Human discovery of cause and effect in perception streams requires reliable online inference in highly unstructured noisy environments with very few (maybe even only one) positive examples. Automated causal discovery research, on the other hand, has typically operated in a qualitatively different setting: Data used for learning is cast into more static scenarios where all prospect observable causes are known in advance. Typically when learning it is assumed that all data is known (i.e., learning is done in batch mode), and little work has been done on the important but hard topic of causal abstraction. In this paper, we present a system that is capable of wading through an unknown set of possible causes (events) in an online fashion to identify a small set of plausible candidates for some effect of interest. As new events are encountered, they can be added to the system in an online fashion. We provide a mechanism to learn arbitrary noisy logic formulae and present a method to do causal abstraction.

This work is an extension of the causal support model presented by Griffiths and Tenenbaum [2005]. That model described human elemental causal explanation in terms of structure learning of a noisy-or parametrized graphical model. Although this model was presented as an explanation for human causal inference of elemental causes, it is also interesting from a automated causal discovery perspective, especially in the context of perception. The causal support model is suitable for learning elemental causes in a batch manner. In this paper, we extend this model in several ways that make it more widely applicable to automated causal discovery in perception:

First, we extend the method to learn arbitrary logic formulae online in (noisy) disjunctive normal form (DNF). We feel this assumption is a reasonable one; it is common in human causal explanations, and is therefore likely to prove useful in automated causal discovery as well. This class of causal models is distinguished, for example, by the class of general Bayesian networks. General Bayesian networks are more expressive, being able to represent arbitrary joint probability distributions; however, they are in turn more difficult to learn, both in terms of computational complexity and (we show empirically) in terms of sample complexity. By restricting our causal hypotheses to logic formulae, we constrain the computational complexity of learning, but we still capture a large class of realistic causal models used by humans. In our method, disjuncts are learned in essentially the same manner as elemental causal factors learned in the causal support model. Conjuncts are found efficiently by annotating an undirected graph with pairwise potentials equal to correlations of the causal support between the two variables. Cliques in this graph comprise the set conjuncts.

Second, we show that our algorithm possesses the desirable property of producing models with incremental complexity. As a few positive instances of the effect is observed, our method can begin to suggest with confidence elemental events that are likely to be the cause. As more positive instances are observed, we are more likely to discover more complex causal factors involving conjunctions of variables. Thus our method learns models with incremental complexity that grows as more data is observed, providing us a natural mechanism to avoid overfitting the data.

Third, our algorithm is able to perform causal abstraction. Given an event as an ISA hierarchy, one can pose hypotheses at different levels of the hierarchy. For example, if drinking Swedish vodka tends to give one a hangover, and drinking Russian vodka does the same thing, one might generalize and conclude that vodka in general causes a hangover. If it is further observed that scotch also gives a hangover, it might be concluded that all hard liquor causes the hangover. When searching for conjunctions of factors using an ISA hierarchy, certain conjunctions of factors don’t make much sense: for example, it doesn’t make sense for hard alcohol and swedish vodka to be the cause of a hangover. To keep track of the different possibilities implied by an ISA hierarchy, we have developed an upward message passing algorithm that provides a dynamic programming solution to finding all possible combinations of consistent factors in the hierarchy.

We have performed two sets of simulated experiments to show the capabilities of our framework. These experiments show that our method can successfully generalize the causal support for the true causes to their super-classes in the perceptual hierarchy. E.g., if observing that Swedish vodka is the true cause of a hangover, given the perceptual hierarchy, our method can concurrently infer that the vodka class has pretty much the same support, although the strength is much less. The second sets of experiments were specially designed to illustrate the previously mentioned incremental complexity nature of our approach. We have compared our approach to a modified Bayesian Structure Learning (BSL) algorithm.

The BSL used the standard BDeu metric [Heckerman et al., 1995] to search for structure, and was modified to learn logic formula in DNF by thresholding the conditional probability table for the best structure learned from data and converting the resulting logic table into DNF. The results show that the the BSL algorithm not only suffers from the exponential computational complexity, but it also needs a fairly large number of positive examples to infer a DNF close the true one. This comes from the fact that, even for inferring a single-term factor, the general BSL algorithm needs to observe sufficient statistics for all combinations involving that single term, while since our method starts with the simplest model (the noisy OR), it can detect at least the simple factors very fast (by observing one or two positive examples). This result reinforces the hypothesis that humans initially try to explain the world with the simplest model. As they fail, they gradually increase the complexity of their model until some measure of error falls below the tolerance level. Our model might thus presents a plausible hypothesis for how humans refine simple hypotheses given more data. This hypothesis remains to be tested using human subject data.

References


* Also adjunct faculty in the Department of Biomedical Informatics, School of Medicine, University of Pittsburgh.