufmann,[3] which purports to predict when a second reading of an indicative conditional exists and its intuitively correct probability. In the first section, I rehearse this theory and generalize it. I argue, however, that Kaufmann's theory is unsatisfactory, both because the conditions it gives on the existence of a second reading are extensionally inadequate and because it leaves something important unexplained. I then offer an original theory that (1) gives correct conditions on the existence of a second reading, (2) gives an account of the process of reasoning used to evaluate this reading, and (3) ties the psychological story with a formula for the probability of conditionals on this reading. I close by offering an explanation of why a second reading of indicative conditionals should exist at all.

1 The phenomenon: locality

1.1 Examples

First, one example of the phenomenon:

[1] I'd like to thank Max Barkhausen, Hartry Field, Daniel Fogal, Dan Waxman, Robbie Williams, an anonymous referee at *The Journal of Philosophy*, and especially Cian Dorr for invaluable comments and discussion.

[2] Some philosophers prefer not to use the terms "Adams thesis" and "Stalnaker thesis" interchangeably, taking the latter to be the version that identifies the *probability* of a conditional with a conditional probability, and the former to identify its "*assertibility*" (or something else less than genuine probability) with the conditional probability.

While I'm on this point, I want to flag a few things. The field of conditionals is laden with contentious issues: for example, whether conditionals have truth-values, and whether they have genuine probabilities or only properties like degrees of assertibility. The triviality results of Lewis and others suggest either that (1) conditionals do not express full-blooded propositions, so cannot be the objects of genuine probabilities and have only degrees of assertibility; or that (2) conditionals do express propositions (for example, "$p \to q$" might just express the truth-functional $p \supset q$), so that they have genuine probabilities, and that the intuitively correct probabilities of conditionals are not their real probabilities, but only something related. For the purposes of this paper, I'll remain neutral on this issue: I'll talk about probabilities of conditionals, but you can take this just to mean degrees of assertibility or otherwise ersatz probability.

[3] Kaufmann (2004).

> *Subway*. You're waiting at the subway stop near your apartment for a train into the city. Standing on the platform, you suddenly realize that you're not certain which day of the week it is. But you have an impression that it's the weekend, which causes you to be *pretty* sure (about 2/3 sure) that it is. You know that on weekends, trains arrive quite infrequently (say, every half hour), but are usually pretty empty; on weekdays, trains arrive once every five minutes and are generally crowded.

How likely is it that if a train arrives in the next five minutes, it will be crowed? The intuitive answer seems to be: not very. After all, you have a highish credence that it's the weekend, in which case the trains are usually pretty empty; you know that there's a chance you're wrong, in which case the train will be pretty crowded, but the chance of that is fairly small. So a low probability assignment to the conditional, "If a train arrives in the next five minutes, then it will be crowded" seems right.

If the Adams thesis is correct, however, then the conditional has a middle-ish or high probability. For suppose that a train does show up in the next five minutes. It's possible that you got lucky and got to the station just in time; but it's more likely that your impression about which day it is was just wrong, and that it's actually a weekday,[4] in which case the train probably will be pretty full.

Subway is an example that, alongside the reading of indicative conditionals that the Adams thesis predicts, there is a second reading that is often more salient. (In fact, there will sometimes be *multiple* other readings). Let's call this phenomenon *locality*, for reasons that will become clear later. Recognition of locality goes back at least as far as Pollock (1981), who gave a (somewhat involved) scenario in which the probability of an indicative conditional, on the salient reading, is not equal to the conditional probability; Van McGee and Stefan Kaufmann have also noticed the phenomenon, and have furnished their own examples.

Kaufmann's example is the following:

> *Bags and balls.* Suppose there are two bags, X and Y, with the following profiles: X has 10 red balls, 9 of which have black spots, and 2 white balls, which are unspotted; Y has 10 red balls, 1 of which has black spots, and 50 white balls (unspotted). One of the bags is in front of you; you have .75 credence that it is bag Y.

What is the probability that, if you draw a red ball (*r*), it will have black spots (*b*)? Kaufmann ran an informal survey on people's intuitions, and found that they mostly assigned a "low" value to

[4] Here's the Bayesian justification:

$$P(\text{weekend}|\text{5 mins}) = \frac{P(\text{5 mins}|\text{weekend})}{P(\text{5 mins})} P(\text{weekend}).$$

We know that $P(\text{weekend}) = 2/3$, $P(\text{5 mins}|\text{weekend}) = 1/6$, and $P(\text{5 mins}|\text{weekday}) = 1$, so the right-hand side becomes

$$\frac{1/6}{1/6 \times 2/3 + 1 \times 1/3} \times 2/3 = 1/4.$$

this probability. The reasoning, formalized, might be as follows:

$$
\begin{aligned}
P(r \to b) &= P(r \to b|x)P(x) + P(r \to b|y)P(y) && \text{(total-probability law)} && (1)\\
&= P(b|x \wedge r)P(x) + P(b|y \wedge r)P(y) && && (2)\\
&= .9 \times .25 + .1 \times .75 = .3. && && (3)
\end{aligned}
$$

Kaufmann writes that "the transition [from (1) to (2)] is not warranted by an 'official' rule of the probabilistic calculus. Instead, it is motivated by the intuition that the [following] equality should hold, at least in this case, if not in general: $P(r \to b|x) = P(b|r \wedge x)$."[5]

Contrast this with the probability that the ball will have black spots conditional on its being red:

$$P(b|r) = P(b|r \wedge x)P(x|r) + P(b|r \wedge y)P(y|r). \qquad \text{(total-probability law)} \qquad (4)$$

We can apply Bayes's rule to figure out $P(x|r)$ and $P(y|r)$; doing the math, $P(b|r)$ comes out to .6. Hence the conditional probability is larger than the probability of the conditional, evaluated this way.

Take another example, due to McGee:[6]

> *Sherlock Holmes.* Suppose Murdoch is found dead. The police believe it was an accident, so you (deferring to their judgment) do too. But later, someone you think is Sherlock Holmes tells you that Brown killed him and that, in any case, if Brown didn't kill Murdoch, then someone else did.

Since you think Holmes is telling you this, and that Holmes is almost certainly right, it seems probable to you that if Brown didn't kill Murdoch, then someone else did. Nonetheless, if you discover that Brown didn't kill Murdoch, then your confidence that the man telling you this is Holmes drops, and you revert to your original low credence in Murdoch's having been killed. So the probability that someone else killed Murdoch conditional on that Brown didn't kill him is low. Again, the Adams thesis seems to predict a reading that's not the most salient.

These examples don't show that the Adams thesis is unqualifiedly *false*, that it does not give the probability of indicative conditionals on *some* legitimate reading of them. Indeed, there's a clear reading of "if p then q" on which the Adams thesis *is* correct: we imagine that we have *observed* that p, and are asking whether q. (In the subway example, we pretend that we've just seen the train arriving, and are asking on this basis whether it's going to be full.) Since updating on an observation proceeds via conditionalization, the probability "if p then q," on this reading, will just be $P(q|p)$.[7] What the examples do show is that this reading isn't always the first that pops into speakers' minds when they see the conditional.

[5] p. 585, notation adjusted. (The principle that, in general, $P(p \to q|r) = P(q|p \wedge r)$ is sometimes called the *generalized Adams thesis*, and can be derived from the ordinary Adams thesis. So this example is an example that the operations of applying the total-probability law and the Adams thesis do not commute.)

[6] MacGee (2000)

[7] The idea that updating proceeds via conditionalization is controversial; see Christensen (1992) for a nice discussion. But the problems for conditionalization do not affect cases, such as the ones here, in which we can unproblematically suppose that the information gained can be represented as a proposition, our credence in which rises to one as a result of the observation.

1.2 The partition theory

Kaufmann explains the difference between the two probabilities as follows. In each of the examples, along with the antecedent and consequent, we have a third proposition that is probabilistically related to (i.e., confirms or disconfirms) the antecedent and consequent, even though causally independent from the antecedent.[8] In the subway example, learning that it's a weekday (the third proposition) confirms both that the next train will show up soon (antecedent) and that it'll be pretty full (consequent). In the bags-and-balls case, knowing which bag we are drawing from (third proposition) changes our credence both that the ball is red (antecedent) and that it has black spots (consequent). In the Sherlock Holmes case, knowing that it was Holmes who spoke to us (third proposition) increases our credence both that Brown killed Murdoch (antecedent), and that someone else killed Murdoch (consequent). Call this third proposition r.

Kaufmann hypothesizes that, in the context of the partition $\{r, \neg r\}$, there are two ways of reading the conditional "if p then q." One is a *global* reading, on which the probability of the conditional is identical to the conditional probability:

$$P(p \rightarrow_g q) = P(q|p) \qquad \text{(Adams thesis)} \quad (5)$$

$$= P(q|p \wedge r)P(r|p) + P(q|p \wedge \neg r)P(\neg r|p). \qquad \text{(total-probability law)} \quad (6)$$

The other is the *local* reading:

$$P(p \rightarrow_l q) = P(p \rightarrow q|r)P(r) + P(p \rightarrow q|\neg r)P(\neg r) \qquad \text{(total-probability law)} \quad (7)$$

$$= P(q|p \wedge r)P(r) + P(q|p \wedge \neg r)P(\neg r). \quad (8)$$

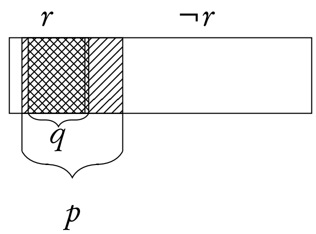
The difference between corresponds to a difference in when one applies the total-probability law: applying the Adams thesis before leads us to replace the marginal probabilities $P(r), P(\neg r)$ in (8) with conditional probabilities $P(r|p), P(\neg r|p)$ in (6). Since, on the global reading, one applies the Adams thesis before applying the partition, the probability is independent of the partition. Not so on the local reading: different partitions may generate different values for $P(p \rightarrow_l q)$. For short, I'll refer to the global and local reading of conditionals simply as "local conditionals" and "global conditionals."

It's not clear if Kaufmann takes probabilistic dependence and causal independence to be conditions on the *availability* of the local reading or just *its having a different probability* from that of the global reading. I take it that Kaufmann intends the probabilistic dependence conditions to be necessary for $P(p \rightarrow_l q)$ to be different from $P(p \rightarrow_g q)$ at all: If r is independent of p, then $P(r|p) = P(r)$, and clearly (6) is identical to (8). And if r is independent of q (conditional on p), then $P(q|p \wedge r) = P(q|p)$; it follows that $P(q|p \wedge r) = P(q|p \wedge \neg r) = P(q|p)$, in which case both (6) and (8) just reduce to $P(q|p)$.[9] Causal independence, on the other hand, seems to be a condition on the availability of the local reading; I'll discuss it later.

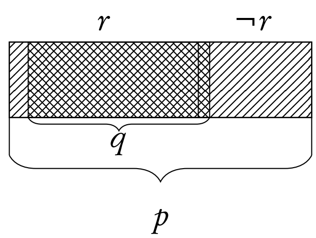
[8] I take it that, by "causally independent," Kaufmann means that the antecedent has no causal influence on the third proposition, rather than the stronger claim that neither has any causal influence on the other. I also take it that Kaufmann has the *de re* reading of the causal-independence claim in mind: that, *of the next ball that I draw*, the bag that I draw it from has no causal influence on its color.

[9] As this shows, we actually need that r be probabilistically related to *q conditional on p*, not just unconditionally.

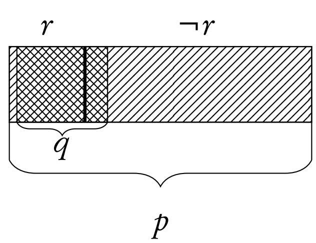
We can represent the whole thing with a diagram, where the probability of some proposition a is represented by the proportion of total space taken up by a-space:



The two readings of "$p \to q$" correspond to two ways of removing the white (non-p-) space and renormalizing. On the global reading, we preserve the proportion of *shaded* r-space to *shaded* $\neg r$-space, stretching out everything shaded by the same amount:



On the local reading, on the other hand, we preserve the proportion of *total* r-space to *total* $\neg r$-space, stretching out the shaded r- and $\neg r$-regions by different amounts (I've made the line between the r- and $\neg r$-regions bold to emphasize that the update doesn't change the size of those regions):



The probability of "if p then q" on each reading is given by the proportion of p-space that's also q-space in each diagram; it's pretty clear that there are cases (as above) where they differ.

Kaufmann takes the two readings to correspond to two ways of reasoning from p. The reasoning that gets us the global reading includes a hypothetical abductive inference from the antecedent to the third variable: in the bags-and-balls case, the best *explanation* for my having drawn a red ball is that this is bag X after all, which *confirms* the hypothesis that this is bag X, and this lowers my credence in the ball's having a black spot. In evaluating the local reading, we ignore this step: we revise our credence in the consequent, but hold fixed our credence in the third proposition. This is why, whereas we have *conditional* probabilities $P(r|p)$ and $P(\neg r|p)$ in equation (6), we have *unconditional* ones $P(r)$ and $P(\neg r)$ in (8). So Kaufmann has a nice story to tell about why the way in which we evaluate local conditionals means that equation (8) gives intuitively correct probabilities for them.

There's also a nice connection, although Kaufmann seems not to notice this, between the line of reasoning that he takes people to employ in evaluating local conditionals and the condition he sets on the availability of a local reading. Since the crucial step in evaluating the local conditional that differentiates it from evaluating a global conditional is that we do not make an abductive inference from p to r, then in order for the local reading to exist at all, r must be such that we can (but do not, when evaluating the local conditional) draw an abductive inference from p to it. Since abductive inferences run from data to hypotheses, and since an effect cannot be a hypothesis that explains a cause, it is impossible to draw an abductive inference from cause to effect: we cannot *abduce* from "there's fire" to "there's smoke." So if r causally depended on p, then we just couldn't run an abductive inference from p to r, and a local reading could not arise. And this means that causal independence of the third proposition from the antecedent has to be a condition on the existence of a local reading.

Equation (8) gives the probability of the local conditionals in the above examples, just as the Adams thesis gives the probability of the global conditionals. Of course, we need to generalize the formula: in the examples above, we had only two-piece partitions (e.g., {bag X, bag Y}), but we can imagine situations in which the antecedent and consequent depend on a variable that can take $n > 2$ possible values (e.g., if there were three bags instead of two), or on several different propositions (e.g., if each bag had a left and right pocket, each having different proportions of red and spotted balls). A suitable generalization is that whenever a partition $\{a_i\}$ generates a local reading,[10] then the probability of the conditional "if p then q" on that reading is

$$P(p \rightarrow_{\{a_i\}} q) = \sum_i P(q|p \wedge a_i)P(a_i). \tag{9}$$

Kaufmann's condition on the availability of $P(p \rightarrow_l q)$, that the third proposition be causally independent of p, becomes the condition that each member a_i of the partition be causally independent of p. (Similarly, his condition on the distinctness of $P(p \rightarrow_g q)$ and $P(p \rightarrow_l q)$, that the third proposition be probabilistically related to the antecedent and consequent, becomes the requirement that, for some member a_i of the partition, $P(p|a_i) \neq P(p)$ and $P(q|p \wedge a_i) \neq P(q|p)$.) Call the theory that conjoins this condition on the availability of a local reading with the claim that equation (9) gives the probability of that reading the *partition theory*. Perhaps a further condition is that, in order for the reading to be salient, the partition must be a salient one. I'll refer to equation (9) as the *partition-theory equation*.

One handy feature of the partition theory is that, when the partition is $\{\top\}$, the right-hand side of (9) becomes $P(q|p)$, which is just the probability on the global reading. (I take it that $\top$ and $\bot$ are causally independent of any antecedent p, so the condition on availability is trivially satisfied.) So the global reading arises as a limiting case of the local readings.

When multiple partitions satisfy these conditions, we get multiple local readings, each with perhaps a different probability. Here's a (slightly intricate) example:

[10] When there are multiple variables $R_1, \ldots, R_n$ that the antecedent and consequent depend on, let each a_i be something of the form $r_1 \wedge \ldots \wedge r_n$, where each r_j is a possible value for R_j.

> *Political betting arena*. You're a member of an online political betting arena. You've taken bets on whether the following three events will occur: (1) the current ruling party will win the election next year (e), (2) national GDP will have grown by next year (g), and (3) military intervention in the Middle East will have ended by next year (m). The bet is such that you'll win $1,000 ($w$) if all three events occur, and nothing if fewer than all three do.
>
> The three events, of course, aren't probabilistically independent. The chance that the ruling party wins the election given that GDP grows and military intervention ends is .9; if the GDP grows but military intervention is ongoing, or if GDP does not grow but military intervention ends, then the chance that the ruling party wins is .3; if GDP does not grow and military intervention is ongoing, then the ruling party has a .1 chance of winning.
>
> The chance that GDP increases is 3/4 and the chance that military intervention ends is 1/2. (Suppose that these two events are causally independent.)

Now, consider the indicative conditional, "If the ruling party wins the elections next year, then I'll win $1,000." We can imagine a range of situations such that a different probability assignment seems intuitively right in each.

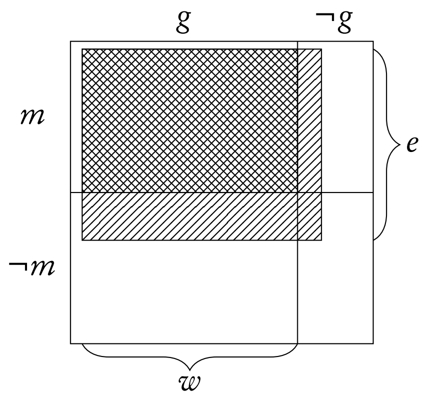
First, suppose that your credence in GDP growth is the result of newly-received economic data. In this situation, you might reason as follows: "If the ruling part re-wins election, then I'll win the money iff both GDP grows and military intervention stops. Now, I've just found out that it's fairly likely that GDP will grow. So if the ruling party wins, then it must be fairly likely that military intervention will have ended. So the chance that if the ruling party wins, then I'll win the $1,000 is the product of two fairly likely events, which comes out to something with middle-ish probability."

Second, imagine another situation, in which you've just heard the president announce that a new plan has been made to withdraw forces from the Middle East, which bumps up your credence that intervention will end to 1/2. Here, you might reason: "It's about even whether military intervention will end. So if the ruling party wins, then it must be very likely that GDP will have grown. So the chance that if the ruling party wins, then I'll win the money is the product of something with middle-ish probability with something with high probability, which comes out to something with middle-ish probability."

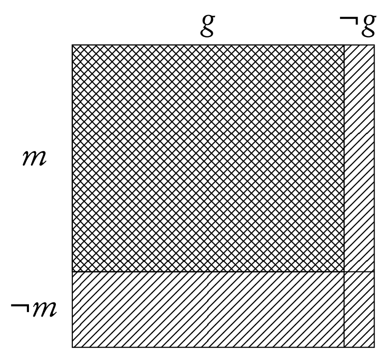
Or imagine a third situation, in which you get the economic and military information pretty much simultaneously. Here, you might reason: "It's fairly likely that GDP will increase, but only 50-50 that the intervention will have ended, so fairly unlikely that both will happen. So it's fairly unlikely that if the ruling party wins, then I'll win the money."

Finally, imagine a situation in which your credences on the matter have been fixed for a while. Here, you might reason, "If the ruling party wins the election, then it must be quite likely that both GDP will have increased and military intervention will have ended. So it's quite likely that if they win the election, then I'll win the money."

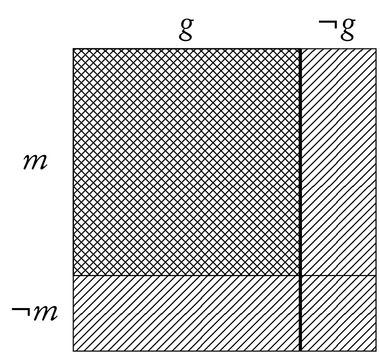
Diagrammatically, we can represent the scenario as follows:



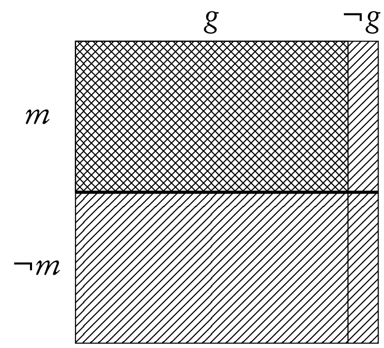
The four readings correspond to four diagrams, each generated from the original by a different way of eliminating the non-e- space (white space) and re-normalizing. We get the global reading by preserving the proportions of shaded $g \wedge m$-, $g \wedge \neg m$-, $\neg g \wedge m$-, and $\neg g \wedge \neg m$-space in the original diagram:



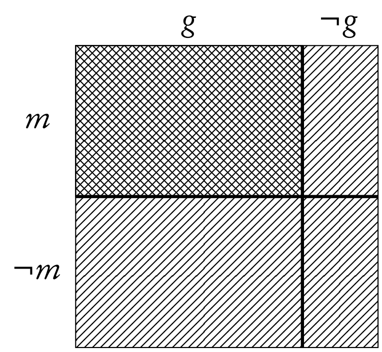
The $\{g, \neg g\}$ reading corresponds to preserving the proportion of g-space to $\neg g$-space, both shaded and unshaded, and preserving the proportion of *shaded* m-space to *shaded* $\neg m$-space:



(I've made the line between g and $\neg g$ extra thick to indicate that we don't change the proportion of total g-space to total $\neg g$-space from that in the original diagram.) Similarly, the $\{m, \neg m\}$ reading corresponds to preserving the proportion of m-space to $\neg m$-space, and preserving the proportion of shaded g-space to shaded $\neg g$-space:



Last, the $\{g \wedge m, g \wedge \neg m, \neg g \wedge m, \neg g \wedge \neg m\}$ reading corresponds to preserving the proportion of space, shaded and unshaded, of each piece of the partition:



The partition theory accounts for the existence of these four readings as follows. In each situation, a different partition is salient to us: $\{g, \neg g\}$, $\{m, \neg m\}$, $\{m \wedge g, m \wedge \neg g, \neg m \wedge g, \neg m \wedge \neg g\}$, and finally (as a degenerate case) $\{\top\}$. Furthermore, each satisfies the availability-condition on a local reading, since whether GDP will have grown, military intervention will have ended, and conjunctions of the two are causally independent of whether the ruling party will win the election. So each of the situations corresponds to a different instance of equation (9); in all, we get:

$$P(e \rightarrow_{\{\top\}} w) = P(w|e) \quad = .68 \tag{10}$$

$$P(e \rightarrow_{\{g,\neg g\}} w) = \sum_{g'=\pm g} P(w|e,g')P(g') \quad = .56 \tag{11}$$

$$P(e \rightarrow_{\{m,\neg m\}} w) = \sum_{m'=\pm m} P(w|e,m')P(m') \quad = .45 \tag{12}$$

$$P(e \rightarrow_{\{g\wedge m, g\wedge\neg m, \neg g\wedge m, \neg g\wedge\neg m\}} w) = \sum_{\substack{g'=\pm g \\ m'=\pm m}} P(w|e,g',m')P(g',m') \quad = .38 \tag{13}$$

You can check that these probabilities accord with the intuitive probabilities in each of the cases.

That the partition theory predicts the sensitivity of the probabilities to which partition is salient is a point in its favor. The theory is not, however, without defects. First, some partitions that satisfy Kaufmann's availability-condition generate strange readings even when salient, so that condition cannot be enough for there to be a natural local reading of a conditional. Take the following example:

> *Socks.* You have 100 pairs of socks in a drawer. 15 of them are red (r), of which 10 are cotton (c) and 5 are nylon (n); 9 of the cotton socks are striped (s) and one is solid; all of the nylon socks are solid. The remaining 85 socks are black, all of which are nylon and solid.

What's the probability of the conditional, "If you pick a red sock, then it will be striped"? Intuitively, something like 9/15, the proportion of red socks that are striped. But suppose that the partition {cotton, nylon} becomes salient. We can suppose that whether a sock is cotton or nylon is causally independent of whether it is red, so according to the partition theory, the partition generates a reading of the conditional on which its probability is

$$P(r \rightarrow_{\{c,n\}} s) = P(s|r \wedge c)P(c) + P(s|r \wedge n)P(n) \tag{14}$$

$$= 9/10 \times 10/100 + 0 \times 90/100 = .09. \tag{15}$$

But a reading that assigns this very low probability just seems bizarre. It seems hard to see why someone would read the conditional this way unless somehow primed.

Second, a perfectly natural local reading of an indicative conditional can arise even when Kaufmann's availability-condition is not satisfied. Consider the following example:

> *Disease.* There's a serious disease such that the death rate without treatment is .9, but the death rate with treatment is .1. Because the death rate without treatment is so high, doctors strongly try to avoid false negatives; but this also leads to many false positives, so that only 1/2 of those treated actually have the disease. Among those not treated, the incidence of the disease is negligible.

Suppose you're 1/3 sure that patient X is being treated for the disease, and that your credence didn't arise from any previous credence in X's having the disease. (Suppose, for example, a doctor told you that exactly one of X, Y, and Z was being treated.) What's the probability that if patient X has the disease, then he'll die? Intuitively, fairly high. This seems to be a straightforward application of (9) on the partition {treated, not treated}:

$$P(h \rightarrow_{\{t,\neg t\}} d) = P(d|h \wedge t)P(t) + P(d|h \wedge \neg t)P(\neg t) \tag{16}$$

$$= .1 \times 1/3 + .9 \times 2/3 = 19/30. \tag{17}$$

The global reading, on the other hand, has a probability of

$$P(h \rightarrow_{\{\top\}} d) = P(d|h) \tag{18}$$

$$= P(d|h \wedge t)P(t|h) + P(d|h \wedge \neg t)P(\neg t|h) \tag{19}$$

$$= .1 \times P(t|h) + .9 \times P(\neg t|h). \tag{20}$$

Using Bayes's rule, we get $P(t|h) \approx 1$ and $P(\neg t|h) \approx 0$,[11] so that the last line works out to about .1. So a natural local reading exists, since the intuitively correct probability of the conditional is different from the probability of the global reading. Yet here, we can suppose that whether the patient is treated *does* causally depend on whether he has the disease; this means that the causal independence of the partition from the antecedent is not necessary for a local reading to arise either.

So far, these are counterexamples to the claim that the partition theory gives correct availability-conditions on a local reading, rather than to the claim that the partition-theory equation assigns the right probabilities to available local readings. Indeed, I take the partition-theory *equation* to be in good shape, at least for fairly simple cases, and I want to preserve it even if other aspects of the partition theory fail. But there is another reason for dissatisfaction with the theory. Recall what Kaufmann says about the line of reasoning we employ in evaluating local conditionals: when a partition $\{a_i\}$ is salient and satisfies the availability-condition, we evaluate the probability of the resulting local conditional "$p \rightarrow_{\{a_i\}} q$" by updating our credences in propositions other than the a_is in accordance with the conditional probabilities, but preserving our original credences in the a_is. What I take to be unsatisfactory about this isn't that the proposed mode of reasoning is *false*, but that it leaves something important *unexplained*: namely, *when* a partition $\{a_i\}$ is salient. It would be nice if we could have an account of the mode of reasoning people employ in evaluating local conditionals that *implies* when a partition is salient, rather than *requires* some prior condition on salience.

A defender of the partition theory might respond that maybe the only thing that we can say about when a partition becomes salient to an agent is that this is subject to all kinds of psychological quirks, so that there's no general story to tell; at any rate, to expect such a story seems to set the bar for a satisfactory theory of conditionals unreasonably high. But I think we can do better, and I aim to.

An improvement on the partition theory, then, should do a few things: first, give correct conditions on *which* partitions generate natural local readings; second, say what kind of *reasoning* we use in evaluating local conditionals, in a way that explains when a partition becomes salient; and third, say how the claim that equation (9) gives intuitively correct probabilities for local conditionals follows from the claim that people are using this kind of reasoning to evaluate them. In the next section, I'll propose a theory that does these things.

1.3 Does the local reading involve a fallacy?

Before going on, I want to address the worry that, in evaluating indicative conditionals on the local reading, people are committing a fallacy. If they are, then such a fallacy would have a lot in common with the probabilistic fallacies that (as psychologists stress) people are so prone to commit. Such a worry, if valid, allows the proponent of the Adams thesis to maintain his position that the thesis predicts the only legitimate reading of indicative conditionals, the global reading, since the local readings all involve some kind of error in reasoning.

The specific charge is as follows. The local reading arises from the same process of reasoning

[11] $P(t|h) = P(t)P(h|t)/P(h) = P(t)P(h|t)/(P(h|t)P(t) + P(h|\neg t)P(\neg t))$. Since $P(h|\neg t) \approx 0$, this last bit comes out to ≈ 1.

that gives rise to the global reading: we pretend to observe the antecedent, and figure out what our new credence in the consequent should be. But in evaluating the conditional in the way that generates the local reading, we commit an error in carrying out this line of reasoning: we *forget* the fact that the antecedent is probabilistically related to the third proposition (or, in the more general case, members of the partition), do not update our credence in the proposition, and end up using that *pre*-update credence in calculating the post-update credence in the consequent. In the subway case, for example, we forget the fact that the next train's arriving within a few minutes confirms its being a weekday; this makes our credence in its being a weekday irrationally low, which in turn makes our credence in the train's being crowded irrationally low as well. (Things look especially bad if the partition theory is right, since the line of reasoning it ascribes to us includes a step explicitly saying we do *not* update certain credences.)

I think, however, that we can say a few things to dispel the idea that people are committing a fallacy here. First, if the explanation of the fallacy is that people are *forgetting* that the antecedent is probabilistically related to the third proposition, then simply bringing this fact to their attention should induce them to revise their judgment. If this is correct, then in the subway case, simply telling people that *the next train arrives in a few minutes* increases the probability of *it's a weekday* should get them to raise their credence in the conditional. Intuitively, however, it doesn't seem that this is the reaction that people would have; it seems that, even after attending to the probabilistic-dependence fact, people would still be disposed to evaluate the conditional in the same way. We can imagine someone saying, "Sure, if the next train arrives in a few minutes, then it's probably not the weekend. But it still seems true that, given that we're pretty sure it *is* the weekend, if the train arrives in a few minutes, then it'll probably be pretty empty." Of course, that people would have this reaction is an empirical claim, but it seems plausible.

Second, if the charge is that the local reading arises out of an *incorrect* way of carrying out the process of pretending to observe the antecedent, then the charge would lose some of its plausibility if we could identify another process of reasoning such that the local reading arises as a *correct* way of carrying out that process. This gives us another reason to provide a different psychological story about how people reason about local conditionals, which is part of the goal of the next section.

2 The epistemic-intervention theory

2.1 When do we get a local reading?

On to the positive proposal. I'll start with the first task, that of giving conditions for a partition to generate a natural reading of a conditional. Consider that, for each of the examples we've given, it sounds intuitively right to say, "whether p depends on whether r," where p is the antecedent and r the third proposition: it sounds right to say, "whether the next train will arrive soon depends on whether it's a weekday," "whether the ball is red depends on which bag it came from," and "whether Brown killed Murdoch depends on whether it was Holmes who said so." Furthermore, the dependence-claims in the other direction are infelicitous: in the given setting, it seems unnatural to say, "whether it's a weekday depends on whether the next train will arrive soon," "which bag the ball came from depends on whether it's red," and "whether it was Holmes who said that Brown killed Murdoch depends on whether Brown killed Murdoch." One *could* say these things, and say them sensibly, but it would involve a shift in cognitive perspective. The take-away of these examples is that, in any such given perspective, "depends" is asymmetric.

Furthermore, such a condition is not met in the socks cases, our example of the absence of a natural local reading: without any special priming, it sounds odd to say that whether the sock is red depends on whether it's cotton or nylon. So it's plausible that it's both necessary and sufficient for a partition to generate a local reading when the antecedent *depends*, in whatever sense is at work in these examples, on some member of the partition. The task now is to find the right sense of "depends."

To go through some candidates:

1. Mere *probabilistic dependence* isn't enough. Probabilistic dependence is symmetric: if $P(x|y) > P(x)$, then $P(y|x) > P(y)$.[12] The examples above, however, show that dependence in our sense is usually asymmetric.

2. Nor does "depends" here just mean *causal dependence* between the propositions themselves: causal relations, after all, are absent in Kaufmann's and McGee's examples.

3. Kaufmann mentions that, in evaluating a global conditional, we make an abductive inference from the antecedent to the third proposition. Talk about abduction suggests another natural idea: that the third proposition must figure into what the speaker takes to be an *explanation* (perhaps the best explanation) of the antecedent. That is, the antecedent must *explanatorily depend* on the third proposition. But this doesn't seem like a necessary condition. If it were, then in the Sherlock Holmes case, we'd be committed to saying that the proposition that it was Sherlock Holmes who spoke to us explains the proposition that Brown killed Murdoch; but it'd be totally bizarre to give this explanation.[13]

4. Nonetheless, it *is* true that *our credence* that the man is Holmes causally explains why *our credence* that Brown killed Murdoch is as high as it is: we *start out* with a credence that the man is Holmes, and can use it (together with our credence that Brown killed Murdoch given that Holmes said so) to *work out* the rational credence that Brown killed Murdoch. Similarly, in the bags and balls case, we start out with a credence that the bag in front of us is X, from which (together with the relative frequencies) we can work out the rational credence that the next ball drawn is red. And in the political betting case, we start out with credences in GDP growth and military withdrawal, and can use those to figure out the probability of the election results. To generalize, in these cases we start out with credences over the partition $\{a_i\}$, and conditional credences in the antecedent p conditional on members of the partition; we can then work out what our credence in p

[12]Proof: By Bayes's rule, $P(y|x) = P(x|y)P(y)/P(x)$; since $P(x|y) > P(x)$, $P(x|y)P(y)/P(x) > P(y)$. (The usual proviso about non-zero probabilities applies.)

[13]Some philosophers, e.g., Van Fraassen in *The Scientific Image*, emphasize that explanations must be given relative to a context, and that, by changing contexts, an *outré* explanation may become completely appropriate. But it seems difficult to say what context could make "that Holmes said so explains why Brown killed Murdoch" appropriate, and why readers would evaluate the conditional in *that* context.

should be by applying the total-probability law, $P(p) = \sum_i P(p|a_i)P(a_i)$.[14] In the cases where the local reading is natural, then, our credences over the partition partly cause our credence in the antecedent. So this causal relation between the relevant credences seems at least a necessary condition for there to be a natural local reading.

Furthermore, such a condition fails in the socks case: there, we're told the probability of the antecedent off the bat. So it seems plausible that it's both necessary and sufficient for a partition to generate a local reading when our credences over the partition partly causally explain our credence in the antecedent. Or, to put it in terms of dependence, a partition generates a local reading when our credence in the antecedent causally depends on (perhaps among other things) our credences over the partition. I'll say that *p epistemically depends on q* (for an agent) iff the agent's credence in q (perhaps together with other things) causally explains his credence in p.

Epistemic dependence, then, is sensitive to the order in which we form our credences in the relevant propositions. You can verify that the following set of probabilities are equivalent to the ones we were given in Kaufmann's example:

$$\begin{aligned} &P(r) = 1/3 \\ &P(x|r) = 5/8 && P(y|r) = 3/8 \\ &P(b|x \wedge r) = 9/10 && P(b|y \wedge r) = 1/10 \end{aligned} \tag{21}$$

But presented with *these* probabilities, now our credence that the bag is X or Y causally depends on our credence that the ball is red. As a result, in this setting, it seems natural to assign a middle-ish or high-ish probability to the conditional, "If I draw a red ball, it will have a black spot."

The direction of causal dependence in our credence can sometimes be the *opposite* of the direction of causal dependence in the world; in these cases, we get local readings where the antecedent *epistemically* depends on the third proposition, even though the third proposition *causally* depends on the antecedent. This is what was going on, for example, in the disease case: there, whether the patient has the disease *epistemically* depends on whether he is being treated (since our credence that he is being treated causally explains our credence in his having the disease), even though whether he is being treated *causally* depends on whether he has the disease. This explains why, contrary to what Kaufmann says, the third proposition need not be causally independent from the antecedent in order for a local reading to arise. I think Kaufmann takes causal

[14] Some might doubt that conditional credences play any explanatory role: $P(q|p)$ is usually just *defined* as the ratio of two unconditional probabilities, $P(p \wedge q)$ and $P(p)$, after all. But Edgington (1995) has stressed that, although $P(q|p)$ is usually so defined, we typically arrive at our credence in q conditional on p without going through any initial credence over the conjunction q and p, or in p: our credence in heads given that the coin is fair is 1/2, although we may not have a prior view on the probability that the coin is fair, much less on the probability that the coin is fair and will land heads. As Edgington writes, "The natural order of human thinking is not the same thing as the most elegant order of mathematical exposition. Humans are not endowed with complete belief-distributions over the finest partitions they need to consider. They need to work out some degrees of belief (as the need arises) in terms of others which are more readily accessible. $b(B|A)$ can be accessible en route to $b(A\&B)$, and can be accessible when $b(A)$ is not" (p. 266–7). So this doesn't seem like good reason to doubt that conditional credences are occasionally explanatory.

independence to be a necessary condition because it's hard to think of cases where we start off with knowledge of the inverse probabilities (of some cause on an effect) rather than the forward ones.

I want to build another condition into the notion of epistemic dependence. Any talk about causation in the context of belief-formation is going to run into the problem of deviant causal chains; here, the problem is that an agent's credence in r can be an indirect cause of his credence in p in a way that should not give rise to a local conditional with p as antecedent. As one example, suppose that Alex discovers that his wife is having an affair, and that this piece of information shocks him so much that he faints, hitting his head in precisely the right way to change his credence in its raining tomorrow. So his credence in his wife's having an affair indirectly caused his new credence in its raining tomorrow, even though it would be absurd to think that this generates a reading of any indicative conditional, "If it rains tomorrow, then ..." not available to Alex before. We need, then, an anti-deviancy condition to rule out such causal chains as constituting genuine instances of epistemic dependence. Obviously it's outside the scope of this paper actually to give such a condition, but we can nevertheless help ourselves to one.

To sum up: the partition $\{a_i\}$ generates a local reading of "if p then q" in a given setting when, in that setting, p epistemically depends on $\{a_i\}$, i.e., when our credences over $\{a_i\}$ (perhaps among other things) causally explain our credence in p in a non-deviant way.

2.2 The connection to counterfactuals and causation

I want to suggest, taking into account this condition on the naturalness of certain readings, a thesis about the reasoning involved in evaluating a local conditional: (1) assign probability 1 to the antecedent, (2) hold fixed our credences in every proposition that the antecedent epistemically depends on, (3) update our credences in other propositions in accordance with the conditional probabilities, and (4) figure out the rational credence in the consequent. (So notice that, on this account, the partition isn't an independent input into this process of reasoning, but instead just given by the dependencies in our credences: $\{a_i\}$ just includes all of the propositions that the antecedent epistemically depends on.)

As it turns out, this way of characterizing the distinction between global and local indicative conditionals reveals some interesting connections between indicative conditionals and counterfactuals. (Kaufmann mentions this connection too, although makes much less of it than he can.) Philosophers have known since Lewis's work on the topic[15] that counterfactuals (or *subjunctive* conditionals) have two readings, a *backtracking* one and a *non-backtracking* one. In evaluating a non-backtracking reading, we simply imagine that the antecedent had been true, holding as much of the *actual* past fixed as possible. In evaluating a non-backtracking reading, on the other hand, we consider *how the past would have needed to be different* in order to accommodate the truth of the antecedent.

As an example, consider the counterfactual, "If the fire department hadn't shown up, the building surely would've burned down." On a non-backtracking reading, the conditional is true: if everything in the history in the world (so including the size of the fire) had been the same up until the fire department's arrival, and something had caused them not to arrive (say, if

[15] See Lewis (1979).

all of the fire engines had simultaneously broken down), then it seems likely that the building would've burned down. On a backtracking reading, however, we might demur. After all, the fire department normally shows up to even quite small fires. In order for the fire department to have not shown up, then any fire must have been quite harmless indeed. And if that had been true, then the building probably wouldn't have burned down.

Lewis wanted to analyze causation in terms of counterfactuals (so that "c caused e" comes out roughly to "if c hadn't happened, then e wouldn't have happened"); flipping the equivalence between counterfactuals and causal notions around, we can avail ourselves of causal notions in characterizing the backtracking/non-backtracking distinction. In fact, there's a well-known distinction in the causality literature that tracks the backtracking/non-backtracking one well: that between *conditioning* and *intervening*.[16] To condition on p is, again, to suppose that one has observed that p, and to revise one's credences in the possible causes or explanations of p, as well as in its effects. To *intervene* on p, on the other hand, is just to make p so (in a way that bypasses the usual causes). In the fire department example, intervening corresponds to the non-backtracking reading: we suppose that something simply made it so that the fire department didn't arrive. Conditioning corresponds to the backtracking reading: we suppose that we observed, contrary to what actually happened, that the fire department did not arrive, and are looking for explanations of this.

This means that there's a nice formal similarity between indicatives and counterfactuals. Consider local indicatives and non-backtracking counterfactuals: in both cases, we hold fixed our credences in every proposition that the antecedent depends on, in the relevant sense, while revising our credences in every proposition that depends, in the same sense, on the antecedent. In the case of indicatives, the relevant sense of "depends on" is epistemic; in the case of counterfactuals, it's causal. And consider global indicatives and backtracking counterfactuals: in both cases, we revise our credences in every proposition that is probabilistically related to the antecedent. In evaluating a local conditional, then, we do something quite similar to pretending to intervene on the antecedent: something formally identical, but epistemic rather than causal. I'll call this account of local indicative conditionals the *epistemic-intervention* theory.

One nice feature of the epistemic-intervention theory is that it helps bridge two topics, indicatives and counterfactuals, that intuitively seem closely related (they use the same two-letter word, after all), but whose intimacy philosophers have had trouble accounting for. Jonathan Bennett, for example, is keen to stress that indicatives and counterfactuals have different functions: indicatives for expressing conditional credences, and counterfactuals for reasoning about nearby possible worlds, where nearness is determined by similarity of the past to the actual past.[17] But the recognition that counterfactuals admit of a backtracking reading, on which we evaluate the conditional by counterfactually supposing that we observed the antecedent, brings counterfactuals closer to the traditional territory of indicatives. And the recognition that indicatives have a local reading, on which we hold fixed our credences about propositions that the antecedent depends on, brings indicatives closer to counterfactual territory. We can even go further and group conditionals by the kind of reasoning used to evaluate them rather than by

[16] See, for example, Meek and Glymour (1994).

[17] Throughout his (2003).

their surface grammar, calling local indicatives and non-backtracking counterfactuals *intervening conditionals*, and global indicatives and backtracking counterfactuals *conditioning conditionals*. Armed with this new terminology, one might be tempted to contend that the distinction between indicatives and counterfactuals masks a more fundamental distinction, that between conditioning and intervening conditionals.

Another advantage of the epistemic-intervention theory is that it lets us exploit the formal similarities between the global/local distinction and the conditioning/intervening one; in particular, we can adapt a formalism devised for the latter to get a formula for the probability of local indicative conditionals that will (I will show) let us derive the partition-theory equation from the thesis about how people evaluate indicative conditionals. I do this next.

2.3 A formalism for intervention

First, a précis of the formalism.[18]

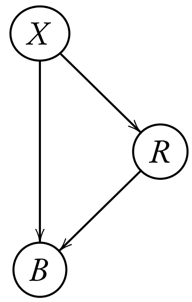
One convenient way to represent the probabilistic dependencies encoded in probability functions is through *undirected graphical models*: graphs in which nodes represent random variables, and edges represent direct probabilistic dependencies between variables: i.e., there is an edge between A and B iff $P(a|b) \neq P(a)$, and there are no other variables C that screen off this dependence.

Most of the work in this area, however, has focused on *directed* graphical models, also known as *Bayesian networks*. The nodes, again, represent random variables; but in place of (undirected) edges, we now have (directed) arrows between the nodes, which represent a specified *asymmetric* relation that implies direct probabilistic dependence. Typically, the relation is direct *causal* dependence, but in our case, we will use epistemic dependence instead.

For example, in Kaufmann's example we have:

$$
\begin{aligned}
&P(x) = .25 && P(\neg x) = .75 \\
&P(r|x) = 5/6 && P(r|\neg x) = 1/6 \\
&P(b|x \wedge r) = .9 && P(b|\neg x \wedge r) = .1. \qquad (22)
\end{aligned}
$$

Let us suppose that, in this setting, A epistemically depends on B iff we are given conditional probabilities for A on (perhaps among other variables) B. Then a graph representing the scenario would look like:



(23)

[18]What follows will be the slimmest of expositions. Unfamiliar readers should consult a text like Pearl (2000). In what follows, I'll use lowercase letters $a, b, c, \ldots$ to represent propositions, and uppercase letters $A, B, C, \ldots$ to represent propositional random variables. A can take a and $\neg a$ as values. Boldface letters will represent sets of such random variables or conjunctions of propositions: if $\mathbf{A} = \{A_1, \ldots, A_n\}$, then $\mathbf{a}$ can be any conjunction of form $a'_1 \wedge \ldots \wedge a'_n$, where each a'_i is either a_i or $\neg a_i$.

In general, we will suppose that each setting is sufficient to orient entirely (i.e., direct each edge in) the graph.

The set of variables with arrows into A are called the *parents* of A, $\mathbf{PA}_A$. Intuitively, the parents of A are all the matter in determining the probability of a. If A is a parent of B, then B is a *child* of A.

This can be combined with the Bayesian-network treatment of causal intervention, which I will adapt to my account. To *intervene causally* on A is to cause it to make a so without thereby causing the (usual) causal parents of A to change their values. Formally, here's what happens:

1. *To the probability function.* When we intervene to make a so, A becomes independent of its usual causes. Suppose our original probability function is $P(\cdot)$; the probability function arising from making a so is called $P_a(\cdot)$ and has the following properties:

   a. $P_a(a) = 1$.
   b. $P_a(b|\mathbf{pa}_B) = P(b|\mathbf{pa}_B)$, for $b \neq a$ and $\mathbf{pa}_B$ consistent with a.

   In words: property (a) says that the effect on a of intervening on it is to make its probability 1; property (b) says that, fixing the values of b's parents, the effect on b of intervening on a is nil (since the values of b's parents jointly determine its probability).

   We can derive a bunch of properties from these two basic ones. For our purposes, one that deserves mention is that conditioning and intervening on the parents of a variable are equivalent: that is, conditioning on $\mathbf{PA}_A$ and intervening on it have the same effect on whether a. In our notation,

   $$P_{\mathbf{pa}_A}(a) = P(a|\mathbf{pa}_A). \tag{24}$$

   This makes sense intuitively: the parents of A are all the variables that determine whether a, so it makes no difference to whether a whether we've *set* the values for its parents or just *observed* them. From this it follows that, for any subset $\mathbf{B}$ of $\mathbf{PA}_A$, the effect of intervening on $\mathbf{B}$ (to make it $\mathbf{b}$, say) while conditioning on the rest of $\mathbf{PA}_A$ is the same as the effect of conditioning on the whole set:

   $$P_{\mathbf{b}}(a|\mathbf{pa}_A \backslash \mathbf{B}) = P(a|(\mathbf{pa}_A \backslash \mathbf{B}) \wedge \mathbf{b}).\text{[19]}$$

[19] Proof:

$$P_{\mathbf{b}}(a|\mathbf{pa}_A \backslash \mathbf{B}) = \sum_{\mathbf{b}'} P_{\mathbf{b}}(a|(\mathbf{pa}_A \backslash \mathbf{B}) \wedge \mathbf{b}') P_{\mathbf{b}}(\mathbf{b}'|\mathbf{pa}_A \backslash \mathbf{B}),$$

where $\mathbf{b}'$ ranges over all possible values of $\mathbf{B}$, by the total-probability law. Since $P_{\mathbf{b}}(\mathbf{b}'|\mathbf{pa}_A \backslash \mathbf{B}) = 1$ when $\mathbf{b}' = \mathbf{b}$, and 0 otherwise, this last line becomes

$$= P_{\mathbf{b}}(a|(\mathbf{pa}_A \backslash \mathbf{B}) \wedge \mathbf{b}),$$

which is just

$$= P(a|(\mathbf{pa}_A \backslash \mathbf{B}) \wedge \mathbf{b})$$

This seems intuitively true too: as long as we know the values of A's parents, it doesn't matter how we came to our knowledge of each.

2. *To the graph.* Suppose that G represents $P(\cdot)$; then removing the arrows leading into the node for A creates a new graph G_A that represents $P_a(\cdot)$. (Since connection by arrows represents causal dependence, doing this mirrors the independence of A from its parents when we intervene on it.)

Since we want a formalism for not causal intervention, but epistemic intervention, we replace causal terms with ones for epistemic dependence in the above description.

The apparatus for epistemic intervention needs to be a bit more flexible than that for causal intervention, since in some cases we have partitions that don't span the entire range of propositions that the antecedent epistemically depends on. (Think about the political betting arena example, where we had the partitions {economy grows, economy shrinks} and {intervention ends, intervention continues}, even though whether the ruling party wins the elections depends on both the economy and the military intervention.) For these cases we need to extend the notion of intervention to *intervention on A with respect to variables* $B_1, \ldots, B_n$: we remove the effect of $B_1, \ldots, B_n$, but not of other variables, on whether a. The gist is that (roughly), for parents of A among the B_i's, intervention with respect to the B_i's will have the same effect as intervention *simpliciter*; for the parents of A not among the B_i's, it will instead have the same effect as conditioning. Let's call the probability function that arises from $P(\cdot)$ by intervening to make a so with respect to $B_1, \ldots, B_n$ $P_{a:B_1,\ldots,B_n}(\cdot)$.

We could give formal properties for this new function, but I suspect that would be overkill for what we're trying to do; throughout the next section, I'll keep the appeals to formal machinery to a minimum.

2.4 How this entails the partition-theory equation

We're now in a position to define the probability of a local conditional in terms of these notions:

$$P(p \rightarrow_{\{\bigwedge_i r_i'\}} q) = P_{p:R_1,\ldots,R_n}(q). \tag{26}$$

where r_i' can be either r_i or $\neg r_i$. In words: the probability of "if p then q" on the reading generated by the partition $\{\bigwedge_i r_i'\}$ is just the probability of q when we intervene on p with respect to the propositions that generate the partition. (The global reading arises as a degenerate case, by intervening on p with respect to *no* variables.)

We can then show that, under some conditions, (26) entails the partition-theory equation,

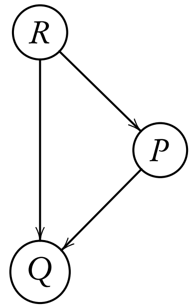
$$P(p \rightarrow_{\{\bigwedge_i r_i'\}} q) = \sum_{r_i' = \pm r_i} P(q|p \wedge r_1' \wedge \ldots \wedge r_n')P(r_1' \wedge \ldots \wedge r_n'). \tag{27}$$

---

by property (b) above, since $(\mathbf{pa}_A \backslash \mathbf{B}) \wedge \mathbf{b}$ is consistent with $\mathbf{b}$.

In this section I'll show that this equation holds for the simple cases we considered in the first section; we can prove a more general version, but I'll leave that proof out here for considerations of space and the reader's attention.

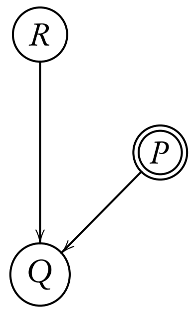
Again, in the simple cases we have a Bayesian network as follows:



(28)

In the subway case, $r =$ it's the weekend, $p =$ the next train will arrive in a few minutes, $q =$ the next train will be crowded; in the bags-and-ball case, $r =$ the bag is X, $p =$ the ball is red, and $q =$ the ball has a black spot; in the Sherlock Holmes case, $r =$ the man is Holmes, $p =$ Brown killed Murdoch, and $q =$ someone killed Murdoch.

Intervening on P with respect to R, we get the following graph,



(29)

in which P and R are now probabilistically independent.

According to (26),

$$P(p \to_{\{r,\neg r\}} q) = P_{p:R}(q) \quad (30)$$

$$= P_{p:R}(q|r)P_{p:R}(r) + P_{p:R}(q|\neg r)P_{p:R}(\neg r), \quad (31)$$

by law of total probability. Now, intervening on p has no effect on r, so $P_{p:R}(r)$ is just $P(r)$, and the last line is

$$= P_{p:R}(q|r)P(r) + P_{p:R}(q|\neg r)P(\neg r). \quad (32)$$

Now we get to appeal to the equivalence stated in equation (25): P and R are the parents of Q, so the effect on Q of conditioning on one and intervening on the other is the same as the effect of conditioning on both. This means that $P_{p:R}(q|r) = P(q|p \wedge r)$, so that the last line is

$$= P(q|p \wedge r)P(r) + P(q|p \wedge \neg r)P(\neg r), \quad (33)$$

which is just what the partition-theory equation predicts.

2.5 Causation and the origin of local conditionals

To close, I want to look at the question of why a local reading of conditionals should exist at all. What is the purpose of having such a way of evaluating conditionals, given that, evaluated this way, their probability does not reflect the credence we should have in the consequent upon observing the antecedent?

Call *causal conditionals* any conditional whose assertibility depends on the existence of a causal relation between the antecedent and consequent: for example, "If the Fed lowers interest rates, then stock prices will go up." Here, as with counterfactuals, there are two natural readings: one on which we pretend to *observe* the antecedent, and one on which we pretend to *intervene causally* to make the antecedent so. It should be clear from what I have said in the last section that the first is just the global reading, and the second the local reading.[20] So in the case of causal conditionals, the local reading has a clear function: it tells us the probability of the consequent when we intervene to make the antecedent so. In the case of non-causal indicative conditionals, however, I am not sure that there is a functional explanation of local readings; in place of that, let me offer a sketch of a non-functional one.

Lakoff and Johnson[21] introduced the idea that we understand entire classes of concepts in terms of others, through a process similar to metaphor; they call this *conceptual metaphor*. As an example, one metaphor we employ is of time as a moving object. Hence, "Time *flies*," "It's hard to *face* the future," "The end of the fiscal year is *approaching*." As another, morality as debt: "What we *owe* to each other," "I am *endebted* to her for her support," "To his *credit*, he admitted his error." This is not the modest thesis that we sometimes use relations among entities in one domain to elucidate those in another domain; rather, it is that our understanding of certain domains altogether depends inextricably on our understanding of other domains.

I want to suggest that we conceive of epistemic relations partly metaphorically, in terms of physical or causal ones. As one piece of evidence, consider the epistemic terms that are derived from physical ones: "depends on," "supports," "weakens," etc. (Even a seemingly sterile word like "confirm" is derived from Latin *confirmare*, to strengthen.) As another piece of evidence, there is a well-documented cross-linguistic tendency for modal verbs expressing epistemic modality to derive from ones expressing physical modality. For example, in English we have "may" (Old English *magan*, to be strong, able), which began as a verb expressing physical possibility; this sense is now obsolete, although attested until the 1600s. The epistemic sense was not attested until the 1200s.[22]

The presence of metaphor here would explain formal similarities between causal/physical and epistemic concepts. If, for example, our concept of epistemic modality is shaped in part by our concept of physical modality, this would explain formal similarities between the two: why "possibly" is equivalent to "not necessarily not" in both the epistemic and physical domains, and so on. Similarly, if our general concept of indicative conditionals is modeled in part on our concept of causal conditionals, this explains the formal similarities between the two: why two

[20]Specifically, the maximally local reading: the one on which we intervene on *all* the causal parents of the antecedent.

[21]Lakoff and Johnson (1980)

[22]See the OED entry on "may," *v*.1, sense II.

readings exist for both, and why the second reading of both involves an asymmetric relation of dependence.

References

Bennett, J. (2003). *A Philosophical Guide to Conditionals*. Oxford University Press.

Christensen, D. (1992). Confirmational holism and Bayesian epistemology. *Philosophy of Science*, 59(4):540–557.

Edgington, D. (1995). On conditionals. *Mind*, 104(414):235–329.

Kaufmann, S. (2004). Conditioning against the grain. *Journal of Philosophical Logic*, 33(6):583–606.

Lakoff, G. and Johnson, M. (1980). *Metaphors We Live by*. University of Chicago Press.

Lewis, D. (1979). Counterfactual dependence and time's arrow. *Noûs*, 13(4):455–476.

MacGee, V. (2000). To tell the truth about conditionals. *Analysis*, 60(1):107–111.

Meek, C. and Glymour, C. (1994). Conditioning and intervening. *British Journal for the Philosophy of Science*, 45(4):1001–1021.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge University Press.

Pollock, J. L. (1981). Indicative conditionals and conditional probability. In Harper, W. L., Stalnaker, R., and Pearce, G., editors, *Ifs: Conditionals, Belief, Decision, Chance, and Time*. Reidel.