Cognitive Psychology 86 (2016) 62–86

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

Causal competition based on generic priors

Derek Powell[a,*], M. Alice Merrick[a], Hongjing Lu[a,b], Keith J. Holyoak[a]

[a] *Department of Psychology, University of California, Los Angeles, United States*
[b] *Department of Statistics, University of California, Los Angeles, United States*

ARTICLE INFO

Article history:
Accepted 1 February 2016

Keywords:
Causal learning
Simplicity
Cue competition
Generic priors
Bayesian models

ABSTRACT

Although we live in a complex and multi-causal world, learners often lack sufficient data and/or cognitive resources to acquire a fully veridical causal model. The general goal of making precise predictions with energy-efficient representations suggests a generic prior favoring causal models that include a relatively small number of strong causes. Such "sparse and strong" priors make it possible to quickly identify the most potent individual causes, relegating weaker causes to secondary status or eliminating them from consideration altogether. Sparse-and-strong priors predict that competition will be observed between candidate causes of the same polarity (i.e., generative or else preventive) even if they occur independently. For instance, the strength of a moderately strong cause should be underestimated when an uncorrelated strong cause also occurs in the general learning environment, relative to when a weaker cause also occurs. We report three experiments investigating whether independently-occurring causes (either generative or preventive) compete when people make judgments of causal strength. Cue competition was indeed observed for both generative and preventive causes. The data were used to assess alternative computational models of human learning in complex multi-causal situations.

* Corresponding author at: Department of Psychology, University of California, Los Angeles, 1285 Franz Hall, Box 951563, Los Angeles, CA 90095, United States.

E-mail address: derekpowell@ucla.edu (D. Powell).

http://dx.doi.org/10.1016/j.cogpsych.2016.02.001
0010-0285/

In real life there are no phenomena that have only one cause and have not been affected by secondary causes. Otherwise we should be living in a world of pure necessity, ruled by destiny alone.

[Spirkin, 1983, p. 90]

1. Coping with causal complexity

We live in a complex and multi-causal world. Consider how every modern commercial airplane is equipped with an in-flight recorder, or "black box." Should the airplane go down, the information provided by the recorder can help investigators diagnose the cause of the crash. Without such an aid, confident diagnosis might prove impossible, as many types of events, individually or in combination, can cause an aircraft crash—flight crew error, mechanical faults, the weather, terrorism, bird strikes, and so on (for a general review of causal thinking, see Lagnado, 2011).

Many more pedestrian causal relationships are no less complex, yet humans must continually learn and make inferences under severe constraints imposed by their limited attentional and memory resources (generally without the aid of external devices). Relative to the complexity of the actual physical and social world, the causal models formed by the human mind are inevitably simplified—more like rough sketches than high-resolution photographs. Causal simplification may play an especially important role in the early stage of learning, when the paucity of data makes it impossible to reliably estimate the strengths of all possible causes of an effect. Especially when data relevant to causal inference must be aggregated over time, memory limitations make it difficult to maintain representations of many alternative possible causal models. For example, when faced with five candidate causes for an effect, there are 32 possible causal models to consider (each including a particular combination of one to five effective causes). Faced with such a bewildering array of alternatives, the learner will likely be forced to make simplifying assumptions.

A general preference favoring simplicity has been proposed as a unifying principle for cognitive science (Chater & Vitányi, 2003). In the case of human causal learning and inference, several distinct simplicity constraints have been proposed. These are of three general types: (1) constraints on causal attribution within a known causal structure, (2) constraints on selection of causal structures, and (3) constraints on causal strengths associated with individual causal links. We will briefly consider each of these in turn, and then focus on the third.

1.1. Causal attribution

There is considerable evidence that people prefer to search their existing causal knowledge for a single explanation of a newly observed effect. As a consequence, multiple alternative explanations compete with one another. This type of competition gives rise to *causal discounting*, whereby the presence of one cause reduces the estimated probability that some other cause was active (Kelley, 1973). For example, if you observe wet grass in the morning, you might suspect it rained overnight. But if you find that there was a sprinkler on, you might attribute the wet grass to the sprinkler and discount the probability that the wet grass was caused by rain. Pearl (1988) showed that causal discounting is a normative consequence of reasoning with causal models. Moreover, the competitive nature of causal explanation follows from Thagard's (1989) theory of explanatory coherence, and is supported by a variety of experimental evidence (Holyoak, Lee, & Lu, 2010; Read & Marcus-Newhall, 1993). In addition to a general preference for single explanations, Read and Marcus-Newhall (1993) showed that people prefer explanations that cover a broader range of data (for additional evidence see Lombrozo, 2007, 2012).

1.2. Causal structure

Causal learners can also make judgments of causal *structure*, assessing which candidate causes in fact generate (or prevent) an effect (Griffiths & Tenenbaum, 2005). Interpreted within a graphical representation, a structure judgment is an evaluation of whether or not a link exists between a node representing a candidate cause to a node representing an effect. Here, simplicity constraints operate during the course of causal learning. The basic machinery of Bayesian inference yields a preference for graphs