Title
Learning the Functional Form of Causal Relationships

Permalink
https://escholarship.org/uc/item/4tk0g04s

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 29(29)

ISSN
1069-7977

Authors
Lucas, Chris G.
Griffiths, Thomas L.

Publication Date
2007

Peer reviewed

Learning the Functional Form of Causal Relationships

Chris G. Lucas and **Thomas L. Griffiths**
Department of Psychology, 3210 Tolman Hall, University of California
Berkeley, CA 94720, USA

Keywords: Causal learning; Bayesian modeling

There is abundant evidence that people use abstract knowledge when reasoning about cause and effect (e.g., Bullock, Gelman, & Baillargeon, 1982; Schultz, 1982). We know things that let us learn quickly from small amounts of data and help us avoid spurious inferences. Our knowledge is often domain or context-specific, and much of it must be learned. Despite the pervasiveness of knowledge effects in causal inference, little is known about exactly how we acquire knowledge that constrains learning, and the best-known accounts of causal induction do not address it. We developed a Bayesian model of the acquisition of abstract knowledge that constrains and facilitates causal inference. The model makes predictions that existing theories of causal inference do not: (1) People can make inferences appropriate to functional relationships beyond those typically used in covariation-based models of causal inference, e.g., the noisy-OR (Cheng, 1997); (2) People use knowledge acquired from learning about one set of objects to inform their inferences about another set.

Experiment

Participants saw training data that suggested causal relationships took one of three functional forms (Table 1) in a particular context. They were they given novel objects and asked to make inferences based on new evidence that was identical across all conditions.

The experiment consisted of two trials. In the first trial, participants were told that some objects were blickets, some were not, and that blickets possess blicketosity while non-blickets possess no blicketosity. Their goal was to determine which objects were blickets, using only the activation of a "blicketosity meter".

Participants saw training data that varied by condition (Table 2) in which the first three objects (A, B and C) were placed singly or in groups on the meter. After the training data, they saw a common test data set with tree novel objects (D, E, and F): D- D- D- E- DF+ DF+. They were then asked to rate the probability that each object was a blicket, from 0 to 10. See Figure 1 for human ratings and model predictions.

Table 1: Functional forms of causal relationships. $\mathrm{P}(d|T,h)$ is the probability of activation with n_c blickets (in h) on it.

Form	$\mathrm{P}(d\|T,h)$
deterministic-OR	$1-0^{n_c}$
noisy-OR	$1-(1-w)^{n_c}$
deterministic-AND	$\min(1,\lfloor\frac{n_c}{2}\rfloor)$

Table 2: Events presented to participants as training data in different conditions.

Condition	**Training data**
AND	A- B- C- AB- AC+ BC-
noisy-OR	A+ B- C- AB- AC+ BC-
deterministic-OR	A+ B- C- AB+ AC+ BC-

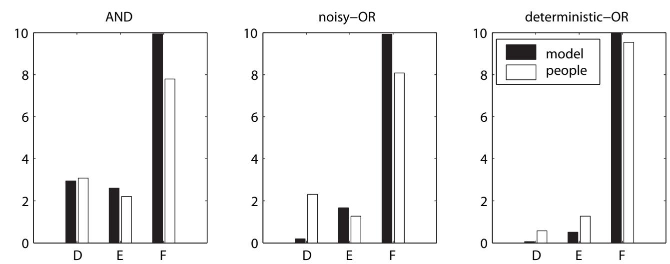

Figure 1: Model predictions versus human ratings.

Results

The prediction that causal knowledge derived from evidence about one set of objects constrains subsequent inference was supported: two-way ANOVAs found a main effect of the Trial 1 data on judgments in Trial 2 for both D and E, ($p < 0.001$).

The prediction that people can reason appropriately and efficiently about conjunctive causal relationships given appropriate evidence-based knowledge was also supported: as the model predicted, given very little evidence in the test condition, participants in the AND condition judged the object D was a blicket more often than in other conditions ($p < 0.001$).

The quantitative predictions of our model were also accurate, correlation strongly with mean judgments ($r = 0.98$). These results are in contrast with the predictions of popular models of causal inference (e.g., Cheng, 1997; Shanks, 1995), which assume a fixed functional form for causal relationships, and cannot account for the kinds of training-based differences in inferences that we observed.

References

Bullock, M., Gelman, R., and Baillargeon, R. (1982). The development of causal reasoning. In W. J. Friedman (Ed.), *The developmental psychology of time* (pp. 209-254). New York: Academic Press.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*, 367-405.

Shultz, T. R. (1982). Rules of Causal Attribution. *Monographs of the Society for Research in Child Development*, *47*, 1-51.