Abduction and indicative conditionals

Stefan Kaufmann
Northwestern University

Background. The notion of causality has made a comeback in Artificial Intelligence (Pearl, 2000; Spirtes et al., 2000) and philosophy (Hausman, 1998; Galavotti et al., 2001).[1] In the theory and practice of Bayesian Networks, observance of causal relations started out as a design guideline (it yields simpler and more intuitive graphs) but has since been exploited to push the use of the models into new territory. This move has potential ramifications for natural-language interfaces: Does sensitivity to causal structure help or hamper the evaluation of sentences against knowledge representation systems?

One area in which the importance of causality and the utility of causal networks are uncontroversial is the interpretation of counterfactuals. Pearl (2000) points out the similarities between his treatment of "structural counterfactuals" and the possible-worlds semantics of Lewis (1973). Kaufmann (2001) integrates causal structure into a probabilistic model for the interpretation of conditionals and argues that this move circumvents certain obstacles to a unified account of indicative and counterfactual conditionals.

The question. Such approaches to counterfactuals raise new questions about *indicative* conditionals, in particular: Should their probabilities be the corresponding conditional probabilities, as is widely assumed in the philosophical literature (Jeffrey, 1964; Adams, 1965; Stalnaker, 1970)? This assumption turns out to be at odds with the otherwise straightforward method of identifying the "global" probability of a sentence under uncertainty about a causally relevant background variable with the expectation of its "local" probabilities for each value of that variable (see below).

The answer. I will discuss cases in which the probabilities of conditionals under this approach are predicted to differ from the conditional probability and conclude that these predictions are not only correct, but of considerable explanatory value, for two reasons: First, there are indeed conditionals whose intuitively correct probability in a given scenario is "ill-behaved" in one of three ways: (i) it is not the conditional probability; (ii) it is not the posterior probability of the consequent upon observing the antecedent; and/or (iii) it is affected by updates which do not affect the conditional probability. The discussion refers to a simple scenario (see below) for (i) and examples from the literature for (ii) and (iii) (McGee, 2000, and Pollock, 1981, respectively).

The possibility of such disparities has been noticed before (cf. Lewis, 1976, on imaging *vs.* conditionalization, and Gibbard, 1981, on "nearness" *vs.* "epistemic" conditionals). Those authors treat both interpretations as fundamen-

[1] No completeness of references is intended throughout this abstract.

tally different. In contrast, my second claim is that (a) conditional probabilities are central to the interpretation of *all* conditionals (with imaging as a limiting case; see Skyrms, 1984, for related discussion); (b) "ill-behaved" conditionals are ambiguous between two readings, and under the less prominent of these the conditional probability is correct; and (c) this less prominent reading is derived from the former by a simple step of *abductive* inference. I illustrate with an "ill-behaved" case of type (i):

An example. *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. In this scenario, is the strength of your belief in (1) best characterized as 'high', 'fifty-fifty', or 'low'?*

(1) *If I pick a red ball, it will have a black spot.*

Speakers generally agree that this probability should be 'low', guided by the intuition that the probability of the conditional is "more likely low than high":[2]

$$P(R \to B) = P(R \to B|X)P(X) + P(R \to B|Y)P(Y) \tag{2}$$
$$= P(B|RX)P(X) + P(B|RY)P(Y) \tag{3}$$
$$= 9/10 \times 1/4 + 1/10 \times 3/4 = .3$$

However, the conditional probability that the ball will have a black spot, given that it is red, is not 'low' but .6:

$$P(B|R) = \frac{P(B|RX)P(X|R)P(R) + P(B|RY)P(Y|R)P(R)}{P(R)} \tag{4}$$
$$= P(B|RX)P(X|R) + P(B|RY)P(Y|R) \tag{5}$$
$$= 9/10 \times 5/8 + 1/10 \times 3/8 = .6$$

But notice that the probabilities in (3) and (5) are derived as expectations of the *same* "local" conditional probabilities; only the weights are different. Intuitively, the difference consists in whether the hypothetical assumption that the ball is red affects one's beliefs about the bag. It *may* do so, for the conditional arguably *also* has the probability in (4), by the following reasoning:

(6) a. I think it is unlikely that I will pick a red ball:
$P(R) = P(R|X)P(X) + P(R|Y)P(Y) = 1/3$
b. But suppose I do.
c. Then this is probably bag X after all:
$P(X|R) = P(R|X)P(X)/P(R) = 5/8$
d. So the ball will probably have a black spot, given that it is red: (4)

The step from (6a,b) to (6c) is 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 a 'low' rating. I hypothesize that its optionality is due to the fact that the contents of the bags are only stochastically, not causally dependent upon the color of the ball.

[2] The transition of from (2) to (3) leads to triviality under the assumption that conditionals denote propositions (Lewis, 1976). I do not make that assumption here.

In general. When the probability of C depends on a variable $\mathbf{X}$ with values $X_1, X_2, \ldots$ which is causally independent of A, the probability of the conditional *If A then C* may be calculated "locally" as in (7) or "globally" as in (8) (assume for simplicity that the "local" conditional probability is defined for all X_i):

(7) $$P_\ell(\text{if } A \text{ then } C) = \sum_{X_i} P(C|A, X_i)P(X_i)$$

(8) $$P_g(\text{if } A \text{ then } C) = P(C|A) = \sum_{X_i} P(C|A, X_i)P(X_i|A)$$

This claim does not question the assumption that true update proceeds by conditionalization. What the examples show is that Ramsey's "hypothetical addition" of the antecedent to one's stock of knowledge need not: *Assuming* that the antecedent is true is not the same as *observing* that it is; the former involves a less sweeping update, leaving background beliefs intact. An account that is sensitive to causal (in)dependencies predicts when and how this difference comes about and how the two readings are related.

Acknowledgments. I am grateful to the Japan Society for the Promotion of Science (JSPS, Grant #P01063) and Satoshi Sato's hospitality at Kyoto University, where much of this work was carried out.

References

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

Fetzer, J. H., editor. 1988. *Probability and Causality*, volume 192 of *Studies in Epistemology, Logic, Methodology, and Philosophy of Science*. D. Reidel, Dordrecht / Boston / Lancaster / Tokyo.

Galavotti, M. C., P. Suppes, and D. Costantini, editors. 2001. *Stochastic Causality*. CSLI Publications, Stanford, CA.

Gibbard, A. 1981. Two recent theories of conditionals. In Harper et al. (1981), pages 211–247.

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

Hausman, D. M. 1998. *Causal Asymmetries*. Cambridge University Press, Cambridge, UK.

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

Kaufmann, S. 2001. Tense probabilism properly conceived. In Stokhof, R. van Rooy. M., editor, *Proceedings of the Thirteenth Amsterdam Colloquium*, pages 132–137, Amsterdam, NL. ILLC/Department of Philosophy.

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

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

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

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

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

Skyrms, B. 1984. *Pragmatics and Empiricism*. Yale University Press, New Haven / London.

Spirtes, P., C. Glymour, and R. Scheines. 2000. *Causation, Prediction, and Search*. MIT Press.

Stalnaker, R. 1970. Probablity and conditionals. *Philosophy of Science*, 37:64–80.