A Framework for Heuristic Reasoning About Uncertainty

Paul R. Cohen
Milton R. Grinberg

Heuristic Programming Project
Department of Computer Science
Stanford University, Stanford CA 94305

Abstract

This paper describes a theory of reasoning about uncertainty, based on a representation of states of certainty called *endorsements* (see Cohen and Grinberg, 1983, for a more detailed discussion of the theory.) The theory of endorsements is an alternative to numerical methods for reasoning about uncertainty, such as subjective Bayesian methods (Shortliffe and Buchanan, 1975; Duda, Hart, and Nilsson, 1976) and the Shafer-Dempster theory (Shafer, 1976). The fundamental concern with numerical representations of certainty is that they hide the reasoning that produces them and thus limit one's reasoning about uncertainty. While numbers are easy to propagate over inferences, what the numbers mean is unclear. The theory of endorsements represents the factors that affect certainty and supports multiple strategies for dealing with uncertainty.

Introduction

Nothing is certain. People's certainty of the past is limited by the fidelity of the devices that record it, their knowledge of the present is always incomplete, and their knowledge of the future is but speculation. Even though nothing is certain, people behave as if almost nothing is uncertain. They are adept at discounting uncertainty—making it go away. This paper discusses how AI programs—specifically, rule-based systems—might be made similarly adept.

There are serious limitations to current AI approaches to reasoning under uncertainty. Many authors have discussed the mathematical, psychological, and practical problems associated with numerical approaches (Quinlan, 1982; Shafer, 1976; Shortliffe and Buchanan, 1976; Duda, Hart, and Nilsson, 1976; Kahneman, Slovic, and Tversky, 1982), and their arguments are not repeated. This section proposes that numerical approaches to reasoning under uncertainty are restricted because the set of numbers is not a sufficiently rich representation to support the considerable heuristic knowledge about uncertainty and evidence.

A common AI method for reasoning under uncertainty is to augment domain inferences with parallel *certainty inferences*. Systems such as EMYCIN (van Melle, 1980) associate *certainty factors* with the conclusions of inference rules. A rule of the form *"IF A and B and C, THEN D"* asserts D when A, B, and C are certain; additionally, a number is associated with D to indicate one's belief that D follows from A, B, and C. If A , B, and C, though certain, suggest but do not confirm D, then the number associated with D will be less than the 1.0 that usually represents certainty in such systems. If A, B, or C are themselves uncertain, then the number associated with D is modified to account for the uncertainty of its premises. These numbers are given different names by different authors; we refer to them as *degrees of belief* (Shafer, 1976). The functions that propagate degrees of belief over inferences are called *combining functions*. Domain rules are assigned a priori degrees of belief and the purpose of the combining functions is to faithfully represent the intent of each in the context in which it is eventually used. Some systems propagate not one degree of belief, but two, indicating a range of certainty. In all cases, one's certainty in a hypothesis is represented only by a numerical degree of belief.

All AI systems that reason this way adhere to the scheme with little or no variation. All evidence is processed in exactly the same way: It propagates through a combining function to the conclusion it supports. Current systems are unable to treat different kinds of evidence differently. They do not take some evidence "with a grain of salt" and other evidence with judicial solemnity, and the reason is obvious: Since evidence is nothing more than a proposition with an associated number, there is no way to tell whether it warrants scrutiny other than by examining the number. This permits only a limited ability to discriminate kinds of evidence. If there are several kinds of evidence, then the numerical representation is inadequate; for example, one cannot expect to discriminate eyewitness evidence from circumstantial, hearsay, or photographic evidence on the basis of their numerical degrees of belief.

Numerical degrees of belief are adequate if one intends only to propagate them over inferences without reflection. But if one desires to treat different kinds of evidence differently, and if the weight of evidence can change in different contexts, then it is necessary to reason *about* one's evidence. Two problems must be solved to achieve this. The first is to represent explicitly many aspects of evidence besides "how much" it is believed. The other problem is to develop additional methods for reasoning with this richer information. The remainder of this paper outlines four components of a model of heuristic reasoning about uncertainty.

A theory of heuristic reasoning about uncertainty.

The theory of endorsements has four components. Endorsements are records of reasons for believing or disbelieving a hypothesis. In addition, the theory proposes heuristics for ranking endorsements, heuristics for propagating endorsements over inferences, and heuristics for discounting uncertainty.

Endorsement

The explicit marking of factors relating to one's certainty is called an *endorsement*. The process is best introduced by an analogy. A piece of work in a bureaucracy proceeds from one stage to the next contingent on the endorsement of a bureaucrat. The job must satisfy certain (typically formal) requirements before it is endorsed at any given level. Imagine, in place of bureaucrats watching over a job, a collection of rules watching over the development of a line of reasoning. Each endorses a step in the argument if it satisfies certain requirements. Whenever a domain rule is used, its conclusion accrues one or more endorsements. Thus, *endorsements* are just records that a particular kind of inference has taken place, and *endorsers* are just the computations that assert the records. Bureaucrats can require a job to be cleared by lesser bureaucrats before they even consider it; for example, a city council won't consider a proposal unless it is cleared by the planning department. Similarly, an endorser may require the conditions of a rule to have a certain level of endorsement before it will endorse the conclusion of the rule. Most conclusions accrue several, more or less stringent endorsements. The certainty of a hypothesis can be represented by its *pedigree*—its endorsements and those of its logical predecessors. In terms of the bureaucracy analogy, one's confidence in a job is proportional to the degree of scrutiny and level through which it has passed.[1]

Ranking endorsements

Implicit in this is the idea that endorsements can be ranked, and that a system needs heuristics to do the ranking. Preference of the endorsement of one hypothesis over the endorsement of another is equivalent to having more confidence in the one hypothesis over the other. For example, in most contexts, eyewitness testimony is preferable to circumstantial evidence, corroboration to contradiction, and inference to assumption. What is meant by one kind of evidence being preferable to another is that a conclusion endorsed as having the one kind of evidence for its support is more certain than a conclusion endorsed as supported by less preferred evidence. Much knowledge is needed to rank endorsements. For example, it is difficult to know whether the eyewitness account of a drunk is more certain than a dozen "respectable" but circumstantial anecdotes, but the ability even to pose the question (if not answer it) is evidence for the role of world knowledge in weighing evidence. If the drunkard's testimony is not comparable to that of the character witnesses (i.e., if neither is preferred to the other) then there would be no rules in the system to rank these kinds of evidence. This is in contrast to numerical approaches to reasoning about uncertainty, which give the erroneous impression that all evidence is comparable.

[1]Endorsement is similar to recording justifications in a truth maintenance system (Doyle, 1979), but with a crucial difference. In Doyle's system, a justification is used to decide whether a conclusion has support, but the kind of support is irrelevant. There are, however, many different kinds of endorsements, corresponding to different kinds of evidence for and against a proposition.

Propagating endorsements over inferences

Just as degrees of belief are propagated over inferences by combining functions, so must endorsements be propagated over inferences by heuristics. Given that one proposition implies another, and the premise has a set of endorsements, what endorsements should the conclusion accrue? A set of rules is needed to propagate endorsements over inferences. These serve the same purpose as combining functions, to make a rule sensitive to the context in which it is used, but the context is now a set of endorsements instead of a single degree of belief. A default rule is that the endorsements of the premise should all propagate to the conclusion. In fact, interactions between endorsements may complicate this picture. For example, if a parameter is derived by taking a *central value* of several others (e.g., an arithmetic mean) and if one of the component values is thought to be an **extreme** value, the endorsement **extreme** usually would *not* apply to the central value. Each domain of expertise is expected to have numerous idiosyncratic rules for propagating endorsements over inferences.

Discounting uncertainty

Endorsements can also be used to resolve uncertain values and discount uncertainty, although this, too, requires masses of general and special knowledge. A general method for resolving uncertain values is to choose a central or representative value in their stead. This is related to the hedging strategy discussed earlier. Alternatively, a highly endorsed hypothesis may be chosen over less well endorsed conflicting hypotheses; for example, the "credentials" of disputants can help decide between their claims—until recently, the word of a congressman outweighed that of a convicted swindler. Other methods are specific to the kinds of uncertainty they resolve. For example, if you can't resolve whether to withdraw $50 or $100 for an evening's entertainment, it's best to get $100.

Discounting uncertainty is not equivalent to certainty: One may be *certain enough* for one purpose but not for another. For example, one might be certain enough of one's prospective income to decide to buy a car, but not certain enough to decide to buy a house. The idea of *complete certainty* is an artifact of numerical representations of degree of belief. It is more productive to think in terms of *certainty with respect to a task*; in fact, it makes little sense to speak of certainty except with respect to a task.

Conclusion

We expect that each domain of expertise has a characteristic set of endorsements and a set of methods that define what they mean. These methods include ranking rules, rules for propagating endorsements over inferences, and rules for discounting uncertainty. The methods interact in characteristic ways. Uncertainty about alternative conclusions can be discounted either by choosing the one with the highest endorsement, or by replacing the uncertain conclusions by another, more certain one. This is not equivalent to saying that the chosen conclusion is certain, only that in the circumstances it is not uncertain enough to warrant other action. But in another context, such a resolution may not be certain enough, and the system will need to backtrack to its original uncertainty and reconsider it. Finally, a system that cannot resolve between uncertain conclusions may propagate both. This makes sense only if information will become available later to resolve the uncertainty. Otherwise, it is most expedient to hedge.

Summary

A heuristic approach to reasoning about uncertainty and a representation that supports it has been presented. The theory includes records of information that affect one's certainty, called endorsements; and rules for propagating endorsements over inferences, for ranking endorsements, and for discounting uncertainty. The rules interact in a classic planning pattern: balancing commitment to resolutions of uncertainty (and the penalty of backtracking) against least-commitment propagation of multiple uncertain hypotheses. For a detailed description of the theory see Cohen and Grinberg (1983).

References

[1] Cohen, P. R., and Grinberg, M. R. 1983. A theory of heuristic reasoning about uncertainty. The AI Magazine. Summer-Fall, 1983.

[2] Doyle, J. 1979. A truth maintenance system. Artificial Intelligence 12:231-272.

[3] Duda, R. O., Hart, P. E., and Nilsson, N. 1976. Subjective Bayesian methods for rule-based inference systems. Technical Note 124, Artificial Intelligence Center, SRI International, Menlo Park, CA.

[4] Kahneman, D., Slovic, P. and Tversky, A. 1982. Judgements under uncertainty: Heuristics and biases. Cambridge University Press.

[5] Quinlan, J. R. 1982. INFERNO: A cautious approach to uncertain inference. RAND Note N-1898-RC, RAND Corporation, Santa Monica, CA.

[6] Shafer, G. 1976. A mathematical theory of evidence. Princeton, NJ.: Princeton University Press.

[7] Shortliffe, E. H., and Buchanan, B. G. 1975. A model of inexact reasoning in medicine. Mathematical Biosciences 23:351-379.

[8] Van Melle, W. 1980. A domain-independent system that aids in constructing knowledge-based consultation programs. Report STAN-CS-80-820, Department of Computer Science, Stanford University.