Heuristic Reasoning about Uncertainty: An Artificial Intelligence Approach

Pitman Advanced Publishing Program
BOSTON · LONDON · MELBOURNE

Paul R. Cohen
Department of Computer and Information Science,
University of Massachusetts at Amherst

Heuristic Reasoning about Uncertainty: An Artificial Intelligence Approach

Pitman Advanced Publishing Program
BOSTON · LONDON · MELBOURNE

PITMAN PUBLISHING LIMITED
128 Long Acre, London WC2E 9AN

PITMAN PUBLISHING INC
1020 Plain Street, Marshfield, Massachusetts 02050

Associated Companies
Pitman Publishing Pty Ltd, Melbourne
Pitman Publishing New Zealand Ltd, Wellington
Copp Clark Pitman, Toronto

First published 1985

Dedication
To **Carole Ruth Beal**, for everything

British Library Cataloguing in Publication Data

Cohen, Paul R.
Heuristic reasoning about uncertainty: an artificial intelligence approach.—(Research notes in artificial intelligence. 2)
1. Artificial intelligence 2. Problem solving
3. Uncertainty 4. Heuristic programming
I. Title II. Series
001.53'9 Q335

ISBN 0-273-08667-7

Library of Congress Cataloging in Publication Data

Cohen, Paul R.
Heuristic reasoning about uncertainty.
Bibliography: p.
1. Uncertainty (Information theory) 2. Heuristic programming. 3. Artificial intelligence. 4. SOLOMON (Computer program) I. Title.
Q375.C64 1985 001.53'9 84-25533

ISBN 0-273-08667-7

Reproduced and printed by photolithography
in Great Britain by Biddles Ltd, Guildford

Acknowledgements

This work has been shaped by many individuals. Some have helped with the rough hewing, and others with the fine points. To all, I wish to express my deep gratitude.

My friend and companion Danny Asher suggested portfolio management as a domain for this work. I was then fortunate to meet Mark D. Lieberman, of Drexel Burnham Lambert, who agreed to be the domain expert for the FOLIO program, and was avid and thoughtful thoughout. The FOLIO effort continued over a year, without complaint from Mark, who was told initially that it would take a few weeks. Professor William F. Sharpe, of the Graduate School of Business at Stanford University, helped me with FOLIO's mathematical algorithms, taught me much about portfolio management and investment, and later agreed to serve on my reading committee. I am very grateful for his insights and enthusiasm for FOLIO and the model of endorsement.

My thanks to the members of my committee, Douglas B. Lenat (Chair), William F. Sharpe, Edward H. Shortliffe, and Amos Tversky. Four academic departments are represented by this committee, and it is due to the efforts of the committee that the work in this dissertation is of general relevance. I wish to thank Doug Lenat, a fountain of ideas, for his enthusiasm and advocacy. Ted Shortliffe invariably and immediately focussed on the weak points of the model of endorsement. I am grateful to him for making the model – and this document – very much stronger by his incisive and perceptive comments, and for his help with the probabilistic end of things. Amos Tversky has cheerfully raised numerous, difficult, and subtle questions. My thanks to him for providing research questions for years to come, and for his encouragement about the potential of this work.

The model of endorsement was developed in a series of conversations with Milton R. Grinberg, of Stanford University. Our collaboration was spirited, productive, and a high point of my stay at Stanford. It has been my good fortune to find in Milt a critic, a colleague, and a friend.

I would also like to thank Avron Barr, Gary Bonitatibus, Tom Dietterich, and Anne Gardner for their thoughtful comments about the developing model of endorsement, and for their camaraderie and kindness during our years at Stanford, especially during the exciting years of the *Handbook of Artificial Intelligence*. I have a great debt to the technical editor of the *Handbook*, Dianne Kanerva, who taught me to write quickly and without fear. This skill, learned during two years of close collaboration, is invaluable, and my gratitude to Dianne for her patient tutelage and friendship is inestimable.

I have a special debt to Professor Edward A. Feigenbaum. He invited me to Stanford to work for just a year on the first volume of the *Handbook of Artificial Intelligence*. Instead, I stayed for five years, through *three* volumes, co-editing the third with Ed. I owe much of what I have, and of what I have done, to Ed's kindness and generosity.

I have been fortunate, in the year since this document was first printed, to have found excellent colleagues at the University of Massachusetts at Amherst. They have helped much with the development of the theory of endorsements. I wish to acknowledge the contributions of Norm Carver, Alvah Davis, David Day, Brian Delaney, Tom Gruber, Philip Johnson, Larry Lefkowitz, Melinda Mesick, and Mike Sullivan.

This research was funded by a grant from NL Industries, for which I am grateful.

My debts to my colleagues are great, but I owe no-one as much as I owe Carole R. Beal. I expect it will take a lifetime to repay her.

Contents

1 INTRODUCTION 1
- 1.1 The Problem 1
- 1.2 The SOLOMON Program 4
  - 1.2.1 Why SOLOMON? 7
- 1.3 Normative and Performance Approaches to Uncertainty 8
- 1.4 Evaluating the Research 10
- 1.5 A Guide for the Reader 11

2 MANAGING UNCERTAINTY 14
- 2.1 Introduction 14
- 2.2 The Representation and Management of Uncertainty in AI Programs 17
  - 2.2.1 The Engineering Approach 17
  - 2.2.2 Diversification 20
  - 2.2.3 Parallel Certainty Inferences: Introduction 21
  - 2.2.4 Bayesian Inference and Related Methods 27
  - 2.2.5 Belief Functions: The Dempster-Shafer Theory of Evidence 37
  - 2.2.6 Fuzzy Logic as an Approach to Vagueness 41
  - 2.2.7 Control Approaches to Uncertainty 42
  - 2.2.8 The Blackboard Model 44
- 2.3 The Model of Endorsement 48
  - 2.3.1 Two Analogies for The Model of Endorsement 49
  - 2.3.2 Design Criteria for the Model of Endorsement 51
  - 2.3.3 Architecture of SOLOMON 54
    - 2.3.3.1 Predicate Calculus Notation 55
    - 2.3.3.2 The Backward Chaining Control Structure 56
    - 2.3.3.3 Control in SOLOMON 57
    - 2.3.3.4 Knowledge in SOLOMON 61
  - 2.3.4 Research Issues for the Model of Endorsement 63
  - 2.3.5 AI Concepts Related to the Model of Endorsement 66

3 BAYESIAN AND ENDORSEMENT-BASED ANALYSES: TWO CASES 71
- 3.1 Introduction 71
- 3.2 Weighing the Evidence of Walker and Leakey 72
  - 3.2.1 Introduction 72

3.2.2 The Problem of the Three Fossil Forms 74
3.2.3 Walker and Leakey's Arguments 77
3.2.4 A Shafer-Tversky Analysis of the Walker-Leakey Evidence 80
3.2.5 An Endorsement-Based analysis of Walker and Leakey's Evidence 84
3.3 Comparison of the Bayesian and Endorsement-Based Approaches 91
3.4 Multiple Dependent Inferences 98
3.4.1 An Analysis Based on Numerical Degrees of Belief 104
3.4.2 An Endorsement-Based Analysis 107
3.5 A Comparison of the Approaches for Multiple Inference Tasks 109
3.6 Conclusion 113

4 SOLOMON: AN AI PROGRAM FOR REASONING ABOUT UNCERTAINTY 114
4.1 Introduction 114
4.2 SOLOMON and the Problem of Investments for Spencer 118
4.2.1 An Example of SOLOMON 119
4.3 Representation of Knowledge in SOLOMON 126
4.3.1 Data Structures: Introduction 126
4.3.2 Knowledge Structures Generated During the Example 128
4.3.3 Procedural Knowledge in SOLOMON 139
4.3.3.1 An Endorsement-ranking Scheme 139
4.3.3.2 Screening Tasks 141
4.3.3.3 Deciding that an Endorsement is Adequate 141
4.3.3.4 Generating Resolution Tasks 142
4.4 Endorsers and Endorsements 143
4.4.1 Rule Endorsements 143
4.4.1.1 Rule Condition Endorsements 143
4.4.1.2 Discussion of Rule-condition Endorsements 145
4.4.1.3 Propmatch-if and Rulematch-if for Rule Condition Endorsements 148
4.4.1.4 Rule Inference-type Endorsements 150
4.4.1.5 Conclusion Rule Endorsements 152
4.4.2 Task Endorsements 153
4.4.3 Conclusion Endorsements 154
4.4.4 Data Endorsements 155
4.4.5 Resolution Endorsements 156
4.4.6 Propagating Endorsements over Inferences 156
4.4.7 Ranking Endorsements 159
4.4.8 The Meaning of Endorsements 161
4.5 Control of processing in SOLOMON 162
4.5.1 SOLOMON's Control Structure: Example 1 164
4.5.2 SOLOMON's Control Structure: Example 2 170

5 FUTURE RESEARCH 175
5.1 Introduction 175
5.2 Analysis of the Endorsement-based Approach 176
5.2.1 The Contribution of this Research 176
5.3 Problems with SOLOMON 177
5.3.1 Resolution, Propagation, and Ranking 180
5.4 Evaluating Model of Endorsement and SOLOMON 185
5.5 What the Future Really Holds 191

6 REFERENCES 198

1 Introduction

1.1 The Problem

The problem addressed by this document is how to represent and reason with heuristic knowledge about uncertainty. We know much about uncertain situations besides our *degree* or strength of belief: We discriminate among many *kinds* of uncertainty, and we know many approaches to resolving or discounting uncertainty. We recognize different kinds of evidence, we prefer some kinds to others, and we recognize that our certainty depends not only on evidence but also on the *importance* of the uncertain situation. We can judge the "utility" of evidence, and decide whether it is worth the effort to obtain it. Remarkably, in a world where almost nothing is certain, we use our knowledge about uncertainty to behave *as if* almost nothing is uncertain. How can we represent and reason with this all-important knowledge about uncertainty? This document proposes an artificial intelligence approach – a model of heuristic reasoning about uncertainty implemented as a computer program.

Uncertainty is not a new problem to academic research. Indeed, its has been studied in great depth by decision analysts, statisticians, philosophers, psychologists, gamblers, insurance companies, as well as practitioners of artificial intelligence (AI). One might legitimately question the utility of more research on uncertainty. The approach to uncertainty of this research is novel in a number of respects: Uncertainty is represented in terms of its effects on reasoning; it is viewed as a structured object that traverses a line of reasoning, an object whose properties change with the style of reasoning, and with the availability, type, and quality of evidence. Particular kinds of uncertainty, when they interfere with problem solving goals, are themselves regarded as problems to be solved. Part of the proposed model of reasoning about uncertainty concerns how these "uncertainty problems" are resolved or discounted. Although this research concentrates on aspects of the model that enhance the reasoning abilities of AI programs, the superordinate goal of the research is to develop a plausible model of human reasoning about uncertain situations. Clearly, to the extent that this goal is achieved, the relevance of this work goes beyond artificial intelligence.

Artificial intelligence is a fortuitous discipline from which to study reasoning about uncertainty. The AI approach regards reasoning as composed of interacting information processes, and as active and constructive. AI methods encourage one to view uncertainty as a problem to be solved by the application of heuristic knowledge. Given this perspective, it is natural to study the *effects* of uncertainty on reasoning processes. For example, to solve a problem, we first state goals, then think about an order in which to achieve them, then search for methods to achieve the goals, and so on. Consider the effects of uncertainty on the process of ordering goals: If the goals are disjunctive, such that solving any one solves the entire problem, then our first attempts should be to solve the goals for which we have the best evidence; but if the goals are conjunctive, such that *all* must be solved, then we should first attempt the goals for which we have the *poorest* evidence, because if these attempts fail, we can save ourselves the effort of working on the other goals. Even this rudimentary reasoning about uncertainty and evidence can obviously have profound effects on the efficiency of problem solving. But it depends on a sufficient representation of knowledge about uncertainty: For the example to succeed, we must represent the strength of evidence for and against each of the goals, and whether the goals are conjunctive or disjunctive. These are simple requirements, easily achieved with current AI technology.

But AI lacks the tools to manage other kinds of knowledge about uncertainty. The previous example was managable because the logical structure of goals (i.e., whether they are conjunctive or disjunctive) is usually explicit in the statement of the goals, and because strength of evidence is easily represented by numbers. But in other cases "strength of evidence" is actually a summary of several factors that pertain to certainty, and if an intelligent reasoner normally discriminates among these factors, then the summary representation is inadequate. For example, you may be equally certain that there are no Venusians and that the Palo Alto municipal swimming pool is not painted green and pink. You believe there are no Venusians because there is no evidence of Venusians, but the conclusion is not entirely certain because if there *were* Venusians, you would not necessarily know about them. You believe that the Palo Alto municipal swimming pool is not green and pink because you have never seen any such swimming pool, but you are not entirely certain because you have never seen the Palo Alto pool. Thus the strength of belief of each statement is composed of a reason to believe and a reason to disbelieve. The difference between the statements is that the nonexistence of Venusians is unprovable (at least in the short run) while the coloration of the Palo Alto municipal pool is demonstrably not pink and green. This difference is not captured in the strengths of belief of the statements, and it is not even necessarily important if we limit our reasoning to assessing evidence for and against the statements. But, if we intend to use the statements in subsequent arguments, then the ability to discriminate among provable and unprovable statements becomes useful. For example, we may decide to limit our reasoning entirely to statements that are disprovable, on the grounds that non-disprovable statements – however credible – reduce the credibility of the arguments they support. This kind of reasoning exceeds the capabilities of current formal approaches to reasoning under uncertainty.

The reasoning of the previous example is qualitatively different, for several reasons, from common numerical approaches to uncertainty in AI and other disciplines. First, the emphasis is on decomposing strengths of belief into reasons for believing and disbelieving. Second, these reasons are not blindly tallied, as strengths of belief often are, but are subject to heuristic reasoning; for example, we may survey our statements of evidence and decide to use non-disprovable statements only if none of our evidence is disprovable. Thus, the weight of all evidence is not calculated in the same way, but depends on the kind of evidence, on the immediate and ultimate problem-solving goals, on the standards one applies to judge arguments, and so on. Third, given *reasons* to believe and disbelieve statements, their *strength* of belief may become irrelevant; for example, evidence that cannot be disconfirmed is less preferable *regardless* of its credibility. These differences lead me to the term "reasoning *about* uncertainty," in preference to the more frequently used "reasoning *under* uncertainty"; my intention is to emphasize that uncertainty is itself an object of reason.

To reiterate, the problem addressed by this research is to represent and reason with the knowledge we have about uncertainty. As the previous paragraphs suggest, the appropriate course is to develop a representation suitable to expressing reasons for believing and disbelieving, and to specify the effects of these reasons on problem-solving processes – on intelligent reasoning.

1.2 The SOLOMON Program

This document describes an artificial intelligence approach to uncertainty called the *model of endorsement* and the structure of a program – called SOLOMON – that implements the model. Endorsements are the structured objects that represent reasons for believing and disbelieving the propositions with which they are associated. SOLOMON uses endorsements to control its reasoning about uncertainty and evidence. Reasons for believing are called *positive* endorsements, reasons for disbelieving, *negative* endorsements. The program uses a modified backward-chaining control structure to prove statements in the predicate calculus. Currently, SOLOMON is a "toy" system that works with just a few inference rules, endorsements, and other kinds of knowledge. It is intended ultimately as a tool for building rule-based expert systems in domains where reasoning about uncertainty is necessary.

The following example illustrates some of the central ideas in the model of endorsement and SOLOMON:

> Consider the question of which of several cars to buy. Test-driving cars is time-consuming and unpleasant, so we want to limit our search to a few likely cars. We search for evidence about numerous cars by posting a bulletin on an electronic bulletin board, asking for information about the candidate manufacturers and models. We get two kinds of responses from drivers who are either satisfied or dissatisfied with their cars. This "evidence" for or against each kind of car has one of two endorsements: Responses from people who *own* the cars they describe get the **long-term** endorsement; and responses from people who comment on cars that they rented, borrowed, or otherwise drove for just a short time, get the **short-term** endorsement. We decide that, for the purpose of *eliminating* cars from consideration, either endorsement is adequate. This is tantamount to saying that a brief but inadequate experience with a car is as compelling evidence as a long and inadequate experience, at least for the purpose of eliminating a car from consideration. Thus, any car that gets a poor review is eliminated. This leaves half a dozen cars that get good reviews. To further narrow the field, we decide to eliminate any recommendation endorsed as **short-term**, on the grounds that a brief affair with a car does not provide the opportunity to observe its subtle strengths or weaknesses. This still leaves two contenders, so we decide to test drive both of them.

One thing to note about the example is that the endorsements, **short-term** and **long-term**, are not general comments about any evidence, but are specific reasons to disbelieve and believe evidence about things that people own, such as cars. It is assumed here that any domain will have a characteristic set of endorsements. However, in an attempt to build a general system, the SOLOMON program works with a small set of very general endorsements.

A second observation is that the *adequacy* of evidence depends on what one intends to use it for. For the purpose of eliminating cars from consideration, evidence endorsed as **short-term** is considered adequate; but for the purpose of including cars in the candidate list, it is considered inadequate. Note that domain knowledge is required to decide in which contexts an endorsement is adequate. The

SOLOMON program uses rules to decide when evidence is "believable enough," and this decision depends on the endorsements of the evidence and on the purpose to which the evidence will be put.

The example also suggests that endorsements can be ranked, at least in the default sense that adequately endorsed evidence is *more believable* than evidence that is inadequately endorsed. Beyond this, SOLOMON also has general rules for ranking adequately endorsed pieces of evidence, although this poses considerable difficulty when evidence is supported by combinations of endorsements (but see Section 5.5). Domain knowledge is required, of course, to rank domain-specific endorsements.

Finally, note that when the available evidence does not support a choice between two options, it is possible to *resolve* the conflict in other ways. The example ends with two options that are equally believable (i.e., they are both supported by evidence endorsed as long-term). The question of which car to test-drive cannot be resolved on the basis of the available evidence, but it can be resolved by driving both cars. Methods to resolve uncertainty are domain-dependent: The general "do both" strategy for resolving conflicts may work for test-driving cars, but it is an expensive resolution for conflicts about which car to *buy*. SOLOMON has knowledge about which general resolution methods are appropriate for which kinds of uncertainty.

In sum, this example illustrates four kinds of knowledge that are required in the model of endorsement to reason about uncertainty. These are endorsements, criteria for when endorsements qualify a statement as certain enough to use as evidence, criteria for ranking endorsements, and rules for generating resolution tasks. (An important kind of knowledge that was not illustrated in the simple example concerns how endorsements propagate over inferences.) Individual problem domains are assumed to afford characteristic kinds of endorsements, as well as the criteria for reasoning with them. SOLOMON, however, is an attempt to capture some fairly general endorsements and reasoning methods. This is partly because of difficulties in obtaining endorsements of inference rules in SOLOMON's test domain, and partly because of a desire to develop a general tool for building expert systems.

1.2.1 Why SOLOMON?

SOLOMON's namesake, King Solomon, was very famous for his wisdom, and so one is hesitant to use his name for an AI program. However, the choice is not inappropriate. A tale of Solomon's wisdom is found in First Kings, Chapter 3: Solomon was approached by two harlots, each of whom claimed to be the mother of a baby. The one argued that her child was stolen by the other, whose own child had died in the night. The scriptures go to lengths to show how little evidence Solomon had to work with; for example, one harlot says "there was no stranger with us in the house, save we two in the house." In other words, there were no witnesses; the king had to judge based on two pieces of equally plausible, conflicting testimony. His solution was to order the child cut in two. When one woman protested, less grieved to relinquish the child than to see it chopped up, the king recognized her as the true mother.

Possibly, this story will seem unrelated to the theory of uncertainty in AI but, in fact, it illustrates three basic precepts of the model of endorsement. First, reasoning about uncertainty requires knowledge, and lots of it. Solomon knew much more than the harlots told him; he used commonsense knowledge to fill in the details, to recognize conflict as a source of uncertainty; and he used wisdom to resolve the conflict by evoking more evidence. Second, the choice of a method to resolve or discount uncertainty depends on explicit representation of the reason for uncertainty. In Solomon's case, the reason was conflicting testimony, and the resolution method was to force the situation to some extreme point where the two women would no longer make identical claims. Other kinds of uncertainty succumb to other methods, provided the type of uncertainty is made explicit. Third, it is imperative to record one's motives for acting, if one is to justify one's decisions later. It is quite possible that Solomon ordered the baby chopped in half because he thought it would be equitable, or because he was irritated by the whole situation. The uncertainty about Solomon's motives no doubt provides good fare for hermeneutics, but reasoning about uncertain situations should not itself introduce uncertainty – by shoddy record-keeping – if it can be avoided.

1.3 Normative and Performance Approaches to Uncertainty

Uncertainty is most often represented in terms of probability. It is natural to speak of the probability of life on Venus or the odds that a football team will win a match. Moreover, probability theory provides a relatively simple calculus for adjusting probabilities in the light of evidence. Once probability is accepted as an appropriate representation of uncertainty, it becomes possible to contrast human performance under uncertainty with performance as specified by probability theory. Descriptions of the first kind are called *performance* or *positive* theories, and theories of reasoning that are consistent with the axioms of probability theory are called *normative*. Human reasoning under uncertainty is clearly not normative (e.g., Edwards, 1968); even expert reasoning violates results from probability theory (e.g., Eddy, 1982; Tversky and Kahneman, 1971). Much has been learned about human judgment and decision by comparing it with normative models. The models serve as a basis for comparison and thus for a vocabulary of effects: Conservatism is conservatism with respect to the outcome of a normative analysis. Systematic deviations from normative models are seen as evidence of *heuristics* in human reasoning about uncertainty.[1]

There are disciplines in which non-normative reasoning is undesirable, for example, calculating odds in poker or blackjack. (The term *prescriptive* is sometimes used for normative theories because they specify what one *should* do if, for example, one wants to play poker or blackjack as well as possible.) Other disciplines, such as cognitive psychology, welcome non-normative behavior as an object of study. There is some question about the attitudes of researchers in artificial intelligence. It has long been recognized in AI that heuristic reasoning is a source of power, that the model of humans as perfect processors of information is not only inaccurate but is also unlikely to lead to efficient and intelligent reasoning. Aspects of cognition that at one time were regarded as "bugs," now appear to be "features"; for example, limited short-term memory serves an important purpose in focusing attention, and is now mimicked by the *agenda* mechanism and other control processes. It is therefore puzzling that AI retains models of reasoning under uncertainty that are derived from normative theories, the more so because the assumptions of the normative approaches are frequently violated, and because the probabilistic interpretation -- and numerical representation -- of uncertainty summarizes and fails to discriminate among reasons for believing and disbelieving. One may speculate that AI uses quasi-probabilistic, numerical methods for reasoning under uncertainty less for any advantages of normative reasoning than for lack of other methods. The research described in this document is an attempt to develop a new method that has obvious advantages over the old.

Two questions that must be asked of a new approach to reasoning about uncertainty are whether the approach is normative and, if so, what does

[1] The main spokesmen for this view were and are Kahneman and Tversky, and their colleagues (see, e.g., Tversky and Kahneman, 1974; Kahneman and Tversky, 1979; Tversky and Kahneman, 1981. See also Kahneman, Slovic, and Tversky for a wide-ranging collection of papers on judgment.)