A Hybrid Learning Model of Abductive Reasoning

Todd R. Johnson[1], Jiajie Zhang[2] and Hongbin Wang[3]
Division of Medical Informatics, Department of Pathology[1]
Department of Psychology[2,3]
Center for Cognitive Science[1,2]
The Ohio State University
Columbus, Ohio 43210
{johnson.25, zhang.52, wang.190}@osu.edu

Introduction

Abduction is the process of generating a best explanation for a set of observations. Symbolic models of abductive reasoning tend to be far too search-intensive, whereas connectionist models have difficulty explaining higher level abductive reasoning, such as the generation and revision of explanatory hypotheses. In addition, abductive tasks appear to have deliberate and implicit components: people generate and modify explanations using a series of recognizable steps, but these steps appear to be guided by an implicit hypothesis evaluation process.

We propose a hybrid learning model for abduction that tightly integrates a symbolic Soar model for deliberately forming and revising hypotheses with Echo, a connectionist model for implicitly evaluating explanations (Thagard, 1989). In this model, Soar's symbolic knowledge compilation mechanism, chunking, acquires rules for forming and revising hypotheses and for taking actions based on the evaluations of these hypotheses. Thus, chunking models the problem solver's shift from deliberate to automatic reasoning. To complement this, Echo learns to provide better hypothesis evaluations by acquiring explanatory strengths based on the frequencies of events from past experience. Since Echo does not have a learning mechanism, we have extended it by adding the Rescorla-Wagner (1972) learning rule.

Motivation for a Hybrid Model

The hybrid model is motivated by several observations and empirical results concerning the relationship between symbolic and connectionist processes and human abductive reasoning.

To successfully solve abductive problems people must learn to quickly generate possible hypotheses for one or more observations, and then integrate these hypotheses into a coherent explanation for the entire set of observations. Symbolic search based approaches have traditionally performed well at modeling hypothesis generation and modification. Likewise, symbolic knowledge compilation can learn explicit rules based on a single problem solving episode, but it cannot easily learn explanatory strengths from previous experience. In contrast, connectionist learning techniques can easily acquire explanatory strengths, but cannot quickly acquire explicit rules.

Research on implicit acquisition and use of event frequencies supports the hybrid Soar/Echo architecture. When conditional probabilities and base rates of occurrence are presented explicitly in terms of numeric values, they are very difficult to learn and utilize (see Kahneman, Slovic & Tversky, 1982). However, when they are presented in terms of real events and occurrences, they can often be learned implicitly and used correctly (e.g., Christensen-Szalanski, & Bushyhead, 1981). A number of studies indicate that the learning of frequency of occurrence is usually implicit (unconscious) and automatic. The Soar/Echo hybrid architecture is consistent with these results, because Echo appears to Soar as an opaque mechanism that automatically and constantly provides confidence values for hypotheses.

Acknowledgments

This research was supported by Office of Naval Research Grant No. N00014-95-1-0241.

References

Christensen-Szalanski, J. J. J., & Bushyhead, J. B. (1981). Physicians' use of probabilistic information in a real clinical setting. *Journal of Experimental Psychology: Human Perception and Performance, 7* (4), 928-935.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: The effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical Conditioning II: Current Research and Theory* (pp. 64-69). New York: Appleton-Century-Crofts.

Thagard, P. (1989). Explanatory Coherence. *Behavioral and Brain Sciences*, 12, 435-502.