

The 2005 AAAI Classic Paper Awards

Tom Mitchell and Hector Levesque

- Mitchell and Levesque provide commentary on the two AAAI Classic Paper awards, given at the AAAI-05 conference in Pittsburgh, Pennsylvania. The two winning papers were “Quantifying the Inductive Bias in Concept Learning,” by David Haussler, and “Default Reasoning, Nonmonotonic Logics, and the Frame Problem,” by Steve Hanks and Drew McDermott.

Quantifying the Induction Bias in Concept Learning

Commentary by Tom Mitchell

David Haussler’s paper “Quantifying the Inductive Bias in Concept Learning,” presented at the AAAI’s Fifth National Conference on AI in Philadelphia in 1986, helped initiate a very fruitful integration of a branch of machine learning theory with machine learning practice. Starting in the 1950s, with work like Samuels’s program that learned strategies for playing checkers, AI researchers had designed and experimented with a variety of learning algorithms and had also developed a number of theoretical results, such as convergence proofs for perceptrons and “learning in the limit” results for grammatical inference. Haussler’s paper helped introduce the AI community to a new line of important theoretical work on what is called PAC learning (probably

approximately correct learning). At the same time, his paper and his subsequent research applied that theoretical framework to analyze the properties of specific machine learning algorithms.

At the time of Haussler’s paper, one informal notion popular in machine learning was the “inductive bias” of a learner; that is, the set of assumptions that, together with the training data, logically entail the hypothesis finally output by the learner. For example, a key part of the inductive bias of most learners is the representation they employ for hypotheses, which defines implicitly the space of hypotheses they can ever consider. The larger the space of hypotheses considered, the more training data needed to reliably learn the target concept—the more constraining the inductive bias, the less training data needed.

During the mid-1980s the new theory of PAC learning was being developed, which allowed deriving quantitative bounds on the probability of successful learning as a function of the number of training examples and the complexity of the learner’s hypothesis space (as measured by its Vapnik-Chervonenkis dimension). What Haussler’s paper did was help introduce this PAC learning theory to the AI community and show that its results made direct contact with the inductive bias of the learner—that in fact the Vapnik-Chervonenkis dimen-

sion offered a quantitative measure of the complexity of the learner’s hypothesis space, characterizing this key aspect of the learner’s inductive bias.

Haussler’s paper was therefore important in linking the new PAC learning theory work with the ongoing work on machine learning within AI. Twenty years later that link is firmly established, and the two research communities have largely merged into one. In fact, much of the dramatic progress in machine learning over the past two decades has come from a fruitful marriage between research on learning theory and design of practical learning algorithms for particular problem classes.

Default Reasoning, Nonmonotonic Logics, and the Frame Problem

Commentary By Hector Levesque

Within knowledge representation and reasoning, the (temporal) projection task is that of determining what does or does not hold after a sequence of actions is performed. It is subject to the frame problem: how do we determine what does not change after an action is performed, when we are only expected to be told what does change? From the mid-1970s or so, this general pattern of drawing conclusions based on a lack of information to the contrary became the focus of much research under the name “nonmonotonic” or “default” reasoning.

The idea is this: Normally, an object is unaffected by an action. If a window is open, then it is reasonable to assume that it remains open after doing an action. There are clear exceptions, however, such as the act of closing the window. A variety of formal systems have been proposed that would allow us to infer in the absence of conflicting information that the window remains open (or that a polar bear is white or that a violin has four strings, and so on). The formal systems do this in various ways, but the main idea is to make the set of exceptions (or abnormalities) as small as possible, given what is known. The sort of reasoning required for projection would then be

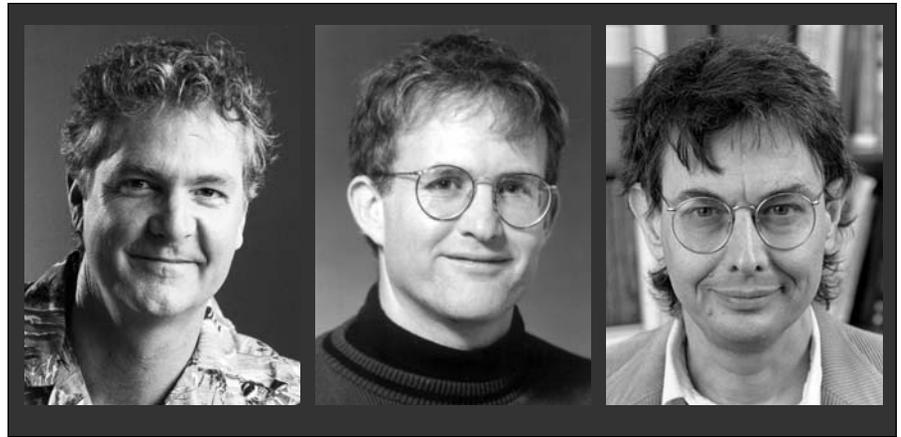
subsumed by this general account of default reasoning.

At least, this was the view held by most researchers until the publication of the paper by Steve Hanks and Drew McDermott in 1986. What this paper showed using a very simple example is that default reasoning in its most obvious form yields conclusions that are too weak for projection.

Here's the example that came to be known as the "Yale Shooting Problem" (YSP). There are three actions: shooting a gun, loading a gun, and pausing for a moment. Loading a gun is exceptional in that it causes the gun to be loaded, when under normal circumstances, a gun would stay unloaded. When the gun is loaded, shooting is also exceptional in causing a target to be hit, when normally the target would stay unhit. No other information is given (for example, about the pause action). Now assume that we load the gun, pause, and then shoot. What happens to the target? Intuitively, we ought to be able to infer by default that the target is hit, even without information about the pause action.

The problem is this: what about the possibility of the gun becoming unloaded as the result of the pause action? This would certainly be an exceptional outcome. But if it happened, the target would then be unaffected by the next action. In other words, something exceptional has to happen: it can either be a shooting that hits the target (preceded by an ineffective pausing) or a pausing that unloads the gun (followed by an ineffective shooting). Default reasoning minimizes the exceptions but has no reason to prefer one or the other. And, as Hanks and McDermott correctly observe, this does not depend on which default reasoning formalism we use to represent the example. Thus they conclude (much more controversially) that the frame problem cannot be solved as a special case of default reasoning.

Agree or disagree with this conclusion, the impact of the paper was undeniable. Many researchers tried very hard to find other ways to deal with the example to avoid the problem. One suggestion was to apply default reasoning chronologically (or pointwise) so that we first conclude by de-



Left to Right: David Haussler, Steve Hanks, Drew McDermott.

fault that the gun remains loaded and then go on from there to conclude that the target is hit. Another approach was to formalize causality explicitly, where causal relations themselves are what are considered to be exceptional, and where pausing is then assumed by default to have no causal consequences. A third approach involved looking at all possible states of the world, and noting that the assumption that pausing changes the gun is otiose, since in other states where the gun is loaded, shooting still changes the target. Finally, a group of researchers tackled the YSP by representing some facts not as sentences but as special "rules" involving the negation-as-failure operator of logic programming (or the default logic equivalent).

All of these techniques solve the YSP to some extent, given certain caveats and restrictions. The most popular solutions today are the causal approach due to Turner, Lifschitz, and others; the solution proposed by Reiter based on ideas of Schubert, Pednault, and Haas; and variants due to Elkan and to Thielscher. Each has become the basis for implemented systems, and Reiter's has the added advantage of using only ordinary monotonic first-order logic through so-called successor-state axioms. A generalization of this solution based on causality due to Lin also addresses the ramification and qualification problems.

An interesting thing about these so-

lutions is that they are only claimed to work for a collection of beliefs (about action effects) of a very specific form. I believe Hanks and McDermott were right: a naive application of default reasoning over a plausible collection of beliefs about a changing world will not solve the frame problem. Perhaps what we have learned in the interim is that it is not worth the effort to find a purely logical solution that applies to beliefs of unrestricted form. At any rate, concern about the frame problem has certainly diminished. In many of the implemented systems today, successor-state axioms are simply encoded directly, bypassing completely the need for default reasoning.



Tom Mitchell is the Fredkin Professor of AI and Learning and director of the Center for Automated Learning and Discovery at the School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania. He was president of AAAI from 2001–03 and is an AAAI fellow.



Hector Levesque is a professor of computer science at the University of Toronto. He cofounded the KR conferences in 1989, was conference chair of the IJCAI conference in 2001, was president of the IJCAI board of trustees from 2001–03, and is an AAAI fellow.