Title
Explanatory Coherence as a Model for Belief Revision

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 18(0)

Author
Carrick, Christina

Publication Date
1996

Peer reviewed

Explanatory Coherence as a Model for Belief Revision

Christina Carrick
Computing Science
Simon Fraser University
Burnaby, BC, Canada V5A 1S6
ccarrick@cs.sfu.ca

Introduction

A belief revision system must be able to decide, in the face of contradictory information, which (if any) beliefs should be abandoned. What general principle should be followed to make this decision? In Martins and Shapiro (1988), several specific tasks were described which must be solved in order for a system to successfully revise beliefs in the face of contradictory evidence: inference, dependency recording, nonmonotonicity, and disbelief propagation. A cognitive theory of belief revision should be able to say which beliefs should be retained or abandoned while allowing solutions to the subproblems to be covered by the same theory.

The ECHO Program and Belief Revision

There are two major epistemic theories of belief revision: the *foundational* theory of belief revision and the *coherence* theory. It is desirable for a model of belief revision to have some of the rationality of the foundational theory, but with a behaviour similar to that produced by the coherence model. Explanatory Coherence (EC) makes such an attempt, encompassing aspects of both theories. EC follows the principle of negative uncertainty, given by foundational theory, by requiring some sort of explanatory relationship between two propositions in order for them to be considered coherent. It also conforms to the principle of conservation, however, which is the basis for coherence theory.

Explanatory Coherence is a property of two or more propositions (Thagard, 1989). The greater a proposition coheres with a set of other propositions, the greater its acceptability; a proposition which is incoherent with the set should be rejected. In his paper, Thagard notes the potential for using explanatory coherence for belief revision. Though EC cannot say how to derive new beliefs from old ones, it could be used to decide whether or not a new belief should enter the system on the basis of coherence with the beliefs already maintained.

Based on the principles of explanatory coherence, the ECHO program is described in detail in Thagard (1989). ECHO is an implementation of an associative artificial neural network. Connections in the network are either excitatory, which represents explanatory relations, or inhibitory, which represents contradiction. Each node represents a proposition or an hypothesis, and the connections are, therefore, the relations between them. Since the purpose of ECHO was to determine coherence and not explanatory relationships, all explanatory and contradictory relationships between propositions are entered as input. A "special evidence" unit provides activation to all propositions of observed data, and this activation spreads throughout the network via the excitatory (explanation) links and is restricted through the inhibitory (contradiction) links.

Experimentation and Results

The ECHO program was extended only slightly for nonmonotonicity experiments. It was made possible to adjust any or all activations and connection strengths of the network, to simulate a "forgetting" effect or a re-initialization. It was also made possible to disconnect nodes from the special evidence unit, to allow propositions which were once true to become false. Two different original data sets were used for the belief revision experiments, each of which were presented to the network in varying manners (certain subsets of evidence were presented, and then others).

It was found that implementing the nonmonotonic ability is not a straightforward matter. The simple EC processes alone were not enough to cause the shifts in belief that should have occurred. Fading all of the activations provided the network with the capability to re-evaluate the beliefs, which is cognitively plausible given the assumption that there exists a momentary suspension of a person's beliefs when he is confronted with new information.

Future work will involve investigation into the issue of epistemic entrenchment in ECHO. Marker-passing will be explored to facilitate shifts between epistemic states without fading all nodes equally. Experimentation with larger data sets is also necessary, to discover ECHO's ability to scale to more complex data.

References

Gardenfors, P. (1992). The Dynamics of Belief Systems: Foundations versus Coherence Theories. In *Knowledge, Belief, and Strategic Interaction* (pp. 377-396). Cambridge: Cambridge University Press.

Goldman, A. I. (1993). Epistemic Folkways and Scientific Epistemology. In *Readings in Philosophy and Cognitive Science* (pp. 95–114). Cambridge, Massachusettes: MIT Press.

Martins, J. P. & Shapiro, S. C. (1988). A Model for Belief Revision. *Artificial Intelligence*, 35(1), 25–79.

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