AI Magazine Volume 12 Number 4 (1991) (© AAAI)

Decision Analysis and Expert Systems

Max Henrion, John S. Breese, and Eric J. Horvitz

Decision analysis and expert systems are technologies intended to support human reasoning and decision making by formalizing expert knowledge so that it is amenable to mechanized reasoning methods. Despite some common goals, these two paradigms have evolved divergently, with fundamental differences in principle and practice. Recent recognition of the deficiencies of traditional AI techniques for treating uncertainty, coupled with the development of belief nets and influence diagrams, is stimulating renewed enthusiasm among AI researchers in probabilistic reasoning and decision analysis. We present the key ideas of decision analysis and review recent research and applications that aim toward a marriage of these two paradigms. This work combines decision-analytic methods for structuring and encoding uncertain knowledge and preferences with computational techniques from AI for knowledge representation, inference, and explanation. We end by outlining remaining research issues to fully develop the potential of this enterprise.

Decision analysis and knowledge-based expert systems share some common goals. Both technologies are designed to improve human decision making; they attempt to do this by formalizing human expert knowledge so that it is amenable to mechanized reasoning. However, the technologies are based on rather different principles. *Decision analysis* is the application of the principles of decision theory supplemented with insights from the psychology of judgment. *Expert systems,* at least as we use this term here, involve the application of various logical and computational techniques of AI to the representation of human knowledge for automated inference. AI and decision theory both emerged from research on systematic methods for problem solving and decision making that first blossomed in the 1940s. They even share a common progenitor, John von Neumann, who was a coauthor with Oscar Morgenstern of the best-known formulation of decision theory as well a key player in the development of the serial computer. Despite common roots, AI soon distinguished itself in its concern with autonomous problem solving and its emphasis on symbolic, rather than numeric, information.

Some of the earliest AI research addressed approximations and heuristics for complex tasks in decision-theoretic formulations (Simon 1955), and early work on expert systems for diagnosis (although before this term became popular) used Bayesian and decision-theoretic schemes (Gorry and Barnett 1968). However, many AI researchers soon lost interest in decision theory. This disenchantment arose, in part, from a perception that it was hopelessly intractable and inadequate for expressing the rich structure of human knowledge (Gorry 1973; Szolovits 1982).

Although similar views are still widespread among AI researchers, there has been a recent resurgence of interest in the application of probability theory, decision theory, and decision analysis to various problems in AI. In

This article was adapted from "Decision Theory in Expert Systems and Artificial Intelligence" by E. J. Horvitz, J. S. Breese, and M. Henrion, *International Journal of Approximate Reasoning*—Special Issue on Uncertainty and Artificial Intelligence, volume 2, pages 247–302. Copyright 1988 by Elsevier Science Publishing Co., Inc. Reprinted by permission of the publisher.

0738-4602/91/$4.00 ©1991 AAAI

this article, we examine some of the reasons for this renewed interest, including an increasing recognition of the shortcomings of some traditional AI methods for inference and decision making under uncertainty and the recent development of more expressive decision-theoretic representations and more practical knowledge engineering techniques.

The potential contributions for tackling AI problems derive from the framework of decision theory and the practical techniques of decision analysis for reasoning about decisions, uncertain information, and human preferences. Decisions underlie any action that a problem solver might take in structuring problems, reasoning, allocating computational resources, displaying information, or controlling some physical activity. As AI moved beyond toy problems to grapple with complex, real-world decisions in such areas as medicine, business, and aerospace, the importance of explicitly dealing with the uncertainty due to partial information and incomplete models became increasingly evident. Real-world applications have also revealed the importance of modeling human preferences and attitudes toward risk, central topics of decision theory to which traditional AI research has paid little attention.

The purposes of this article are to provide an introduction to the key ideas of decision analysis for an AI-oriented reader and to review recent research that is applying these ideas to the development of a new generation of expert systems that treat uncertainty, preferences, and decision making on a more principled basis. In particular, we concentrate on the use of influence diagrams and belief nets and their role in representation, knowledge engineering, tractable inference, and explanation. Although we believe that decision theory and decision analysis can also make valuable contributions to broader issues in AI, including planning, reasoning under constrained resources, autonomous agents, and a variety of other topics, these subjects are beyond the scope of this article. Our focus here is specifically on their contributions to knowledge-based expert systems.

Foundations of Decision Analysis

The foundations of probability theory extend at least as far back as the seventeenth century in the works of Pascal, Bernoulli, and Fermat. *Probability* is a language for expressing uncertainty about propositions and quantities in terms of degrees of belief. *Decision theory* extends this language to express preferences among possible future states of the world and, hence, among alternative actions that might lead to them. Probability and decision theory provide a set of principles for rational inference and decision making under uncertainty. By themselves, however, these mathematical theories are insufficient to address real problems. *Decision analysis* is the art and science of applying these ideas to provide practical help for decision making in the real world.

We start by introducing the essential ideas of decision theory and decision analysis. We will not detail the axioms of decision theory or present decision analysis techniques, such as decision trees. Many excellent texts present this material (Howard and Matheson 1984; Raiffa 1968). Here, we focus on the essential underlying ideas and rationale, which are often misunderstood even among those who have mastered the technical material.

Subjective Probability

A *probability,* of course, is simply a number expressing the chance that a proposition is true or that some event occurred, with a value in the range from 0 (certainly false) to 1 (certainly true). A *subjective probability* is a probability expressing a person's degree of belief in the proposition or occurrence of an event based on the person's current information. This emphasis on probability as a personal belief depending on available information contrasts with the propensity and frequency views of probability as something existing outside any observer. In the *propensity view*, probability is a physical property of a device, for example, the tendency of a particular coin to come up heads. In the *frequency view*, probability is a property of a population of similar events, for example, the fraction of heads in a long sequence of coin tosses.

A subjectivist might start with some prior belief about the fairness of the coin, perhaps based on experience with other coins, and then update this belief using Bayes's rule as data become available from experimental tosses. After many coin tosses, the belief of the subjectivist will, in general, converge to the observed frequency as the data overwhelm the prior belief. Thus in the long run, the subjectivist and the frequentist will tend to agree about a probability. The key distinction is that the subjectivist is willing to assign probabilities to events that are not members of any obvious repeatable sequence, for example, the discovery of room temperature superconductivity before the year 2000, but the frequentist is not. Almost all real-world prob-

Decisions underlie any action that a problem solver might take…

The appeal of subjective probability is that it is applicable whether there is little or much data available.

lems of interest to decision analysts and expert system builders involve at least some uncertain events or quantities for which empirical data are unavailable or too expensive to collect; so, they must resort to the use of expert opinion. The appeal of subjective probability is that it is applicable whether there is little or much data available.

Many people find the subjectivist (also known as Bayesian or personalist) view of probability natural. However, a few find the use of probability a stumbling block, often because the distinction between subjective and classical views is not made clear, as it rarely is in introductory courses on probability and statistics. An indication of this confusion is the common misapprehension that probabilistic methods are only applicable when large amounts of data are available. Subjective probabilities are often used to encode expert knowledge in domains where little or no direct empirical data are available. However, if and when data do become available, Bayesian reasoning provides a consistent framework to combine data with judgment to update beliefs and refine knowledge.

If a probability depends on who is assessing it, then it does not make sense to talk about "the" probability. Therefore, the subjectivist talks about "your" probability or "expert A's" probability. To draw attention to the fact that the probability is based or conditioned on the information available to the assessor, it is often explicitly specified. $P(X \mid s)$ is used to notate the probability of a proposition or event X conditioned on s, the assessor's prior state of information or background knowledge. If the assessor gets a new piece of evidence E, the revised probability of X is written $P(X \mid E, s)$, where the comma denotes the conjunction of evidence E and prior knowledge s. We often call $P(X \mid s)$ the prior probability, that is prior to observing E, and $P(X \mid E, s)$ the posterior probability.

Qualitative Structuring and Conditional Independence

Probability is best known as a way to quantify uncertain beliefs, but at least as important is its role in providing a principled basis for qualitative encoding of belief structures. This qualitative role is based on the notion of conditional independence. Event A is judged independent of event B conditional on background knowledge s if knowing the value of B does not affect the probability of A:

$$P(A \mid s) = P(A \mid B, s) \ .$$

Conditional independence is symmetric, so the previous equation is equivalent to stating that B is conditionally independent of A:

$$P(B \mid s)= P(B \mid A, s) \ .$$

Conditional independence formalizes the qualitative notion that A and B are irrelevant to each other. Conditional independence and, conversely, dependence, provide the basis for expressing the qualitative structure in graphic form as a belief network or influence diagram. Practicing decision analysts, like knowledge engineers building conventional expert systems, often claim that designing the qualitative structure is far more important than precision in the numeric parameters. In Knowledge Representations from Decision Analysis and in Knowledge Engineering and Decision Analysis, we discuss structuring techniques using belief nets and influence diagrams as well as methods for eliciting numeric probabilities.

Decision Theory

Decision theory is based on the axioms of probability and utility. Where probability theory provides a framework for coherent representation of uncertain beliefs, utility theory adds a set of principles for consistency among beliefs, preferences, and decisions. A *decision* is an irrevocable allocation of resources under the control of the decision maker. *Preferences* describe a decision maker's relative ordering of the desirability of possible states of the world. The key result of utility theory is that, given fundamental properties of belief and action, there exists a scalar function describing preferences for uncertain outcomes. Utility theory also provides ways to express attitudes toward uncertainty about outcome values, such as risk aversion. The valuation of an outcome can be based on the traditional attributes of money and time as well as on other dimensions of value, including pleasure, pain, life-years, and computational effort. Multiattribute utility theory provides ways to combine all these elements to produce a single scalar utility to represent the relative desirability of any certain outcome. Based on the axioms of decision theory, it is relatively easy to show that one should select the decision that maximizes the expected utility over a set of decisions with uncertain outcomes.

Normative versus Descriptive Theories

The axioms of probability and decision theory are fairly simple and intuitive. Many but not all who examine them find them compelling as principles for rational action under uncertainty. The theory is often referred to as *normative*, in the sense that it provides a set of criteria for consistency among beliefs, preferences, and choices that it claims should be adhered to by a rational decision maker. Given a set of beliefs and preferences, the theory prescribes which decisions should be chosen, namely, those that maximize expected utility.

Decision theory is not proposed as a *descriptive* theory; it does not purport to provide a good description of how people actually behave when making choices under uncertainty. Indeed, it has provoked a large body of empirical research examining the differences between how decision theory suggests we ought to behave and how we actually do (Kahneman, Slovic, and Tversky 1982). These psychological studies have found qualitative similarities between human intuitive judgment under uncertainty and the prescriptions of decision theory, but they have also demonstrated pervasive and consistent biases and inconsistencies. These differences seem to result from various mental heuristics that we use to render complex reasoning tasks amenable to our cognitive capacities.

Proponents of alternative schemes for uncertain reasoning used in expert systems, such as certainty factors and fuzzy set theory, have sometimes used these findings from psychology to justify using these formalisms on the grounds that they can be better models of human reasoning than probability and decision theory. However, virtually no empirical evidence shows that these alternative schemes provide better psychological models. In any case, for decision analysis, the goal is to improve human reasoning rather than replicate it. The observed biases and inconsistencies in unaided human reasoning are central to the justification for using normative aids, just as our limited capacity for mental arithmetic is the reason we find electronic calculators so useful.

Divide and Conquer

Decision analysis does not aim to avoid subjective judgments. That would be impossible. Rather, its strategy is divide and conquer: It replaces complex subjective judgments about which decisions are best with simpler subjective judgments about the probabilities of component events and relative preferences for elements of possible outcomes. The psychological literature tells us what kinds of judgments people find simplest and how to minimize cognitive biases in obtaining them. The components of the decomposed model are then reassembled to obtain recommendations about the complex decisions implied by the simpler judgments. Decision theory justifies the technical operations involved in the reassembly, such as applying Bayes's rule and selecting choices to maximize expected utility. Thus, decision analysis does not seek to eliminate human judgment but, rather, to simplify and clarify it.

The Focusing of Attention

Real-world situations have unlimited complexity in terms of the number of conceivable actions, states of the world, and eventual outcomes. Models are necessarily incomplete and uncertain. The key issue in modeling, whether for decision analysis or knowledge engineering using other representations, is the focus of attention: how to identify what matters and ignore the rest. Decision analysis provides a variety of techniques for focusing attention by performing sensitivity analysis to help identify those uncertainties and assumptions that could have a significant effect on the conclusions. The focus on decisions is of particular importance: The question about a possible model elaboration is not simply, Might it be relevant? but instead, Is it likely to change the resulting decision recommendations? This method turns out to be an extremely effective way to guide the modeling process and pare away the inessential. Resources can then be directed to modeling and analyzing the most sensitive aspects of the problem.

Insight, Not Numbers

A decision analyst that simply presents a decision maker with the numeric expected utility of each decision strategy, along with the injunction to choose the highest expected utility, is unlikely to be effective. Practicing

The question …is not, Might this be relevant? but instead, Is it likely to change the resulting decision…

decision analysts, like builders of expert systems for decision support, discovered early that the most important product of the analysis is not the numbers or even the recommended decision but the improved insights for the decision makers. These insights come from understanding why one decision is recommended over another and which assumptions and uncertainties are most critical to this conclusion. The process is generally fostered by the close involvement of decision makers in the modeling and analysis process. Without a basic understanding of the analysis, decision makers are unlikely to accept the results as a trustworthy basis for action. This perspective emphasizes the importance of clear explanations of the model assumptions and analysis if they are to be used and useful.

The Practice of Decision Analysis

Just as success in building conventional expert systems requires a great deal more than understanding the relevant AI research, decision analysis involves a great deal more than decision theory. Among other things, it includes techniques for structuring problems, encoding probabilities and utilities, computing implications, analyzing sensitivities, and explaining results to highlight insights, as we discuss later. Decision analysis emerged in the 1960s from the recognition that probability and decision theory could be applied to real-world decision problems (Howard and Matheson 1984; Raiffa 1968; von Winterfeldt and Edwards 1986). Over the last 20 years, it has grown into an established professional discipline. A number of commercial consulting and research firms perform decision analyses in business, government, and medicine. The number of professional decision analysts is comparable to the number of professionals building expert systems for real-world problems. Many large corporations routinely apply decision analysis to scheduling, capital expansion, and research and development decisions. The emphasis has been on assisting people and organizations faced with high stakes and complex resource-allocation problems.

Early Bayesian Expert Systems

By expert system, we mean a reasoning system that performs at a level comparable to or better than a human expert does within a specified domain. It is useful to classify tasks for which expert systems have been constructed as analytic or synthetic. In systems dedicated to *analytic* tasks, a set of alternatives such as possible diagnoses or decisions is either explicitly enumerated or relatively easy to enumerate; the central task is the valuation of the alternatives. With *synthetic* tasks, the space of alternatives (for example, the set of possible configurations or plans) can be extremely large, and the main problem is constructing one or more feasible options. Analytic tasks include prediction, classification, diagnosis, and decision making about a limited set of options. Synthetic tasks include the generation of alternatives, design, configuration, and planning.

The area of AI in which decision theory has had the most obvious influence is analytic expert systems, particularly diagnostic systems. Probability and decision analysis provide an appealing approach to analytic tasks because of the central role of inference and decision making under uncertainty. Consequently, we focus here on expert systems for analytic tasks. Decision theory can also be relevant to synthetic tasks, where probabilistic or preference value functions can guide search among large numbers of options. The pioneering work in analytic expert systems involved medical applications, although much recent work has addressed fault diagnosis in electronic components and mechanical devices (de Kleer 1991; Genesereth 1984).

In general, three kinds of tasks are involved. The first task is diagnosis: How can we infer the most probable causes of observed problems (for example, diseases or machine faults) given a set of evidence (for example, symptoms, patient characteristics, operating conditions, or test results)? The second task is information acquisition: What additional information or tests should we request? This choice involves weighing the benefits of achieving a possibly more accurate diagnosis against the costs of obtaining the information. The third task is making treatment decisions: What can we do to fix or treat the problem?

The earliest work on diagnostic expert systems used explicitly Bayesian and decision-analytic approaches (Gorry and Barnett 1968; Ledley and Lusted 1959). The general Bayesian formulation for diagnostic inference is as follows: Suppose we are considering a set H of n possible hypotheses,

$$H = \{h_1, h_2, \dots h_n\} ,$$

and a set F of m findings,

$$F = \{f_1, f_2, \dots f_m\} .$$

In a medical application, the hypotheses are possible diseases, and the findings can include patient history, physical signs, symptoms, and laboratory results. We assume for the simplicity of the presentation that all hypotheses and pieces of evidence are two-valued, logical variables, each either true or false. A diagnosis or explanation D is a subset of H. It is a set of hypotheses believed to be present, implying all others are absent. We represent initial beliefs about the prevalence of the diseases as a prior probability distribution over all diagnoses $P(D \mid s)$ for each diagnosis D, conditioned on the expert's knowledge s. Suppose E is observed evidence that some findings are present, others absent, and the rest unobserved. Knowledge about the uncertain relationships between diagnosis and evidence is represented as the conditional probability distribution $P(E \mid D, s)$. We can apply Bayes's theorem to compute the posterior probability of each diagnosis after observing evidence E:

$$P(D \mid E, s) = \frac{P(E \mid D,s)\, P(D \mid s)}{\sum_{\forall D \subseteq H} P(E \mid D,s)\, P(D \mid s)} \; .$$

This most general formulation is complex both to assess and to compute. Because a patient can have more than one disease out of n possible diseases, the number of possible diagnoses (that is, disease combinations) is 2^n. Thus, the number of independent parameters necessary to specify the complete prior distribution is $2^n - 1$. For m pieces of evidence, the general conditional distribution has $2^m - 1$ independent parameters given each hypothesis, requiring the specification of $2^n (2^m - 1)$ total independent parameters for all diagnoses. Clearly, this approach becomes impractical for more than two or three hypotheses and pieces of evidence without some kind of simplification.

Two simplifying assumptions were often made: First (A1) is that the hypotheses in H are mutually exclusive and collectively exhaustive; for example, each patient has no more than one disease. Second (A2) is that there is conditional independence of evidence; that is, given any diagnosis, the occurrence of any piece of evidence f_i of the component hypotheses is independent of the occurrence of any other piece of evidence f_j:

$$P(f_i \mid s) = P(f_i \mid f_j, s) \; .$$

Figure 1 shows a belief network expressing these two assumptions. With assumption A1, the only diagnoses we need to consider are the n singleton hypotheses, h_i. With assumption A2, the conditional probability distribution of the evidence E given a disease h_i (as required for Bayes's theorem) can be decomposed into the product of the conditionals for individual findings, as follows:

$$P(E \mid h_i, s) = \prod_{\forall f_i \in E} P(f_i \mid h_i, s) \; .$$

Under the assumptions A1 of mutually exclusive hypotheses and A2 of conditionally independent findings, only m x n conditional probabilities and n - 1 prior probabilities are required. The great simplicity of probabilistic systems based on these two assumptions made the approach popular. Several medical diagnostic systems have been constructed based on the simplified probabilistic scheme (Szolovits and Pauker 1978), including systems for the diagnosis of heart disease (Gorry and Barnett 1968), acute abdominal pain (de Dombal et al. 1972), and surgical pathology (Heckerman et al. 1991). Despite the apparent simplicity of the assumptions, some of these systems performed at or above the level of experts. For example, the system of de Dombal and his colleagues (1974) averaged over 90-percent correct diagnoses of acute abdominal pain such as appendicitis, whereas expert physicians were averaging 65 percent to 80 percent.

Despite the success of this simple Bayesian scheme in several of these early applications, enthusiasm for this approach began to fade in the early 1970s. One reason might have been the poor user interfaces of many early systems and the general lack of attention to integrating systems with the habits and environment of the diagnostic practitioner. An important lesson from this experience is that superior diagnostic performance alone is not sufficient for acceptance.

A second and more often-cited reason is the restrictiveness of the assumptions of mutual exclusivity and conditional independence. This scheme is sometimes termed "Idiot's Bayes." More generally, critics have pointed out the limited expressiveness of this formulation and the apparent mismatch between

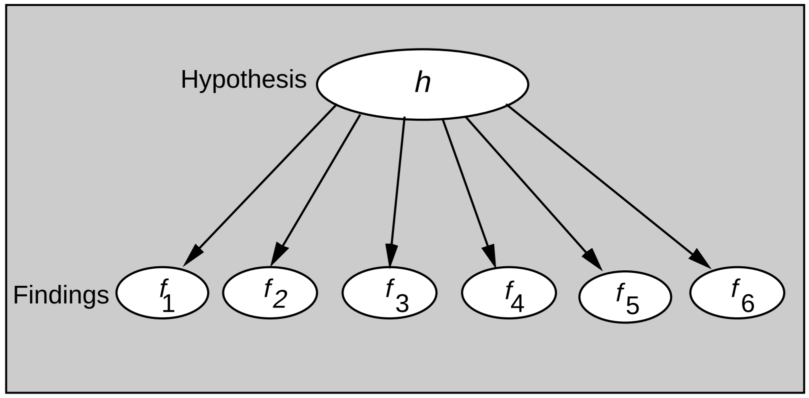


Figure 1. The Simplified Bayes's Formulation.

The fact that there is a single hypothesis node, h, indicates that the hypotheses are mutually exclusive and exhaustive. The lack of influence arcs between the finding nodes indicates that the nodes are conditionally independent given any hypothesis.

. . . developers of . . . MYCIN and PROSPECTOR originally intended their schemes as approximations to the probabilistic ideal, which they saw as unattainable . . .

the rigorous, formal, quantitative approach of probabilistic inference and the informal, qualitative character of human reasoning. This mismatch leads to difficulties in encoding expertise and explaining results so that users can understand and trust them (Davis 1982; Gorry 1973; Szolovits 1982).

AI Approaches to Expert Systems

Perhaps the decisive blow to early Bayesian schemes was the appearance of an appealing alternative approach using logical and rule-based representations derived from AI. This approach focused more on the representation and use of large amounts of expert knowledge and less on questions of normative optimality. Many researchers in this area had had little exposure to, or interest in, probability and decision theory.

A key feature of the new expert system paradigm was the application of the production-rule architecture to real-world diagnosis. The appeal of production rules lay in their apparent capacity to represent expert knowledge in a flexible declarative and modular form (Buchanan and Shortliffe 1984). The production rule has the form of logical implication: To handle the uncertainty in real-world diagnosis, investigators simply extend the production-rule representation to allow intermediate degrees of truth between true and false for both propositions and for the applicability of each rule. The two best-known approaches that represent uncertainty as an extension of deterministic rule-based expert systems are MYCIN (Buchanan and Shortliffe 1984) and PROSPECTOR (Duda, Gaschnig, and Hart 1979).

MYCIN, the expert system to aid physicians in the diagnosis of bacterial infections, introduced the *certainty factor*, a number representing the degree of confirmation (between 0 and 1) or disconfirmation (between 0 and -1) of each proposition or rule. The basic MYCIN scheme was made available for other applications as EMYCIN and it is used in several commercially available expert system shells. PROSPECTOR was constructed to aid geologists in the identification of commercial mineral deposits. PROSPECTOR uses probabilities to represent degrees of belief in propositions and quantities related to likelihood ratios to quantify rule strengths, although its updating rules are not exactly consistent with a coherent probabilistic interpretation.

The developers of both MYCIN and PROSPECTOR originally intended their schemes as approximations to the probabilistic ideal, which they saw as unattainable for the reasons we discussed. Recent theoretical and experimental work examined these and other heuristic schemes for uncertain reasoning with production rules and found a number of inherent problems related to assumptions about priors, irreversibility, and the myth of modularity.

A common objection to probabilistic approaches is the difficulty of assessing prior probabilities, degrees of belief in hypotheses before evidence is available. Empirical data are often hard to obtain, and subjective judgments can be unreliable. MYCIN (although not PROSPECTOR) appears to evade this problem by not requiring prior beliefs. Contrary to many popular interpretations, the certainty factor was originally intended to represent an update or change in belief induced by the evidence, not an absolute degree of belief (such as a probability) (Heckerman 1986; Heckerman and Horvitz 1987). Thus, it aggregates the overall change in belief given the evidence without having to explicitly represent the prior or posterior belief in each hypothesis.

When MYCIN suggests a treatment for an infection, it effectively uses the certainty factors for the diseases as a proxy for their relative probability. Because it avoids explicit reference to priors or prevalence rates, it is, in effect, treating all infections as having equal prior probabilities. In fact, diseases often differ in prevalence rates by many orders of magnitude, and although physicians might find them difficult to precisely quantify, they usually have approximate knowledge about them. In addition, even approximate priors can have a substantial effect on diagnosis and treatment. For example, the fairly prevalent mononucleosis and relatively rare Hodgkin's disease can appear similar in a lymph node biopsy; the differences in the prior probabilities can be essential in diagnosis and treatment. Given that experts have even a rough knowledge about priors, it seems important to explicitly represent this knowledge.

An inherent problem of rule-based representations is their irreversibility for uncertain inference. Rules for diagnosis are generally specified in the direction from possible evidence to the hypotheses. This approach supports diagnostic reasoning, but it is hard to reverse the direction to support predictive inference, for example, to predict the likely effects of a given disease. (Note that the issue here is the flow of evidence, not the flow of control, so the applicability of forward or backward chaining is irrelevant.) In many cases, it seems easier to assess uncertain dependence in the *causal direction*, for example, the propensity of a given disease to cause a symptom, than in the *diagnostic direction*, the degree to which the symptom suggests the presence of the disease. Causal dependencies are often more invariant over different situations because they reflect basic properties of the mechanisms, whereas diagnostic dependencies depend on the prevalence of alternative possible explanations of the effect. Hence, it is often desirable to encode expert knowledge about causal dependencies but reverse the direction for diagnostic inference (Shachter and Heckerman 1987). This is precisely what Bayes's theorem does for coherent probabilistic representations, but it is generally difficult to do with rule-based representations.

It is also hard to support *intercausal inference* (Henrion 1987; Wellman and Henrion 1991), that is, increasing belief in one possible cause of an observed effect because of new evidence against another cause (and conversely). For example, given that an incipient cold and an allergy attack are both possible causes of sneezing, and given a person is sneezing, the observation of an allergen (for example, a cat) should reduce belief in the cold (figure 2). Although one can add special-purpose rules to achieve this effect, this approach defeats the goal of having the rules encode only domain knowledge and not general knowledge about how to perform inference. Again, this kind of inference arises naturally in probabilistic representations but is awkward and impossible to do in general for rule-based schemes. Rules primarily support reasoning in the direction from condition to action: An ideal knowledge representation is isotropic in that it encodes knowledge in whatever way is most natural but supports reasoning in any direction required: predictive, diagnostic, or intercausal.

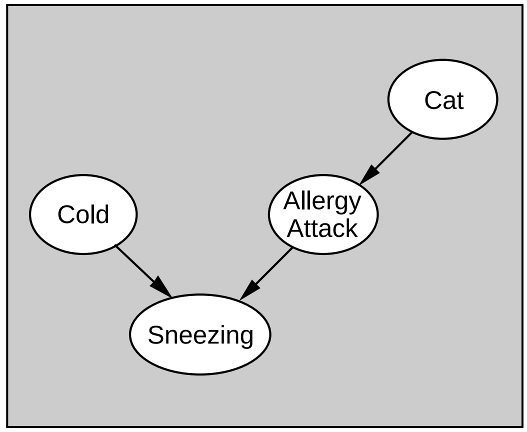


Figure 2. Explaining Away.

If sneezing has been observed, then independent evidence for the allergy attack, that is, a cat in the vicinity, explains away the sneezing, and so reduces support for a cold.

An often-cited advantage of the rule-based representation scheme is the ability to add or remove rules from a knowledge base without modifying other rules (Davis 1983). This property has been referred to as *modularity*. The modularity of rules in a logical production system is a consequence of the monotonicity of logic: Once asserted, the truth of a proposition cannot be changed by other facts. Unfortunately, it has been shown that this property does not carry over in any straightforward manner to uncertain reasoning with rules. Uncertain beliefs are intrinsically less modular than beliefs held with certainty, frequently making the rule-based calculi inefficient for reasoning with uncertainty (Heckerman and Horvitz 1987). The traditional assumption of modularity in rule-based approaches for reasoning under uncertainty has implications that had not previously been appreciated.

Thus, like the early probabilistic systems, rule-based methods impose strong restrictions on the kinds of dependence that can effectively be represented. Unlike the explicit assumptions of the simplified probabilistic systems, the restrictive assumptions in the heuristic approaches have been less apparent. The implicit nature of the assumptions in rule-based systems has tended to promote a dangerous "myth of modularity": Rule-based approaches, like the simple probabilistic approaches, do not have the expressiveness necessary to coherently represent the relationships among uncertain beliefs.

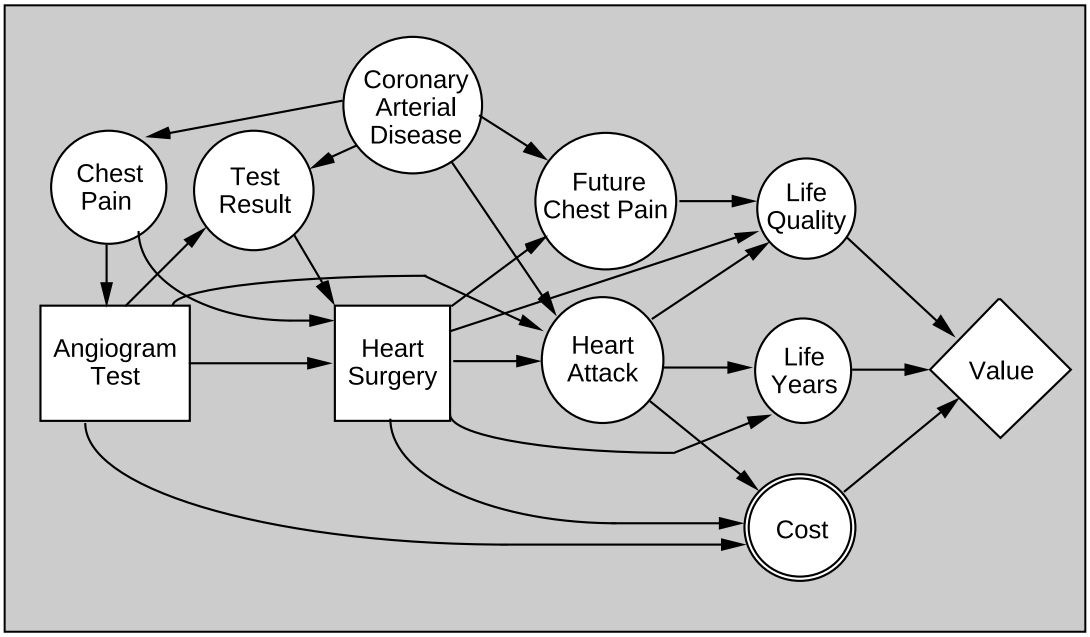


Figure 3. An Influence Diagram for a Patient with Heart Disease.

Circle nodes denote chance variables; rectangles denote decisions; the diamond denotes the value or utility to the decision maker; the double circle is a deterministic variable. Arrows into chance nodes and the value node represent influence arcs, that is conditional dependence of the destination on the origin. Arrows into decision nodes represent informational arcs, that is the origin variable will be known when the destination decision is made.

Knowledge Representations from Decision Analysis

As we saw, there has been justified criticism of the restrictive assumptions of both the simplified Bayesian scheme and the heuristic rule-based approaches to uncertain reasoning. Some have been led to believe that the assumptions of mutual exclusivity and conditional independence of the Idiot's Bayes's scheme are essential to any Bayesian scheme. However, this belief is a misconception. In the last decade or so, much richer knowledge representations have been explored, still based in a principled way on probability and decision theory but capable of expressing a wider range of both qualitative and quantitative knowledge in a flexible and tractable manner. Much of this work has centered on the use of acyclic-directed graphs to represent uncertain relationships, including belief networks and influence diagrams. These representations facilitate the assessment of coherent prior distributions and make it easier for knowledge engineers and experts to express and understand more general kinds of dependence and independence assumptions.

Influence Diagrams

Influence diagrams are a graphic knowledge representation language to display decision problems, including decision variables, state variables, and preference or value variables and their relationships (Howard and Matheson 1981). As well as having a rigorous formal interpretation, they have a perspicuous qualitative structure that facilitates knowledge acquisition and communication. Influence diagrams offer an important complement to more traditional representations, such as decision trees and tables of joint probability distributions and outcome values for each action and state (Raiffa 1968). Unlike these models, influence diagrams provide an explicit representation of probabilistic dependence and independence in a manner accessible to both human and computer.

The influence diagram is an acyclic-directed graph. The nodes represent propositions or quantities of interest, including decision variables, states of the world, and preference values. The arcs represent influence or relevance, that is, probabilistic or deterministic relationships between the variables. An influence diagram for a medical decision problem is shown in figure 3. The diagram encodes a decision problem about whether to undergo coronary artery bypass graft surgery. The danger in this situation is the risk of myocardial infarction, that is, heart attack.

. . . rule-based methods impose strong restrictions on the kinds of dependence that can be represented effectively.

The example demonstrates the four different kinds of nodes in an influence diagram. A *decision node,* depicted as a rectangle, represents a set of possible alternative actions available to a decision maker. Decisions are the control variables or policy variables under the direct control of a decision-making agent. In the example, the angiogram test node represents the decision of whether to perform an artery-imaging procedure that provides information about the extent of coronary artery disease in the patient. The heart surgery node is the decision about whether to undergo a coronary bypass operation.

The arcs into a decision node are *informational arcs*. They indicate what information is available, that is, the values of uncertain variables or previous decisions that will be known at the time the decision is made. The diagram indicates that when he/she makes the surgery decision, the decision maker will know whether the patient has chest pain and the outcome of the angiogram test if it was performed.

Chance nodes represent states of the world that are uncertain. They are depicted by circles or ovals. Uncertain belief about a chance node is specified as a probability distribution conditioned on the outcomes of its predecessor nodes. For example, the probability distribution over the values of *life-years* (years of life remaining to the patient) depends on whether the patient has a heart attack and whether heart surgery was performed because there is a risk of death from the surgery itself.

A *deterministic node* is depicted by a double circle and represents a state of the world that is a deterministic function of its predecessor nodes. In the example, the *cost* is simply the sum (a deterministic function) of the monetary expenses of the angiogram test, the surgical procedure, and the hospitalization following a heart attack. Note that we can actually be uncertain about a deterministic node because its predecessors are uncertain. In this case, we are uncertain about the cost because the heart attack node is uncertain, even though it depends deterministically on these predecessors.

Finally, the *value node* is depicted as a diamond and represents the preferences or utilities of a decision maker for alternative outcomes. Generally, each influence diagram has only one value node. Its predecessors indicate those outcomes or attributes that are included in the evaluation of a choice or plan. For the heart disease example, the attributes are life quality, life-years, and cost. This multiattribute utility function expresses trade-offs among these attributes for an individual patient as well as attitudes toward risk and time.

Any variable can be represented by a continuous scalar or a set of discrete values. It is usual for the value node to be continuous. Some variables can be inherently discrete, such as the heart surgery decision. It is either performed or not, with no intermediate possibilities. In other cases, the variable can be inherently continuous but treated as discrete for representational and computational convenience. For example, in response to a particular level of exertion, the node chest pain has the values none, mild discomfort, and crushing sensation. It is important that the set of outcomes for each variable be defined unambiguously. They must be mutually exclusive and exhaustive in the sense of covering all possible values.

Belief Networks

Much of the research on representation and inference with these graphic representations has focused on specializations of influence diagrams that contain only chance nodes (Cooper 1984; Kim and Pearl 1983; Lemmer 1983; Pearl and Verma 1987). These specialized representations exclusively express probabilistic relationships among states of the world without explicit consideration of decisions and values. Several different terms have been used for these representations, including causal probability networks and Bayesian nets (Pearl 1988). We use belief networks, which seems to be the most popular.

Three Levels of Representation

The representation of a decision problem can be seen at three levels of specification: relation, function, and number (Howard and Matheson 1981). We can define a model at each level without defining information at more specific

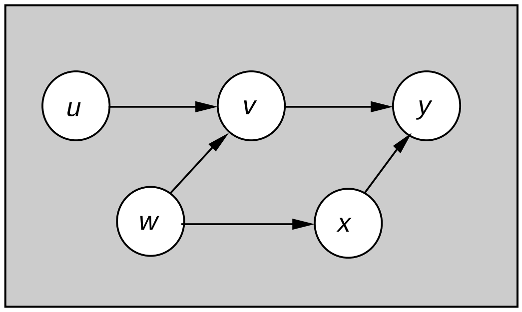


Figure 4. Expressing Independence.

The arcs, and lack of arcs, in a belief network express independence. For example, variables u and w, having no common predecessors, are marginally independent. Variables v and x are conditionally independent, given their common predecessor x. Variable y is conditionally independent of u and w, given its direct predecessors, v and x. In general, each variable is independent of all other variables given its Markov blanket, that is its direct predecessors, successors, and predecessors of its successors.

levels. The *relation level* captures the qualitative structure of the problem, as expressed in the topology of the influence diagram. At this level, the arcs specify dependence and independence between propositions or variables (nodes). Influence diagrams at the relation level are similar to several common representations in modeling and AI research, such as semantic nets.

The level of function specifies the qualitative functional form of the probabilistic and deterministic relationships among nodes. For example, Wellman (1988a) defines monotonic and synergistic influences between variables in qualitative probabilistic terms and presents methods of qualitative probabilistic reasoning based on them.

Finally, the level of number quantifies the numeric values in the functions and conditional distributions. For example, at the level of number, we might specify that *P(chest pain = mild discomfort* | *coronary artery disease* = 1 *vessel*) = 0.25. Chance nodes without predecessors can be specified at the level of number with marginal (prior) probability distributions.

Conditional Independence

As we mentioned previously, at root, probabilistic independence is a qualitative relationship among variables. It captures the intuitive notion of irrelevance. A belief network expresses independence graphically. The arrows or arcs—or, more precisely, the lack of arcs between variables—express probabilistic independence. Several kinds of independence are illustrated in figure 4. *Source variables* (that is, those variables with no predecessors) are marginally independent. Where two variables have one or more common parents but no arc between them, they are conditionally independent of each other given their common parent(s). Finally, a variable is conditionally independent of its indirect predecessors given all the variable's immediate predecessors (that is, those nodes from which it directly receives an arc).

The influence diagram, or belief net, provides knowledge engineers the flexibility to specify and reason about dependencies—or, more important, independencies—at a purely qualitative level before progressing to the level of function or number. Thus, they can capture expert beliefs in their full richness (or simplicity as the case might be) without arbitrary restrictions. The representation provides explicit control over modularity assumptions. The independencies in an influence diagram are a formal expression of the locality of effect among variables. The effects of one variable on a distant variable can only propagate along the influence arcs. More precisely, a variable is screened from the effects of distant variables (is conditionally independent of them) given its *Markov blanket*, that is, its direct predecessors, direct successors, and the direct predecessors of these successors (that is, parents, children, and spouses). In figure 4, the Markov blanket of y is $\{v, x\}$, and the Markov blanket of w is $\{u, v, x\}$.

One way to look at a belief net is as a decomposition of the joint-probability distribution over all its variables. The joint probability is equal to the product of the probability for each variable conditional on its predecessors (if any). The belief net in figure 4 specifies the following decomposition:

$$P(u,v,w,x,y \mid s) = P(y \mid v,x, s)\, P(v \mid u,w, s)\, P(x \mid w, s)\, P(w \mid s)\, P(u \mid s) \quad (1)$$

If we assess the probability distribution for each variable conditional on its predecessors, provided this network has no directed cycles, the product of the variables is guaranteed to be a coherent joint distribution. That is, the joint will be *complete*, having no unspecified parameters, and *consistent*, containing no conflicts. In this way, the belief network provides a simple solution to the problem that was unsolved in PROSPECTOR and related systems, namely, how to assign probabilities to variables and links without creating incoherence.

We can now more carefully interpret the heart surgery influence diagram (figure 3) in terms of conditional independence. In the diagram, the primary expression of conditional independence involves coronary artery disease and its effects. The diagram asserts that the probabilities over the values of chest pain (both current and future), angiogram test, and myocardial infarction are all dependent on the value of coronary artery disease. Furthermore, given knowledge of coronary artery disease, the effects of the disease are conditionally independent of one another. Once we know the precise extent of a patient's heart disease, then the presence of chest pain does not change our belief that he/she might have a heart attack at some time in the future. The knowledge of coronary artery disease as the causal agent tells us all the information available about the interaction of its effects.

Knowledge Engineering and Decision Analysis

Knowledge engineering is the process by which knowledge is elicited from experts in the domain, structured, encoded, installed, and refined in expert systems. Although this term has not traditionally been used in the decision analysis field, the work of a decision analyst in constructing a decision model is fundamentally similar to the activities of a knowledge engineer. Both the knowledge engineer and the decision analyst work with a decision maker or a domain expert to create a formal representation of aspects of his/her knowledge, often supplemented with data from other sources, such as texts and policy documents. In conventional expert systems, the knowledge engineer typically uses rule-based and object-based representations coupled with some type of deductive inference method.

The decision analyst uses influence diagrams and decision trees to express qualitative knowledge about the situation. Influence diagrams have proved particularly useful for expressing qualitative judgments about uncertain dependence and independence relations in a way that is intuitive but also with a principled probabilistic basis. The qualitative diagram is then used as a framework for assessing quantitative probabilistic dependencies. Decision analysts differ in putting more emphasis on decisions and building an explicit quantitative preference or utility model, which is only implicit or absent in rule-based representations (Henrion and Cooley 1987; Langlotz et al. 1986). Like knowledge engineers for expert systems, practicing decision analysts have developed and refined a battery of practical techniques for eliciting and quantifying knowledge and beliefs.

The Encoding of Probabilities

How to obtain the necessary probabilities is frequently a major concern of those contemplating decision analysis for the first time. For some events, there can be relevant empirical data to guide probability assessment, but for many real problems, most or all probabilities will need to be obtained from expert judgment. Indeed, even where data are available, it is a matter of expert judgment about how relevant they are and whether adjustments are needed to fit the situation at hand. To the strict subjectivist, one should be able to express one's belief in any proposition as a single probability number no matter how little or much one knows about it. This is a consequence of the axioms of decision theory, such as the ability to order the relative probability of any set of events. The objection What if I don't know the probability? loses force once one realizes that probability is not a physical characteristic, such as mass or length, that one is trying to estimate; it is just a way to express one's degree of knowledge or ignorance about the proposition. Nonetheless, expressing one's knowledge in terms of probabilities is often a demanding task.

However, decision analysts have developed a variety of techniques to make it as easy as possible, even for assessors who have little technical understanding of probability (Morgan and Henrion 1990; Spetzler and Stael von Holstein 1975). Methods are available for assessing discrete-event probabilities and continuous probability distributions. The simplest methods require the assessor to make only qualitative judgments about which is more probable, the event of interest or some reference event of agreed probability. The *probability wheel* is a popular method for providing a reference event. It is a simple graphic device consisting of a disk with a colored sector whose angle visually represents the probability of the reference event. According to whether the probability of the event is judged greater or lesser than the relative size of the sector, its angle is adjusted larger or smaller until the expert is indifferent. Thus, a probability can be obtained without explicitly mentioning a number. As they gain experience with probability assessment, many experts find they prefer directly giving numeric probabilities. For extremely low or high probabilities, techniques that use odds or log-odds scales have been shown to be useful (von Winterfeldt and Edwards 1986).

An extensive literature on human judgment has identified cognitive biases and mental heuristics that tend to distort human judgments about uncertain events (Kahneman,

What if I don't know the probability?

…representing what is not explicit in the knowledge base… mitigating the effect of the closed-world assumption.

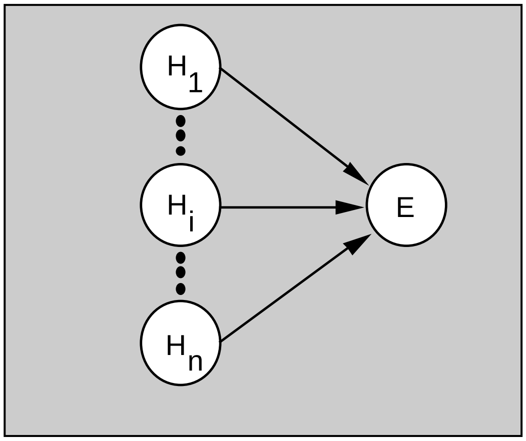


Figure 5. The Noisy-Or Influence.

This is a prototypical dependency in which each input H_i , if it occurs, has a probability that it is sufficient to cause output E to occur. The event that H_i is sufficient is independent from the occurrence or sufficiency of each other input. The Noisy-Or requires specification of only n parameters for n inputs.

Slovic, and Tversky 1982). One common bias is the tendency to underestimate uncertainty, assessing probabilities that are nearer 1 or 0 than is appropriate. Decision analysts have drawn on this research to develop methods to counteract the effects of these biases. *Debiasing techniques* include attempts to make all assumptions explicit, encouraging assessors to consider extreme possibilities and unexpected outcomes. For probabilities or distributions that are particularly important, decision analysts often use an extensive protocol to ensure the quantity is clearly defined and understood, make explicit all important conditioning events, and counteract possible cognitive biases (Morgan and Henrion 1990; Spetzler and Stael von Holstein 1975).

Of course, there are limits to the precision with which a person can provide a probability, depending on the skill and knowledge of the assessor and the complexity of the domain, as well as cognitive biases. You might be hard put to say whether your degree of belief in X is better expressed as 0.7 or 0.8, but approximate numbers are often sufficient. The response of the decision analyst is to do a sensitivity analysis to see whether a change over this range really matters. If not, which is often the case, then there is no need to worry about more precision. If it does matter, then it might be worth trying to develop a more elaborate model that is conditioned on other events to better assess its probability. In this way, we can apply decision-analytic tools to identify the most important probabilities, so as to best apportion the probability assessment effort (Heckerman and Jimison 1989).

Prototypical Influences

To fully express the influence on a binary variable E dependent on n binary hypotheses $H_1, H_2, \ldots H_n$ requires a conditional probability distribution with 2^n independent parameters. However, in practice, influences can often be specified by a prototypical function incorporating independencies that greatly simplify the assessment. A common example of a prototypical influence is the noisy-Or gate, which is a probabilistic generalization of a standard Boolean Or (figure 5). With a Boolean Or, the occurrence of any single one of the input events is sufficient to cause the output event. In a noisy-Or, each input has some probability of being sufficient to cause the output. The processes that prevent the signal from being sufficient are independent of each other. This structure is useful in representing many causal relationships, for example, where several different faults can each cause the same device failure mode, or several different diseases can each cause a common symptom.

The noisy-Or relationship requires the specification of only one parameter for each input, the probability of the effect given only that input is present, $p_i = P(E \mid H_i \text{ only}, s)$. The probability of the effect given any combination of input can simply be derived from the individual parameters:

$$P(E \mid H_1, H_2, \ldots H_n, s) = 1 - (1 - p_1)(1 - p_2) \ldots (1 - p_n) \; .$$

Thus, the complete conditional distribution requires the specification of only n parameters (Good 1950; Pearl 1986b).

It is often useful to also introduce a *leak*, that is, the probability that the effect can occur in the absence of any cause explicitly represented in the knowledge base. The leak constitutes a kind of residual "all others" category, representing what is not explicit in the knowledge base and substantially mitigating the effects of the closed-world assumption. For example, in the QMR-BN knowledge base for internal medicine, each finding or symptom is, on average, associated with about 80 diseases (Shwe et al. 1991). The leaky noisy-Or assumption allows it to be encoded by a single causal strength for each disease-finding pair (about 40,000 such pairs exist in the knowledge base). Each of the approximately 4,000 findings also has a *leak probability*, representing the chance the finding is observed because of an unmodeled cause or a false positive test result.

Researchers are seeking techniques for explicitly acquiring and representing other forms of independence. Heckerman (1990) describes the use of partitions that can substantially reduce the effort to assess the disease-finding links. For each finding, its associated diseases are classified into groups that have similar causal strengths. This strength need only be assessed then once for many diseases. The approach, and the related *similarity networks,* are efficient because experts find it much easier to judge whether two probability distributions are similar than to assess the distributions themselves.

Value Structuring and Utility Modeling

The modeling of human preferences in the form of utility functions is a major concern for decision analysts. In important decisions, there are often several objectives that conflict, for example, in the heart surgery case, life duration, life quality, pain, and monetary cost. Decision theorists have developed the multiattribute utility theory as a way to model human preferences in such situations (Keeney and Raiffa 1976; von Winterfeldt and Edwards 1986). The most important stage in preference modeling is generally the first in which these objectives or attributes are identified. They are often organized into a value tree or attribute hierarchy. It is important not to omit key attributes and avoid attributes that overlap.

For each attribute, a value function can be assessed that maps from the quantity of interest, for example, dollars of income or years of life expectancy, into a value scale whose increments are equal in significance to the decision maker. A function must be assessed then that combines the values for each attribute into a single utility number to compare alternative options. Depending on whether the attributes are judged independent or whether there are interactions among them, this can be a simple additive model or something more complex. Again, a variety of elicitation techniques are available to discover the model form and assess weights or trade-offs between the attributes (Keeney and Raiffa 1976; von Winterfeldt and Edwards 1986). Several researchers have examined computer aids for value structuring and preference modeling (Holtzman 1988; Klein 1990; and Wellman 1985).

Utility functions also represent attitudes to uncertainty or risk. For example, a risk-averse person will prefer a certain prize of a $500 to a 50-percent chance of $1000 and a 50-percent chance of 0. A variety of elicitation techniques have been developed to help decision makers assess their risk attitudes by asking their relative preferences among gambles or lotteries with varying probabilities and outcomes (Howard 1970; Keeney and Raiffa 1976; von Winterfeldt and Edwards 1986). Assessment protocols generally involve making judgments in a variety of different ways to provide cross-checks.

Model Refinement and Sensitivity Analysis

As we noted in Early Bayesian Expert Systems, no model is complete: It is necessarily a simplification of the knowledge, beliefs, and preferences of the expert or decision maker, which are themselves a simplification of reality. It is a compromise between simplicity for ease of knowledge engineering and computational tractability and completeness for maximum precision. During model construction, we can elaborate it and simplify it as we explore trade-offs between accuracy and tractability. We refer to this process as *completeness modulation* (Horvitz 1987). This process is also used to some extent by knowledge engineers, but decision analysts have particular perspectives and tools to help them.

In particular, the decision analysts use a variety of sensitivity analysis methods to examine the importance of various model parameters and structures, including specific probabilities, risk tolerance (expressing risk attitude), and multiattribute trade-off weights (Howard 1968). In judging the sensitivity of the model to a particular parameter or model feature, the question is not just, Could it affect model predictions? but more specifically, Could it affect the predicted expected utility in such a way as to change the recommended decision? This decision-oriented perspective provides a stronger measure of importance that allows pruning many more parameters as unimportant.

A well-known decision-oriented sensitivity measure is the *expected value of perfect information* (EVPI), which is the increase in expected value from improving the decision should the true value of an uncertain quantity become known. Another such measure is the *expected value of including uncertainty* (EVIU), which assesses the importance of the explicit representation of uncertainty in a variable (Morgan and Henrion 1990, chapter 12). Because the probabilistic representation of a variable exacts costs in elicitation, representation, and inference, it is desirable to include only those uncertainties that matter. A third decision-oriented measure is the *expected*

A comparison . . . found that the Bayesian version significantly . . . outperformed formulations using certainty factors and Dempster-Shafer belief functions.

value of computation (EVC) (Horvitz, Cooper, and Heckerman 1989). Where additional computation gives improved precision, but at additional cost, the EVC can help decide how much computation is worthwhile.

To date, few people have investigated the automation of sensitivity analysis for probabilistic reasoning. A promising area in this regard is the development of better error theories, improving our ability to predict the effects of various kinds of errors and simplifications in input in different model classes. For example, the effects of errors in the assessment of conditional probabilities can only be attenuated in predictive inference but can be amplified in their effect on posterior probabilities in diagnostic inference (Henrion 1988a). General theory of this type could lead to more intelligent automated sensitivity aids to guide knowledge engineering in the most profitable directions.

Computer Support for Problem Structuring

Conventionally, decision analysis has been applied to specific decision problems, where expert systems are designed to support a whole class of decision problems within a domain. Of course, the cost in constructing it can be amortized over a large number of applications. The usefulness of decision-focused sensitivity analysis for guiding model construction in decision analysis points to a fundamental difference in goals between conventional decision analysis and knowledge-based expert systems. Decision-focused sensitivity analysis seems to work well in cutting a model down for a specific decision problem. However, when constructing a knowledge base for a wide class of decisions, determining what might be relevant is much harder. Almost any hypothesis or evidence might potentially be of importance to support decision making for, say, internal medicine. A vast knowledge base is necessary because little can be ruled out a priori as irrelevant.

Given a large knowledge base of potentially relevant information, can we automate the task of constructing a tractable model focused on a particular case? Holtzman (1988) approached this problem by encoding the expertise of the decision analyst as a set of rules for constructing influence diagrams within a particular domain. Holtzman's prototype, RACHEL, addresses decision counseling on problems of infertility. It contains deterministic rules that construct and analyze an influence diagram that responds to the specific information, preferences, and situation of a particular case.

Breese (1990) takes a different approach to this problem. His knowledge base represents specific and general facts about variables and their probabilistic relations. His system, ALTER ID, constructs a belief net or influence diagram from components in the knowledge base using a refinement process that is analogous to the human decision analyst. ALTER ID adds variables to the model in stepwise fashion and uses deterministic sensitivity analysis to guide the direction and extent of the refinements. These approaches raise the more general question of what kind of structures are appropriate for representing uncertain knowledge for decision making if belief nets and influence diagrams are not the best primitive representation.

Reasoning with Belief Networks and Influence Diagrams

Suppose we observe the actual value of one or more variables in a belief network or influence diagram, which might have implications for other chance variables whose probabilities will change as a result. A variety of reasoning types are possible. For example, in figure 3, knowledge that a patient has coronary artery disease allows us to infer *predictively* (that is, in the direction of the influence arcs, generally from cause to effect) the changed probability that he/she will suffer future chest pain or a heart attack. Conversely, given that the patient has a specified level of chest pain and angiogram test outcome, we can infer *diagnostically* (that is, in the reverse direction, from effect to cause) the probability that the patient has coronary artery disease. Given evidence about an effect, for example, sneezing in figure 2, the independent evidence eliminating one cause, allergy, allows us to reason *intercausally* that the other possible cause, a cold, is more likely.

We can also calculate the expected value or utility of alternative decisions, thus obtaining recommendations for which a decision is preferable. These decisions can include primary decisions, for example, whether heart surgery is recommended given current knowledge, and information-acquisition decisions, for example, whether doing an angiogram test is worthwhile given its costs and the chance it might improve the primary decision. We can also do a variety of sensitivity analyses, answering such questions as, How probable must coronary artery disease be for surgery to be worthwhile? or for a given probability of coronary artery disease, What trade-off between life quality (freedom from pain) and life-years of survival would make surgery worthwhile?

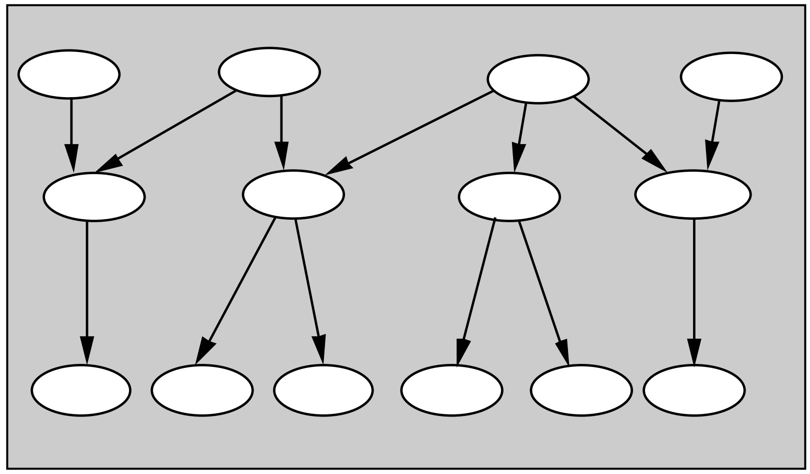


Figure 6: A Singly Connected Belief Network, or Polytree.

Exact Inference in Belief Networks

Conceptually, the simplest way to perform inference and compute the probability of an outcome that is conditional on any observations is simply to generate the joint distribution for all the variables as the product of all the component conditionals, as illustrated in equation 1. Given a belief network, as in figure 4, the probability of each instantiation of the variables (each scenario) can simply be computed as the product of the conditional probabilities for each variable given the specified instantiations of its predecessors, as illustrated in equation 1. From this computation, one can figure the probability of any set of observations $P(E \mid s)$ by summing over the irrelevant variables and compute the conditional probability by dividing the joint probability by the probability of the evidence:

$$P(A \mid E, s) = \frac{P(A, E \mid s)}{P(E \mid s)}.$$

Of course, the snag with this approach is that for n variables, there are 2^n scenarios. Thus, computational effort is exponential in the number of variables and is infeasible for more than a dozen or so variables. The key to computational efficiency for inference in belief networks is the exploitation of the specified independence relations to avoid having to explicitly calculate the full joint distribution. Most of these decompose the network into smaller pieces. A variety of methods have been developed, each focusing on particular families of belief network topology. The simplest method is the simplified Bayes's scheme, which we discussed earlier and is illustrated in figure 1. The next simplest applies when the network is *singly connected*, that is, a polytree with no more than one path between any pair of nodes, as illustrated in figure 6 (Kim and Pearl 1983). Both these algorithms are linear in the number of network variables.

Unfortunately, most real networks are multiply connected; so, more complex methods are required. A variety of approaches have been explored (Lauritzen and Spiegelhalter 1988; Pearl 1986a; and Shachter 1986) whose efficiency varies according to the network's characteristics. All these approaches can have problems if the network contains many intersecting loops (ignoring directionality), as illustrated in figure 7. Cooper (1991) shows that the general problem of exact probabilistic inference in a belief network is NP-hard; so, we should not expect to find an exact method that is computationally efficient for arbitrary networks. Nevertheless, exact methods have proved practical for sizable multiply connected networks. For example, the HUGIN system can perform diagnostic inference in under five seconds on the MUNIN network for neuromuscular disorders, containing about 1000 chance nodes with up to 7 states each (Andreassen et al. 1987).

Approximate Inference in Belief Networks

Despite these successes, there remain networks such as the QMR-BN belief network (Shwe et al. 1990) with about 4500 nodes with many intersecting loops, an arrangement that generally appears intractable for exact methods. Concern about the tractability of exact methods has provoked research into approximate

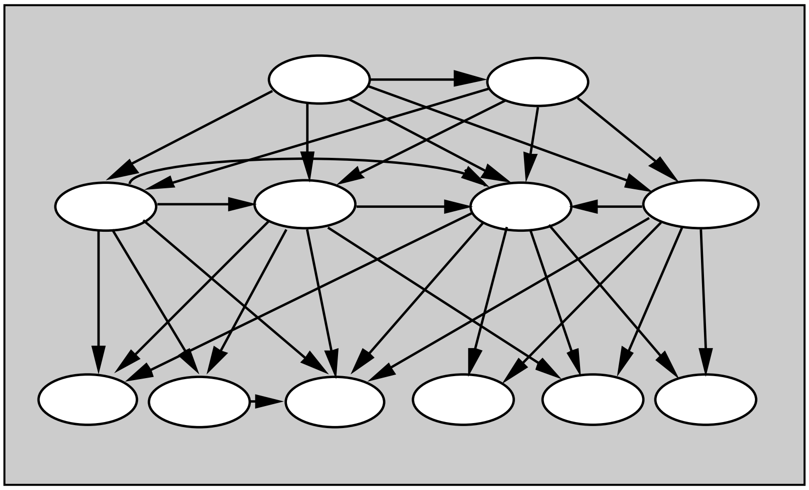


Figure 7. A Multiply Connected Belief Network

methods. Most approximate methods achieve their efficiency by considering only a small subset of the 2^n scenarios and estimating or bounding the probabilities of events from this subset. These methods include random sampling schemes, bounded conditioning, and search schemes (Henrion 1990b).

Random sampling, or simulation-based approaches, generates a random sample of network instantiations and estimates resulting probabilities from the sample. In logic sampling (Henrion 1988b), each instantiation is generated following the belief network arcs, selecting a value for each variable using the probability conditional on its instantiated parents. Logic sampling for predictive inference is a version of Monte Carlo simulation, applied to discrete variables. For purely predictive inference, that is in the direction of the influence arrows, such random sampling methods are efficient, with computational effort linear in the number of uncertain variables and quadratic in the level of precision required. These methods work equally well for continuous or discrete variables. For purely continuous variables there are also a variety of analytic methods of error propagation available (see chapter 8 of Morgan and Henrion [1990] for a review).

For diagnostic reasoning, the complexity of pure logic sampling is exponential in the number of observed variables because all instantiations inconsistent with observed evidence are simply dropped from the sample. *Likelihood weighting* (Shachter and Peot 1990) mitigates this problem. Instead of instantiating the observed variables, it computes the joint likelihood that is conditional for each scenario, using this number to weight the scenario in the sample. A virtue of logic sampling and derived approaches is that successive samples are independent, so convergence is guaranteed, and the precision of estimated probabilities can be estimated by standard statistical methods.

An alternative simulation-based approach is stochastic, or Markov, sampling (Pearl 1987). In this case, variables are instantiated in random sequence, using the probability conditional on their entire Markov blanket. This approach neatly avoids problems from observed variables, which are simply clamped to their observed value. However, unlike logic sampling, successive instantiations are not independent, and the process can suffer from convergence problems when the network contains near-deterministic influences (Chin and Cooper 1987).

Cutset conditioning (Pearl 1988) involves instantiating variables in the cutset of a network to render the remaining network singly connected, thus amenable to the efficient polytree method. *Bounded conditioning* improves the efficiency of complete cutset conditioning by examining the most probable instantiations of the cutset from which it computes bounds on the posterior probability (Horvitz, Suermondt, and Cooper 1989).

A different class of approximate methods uses heuristic search of the space of possible scenarios to find the hypotheses that best explain the observed findings (Cooper 1984; Henrion 1990a; Peng and Reggia 1987; de Kleer 1991). Because a small fraction of the hypotheses often accounts for the bulk of the probability, this approach can be efficient. Admissibility heuristics can guarantee finding the most probable hypotheses (or *n* most probable hypotheses). The most powerful heuristics depend on specific assumptions, such as the bipartite network with noisy-Or–combining functions. Versions of these search methods can compute bounds on the probabilities of the hypotheses. These methods seem promising. For example, the TOP N algorithm was applied to the QMR-BN network, and can find the most probable multiple-disease hypotheses in seconds (Henrion 1991).

All these approximate algorithms are *flexible* or *anytime* algorithms, that is they can be halted at any time to give an answer, and additional computation continually improves results. Sampling algorithms decrease the variance of their estimates as the sample size increases. Bounded conditioning and search-based methods, such as TOP N, monotonically narrow the bounds on the probability of the hypothesis as the search continues. Such flexibility is valuable when there is uncertainty about deadlines or the cost of delaying a decision, and can be shown to yield increases in the expected value of reasoning (Horvitz 1988). Given a model of delay costs, metelevel decision analysis can be applied to decide when to stop computing and act (Horvitz, Cooper, and Heckerman 1989; Horvitz and Rutledge 1991).

Explanation

Although decision analysts have long understood the importance of explaining decision analysis, as encapsulated in the slogan "insight not numbers," AI researchers have frequently criticized probabilistic and decision-theoretic reasoning as inherently difficult to explain (Davis 1982; Politser 1984; Szolovits 1982). Teach and Shortliffe (1981) identified the ability of an expert system to explain its reasoning strategies and results as key to its acceptance by users. Automating the generation of explanations has long been a focus of attention in expert system research (Shortliffe 1984; Winograd 1971; Swartout 1983; Wallis and Shortliffe 1982). More recently, the development of decision analysis–based systems has led to increased interest in automated methods to generate explanations of probabilistic and decision-analytic reasoning.

Evidence Weights

One approach to explaining probabilistic inference is to examine the individual impact of each piece of evidence on the overall conclusion. A classic technique is the *evidence weight*, that is, the logarithm of the likelihood ratio, of a hypothesis H with respect to each piece of evidence, e_i:

$$W(H, e_i) = \mathrm{Log}\left(\frac{P(e_i \mid H, s)}{P(e_i \mid H, s)}\right) .$$

Weights of evidence have the useful property of being additive provided the pieces of evidence, e_i, are conditionally independent. That is, the evidence weight for combined evidence is just the sum of the weights of the components:

$$W(H, e_1 \,\&\, e_2 \,\&\, \ldots\, e_n) = W(H, e_1) + W(H, e_2) + \ldots + W(H, e_n) .$$

The convenience of weights of evidence for performing and explaining probabilistic inference has an ancient and venerable history, having first been pointed out by Peirce (1956) in 1878 and rediscovered by Turing (Good 1950) and Minsky and Selfridge (1961), among others. Several probabilistic expert systems for medical diagnosis have made use of likelihood ratios and weights of evidence for explaining the importance of evidence to hypotheses (Ben-Bassat et al. 1980; Heckerman, Horvitz, and Nathwani 1991; Reggia and Perricone 1985; Spiegelhalter and Knill-Jones 1984). The additivity of evidence weights has led to perspicuous explanations using *evidence ledger sheets*, which sum the weights for and subtract the weights against a hypothesis (Spiegelhalter and Knill-Jones 1984).

The Modulation of Completeness and Abstraction

A good explanation should be as simple as possible, consistent with communicating the essential information. One approach to creating simpler explanations of probabilistic diagnosis is to control the level of abstraction to suit the user's needs. For example, PATHFINDER can generate explanations at the level of evidence for and against general classes of disease, such as inflammatory, infectious, and malignant diseases, rather than about each individual disease in the classes (Horvitz et al. 1989). In PATHFINDER and related work by Ben-Bassat and Teeni (1984), the level of abstraction can be controlled with heuristic abstraction hierarchies, classifying diseases into groups that depend on the diagnostic problem at hand. Generating a good explanation involves trade-offs between various goal attributes, particularly simplicity and completeness. The use of multiattribute utility models supports metareasoning about these trade-offs to construct more effective explanations (Horvitz 1987; McLaughlin 1987).

Qualitative Explanation

Because people generally seem to find qualitative descriptions easier to understand than quantitative results, several projects have sought to explain probabilistic and decision-analytic reasoning in qualitative terms. One approach has been to map numeric probabilities into phrases such as "unlikely" or "almost certain" or value outcomes into phrases such as "slight risk" or "severe danger." Users often find this approach appealing both for encoding and explanation. Elsaesser (1988) uses a related approach for the qualitative explanation of Bayesian updating. Sember and Zukerman (1989) provide a linguistic explanation of reasoning between neighbors in a belief network. Klein (1990) demonstrates how to generate qualitative explanations of choices based on hierarchical additive multiattribute utility models. This approach has also been used to explain Bayesian updates; for exam-

For decision analysis, the goal is to improve human reasoning rather than to replicate it.

ple, "Observation of *A* makes *B* a great deal more likely" (Elsaesser and Henrion 1990). The literature on the numeric interpretation of probability phrases is considerable, showing a degree of consistency but significant context effects that can cause misinterpretations (Wallsten et al. 1986).

A related approach focuses on the relative sizes of probabilities, values, or expected utilities. Langlotz, Shortliffe, and Fagan (1986b) constructed a system called QXQ that explains why one decision has higher expected value than another using qualitative comparisons of the probabilities and values of the most important possible outcomes of each decision. Henrion and Druzdzel (1990) developed a scenario-based scheme in which posterior probabilities are explained in terms of the relative plausibility of alternative scenarios, or deterministic stories, that can explain the observed findings. Explanations are simplified by considering only the most likely scenarios. A third qualitative approach uses *qualitative probabilistic networks* (QPNs). QPNs are belief nets and influence diagrams in which influences are represented as purely qualitative; for example, observing *A* increases the probability of *B* (Wellman 1988b). QPNs have been used as the basis for explaining probabilistic reasoning by qualitative belief propagation, tracing the directions of the impacts of evidence through a network (Henrion and Druzdzel 1990; Wellman and Henrion 1991).

Comparing Decision Analysis and Rule-Based Expert Systems

There has been a common perception that even if commonly used calculi for uncertain reasoning have deficiencies, it does not really make much practical difference what method you use. This perception is partly based on an early experiment with MYCIN that found its conclusions were fairly robust with respect to the granularity of the certainty factors (Buchanan and Shortliffe 1984). However, as the MYCIN authors originally warned, it is dangerous to overgeneralize from these results. An empirical comparison of certainty factors with probabilistic inference in a simple case showed that the certainty factor–based inference underresponded to the true diagnosticity of the evidence by a factor of 2 and in 25 percent of the cases actually responded in the wrong direction (Wise 1986). However, MYCIN itself was a forgiving application in which this behavior did not have serious results on recommended treatments.

Comparisons of Uncertainty Calculi

Experimental comparisons of certainty factor and other rule-based schemes with probabilistic inference have shown that the differences can be significant under other circumstances. A comparison of alternative formulations of PATHFINDER found that the Bayesian version significantly outperformed formulations using certainty factors and Dempster-Shafer belief functions (Heckerman 1988). A systematic comparison of six common uncertainty calculi with a variety of small- and medium-sized rule sets found that the difference between schemes depended strongly on the situation (Wise 1986). When the evidence is strong and consistent, most schemes perform well. However, if the evidence is weak or conflicting, heuristic schemes can perform poorly relative to Bayesian schemes, sometimes doing no better than random guessing.

Differences in Knowledge Engineering

The studies just mentioned compared the performance of different uncertain reasoning schemes within the same basic structure. In practice, the decision-analytic approach to knowledge engineering can focus on different aspects of a domain and result in a significantly different structure than rule-based and other expert system paradigms. The resulting structural differences can cause greater differences in performance than the differences in inference mechanism. To investigate differences in the knowledge engineering process, Henrion and Cooley (1987) examined the construction of two diagnostic systems for the same problem, namely, the diagnosis and the treatment of disorders of apple orchards. The systems were based on decision analysis and a rule-based paradigm, respectively. Part of the inference network for the rule-based system is shown in figure 8a, and the corresponding part of the influence diagram for the decision-analytic system is shown in figure 8b.

The arrows in figure 8a indicate dependencies in the diagnostic rules, in the direction of the flow of evidence from findings to diseases to the treatment decision. The arrows in figure 8b represent influence, coded as conditional probabilities, from cause to effect. For example, the diagnostic rule goes from Phytophthora lab test to Phytophthora infection. (Phytophthora is a fungus that attacks apple trees). However, the influence is coded in the other direction, from disease to test result. A diagnostic decision rule directly recommends treatment depending on the amount of evidence for the disorders, the abiotic stress, and the Phytophthora infection. Given the full diagram, Bayesian inference algorithms allow reasoning in either direction. The influence diagram contains an extra residual node, "other root disorder." This node is useful for handling other disorders that can explain away some of the observtions.

The most obvious difference between the models is that the influence diagram contains an explicit model of the orchardist's costs and preferences (the diamond node) as a function of the potential costs and benefits of fungicide treatment. Where the rule-based system simply asserted that treatment was recommended if a fungus infection was suspected, the decision model explicitly weighed the costs and probabilities. The influence diagram being somewhat larger took somewhat more effort to build and quantify. However, the decision-analytic model could cover a wider range of cases and types of inference.

Figure 8. Comparing Rule-based and Influence Diagram Representations.

These two figures show the same part of a system for diagnosing apple trees and recommending treatment (Henrion & Cooley, 1987).

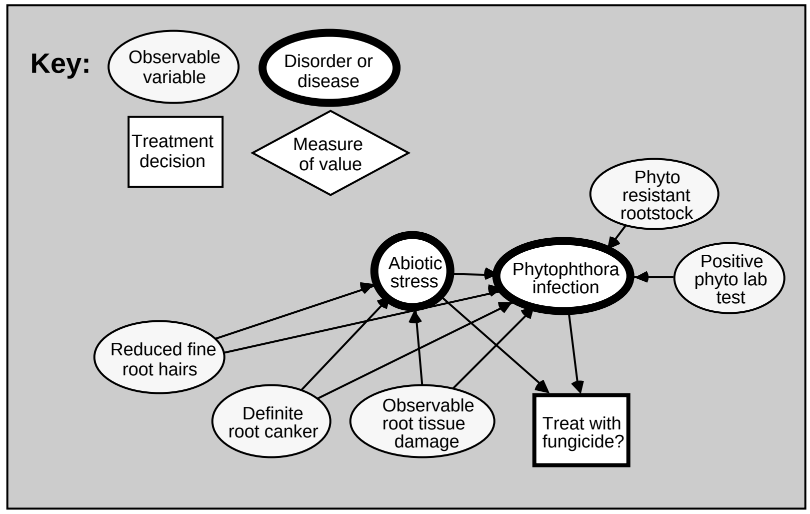


Figure 8a is part of an inference network for a rule-based system. The arrows indicate direction of diagnostic rules, from observables to disorders, to the treatment decision.

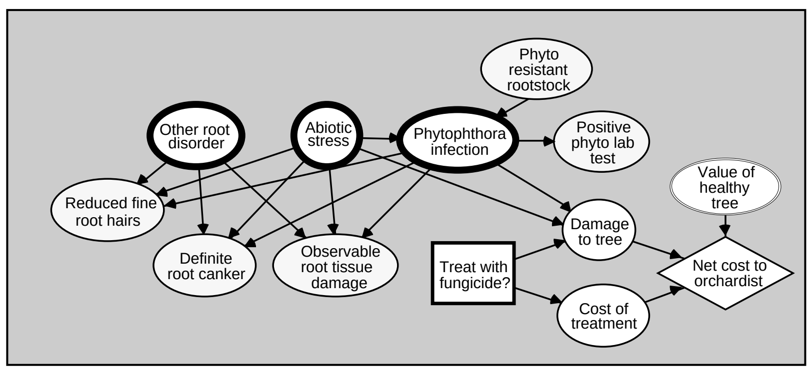


Figure 8b shows the influence diagram for the corresponding part. The arrows here denote influences representing uncertain causal relations. The influence diagram includes an explicit cost model to evaluate outcomes.

The Perfect versus the Fallible Expert

The difference in the treatment of discrepancies between system performance and the expert's expectation illustrates a fundamental difference between the two paradigms. The rule-based system was tuned so that its behavior coincided as closely as possible with the expert's judgments. Thus, the expert was the "gold standard" for performance. In contrast, when a result of the decision analysis (that the fungicide treatment was barely worthwhile in a particular case) ran counter to the expert's intuition, the decision analyst explored with the expert the possibility that the result inferred by normative methods from the expert's more strongly held opinions might be correct. Indeed, it turned out that the analysis of this apparent discrepancy led the expert to a new insight about the nature of the problem, and he changed his intuitions. This experience illustrates the way in which decision analysis, by formalizing, aims at clarifying and improving human intuition. This approach is based on a view of human reasoning under uncertainty as being fallible, which is supported in the psychological literature (Kahneman, Slovic, and Tversky 1982), and contrasts with paradigms that treat the "expert as gold standard."

A second study—troubleshooting in motorcycle engines—also illustrates this contrast, having examined the differences between expert rules and decision-analytic inference. This study compared decision-analytic meth-

…discrepancy led the expert to a new insight about the problem and he changed his intuitions.

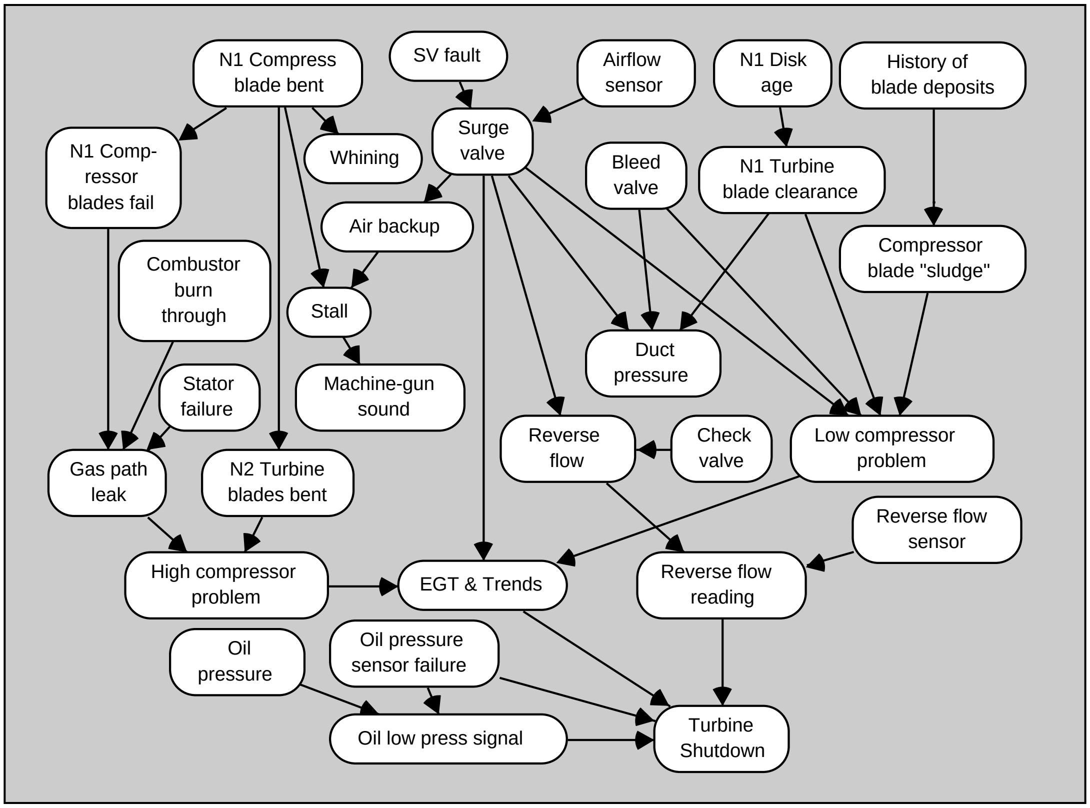


Figure 9: Belief Net for Gas Turbine Diagnosis.

This is a fragment of a belief net for a system for diagnosing faults in a gas turbine for an auxiliary power unit for a commercial aircraft. (Courtesy of Knowledge Industries).

ods with sequencing rules of expert mechanics for selecting the next component to test (Kalagnanam and Henrion 1990). The optimal decision-analytic strategy that minimizes average costs (time to identify the fault) is to test the next untested component that has the minimum ratio of test cost to probability of failure. In a simulation, this strategy, using the subjective probabilities and costs estimated by expert mechanics, performed about 20 percent better in terms of expected cost than the test sequences suggested by the same mechanics. This improvement in performance was robust to substantial errors in the assessed costs and probabilities. This study is another empirical indication of how normative inference methods can result in significant improvement over unaided human expert judgment.

Recent Applications

No matter how great the theoretical appeal of a set of principles of rationality, most practitioners that build knowledge-based systems are unlikely to be convinced that the approach is worth trying in the absence of substantial successful applications. Although some aspects of the technology, including knowledge engineering tools, inference algorithms, and explanation techniques for large networks, have yet to reach full maturity, a few significant applications are already in practical use, with several more on the way to completion.

Of course, decision analysis has been applied to practical decision making for over 20 years in business, engineering, medicine, and other domains. In many cases, the product has been a computer-based tool to support decisions. However, these tools have typically focused on a narrowly specified class of decisions for which only a limited number of uncertain variables (a few tens at most) are used. Our focus here is on systems that can legitimately be termed knowledge based, that is, systems that incorporate belief networks of significant size. Most such systems to date aim at diagnosis in medicine or engineered systems. We will describe the PATHFINDER and INTELLIPATH project, and briefly mention several others.

The PATHFINDER project has used a probabilistic and decision analytic approach to develop

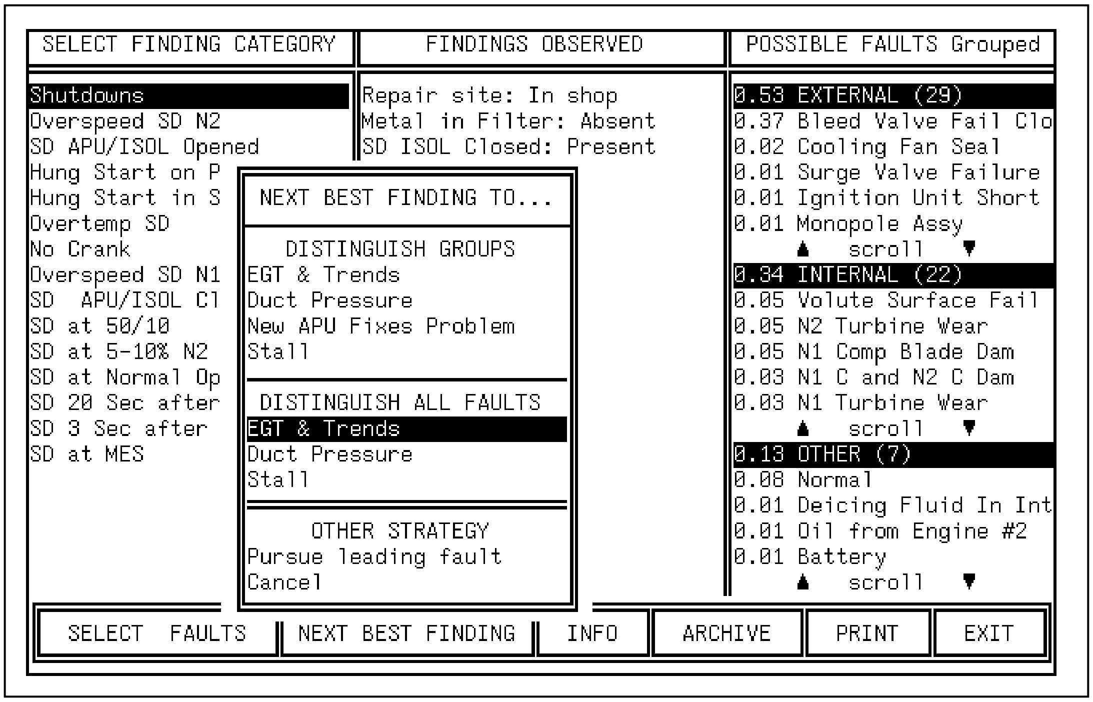


Figure 10: End-user Interface for Turbine Diagnosis.

This screen is from the gas turbine diagnosis system illustrated in figure 9. It is used during diagnosis for choosing findings to observe and displaying current most probable diagnoses. The left column shows categories of findings; the center shows findings observed so far; and the right column shows the list of possible faults, grouped by type, and ordered by probability. The popup box shows specification of the observed value of selected finding. This screen is from the IBM-PC delivery system. (Courtesy of Knowledge Industries.)

diagnostic systems for lymph node pathology. The academic project has given rise to a commercial system, INTELLIPATH, which provides diagnostic support for surgical pathologists in several domains of pathology. Over 250 INTELLIPATH systems have been sold worldwide, and it represents one of the few commercially successful applications of knowledge- based technology for medical diagnosis. The history of PATHFINDER and INTELLIPATH provides a number of interesting lessons.

Surgical pathologists make diagnoses based on the observation of features in a section of tissue under the microscope. The diagnosis of diseases from lymph node biopsies is one of the most difficult tasks of surgical pathology. Although experts show agreement with one another, diagnoses by community hospital pathologists have to be changed by expert hematopathologists in as many as 50 percent of the cases (Velez-Garcia et al. 1983). It is critical to differentiate benign from malignant conditions and to classify precisely malignant lymphoma so that the most appropriate form of surgery, radiation therapy, or chemotherapy can be chosen. For these reasons, improvements in diagnostic performance have the potential for significant benefits.

Early work on PATHFINDER at the Stanford University Medical Computer Science Group explored a variety of nonprobabilistic and quasiprobabilistic schemes (Heckerman 1988). These schemes included a rule-based production system using propositional logic, EMYCIN certainty factors, and Dempster-Shafer belief functions (Shafer 1976). Finally, PATHFINDER investigators tried a Bayesian probabilistic scheme, reinterpreting the numeric beliefs as probabilities and assuming the conditional independence of findings. The switch to the probabilistic interpretation resulted in a qualitative improvement in performance, that is, agreement with expert diagnosticians. This improvement was immediately apparent and remarked on by the expert in a blinded study, where the expert did not know which representation was being used.

The current PATHFINDER belief network includes about 30 different types of primary and secondary malignant hematopoietic diseases of the lymph node and about 30 different benign diseases that are easily confused with malignant lymphomas (Heckerman, Horvitz, and Nathwani 1991). The diseases are assumed to be mutually exclusive, an appropriate assumption in pathology. The

…no scheme for reasoning …under uncertainty can avoid making assumptions about prior beliefs and independence…

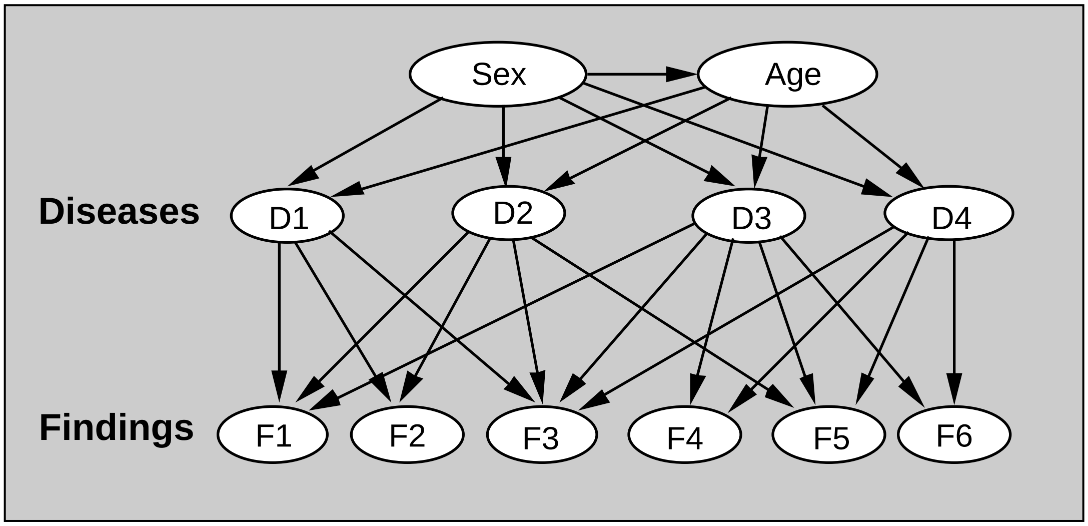


Figure 11: The Qualitative Structure of the QMR-BN Belief Network.

Disease probabilities are conditioned on the sex and age of the patient. Findings are conditionally independent given the diseases. Multiple connectedness (multiple undirected paths between nodes) is apparent.

network also includes information about the dependencies among the diseases and about 100 morphologic and nonmorphologic features visible in lymph node tissue. These findings are not assumed to be conditionally independent, and their dependencies are encoded explicitly, resulting in a multiply connected network. The network makes use of over 75,000 probabilities in performing inference. The use of similarity networks and partitions have made it practical to encode this amount of expert knowledge without unreasonable effort (Heckerman 1990). PATHFINDER performs exact diagnostic inference in this network in under a second on a 486-based computer.

Although PATHFINDER does not make treatment recommendations, it does use approximate expected value of information calculations to guide the sequencing of questions to those that will be most valuable in differentiating among the most probable and important diagnoses. As a measure of importance, it uses estimates of the relative costs of misdiagnosis, that is, the potential cost of treating for disease *A* when disease *B* is present. Currently, a study funded by the National Cancer Institute on the clinical efficacy of using the PATHFINDER belief network is under way at the University of Southern California.

INTELLIPATH is a commercial system based on the ideas in PATHFINDER and is distributed by the American Society of Clinical Pathologists. It provides a practical diagnostic aid for surgical pathologists. INTELLIPATH integrates probabilistic reasoning with a set of supportive tools, including an analog videodisc library of microscope images to illustrate typical features, text information on diseases and microscopic features, references to the literature, and an automated report generator. Hitherto, knowledge bases encoding the knowledge of teams of expert surgical pathologists are available for 10 of 40 different tissue types.

Another significant medical application is MUNIN, a system for diagnosis of neuromuscular disorders developed by a team at the University of Aalborg, Denmark (Andreassen et al. 1987). MUNIN is implemented in HUGIN, a general tool for construction and inference on belief networks (Andersen 1989). MUNIN contains about 1000 chance nodes with as many as 7 states each and has significant multiply connected portions. Exact diagnostic inference takes under 5 seconds.

A third medical application, SLEEP CONSULTANT, diagnoses sleep disorders and related pulmonary disorders (Nino-Marcia and Shwe 1991). It was developed by Knowledge Industries (KI), a Palo Alto, California, company that specializes in the development of probabilistic and decision-analytic systems. Data are automatically obtained from a digital polysomnograph monitor that records dynamic physiological data while the patient is sleeping (or trying to). In the physician's office, the data are transferred to the diagnostic system and combined with other data obtained from patient interviews and records to obtain probabilities for possible disorders. The system is being distributed by CNS, a Minneapolis company specializing in sleep diagnosis equipment.

A number of probabilistic systems are being developed for the diagnosis of capital equipment. Knowledge Industries has developed a diagnostic system for a major airline that deals with jet engine problems (figures 9 and 10). Another KI system is under development for the Electric Power Research Institute to diagnose problems with large gas turbines used by electric utilities. SRI International, together with Ontario Hydro, a Canadian power utility, developed GEMS (generator expert monitoring system), a prototype system for the diagnosis of electric power generator problems. This project is interesting in that it was originally developed using a forward-chaining rule-based system with certainty factors. Because of difficulties with this representation, the authors, who had no prior experience with probabilistic schemes, converted it to a belief network representation. They found this approach a far more natural and convenient representation of the causal knowledge of the experts. The resulting network contains about 1000 diagnostic nodes and 1300 links, primarily noisy-Ors (Klempner, Kornfeld, and Lloyd 1991).

Perhaps the largest Bayesian belief network currently in existence was developed as part of the QMR-DT Project (Shwe et al. 1991). QMR (Quick Medical Reference) is a knowledge base and diagnostic aid for internal medicine, a development of the INTERNIST-1 system. (Miller, Masarie and Myers 1986). QMR contains carefully encoded knowledge of almost 600 diseases, about 4,000 findings (signs, symptoms, lab results, and so on), and 40,000 links between them. QMR uses a heuristic, numeric representation of the uncertain relationships among diseases and findings. The QMR-DT (for QMR Decision Theory) Project—at Stanford University, Carnegie Mellon University, and the University of Pittsburgh—has developed a probabilistic reformulation of QMR. This reformulated system provides a rigorous probabilistic interpretation of the independence assumptions and a mapping from the numeric associations to conditional probabilities. The assumptions in the first version, QMR-BN (QMR Belief Net) (figure 11), include marginal independence of diseases given age and sex of the patient, conditional independence of findings, and noisy-Or influences. The system also includes *leak rates*, that is, probabilities that findings will be observed, in the absence of causal diseases. QMR-BN adds prior probabilities for diseases that are obtained from empirical data on prevalence rates.

The purpose of the QMR-DT Project is to explore the practicality and value of a coherent probabilistic interpretation compared to the existing heuristic QMR scheme. It has also stimulated the development of new inference algorithms: Because multiple diseases are possible and because the network is highly multiply connected, standard exact algorithms are rendered impractical. Random Sampling with likelihood weighting (Shwe et al. 1991) and the search-based TOP N algorithm (Henrion 1991) have both proved practical. Initial results have demonstrated comparable performance to the original QMR on a number of test cases (Middleton et al. 1991).

Conclusions

We introduced the basic ideas of probabilistic reasoning and decision analysis and reviewed their application to work on knowledge-based expert systems. Historically, interest in heuristic uncertainty calculi and rule-based representations arose partly in response to the computational difficulties and restrictive expressiveness of the early probabilistic expert systems. However, recent work has revealed that these heuristic schemes have problems of their own: restrictive assumptions and inconsistencies. A decision-theoretic perspective makes it clear that no scheme for reasoning and decision making under uncertainty can avoid making assumptions about prior beliefs and independence, whether these assumptions are implicit or explicit.

Recognition of the difficulties of the heuristic approaches, coupled with the recent development of more efficient and expressive representations from decision analysis, has stimulated a renewed interest in probabilistic and decision-theoretic approaches. In particular, belief networks and influence diagrams provide an appealing knowledge representation that can express uncertain knowledge, beliefs, and preferences in both qualitative and quantitative forms in a flexible yet principled fashion. Among the advantages of these representations, we have discussed the following. Knowledge, whether expert judgment or from empirical data, may be expressed in whichever direction is most natural, generally, but not necessarily in the causal direction. Probabilistic algorithms use this form to reason in whichever direction is required—causal, diagnostic, or intercausal. Thus, unlike rule-based representations, there is no essential link between the form of encoding and the form of reasoning for which it is used. The influence diagram also includes an explicit model of decisions and the utilities of their possible outcomes. This modularity maintains separation between judgments of likelihood and judgments of preferences. These representations provide the basis for a rapidly increasing array of developments.

A number of diagnostic systems based on probabilistic and decision-analytic ideas have achieved expert system status in that they have achieved expert-level performance. Indications are increasingly positive that these methods will scale up for large knowledge bases. A variety of general tools, or shells, for building and reasoning with belief networks and influence diagrams are now available (Andersen 1989; Arnold and Henrion 1990; Srinivas and Breese 1990).

Current research focuses on the development of improved tools for knowledge acquisition, the search for more efficient inference algorithms for complex multiply connected belief networks and influence diagrams, and more effective techniques for explaining decision-analytic reasoning. Some topics that deserve particular attention for further research include (1) the extending of the integration of influence diagrams with the representation of other kinds of semantic network notions, including causal relations, inheritance and abstraction hierarchies; (2) improved approximate inference algorithms for large networks; (3) compilation techniques for influence diagrams to support rapid inference; (4) a wider array of prototypical influences to facilitate knowledge engineering; (5) intelligent sensitivity analysis and more sophisticated knowledge-based aids to guide problem structuring; and (6) better error theories of the effects of simplifications and approximations. Of particular benefit for knowledge encoding, inference, and explanation would be the development of hybrid qualitative and quantitative approaches, including ranges of probabilities and values and order-of-magnitude estimates, to allow a wider array of techniques for handling partial and vague specifications that are qualitatively consistent with the principles of decision analysis.

In summary, we believe that the synthesis of ideas from decision analysis and expert systems and, more generally, from decision theory and AI is starting to generate considerable mutual benefit. Probability and decision theory provide a principled basis for automated reasoning and decision making under uncertainty. Decision analysis provides practical ideas and techniques for knowledge engineering. AI provides a variety of computational and representational techniques to help automate various aspects of decision analysis and, thus, render it more widely available. This interaction between decision analysis and expert systems has already produced some important contributions to both reasoning theory and knowledge representation and the development of a new generation of practical technology. In the long run, we anticipate the emergence of a combined enterprise in which the current separation between the two technologies disappears.

Acknowledgments

This work was supported by grant IRI-8703710 to Carnegie Mellon University from the National Science Foundation, internal research funds from Rockwell International Science Center, a National Aeronautics and Space Administration fellowship to Eric Horvitz under grant NCC-220-51 to Stanford University, and National Library of Medicine grant R01LM04529. We thank Greg Cooper, Renwick Curry, Lyn Dupre, Michael Fehling, Kenneth Fertig, Michael Genesereth, David Heckerman, Edward Herskovits, Harold Lehmann, and Peter Szolovits for critiquing drafts of this manuscript.

References

Andersen, S. 1989. HUGIN—A Shell for Building Bayesian Belief Universes for Expert Systems. In Proceedings of the Eleventh International Joint Conference for Artificial Intelligence, 1080-1085. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.

Andreassen, S.; Woldbye, M.; Falck, B.; and Andersen, S. 1987. MUNIN—A Causal Probabilistic Network for Interpretation of Electromyographic Findings. In Proceedings of the Tenth National Conference on Artificial Intelligence, 121–123. Menlo Park, Calif.: American Association for Artificial Intelligence.

Arnold, B., and Henrion, M. 1990. Demos Interface Manual, Technical Report, Dept. of Engineering and Public Policy, Carnegie Mellon Univ.

Ben-Bassat, M., and Teeni, D. 1984. Human-Oriented Information Acquistion in Sequential Pattern Classification, Part 1—Single Membership Classification. *IEEE Transactions on Systems, Man, and Cybernetics* 14:131–138.

Ben-Bassat, M.; Carlson, V.; Puri, V.; Davenport, M.; Schriver, J.; Latif, M.; Smith, R.; Lipnick, E.; and Weil, M. 1980. Pattern-Based Interactive Diagnosis of Multiple Disorders: The MEDAS System. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2:148–160.

Breese, J. 1990. Construction of Belief and Decision Networks, Technical Memorandum 30, Rockwell International Science Center, Palo Alto, California.

Buchanan, B., and Shortliffe, E., eds. 1984. *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Reading, Mass.: Addison-Wesley.

Chin, H., and Cooper, G. 1987. Stochastic Simulation of Bayesian Belief Networks. In Proceedings of the Third Workshop on Uncertainty in Artificial Intelligence, 106–113. Mountain View, Calif.: Association for Uncertainty in Artificial Intelligence.

Cooper, G. 1990. The Computational Complexity of Probabilistic Inference Using Belief Networks. *Artificial Intelligence* 42:393-405.

Cooper, G. 1984. NESTOR: A Computer-Based Medical Diagnostic Aid That Integrates Causal and Probabilistic Knowledge. Ph.D. diss., Dept. of Computer Science, Stanford Univ.

Davis, R. 1983. Diagnosis via Causal Reasoning: Paths of Interaction and the Locality Principle. In Proceedings of the Eighth International Joint Conference for Artificial Intelligence, 88–94. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.

Davis, R. 1982. Consultation, Knowledge Acquisition, and Instruction. In *Artificial Intelligence in Medicine*, ed. P. Szolovits, 57–78. Boulder, Colo.: Westview.

de Dombal, F.; Leaper, D.; Horrocks, J.; Staniland, J.; and McCann, A. 1974. Human and Computer-Aided Diagnosis of Abdominal Pain: Further Report with Emphasis on Performance. *British Medical Journal* 1:376–380.

de Dombal, F.; Leaper, D.; Staniland, J.; McCann, A.; and Horrocks, J. 1972. Computer-Aided Diagnosis of Acute Abdominal Pain. *British Medical Journal* 2:9–13.

de Kleer, J. 1991. Focusing on Probable Diagnosis. In Proceedings of the Ninth National Conference on Artificial Intelligence, 842-848. Menlo Park, Calif.: American Association for Artificial Intelligence.

Duda, R.; Gaschnig, J.; and Hart, P. 1979. Model Design in the PROSPECTOR Consultant System for Mineral Exploration. In *Expert Systems in the Microelectronic Age,* ed. D. Michie, 153–167. Edinburgh: Edinburgh University Press.

Elsaesser, C. 1988. Explanation of Probabilistic Inference for Decision Support Systems. In *Uncertainty in Artificial Intelligence 3*, eds. T. Levitt, J. Lemmer, and L. Kanal, 394–403. Amsterdam: North Holland.

Elsaesser, C., and Henrion, M. 1990. Verbal Expressions for Probability Updates: How Much More Probable Is "Much More Probable"? In *Uncertainty in Artificial Intelligence 5,* eds. M. Henrion, R. Shachter, L. N. Kanal, and J. Lemmer, 319–330. Amsterdam: Elsevier.

Genesereth, M. 1984. The Use of Design Descriptions in Automated Diagnosis. *Artificial Intelligence* 4:311–319.

Good, I. 1950. *Probability and the Weighing of Evidence*. New York: Hafners.

Gorry, G. 1973. Computer-Assisted Clinical Decision Making. *Methods of Information in Medicine* 12:45–51.

Gorry, G., and Barnett, G. 1968. Experience with a Model of Sequential Diagnosis. *Computers and Biomedical Research* 1:490–507.

Heckerman, D. 1990. Probabilistic Similarity Networks. *Networks* 20:607–636.

Heckerman, D. 1988. An Empirical Comparison of Three Inference Methods. In Proceedings of the Fourth Workshop on Uncertainty in Artificial Intelligence, 158-169. Mountain View, Calif.: Association for Uncertainty in Artificial Intelligence.

Heckerman, D. 1986. Probabilistic Interpretations for MYCIN's Certainty Factors. In *Uncertainty in Artificial Intelligence,* eds. L. Kanal and J. Lemmer, 167-196. Amsterdam: North Holland.

Heckerman, D., and Horvitz, E. 1987. On the Expressiveness of Rule-Based Systems for Reasoning under Uncertainty. In Proceedings of the Sixth National Conference on Artificial Intelligence, 121–126. Menlo Park, Calif.: American Association for Artificial Intelligence.

Heckerman, D.; Horvitz, E.; and Nathwani, B. 1991.Toward Normative Expert Systems: The PATHFINDER Project. *Methods of Information in Medicine*. Forthcoming.

Heckerman, D. and Jimison, H. 1989. A Perspective on Confidence and Its Use in Focusing Attention During Knowledge Acquisition. In *Uncertainty in Artificial Intelligence 3*, eds. L. Kanal, T. Levitt, and J. Lemmer, 32-131. New York: North Holland.

Henrion, M. 1991. Search-Based Methods to Bound Diagnostic Probabilities in Very Large Belief Networks. In Proceedings of the Seventh Workshop on Uncertainty in Artificial Intelligence, 142–150. San Mateo, Calif.: Morgan Kaufmann.

Henrion, M. 1990a. Toward Efficient Inference in Multiply Connected Belief Networks. In *Influence Diagrams, Belief Nets, and Decision Analysis,* eds. R. Oliver and J. Smith, 385–407. Chichester, England: Wiley.

Henrion, M. 1990b. An Introduction to Algorithms for Inference in Belief Networks. In *Uncertainty in Artificial Intelligence 5,* eds. M. Henrion and R. Shachter,129-138. Amsterdam: North Holland.

Henrion, M. 1988a. Practical Issues in Constructing a Bayes' Belief Network. In *Uncertainty in Artificial Intelligence 3*, eds. T. Levitt, J. Lemmer, and L. Kanal, 132–139. Amsterdam: North Holland.

Henrion, M. 1988b. Propagation of Uncertainty by Probabilistic Logic Sampling in Bayes' Networks. In *Uncertainty in Artificial Intelligence 2*, eds. J. Lemmer and L. Kanal, 149–164. Amsterdam: North Holland.

Henrion, M. 1987. Uncertainty in Artificial Intelligence: Is Probability Epistemologically and Heuristically Adequate? In *Expert Judgment and Expert Systems,* ed. J. Mumpower, 105–130. Berlin: Springer-Verlag.

Henrion, M., and Arnold, B. 1991. Demos Graphic Manual, Technical Report, Dept. of Engineering and Public Policy, Carnegie Mellon Univ.

Henrion, M., and Cooley, D. 1987. An Experimental Comparison of Knowledge Engineering for Expert Systems and for Decision Analysis. In Proceedings of the Sixth National Conference on Artificial Intelligence, 471–476. Menlo Park, Calif.: American Association for Artificial Intelligence.

Henrion, M., and Druzdzel, M. 1990. Qualitative and Linguistic Explanations of Probabilistic Reasoning in Belief Networks. In Proceedings of the Sixth Workshop on Uncertainty in Artificial Intelligence, 10–20. Mountain View, Calif.: Association for Uncertainty in Artificial Intelligence.

Holtzman, S. 1988. *Intelligent Decision Systems*. Reading, Mass.: Addison-Wesley.

Horvitz, E. 1988. Reasoning Under Varying and Uncertain Resource Constraints. In Proceedings of the Seventh National Conference on Artificial Intelligence, 111-116. Menlo Park, Calif.: American Association for Artificial Intelligence.

Horvitz, E. 1987. A Multiattribute Utility Approach to Inference Understandability and Explanation, Technical Report, KSL-87-28, Depts. of Computer Science and Medicine, Stanford Univ.

Horvitz, E. and Rutledge, G. 1991. Time-Dependent Utility and Action Under Uncertainty. In Proceedings of the Seventh Conference on Uncertainty in Artificial Intelligence, 151-158. San Mateo, California; Morgan Kaufman.

Horvitz, E.; Cooper, G.; and Heckerman, D. 1989. Reflection and Action Under Scarce Resources: Theoretical Principles and Empirical Study. In Proceedings of the Eleventh International Joint Conference on Artificial Intelligence, 1121-1127. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.

Horvitz, E.; Suermondt, H.; and Cooper, G. 1989. Bounded Conditioning: Flexible Inference for Decisions under Scarce Resources. In Proceedings of the Fifth Workshop on Uncertainty in Artificial Intelligence, 182–193. Mountain View, Calif.: Association for Uncertainty in Artificial Intelligence.

Horvitz, E.; Heckerman, D.; Ng, K. and Nathwani, B. 1989. Heuristic Abstraction in the Decision-Theoretic PATHFINDER System. In Proceedings of the Symposium on Computers in Medical Care, 178-182. Los Angeles: IEEE Computer Society Press.

Howard, R. 1970. Risk Preference. In *Readings on the Principles and Applications of Decision Analysis,* eds. R. Howard and J. Matheson, 629–663. Menlo Park, Calif.: Strategic Decisions Group.

Howard, R. 1968. The Foundations of Decision Analysis. *IEEE Transactions on Systems, Science, and Cybernetics* 4:211–219.

Howard, R., and Matheson, J. 1984. *Readings on the Principles and Applications of Decision Analysis*. Menlo Park, Calif.: Strategic Decisions Group.

Howard, R., and Matheson, J. 1981. Influence Diagrams. In *Readings on the Principles and Applications of Decision Analysis*, eds. R. Howard and J. Matheson, 721–762. Menlo Park, Calif.: Strategic Decisions Group.

Kahneman, D.; Slovic, P.; and Tversky, A. 1982. *Judgement under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

Kalagnanam, J., and Henrion, M. 1990. Comparing Expert Rules and Decision Analysis for Sequential Diagnosis. In *Uncertainty in Artificial Intelligence 4,* eds. T. Levitt, R. Shachter, L. Kanal, and J. Lemmer, 271–282. Amsterdam: Elsevier.

Keeney, R., and Raiffa, H. 1976. *Decisions with Multiple Objectives: Preferences and Value Trade-Offs*. New York: Wiley.

Kim, J., and Pearl, J. 1983. A Computational Model for Causal and Diagnostic Reasoning in Inference Engines. In Proceedings of the Eighth International Joint Conference on Artificial Intelligence, 190–193. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.

Klein, D. 1990. Explaining and Refining Decision-Theoretic Choices. Ph.D. diss., Dept. of Computer Science, Univ. of Pennsylvania.

Klempner, G.; Kornfeld, A.; and Lloyd, B. 1991. The EPRI Generator Expert Monitoring System: Experience with the GEMS Prototype. Presented at the American Power Conference, May, Chicago, Illinois.

Langlotz, C.; Shortliffe, E.; and Fagan, L. 1986. A Methodology for Computer-Based Explanation of Decision Analysis, Technical Report, KSL-86-57, Depts. of Computer Science and Medicine, Stanford Univ.

Langlotz, C.; Fagan, L.; Tu, S.; Sikic, B.; and Shortliffe, E. 1986. Combining Artificial Intelligence and Decision Analysis for Automated Therapy Planning Assistance. In Proceedings of Medinfo-86. Washington, D.C.

Lauritzen, S., and Spiegelhalter, D. 1988. Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems. *Journal of the Royal Statistical Society B*50:157–224.

Ledley, L., and Lusted, R. 1959. Reasoning Foundations of Medical Diagnosis. *Science* 130:9–21.

Lemmer, J. 1983. Generalized Bayesian Updating of Incompletely Specified Distributions. *Large-Scale Systems* 5:51-68.

McLaughlin, J. 1987. The Utility-Directed Presentation of Graphical Simulation, Technical Report, KSL-87-59, Depts. of Computer Science and Medicine, Stanford Univ.

Middleton, B.; Shwe, M.; Heckerman, D.; Henrion, M.; Horvitz, E.; Lehmann, H.; and Cooper, G. 1991. Probabilistic Diagnosis Using a Reformulation of the INTERNIST-1/QMR Knowledge Base, Part 2—Evaluation of Diagnostic Performance. *Methods of Information in Medicine.* 30:256-267.

Miller, R.; Masarie, F.; and Myers, J. 1986. Quick Medical Reference (QMR) for Diagnostic Assistance. *Medical Computing* 3(5):34-48.

Minsky, M., and Selfridge, O. 1961. Learning in Random Nets. In *Information Theory*, ed. C. Cherry, 335–347. London: Butterworths.

Morgan, M., and Henrion, M. 1990. *Uncertainty: A Guide to the Treatment of Uncertainty in Policy and Risk Analysis*. Cambridge: Cambridge University Press.

Nino-Murcia, G. and Shwe, M. 1991. An Expert System for Diagnosis of Sleep Disorders. In *Sleep Research*, vol. 20, eds. M. Chase, R. Lydic, and C. O'Connor, 433. Los Angeles: Brain Information Service.

Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems.* San Mateo, Calif.: Morgan Kaufmann.

Pearl, J. 1987. Evidential Reasoning Using Stochastic Simulation of Causal Models. *Artificial Intelligence* 32:245–257.

Pearl, J. 1986a. A Constraint-Propagation Approach to Probabilistic Reasoning. In *Uncertainty in Artificial Intelligence*, eds. L. N. Kanal and J. F. Lemmer, 357–369. Amsterdam: North Holland.

Pearl, J. 1986b. Fusion, Propagation, and Structuring in Belief Networks. *Artificial Intelligence* 29:241–288.

Pearl, J., and Verma, T. 1987. The Logic of Representing Dependencies by Directed Graphs. In Proceedings of the Sixth National Conference on Artificial Intelligence, 374–379. Menlo Park, Calif.: American Association for Artificial Intelligence.

Peirce, C. 1956. The Probability of Induction. In *The World of Mathematics*, 1341–1354. New York: Simon and Schuster.

Peng, Y., and Reggia, J. 1987. A Probabilistic Causal Model for Diagnostic Problem Solving, Part 1—Integrating Symbolic Causal Inference with Numeric Probabilistic Inference. *IEEE Transactions on Systems, Man, and Cybernetics* SMC-17(2): 146–162.

Politser, P. 1984. Explanations of Statistical Concepts: Can They Penetrate the Haze of Bayes? *Methods of Information in Medicine* 23:99–108.

Raiffa, H. 1968. *Decision Analysis: Introductory Lectures on Choice under Uncertainty.* Reading, Mass.: Addison-Wesley.

Reggia, J., and Perricone, B. 1985. Answer Justification in Medical Decision Support Systems Based on Bayesian Classification. *Computers in Biology and Medicine* 15:161–167.

Sember, P., and Zukerman, I. 1989. Strategies for Generating Micro Explanations for Bayesian Belief Networks. In Proceedings of the Fifth Workshop on Uncertainty in Artificial Intelligence, 295–302. Mountain View, Calif.: Association for Uncertainty in Artificial Intelligence.

Shachter, R. 1986. Intelligent Probabilistic Inference. In *Uncertainty in Artificial Intelligence,* eds. L. Kanal and J. Lemmer, 371–382. Amsterdam: North Holland.

Shachter, R., and Heckerman, D. 1987. Thinking Backward for Knowledge Acquisition. *Artificial Intelligence* 8:55–63.

Shachter, R., and Peot, M. 1990. Simulation Approaches to General Probabilistic Inference on Belief Networks. In *Uncertainty in Artificial Intelligence 5*, eds. M. Henrion et. al., 221-231. Amsterdam: North Holland.

Shafer, G. 1976. *A Mathematical Theory of Evidence*. Princeton, N.J.: Princeton University Press.

Shortliffe, E. 1984. Explanation Capabilities for Medical Consultation Systems. In Proceedings of AAMSI 1984, Third Spring Joint National Conference, eds. D. A. B. Lindberg and M. F. Collen, 193–197.

Shwe, M.; Middleton, B.; Heckerman, D.; Henrion, M.; Horvitz, E.; Lehmann, H.; *et al.* 1991. Probabilistic Diagnosis Using a Reformulation of the INTERNIST-1/QMR Knowledge

Base: I. The Probabilistic Model and Inference Algorithms. *Methods of Information in Medicine*. 30:241-255.

Simon, H. 1955. A Behavioral Model of Rational Choice. *Quarterly Journal of Economics* 69:99–118.

Spetzler, C., and Stael von Holstein, C. 1975. Probability Encoding in Decision Analysis. *Management Science* 22:340–358.

Spiegelhalter, D., and Knill-Jones, R. 1984. Statistical and Knowledge-Based Approaches to Clinical Decision Support Systems, with an Application in Gastroenterology. *Journal of the Royal Statistical Society* 147:35–77.

Srinivas, S., and Breese, J. 1990. IDEAL: A Software Package for the Analysis of Influence Diagrams. In Proceedings of the Sixth Workshop on Uncertainty in Artificial Intelligence, 212-219. Mountain View, Calif.: Association for Uncertainty in Artificial Intelligence.

Swartout, W. 1983. XPLAIN: A System for Creating and Explaining Expert Consulting Systems. *Artificial Intelligence* 21:285–325.

Szolovits, P. 1982. Artificial Intelligence in Medicine. In *Artificial Intelligence in Medicine,* ed. P. Szolovits, 1-19. Boulder, Colo.: Westview.

Szolovits, P., and Pauker, S. 1978. Categorical and Probabilistic Reasoning in Medical Diagnosis. *Artificial Intelligence* 1:115–144.

Teach, R., and Shortliffe, E. 1981. An Analysis of Physician Attitudes Regarding Computer-Based Clinical Consultation Systems. *Computers and Biomedical Research* 14:542–558.

Velez-Garcia, E.; Durant, J.; Gams, R.; and Bartolucci, A. 1983. Results of a Uniform Histopathological Review System of Lymphoma Cases: A Ten-Year Study of the Southeastern Cancer Study Group. *Cancer* 52:675–679.

von Winterfeldt, D., and Edwards, W. 1986. *Decision Analysis and Behavioral Research*. New York: Cambridge University Press.

Wallis, J., and Shortliffe, E. 1982. Explanatory Power for Medical Expert Systems: Studies in the Representation of Causal Relationships for Clinical Consultations. *Methods of Information in Medicine* 21:127–136.

Wallsten, T.; Budescu, D.; Rapoport, A.; Zwick, R.; and Forsyth, B. 1986. Measuring the Vague Meanings of Probability Terms. *Journal of Experimental Psychology: General* 115(4): 348–365.

Wellman, M. 1988a. Formulation of Trade-Offs in Planning under Uncertainty. Ph.D. diss., Dept. of Computer Science, Massachusetts Institute of Technology.

Wellman, M. 1988b. Qualitative Probabilistic Networks for Planning under Uncertainty. In *Uncertainty in Artificial Intelligence 2*, eds. J. Lemmer and L. Kanal, 197–209. Amsterdam: North Holland.

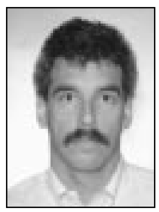
Wellman, M. 1985. Reasoning about Preference Models. Master's thesis, Dept. of Computer Science, Massachusetts Institute of Technology.

Wellman, M., and Henrion, M. 1991. Qualitative Intercausal Relations, or Explaining Explaining Away. In *KR-91: Principles of Knowledge Representation and Reasoning—Proceedings of the Second International Conference,* 535–546. San Mateo, Calif.: Morgan Kaufmann.

Winograd, T. 1971. A Computer Program for Understanding Natural Language, Technical Report, TR-17, AI Lab, Massachusetts Institute of Technology.

Wise, B. 1986. An Experimental Comparison of Uncertain Inference Systems. Ph.D. diss., Departments of Engineering and Public Policy, Carnegie Mellon University.

Max Henrion is a research scientist at the Rockwell International Science Center, Palo Alto Laboratory, and a consulting associate professor at Stanford University. Until 1990, he was an associate professor at Carnegie Mellon University. He has an M.A. from Cambridge University, an M.Des. from the Royal College of Art, and a Ph.D. from Carnegie Mellon University. His research interests are in decision theory and AI, particularly in reasoning under uncertainty. He is the author of DEMOS, software for probabilistic modeling and uncertainty analysis. He has developed applications in energy and environmental systems, aerospace, and medicine. He is coauthor of *Uncertainty* (Cambridge University Press, 1990). He was the founding president of the Association for Uncertainty in Artificial Intelligence. He is a member of the American Association for Artificial Intelligence, and the Institute for Management Science.



John Breese is a researcher at the Rockwell International Science Center, Palo Alto Laboratory. He received his Ph. D. from Stanford in 1987. His interests since 1982 have centered on the practical and theoretical issues at the interface between decision theory and artificial intelligence. Recent projects have examined decision-theoretic control of pilot displays, knowledge-based construction of decision models, and planning. He is also a principal in Knowledge Industries, a company specializing in decision-analytic software and services. He is a member of the American Association of for Artificial Intelligence and the Institute for Management Science.

Eric Horvitz is a research scientist at the Rockwell Palo Alto Laboratory. He completed his PhD in 1990 at Stanford University, specializing in models of rational computation and action under scarce resources. His research interests include the utility-directed control of inference in large probabilistic knowledge bases, flexible procedures for determining ideal actions under varying resources, and techniques for exploiting the relationships between deliberative and compiled reasoning. He is a cofounder and principal of Knowledge Industries, a Palo Alto company dedicated to developing reasoning systems based on decision-theoretic principles. He is a member of the American Association for Artificial Intelligence, the Institute for Management Science, the Institute of Electrical and Electronic Engineers, and the American Medical Informatics Association.