Background to Qualitative Decision Theory

Jon Doyle and Richmond H. Thomason

■ This article provides an overview of the field of qualitative decision theory: its motivating tasks and issues, its antecedents, and its prospects. Qualitative decision theory studies qualitative approaches to problems of decision making and their sound and effective reconciliation and integration with quantitative approaches. Although it inherits from a long tradition, the field offers a new focus on a number of important unanswered questions of common concern to AI, economics, law, psychology, and management.

The field of decision theory and its companion methodology of decision analysis deal with the merits and making of decisions. As developed by philosophers, economists, and mathematicians over some 300 years, these disciplines have developed many powerful ideas and techniques, which exert major influences over virtually all the biological, cognitive, and social sciences. Their uses range from providing mathematical foundations for microeconomics to daily application in a range of fields of practice, including finance, public policy, medicine, and now even automated device diagnosis.

In spite of these remarkable achievements, the tools of traditional decision theory have not proven fully adequate for supporting recent attempts in AI to automate decision making. The field of qualitative decision theory aims to provide better support for automation efforts by developing qualitative and hybrid qualitative-quantitative representations and procedures that complement and improve the quantitative approach's ability to address the full range of decision-making tasks in the way such tasks appear within larger activities of planning, learning, and collaboration.

The following brief survey of qualitative decision theory seeks to stimulate new work in the area and alert researchers in other areas to topics of mutual interest. We first illustrate some of the motivations for pursuing more qualitative approaches and continue by examining the nature of traditional decision theory and analysis. We then identify a number of technical issues and topics for investigation. We provide sketches of representative results and work concerning these matters. Much of this work is incomplete and preliminary, providing many opportunities for further research. The concluding remarks seek to reflect on the available results to help set the context for future studies.

A Challenging Example

Naive and expert humans regularly and routinely solve decision problems that challenge the formalizations and methods of decision making that predominate in economics and AI.

To illustrate this point, consider someone, call him Aethelred, who goes to meet with a financial planner. Aethelred brings a number of preferences to the meeting, some of them implicit and unexamined, some pretty abstract or generic, and many of them competing. He feels restless and dissatisfied and would rather retire early, although he will not be forced to retire on account of age. He is more timid about financial risk than he was when he was younger and unsure about the financial markets. He feels increasingly uneasy about his health (one reason he wants to retire early).

When Aethelred meets with the planning expert, he somehow focuses on one part of this mix of preferences and beliefs and produces a goal. He says, "I want to retire at age 60." This opening remark clearly does not provide a total picture of his state of mind, nor does traditional decision theory provide a way of expressing the goal formally. A good advisor might explore what happens through planning from the supposed goal but will also be quite willing to challenge the goal itself. Suppose, for example, that the expert describes some scenarios based on the announced goal, and Aethelred is

unhappy with them all. The expert asks him why he wants to retire at 60. Aethelred finds he can't produce any very compelling reason. He chose the number 60 arbitrarily. His reasons for preserving his salary and employee benefits are much more compelling than his reasons for early retirement. The expert points out that a large proportion of early retirees are restless and dissatisfied with their retirement. At the end of the discussion, Aethelred's preferences (or the degrees of importance he attaches to preference factors) have changed. He decides to do nothing for at least five years and to try harder to enjoy his job before rethinking the matter. The expert has neither brainwashed nor compelled him but has instead helped Aethelred to reexamine and readjust his preferences.

Naive and expert humans regularly and routinely solve decision problems that challenge the formalizations and methods of decision making that predominate in economics and AI.

This scenario exhibits features typical of many formal and informal advising situations. The advisor initially knows nothing about the advisee's preferences, possessing only expectations developed from stereotypes and evidence provided by the advisee's current financial situation. The advisor develops better information about these preferences in the course of the advisory dialog, but some uncertainties might remain even at the end of the conversation. The same situation obtains concerning the advisee's beliefs and subjective probabilities, for which the advisor refines or corrects initial ignorance or expectations in the course of the interview. The advisor might also treat information about the advisee's beliefs differently from information about preferences; rather than simply accepting both types of information as parameters defining the financial planning problem, the advisor might challenge or seek to correct the client's views, for example, on the likelihood of inflation, the riskiness of certain investments, or the permanence of current government policies. In discussing both preference and probability information, the advisor and advisee mainly stick to generalities and explanations of reasons and assumptions; they do not descend, at least in the preliminary stages, to talk about specific probability or utility values.

The communications and actions illustrated in this scenario involve nothing out of the ordinary but challenge both the theory of expected utility and the formalisms used in AI planning systems because these make no provision for making decisions in cases where there is uncertainty about goals and preferences. Neither do they provide a satisfactory model for rational revision of goals and preferences, for using stereotypes concerning values and beliefs in modeling an advisee, or for the incremental

extraction of information about either values or beliefs through a conversational medium. The scenario suggests the importance of these issues in advising tasks, but the issues also arise, perhaps even more strongly, in tasks involving planning and acting, for example, should Aethelred hire someone to manage his estate in accord with his preferences.

To better understand how to respond to these challenges, we first reexamine the nature of decision theory and its standard methods of application.

Decision Theory and Its Discontents

Decision theory and decision analysis provide answers to four main questions:

1. What Is a Decision? A decision is a choice made by some entity of an action from some set of alternative actions.

Decision theory has nothing to say about actions—either about their nature or about how a set of them becomes available to the decision maker. The main branch of decision theory treats decisions of individual actors, and other branches treat decisions made by groups of individuals or groups of groups. We focus here on choices made by individual decision makers.

2. What Makes a Decision Good? A good decision identifies an alternative that the decision maker believes will prove at least as good as other alternative actions.

We here recount the standard answer about ideal decisions. Variant decision theories, touched on later, provide different answers, such as calling a decision good if it identifies an alternative that the decision maker considers good enough, even if not as good as other alternatives that might have been found through further deliberation.

3. How Should One Formalize Evaluation of Decisions? Good decisions are formally characterized as actions that maximize *expected utility,* a notion involving both belief and goodness.

Decision theory develops this notion in stages. Let *A* stand for the set of actions or alternatives. Decision theory first presumes an association of a set of *outcomes* with each action. As with actions, the theory says little about the nature of outcomes (for example, whether they range over long or short intervals of time) or about the factors that determine the association, leaving it to the decision analyst to decide these things. Let Ω stand for the set of all outcomes identified with any actions (the union of those associated with each action).

Second, the theory presumes a measure *U* of outcome value that assigns a *utility U(ω)* to each outcome $\omega \in \Omega$. To make this assignment possible, the theory requires that outcomes be identified so as to have some determinate value or utility in the context of the decision under consideration to ensure that a single outcome cannot come about in ways that differ in value. Given this requirement, the theory takes outcomes as unrefinable because no refinements matter to the model of the decision.

Third, the theory presumes a measure of the probability of outcomes conditional on actions, with *Pr(ω | a)* denoting the probability that outcome ω comes about after taking action $a \in A$ in the situation under consideration.

Using these elements, the theory defines the expected utility *EU(a)* of an action *a* as the average utility of the outcomes associated with the alternative, weighting the utility of each outcome by the probability that the outcome results from the alternative, that is, $EU(a) = \int_{\Omega} U(\omega) \, Pr(\omega \mid a) d\omega$.

Decision theory defines *rational* decision makers as those that always maximize expected utility. Again, we have described only the orthodox account; some approaches employ nonprobabilistic notions of belief, different notions of worth, or criteria other than expected utility maximization.

4. How Should One Formulate the Decision Problem Confronting a Decision Maker? One identifies alternatives, outcomes, probabilities, and utilities through an iterative process of hypothesizing, testing, and refining a sequence of tentative formulations.

One identifies the alternatives and outcomes through direct queries or knowledge of the circumstances. (How can you intervene in this situation? How do you evaluate the success of an intervention?) One assesses probabilities and utilities either through direct queries (How likely do you think this outcome is? How much would you be willing to pay for this outcome?) or indirectly by inferring them from patterns of hypothetical choices. The indirect approach exploits fundamental theorems of decision theory showing that certain patterns uniquely determine the probabilities and utilities. (In particular, Savage's [1972] representation theorem provides an isomorphism between rational preferences over uncertain outcomes and quantitative representations in terms of expected utility.) One repeatedly improves tentative identifications of alternatives, outcomes, probabilities, and utilities by evaluating their completeness, consistency, conceptual convenience, sensitivity to errors in assessment, and relation to intuitive judgments, if any. Complicated cases can require many iterations and substantial analytic effort to obtain a satisfactory decision formulation.

These central questions by no means exhaust the important practical or philosophical issues. Philosophical criticisms of utility-based approaches to good decisions go back at least to the debate over utilitarianism in the late eighteenth and early nineteenth centuries. They include doubts about the existence of any universal currency for reconciling very different sorts of goods and evils (see, for example, chapter 7 of Hare [1963] and Keeney and Raiffa [1976]), the possibility of estimating the outcomes of actions (especially over long periods of times), and the possibility of defining a legitimate measure of group welfare (especially in the context of actions that can change the membership of the group whose welfare is at stake).

The issues that arise in automated decision making recall some of the traditional philosophical questions but locate the debate in a new setting by focusing on the need for efficient reasoning in the very complex decision-making cases that can arise in human affairs. Quantitative representations of probability and utility, and procedures for computing with these representations, do provide an adequate framework for manual treatment of very simple decision problems, but they are less successful in more complicated and realistic cases.

For example, virtually every autonomous system, whether it be a robotic spacecraft exploring the solar system or a housecleaning robot exploring under a third-grader's bed, must possess the ability to formulate decisions in unforeseen settings. Performing these tasks might require unforeseen preferences and call for the structuring of the new decision problem by applying general guidelines to the particulars of the situation. Such creativity in individual decision making requires reasoning quite similar in quality to the advisory reasoning that we found in the "challenging example" presented earlier.

Simply attempting to mechanize elements of the traditional decision-analytic approach is challenging in itself and provides many highly useful projects. (See Linden, Hanks, and Lesh [1997], for example, for an application of multiattribute utility theory to infer user preferences over fully determinate outcomes.) One can get even further by seeking to apply standard knowledge representation methods for structuring and representing decision-analytic information, as can be seen in the developing field of automated decision-model construction. (See Wellman, Breese, and Goldman [1992] for an overview of this area.)

The traditional decision-theoretic approach, however, fails on many counts to provide useful guidance for decision-making activities. It does not address making decisions in unforeseen circumstances or changing the assumptions that underlie decisions. It offers no means for capturing generic preferences (that is, preferences among classes, such as "I'd rather eat Italian than Chinese tonight") and other common human expressions of decision-making guidelines in convenient formal terms, even though such expressions compactly communicate large amounts of information at comfortable levels of detail. The traditional approach provides little help in modeling decision makers who exhibit discomfort with numeric trade-offs or who exhibit outright inconsistency. Finally, the traditional approach provides little help in effectively representing and reasoning about decisions involving broad knowledge of the world and in communicating about the reasons for decisions in ways that humans will find intelligible. These failures severely limit the utility of the unadorned traditional concepts for automating autonomous decision making and, more generally, for eliciting decision-making information and guidelines from decision makers or communicating such information to decision makers or decision-making aids, such as emergency room physicians, executors of wills, governmental representatives, or even (one might hope) everyday software systems. Automating these abilities can require rethinking not only the classical approach to decision theory to take into account qualitative goals, generic preferences, and the changeability of preferences but also the usual approaches to planning and learning to relate their goals and methods to preferences and to base their decisions on decision-theoretic information rather than ad hoc rules (see Dean and Wellman [1989]).

Qualitative decision theory seeks to adapt or extend traditional decision-theoretic and analytic concepts to a broader context, building whenever possible on existing strengths. Doing this requires facing up to the limitations of traditional decision-theoretic conceptual tools. In some cases, such as how to best represent actions and outcomes, these limitations correspond to genuine unanswered questions about knowledge representation, questions reaching far beyond qualitative decision theory. In other cases, the limitations represent weaknesses of the quantitative representations and methods used, not of the underlying decision theory. Recent research has begun to tackle some of these problems, but others remain essentially unexplored.

Traditional decision theory provides an account of the information that, in principle, suffices for making a rational decision. In practice, however, the decision maker might have never considered the type of choice in question and so might not happen to possess this information.

Information-Limited Rationality

Traditional decision theory provides an account of the information that, in principle, suffices for making a rational decision. In practice, however, the decision maker might have never considered the type of choice in question and so might not happen to possess this information. Groups seeking to come to decisions, as well as individual decision makers possessing group-like internal organizations (for example, "society of mind" agents), can harbor competing preferences and so might suffer additional limitations illustrated by Arrow's theorem, which shows the impossibility of combining group-member preferences in ways that meet conditions of ideal rationality (see Doyle and Wellman [1991] and Arrow [1959]). Even when a decision maker with coherent preferences knows how to obtain the information needed for a decision, the process can take an amount of time that, even if known, can exceed what is available or reasonable for deliberation, leading the deliberation to proceed without complete information. In other cases, the decision maker may feel quite sure that the decision in question requires only a few preferences already possessed, not the large number needed to relate all possible contingencies.

The need for an approach to decision making based on partial information provides a powerful motivation for a qualitative decision theory. Common experience shows people offering partial, abstract, generic, tentative, and uncertain information in explaining their decisions. This information includes qualitative probability ("I'm likely to need $50,000 a year for retirement"), generic preference information ("I prefer investments in companies that respect the environment"), and generic goals ("I want to retire young enough to enjoy it"). If we wish a more direct model of the way people seem to think about decisions, we need to deal with such information and need models of deliberative reasoning that can make use of it in formulating and making decisions.

Because qualitative decision theory seeks to formalize reasonable decision making by relaxing the decision preconditions of classical decision theory, the study of qualitative decision making is closely related to the study of limited rationality (see, for example, Doyle [1992]). The recent work in qualitative decision theory, however, has not yet developed to the point where it has a distinct contribution to make to the field of limited rationality. Using notions such as expected value of information (Howard 1966) to guide the search for better or more complete decision-making information seems an obvious topic for pursuit here.

Formalization Tasks

Developing qualitative representations and methods for automated decision making and for direct models of the way people think about decisions leads to a variety of challenging formalization tasks. Some of these tasks concern the notions of alternative actions, outcomes, probabilities, and preferences that have already received substantial attention. For example, although decision theory takes outcomes as unrefinable, practical approaches must take the opposite view and seek to reveal an internal structure for outcomes that facilitates computation of probability and utility information (see, for example, Boutilier, Dean, and Hanks [1999]). This means interpreting these nominally unrefinable outcomes as constructs from sets of properties or attributes, a task that shades off into the much larger area of general knowledge representation. Beyond these well-developed topics, qualitative decision theory must also look for formalizations of a number of other notions.

Fortunately, AI and philosophical logic have developed formalizations of some of these notions already, although much remains undone. Of course, the usual standards of formalization apply. In particular, representations should have a sound semantics. The formalization should illuminate concepts and methods with deep theorems that aid mathematical analysis, computation, and reasoning. If possible, the formalization should also reflect familiar concepts and constructs and relax classical decision-theoretic concepts in reasonable ways.

Generic and Structural Information

Standard decision theory rests on qualitative foundations, expressed in terms of qualitative preference orders and comparisons among alternatives. Unfortunately, these foundations do not provide any ready way of expressing generic probabilities and preferences. Recent work has begun to address these gaps with active investigations of logics of probability and preference; work on multiattribute utility structures; partial probability theory; and variations from the standard decision-theoretic model, including possibilistic logic.

Work on qualitative representation of probabilistic information has explored several paths, notably qualitative probabilistic networks, logics of probability, and partial probabilities. *Qualitative probabilistic networks,* introduced by Wellman (1990b), directly represent generic probabilistic influences, for example, "The more likely that an investment has done well recently, the more likely it will do well in the future." Formal semantics for such generic

probabilistic relationships bear close similarities to some logics of preferences discussed later. Direct logics of probability—as in the work of Haddawy (1991); Halpern (1990); and Bacchus, Grove, Halpern, and Koller (1996)—permit specification of properties of probability distributions without necessarily specifying particular values of the distribution. *Partial probability theory* relaxes classical probability by replacing a single probability measure by a family of such measures. A generalization of this idea to utility theory is discussed in Voorbraak (1997). Attractive as this idea might be as a mathematical approach to probabilistic uncertainty, the problem of providing a theory of reasoning with families of pure probabilities is challenging (see, for example, Ha and Haddawy [1998a]) and is likely to be even more challenging when utilities are brought into the picture.

Direct *logics of preference* have been studied in the philosophical literature; see, for example, Hansson (1996), Hansson (1990), and Packard (1975). Papers such as Wellman and Doyle (1991), Doyle, Shoham, and Wellman (1991), Boutilier (1994), Doyle and Wellman (1994), Mantha (1994), Tan and Pearl (1994a, 1994b), Doyle (1995), and Bell and Huang (1997) present variations on this approach. Logics of this kind are compatible with most of the other approaches described here and could serve as a useful way of organizing and comparing them.

It has recently occurred to many members of the AI community that *multiattribute utility,* developed in works such as Keeney and Raiffa (1976) and Fishburn (1977), can be an important component of a qualitative approach to decision theory. Multiattribute utility theory provides methods for constructing utility functions over several attributes of states by means of functional composition of subutility functions over subsets of the attributes. For example, a simple utility function might simply sum (or multiply) together subutility measures over each individual attribute. To our knowledge, Wellman (1985) presents the first detailed attempt to build on this work in the literature on qualitative decision theory (and even perhaps the first citation to it). Other papers that make use of multiattribute utility are Haddawy and Hanks (1998, 1992, 1990); Doyle, Shoham, and Wellman (1991); Wellman and Doyle (1992); Draper, Hanks, and Weld (1994); Linden, Hanks, and Lesh (1997); Boutilier, Brafman, and Geib (1997); Bacchus and Grove (1997, 1996, 1995); and Shoham (1997a, 1997b, 1997c). Some of these works begin development of languages for composing utility measures and construction of

libraries of standard forms for utility functions. Multiattribute decision theory identifies some basic structural forms here, but investigation of decision-specific forms and ways of organizing them holds much promise.

Now, multiattribute utility theory in itself merely provides a way of representing and calculating the classical utilities of outcomes by decomposing them into the utilities of the value-relevant factors that make up the outcomes. The theory becomes relevant to qualitative (or at least nonclassical) approaches to decision analysis only by combining the decomposition of utilities into factors with fixed assumptions about the independence of the values assigned to different factors and with defeasible (nonmonotonic) assumptions about the magnitudes of these values. The papers mentioned earlier explore this idea; perhaps the most ambitious application of this idea appears in Yoav Shoham's work. At the moment, much of this work remains unpublished, but it gives reason to hope that a formalism for decision theory might emerge from it that extends the Bayesian network representation of the structure of probability distributions to represent the structure of utility as well.

Possibilistic logic is an extension of possibility theory (Dubois and Prade 1986), a theory inspired in part by fuzzy logic. Possibility theory takes utilities into account, as well as probabilities, but provides a qualitative approach to decision problems by replacing numeric preferences and probabilities with linear orderings. To obtain nontrivial recommendations of actions and useful algorithms, additional assumptions need to be made: the scales for preferences and belief need to be interrelated, and a decision-making policy needs to be adopted, for example, a minimax policy. For discussion of this approach to utilities, see Dubois, Godo, et al. (1998); Dubois, Le Berre, et al. (1998); and Dubois, Prade, and Sabbadin (1997).

Properties of Decision Formulations

The decision-analytic process of formulating decisions naturally involves analyzing a number of properties of tentative formulations, such as stability under perturbations and agreement with intuitions. The set of properties grows significantly when formulations can involve partial information about alternatives, outcomes, probabilities, and utilities.

The most obvious properties of interest regarding probabilities and utilities concern the logical consistency and completeness of the partial specifications, especially when we think of the partial specifications as axioms that partially characterize the ultimate proba-

bility and utility measures. Decision analysis already involves taking new information provided by the decision maker being modeled and checking this information for consistency with past answers and for completeness to provide comparisons of all alternatives. The setting of partial formulations complicates this task here. In the usual setting, one easily ensures the comparability of all alternatives by simply hypothesizing a multiattribute utility function combining subutility functions over the different attributes of alternative actions. With partial formulations, one deals with families of utility functions that apparently force the reasoning required for user modeling to depend on special assumptions, especially assumptions concerning independence, making it crucial to identify plausible assumptions of this kind and analyze their relationship to the reasoning. See Boutilier et al. (1997) and Ha and Haddawy (1998b, 1997) for explicit investigations in this spirit of incremental preference representation methods.

The extra analysis induced by partial formulations is not a mere artifact of the formalism because one can view it as a more direct way of performing the standard perturbation analysis of decision analysis. One advantage of starting with qualitative formulations is that they force the analyst to capture the stable, qualitative information about comparisons directly, only proceeding to detailed quantitative comparisons once the stable core has been identified.

The most obvious properties of alternatives for analysis concern their relations to each other. Standard decision analysis already considers aspects of this issue in Savage's (1972) notion of "small worlds," that is, in choosing how one individuates alternatives and outcomes to avoid unnecessary complexity and to reduce the complexity to intelligible levels. The knowledge-structuring techniques of AI can help ensure intelligibility even when the formulation must contain great numbers of alternatives and outcomes. These techniques include the standard multiattribute techniques, in which one represents alternatives and outcomes as sets of attribute values, but also include methods for representing abstract attributes and taxonomic or other relations among these abstractions and the most concrete attributes.

The true power of abstractions of alternatives and outcomes comes to the fore when one considers similarity or distance relations among alternatives and outcomes that permit the decision analyst to group these objects into qualitatively different classes. These groupings satisfy one of the principal needs of decision-support systems—the need to provide a human deci-

sion maker with a set of meaningful options rather than an incomprehensibly large number of alternatives, most of which represent unimportant variations on each other. See Young (1997) for a detailed treatment of this issue, which uses similarity of utility as a grouping criterion. Similarity orderings play a prime role in case-based reasoning. It might prove very interesting to study the relation between similarity relations useful for decision formulation and those useful for case-based decision making, as in Chajewska et al. (1998) and Ha and Haddawy (1998b). Similarity orderings among alternatives also suggest other interesting decision-theoretic notions, such as formalizing notions of qualitatively optimal alternatives or probably approximately optimal decisions.

Reasons and Explanations

People tend to recommend, explain, criticize, and excuse decisions in terms of generic preferences (for example, "He bought the Jaguar because he likes sports cars"). This tendency makes generic preferences important for communicating about decisions as well as for coming to decisions on the basis of partial information. The problem of formalizing generic preferences bears close connections to the problem of formalizing obligations, especially *conditional obligations.* Recently, the logic of conditional obligation has even been used as a basis for formalizing generic statements in natural language (see Asher and Morreau [1995]).

Conditional obligation forms the subject of a long-standing and increasingly sophisticated tradition in modal logic, begun in the first half of this century by Georg H. von Wright (1963). The more recent forms of this study draw on ideas from nonmonotonic logic (see Horty [1994, 1993]). The AI literature offers a corresponding argument-based approach to decision making in which reasoning identifies explicit reasons for alternatives, generic preferences, and other elements of decision formulations. These explicit reasons themselves can enter into reasoning about the decision formulation through comparisons of reasons with respect to importance measures and through a variety of methods for making decisions on the basis of the arguments represented by the reasons. Some papers in this tradition include Doyle (1980), Thomason (1993), Fox and Parsons (1997), van der Torre (1997), and van der Torre and Tan (1997).

Studies of argument-based decision making have also taken the opposite tack, using preferences to formalize the meanings of reasons and arguments, as in Doyle (1983), Shoham (1988), Doyle and Wellman (1991), and Doyle (1994).

Revision of Preferences

Developing a satisfactory theory of decision capable of providing an account of rational preference revision—and, more generally, rational guidelines for decision-theoretic reasoning about novel circumstances—requires much foundational work.

The scarcity of work devoted to the rational revision of preferences seems remarkable in view of the attention that has been paid to rational belief revision and the many analogies between belief and preference (but see, for example, Jeffrey [1977]; Doyle [1990, 1983, 1980]; Georgeff and Rao [1995]; Shapiro, Lespérance, and Levesque [1995]; and Bell and Huang [1997]).

To some extent, lack of attention to preference revision might reflect a long-standing attitude in economics that preferences change very slowly compared to beliefs as well as a more general empiricist conception that beliefs and belief changes have something like a rational (logical) basis in experience but that desires and preferences stem from irrational or physiological processes. Such attitudes dissipate quickly as the topic turns to planning and acting in novel circumstances, which forms a requirement for all but the most circumscribed robots or human proxies. Novel circumstances call for creating preferences prior to revising or refining them. For example, on finding oneself living in or traveling through a strange town, it is necessary to reinvent the preferences needed for everyday life. Daily shopping decisions, for example, require standing preferences about where to shop and what brands to buy. Shopping preferences from one's previous life do not apply automatically to the new circumstances of different stores and different selections of goods on the shelves.

In such cases, theoreticians often refer to high-level standing goals, such as the goal to lead a good life. However, such references rarely say anything about the reasoning that could lead from such uninformative general goals to the specific preferences needed for daily activities. Generic preferences offer a more natural and more informative starting point for such reasoning. A generic preference, for example, for food stores that minimize driving time over those that minimize cost, could provide the sort of specific guidance needed for shopping decisions. Of course, once created in this fashion, specific preferences might require revision in light of experience. Indeed, if one interprets goals as generic preferences along the lines of Wellman and Doyle (1991) and Doyle, Shoham, and Wellman (1991), then standard deliberative methods such as sub-

Developing a satisfactory theory of decision capable of providing an account of rational preference revision— and, more generally, rational guidelines for decision- theoretic reasoning about novel circumstances —requires much foundational work.

goaling constitute reasoned adoption and abandonment of preferences. Ideas from case-based planning can prove useful in providing more complex forms of preference revision, although work in this area has not, as far as we know, dealt with the problem of goal modification (see, for instance, Hanks and Weld [1992] and Koehler [1996]).

Practicable Qualitative Decision-Making Procedures

An adequate qualitative decision theory should go beyond the needs of formalizing a fairly broad spectrum of decision problems to provide decision-making methods that obtain solutions in many cases without numeric calculation. Some qualitative decision-making procedures have long standing in the literature, but others might await discovery.

Dominance relationships provide an important and common type of qualitative decision-making method. Speaking informally, a plan *p* *dominates* a plan *q* when one can be sure that adopting *p* would yield better outcomes than adopting *q*. Leonard Savage (1972) named the rule of preferring a dominating action *the sure-thing principle*. Michael Wellman's (1990a) work showed how to formalize purely qualitative dominance reasoning and use it profitably in planning. Recent work by Brafman and Tennenholtz (1996) attempts to study such reasoning axiomatically to assess its usefulness as a practical decision-making method.

Work by Thomason and Horty (1996) provides at least one attempt to provide a formal foundation for dominance-based *causal* decision making.[1] It develops a temporal formalism similar to those that have been proposed for nondeterministic, multiagent planning. The models of this system resemble the decision trees of Shafer (1997) but compare outcomes by means of a partial order instead of assigning quantitative utilities to outcomes. In this approach, the preference relation compares *histories,* each of which records the states that ensue as an agent executes the steps of a plan. In the simplest case, one can identify a history with the sequence of actions that make up the plan. In the nondeterministic case, a history also records acts of nature and other agents.

By making independence assumptions concerning the actions of different agents, one can define counterfactual selection functions in these models and consider perturbations of a history *h* in which a planning agent follows a plan *p* to obtain a set of *maximally close* histories in which the agent counterfactually follows another plan *q*. (Think of these perturba-tions as histories in which other agents act as they do in *h*.) These counterfactuals enable one to make sense of the phrase "would yield better outcomes" in the characterization of dominance. Formally, one can extend the preference relation over outcomes to a dominance relation over histories and prove the soundness of this notion of dominance, that is, that one can soundly embed any qualitative dominance model in a quantitative decision-theoretic model. The resulting quantitative theory represents an alternative version of decision theory that, although attracting some attention in the foundational decision-theoretic literature, appears unfamiliar to most computer scientists. For information about this causal version of decision theory, see Armendt (1988), Gibbard and Harper (1978), Eels (1982), and Lewis (1981).

Agent Modeling

Formalized decision-making information should not exist in a vacuum but should be integrated in coherent ways into agent models developed in the broader fields of AI and distributed systems. The distributed agent models of Brafman and Tennenholtz (1996) provide an example of such integration at a theoretical level, and the market-oriented programming methodology of Wellman (1996) provides some practical foundations for such integration. Many practical agent architectures include elements of decision-making information, but a full integration tends to be lacking.

Hybrid Representation and Reasoning

One need not limit qualitative decision-making techniques to explicit reasoning with qualitative information. Such reasoning can adequately address some tasks, but one should not expect purely qualitative information to settle the trade-off questions likely to be involved in the majority of practical decisions.

More generally, one should expect the need to represent every sort of decision-making information in different forms for different computational purposes, even within the same decision-making process. In the case of preferences, generic representations might provide the best basis for the reasoning involved in elicitation, revision, and reuse of preferences, but numeric re-representations of this same information might provide the best basis for computing expected utilities. These comparative advantages call for integrating qualitative and quantitative information even when the mechanisms one obtains in the end for mak-

ing actual decisions do not differ in computational efficiency from hand-crafted quantitative mechanisms.

Qualitative Specification of Quantitative Models

Even when quantitative trade-offs prove necessary, the decision-making process can benefit from indirect use of the qualitative information. More specifically, the generic information might provide constraints that guide selection or construction of a quantitative utility measure. This approach draws on a lesson learned from the study of qualitative reasoning about physical systems (Berleant and Kuipers 1997; Sacks and Doyle 1992).

In this approach, the decision maker would perform most reasoning using ordinary quantitative utility functions in ways well known from decision analysis, especially in computing optimal choices and assessing the sensitivity of these choices to variations in probabilities. When this analysis reveals a deficiency in the current utility model, the analyst formulates the deficiency as a change in the qualitative specification of the utility function and uses the updated specification to compute a new utility function, either directly or as a modification to the previous one.

This approach could offer some attractive benefits: obtaining the efficient computations possible with numeric utility representations, yet preserving the intelligibility and easy modifiability of the qualitative constraints underlying the raw numbers. Making this idea work requires methods for constructing utility functions from generic constraints on preferences, a topic hardly explored to date.

Graphical Representations

Graphical representations of probability distributions, such as *Bayesian networks,* constitute perhaps the best-known hybrid representation. Bayesian networks simplify the specification of many useful probability distributions by using a network structure to encode relationships of probabilistic dependence and independence. The network specifies direct probabilistic dependence explicitly and probabilistic independence relationships implicitly. Each node in the network bears quantitative annotations that indicate the conditional probabilities of the node given the probabilities of the nodes on which it depends. Details of these representations can be found in many sources, for example, in Pearl (1988) and the article by Bruce D'Ambrosio (1999), also in this issue.

The success of Bayesian networks in representing probability information suggests extending these graphical methods to represent preference and utility information as well as to model and implement reasoning about expected utility. The most direct such extensions include *influence diagrams* (Howard and Matheson 1984), which add "value" nodes to Bayesian networks to indicate the dependence of utility on conditions described in the network, and the closely related representations found in Shafer's (1997) decision trees. Influence diagrams, however, constitute hybrid qualitative-quantitative representations of utility only in a weak sense because they only represent concrete utility functions, not qualitative or generic information about utility or preference. The more recent developments discussed earlier (Bacchus and Grove 1997, 1996, 1995; Moorhouse and McKee 1997; Shoham 1997a, 1997b, 1997c; Wellman and Doyle 1992; Wellman 1990a) specifically describing the structure of preference and utility functions do a better job of pointing the way toward hybrid representations for a qualitative decision theory.

Decision Making in Context

Other issues for investigation relate to the roles and forms of qualitative decision-theoretic representations and reasoning as they enter into planning, collaboration, and learning of decision-making information. The obvious roles here concern the decisions made in the course of planning, collaboration, or learning, but the more important questions for the development of qualitative decision theory concern the identification of requirements or constraints placed on decision-theoretic representations by the informational and computational needs of these processes.

For example, several works have explored the role of utility in guiding learning, both in the elementary case of search and in higher-level processes of learning concepts and procedures (see, for example, Russell and Wefald [1991]; Doyle [1988]; Minton [1990]; and Gatch and DeJong [1996]). Are there specific types of preference especially useful in guiding learning? Going the other direction, how can one learn or adapt different types of generic preference and probability information through experience gained during the formulation process or retrospectively from the decision results?

We cannot provide a full discussion of these contextual issues here. Instead, we provide brief discussions of how they arise in planning and collaboration.

More generally, one should expect the need to represent every sort of decision-making information in different forms for different computational purposes, even within the same decision-making process.

Decision-Theoretic Planning

Decision-theoretic planning has already grown into a substantial topic in its own right, as can be seen from the companion article on this topic (see the Blythe [1999] article, also in this issue). Much attention concerns how to structure the state spaces involved to make the problem feasible and how probability and utility information enter into planning decisions. In addition, growing attention has been directed toward determining the representations of probability and preference information most appropriate for planning deliberations. We have already mentioned Wellman's (1990a) work on dominance-based planning, which used qualitative probabilistic networks. A genuine departure from this work would seek to use qualitative or generic preference information as well (see, for example, Tan and Pearl [1994a, 1994b]). A more specific question is whether one can obtain such constraints by analyzing the particular objectives and underlying assumptions of a planning task. Complementary issues concerning identification of qualitatively different alternatives have already been discussed.

Another rich topic concerns how one determines the preferences and beliefs appropriate for specific planning tasks. The most familiar methods involve ordinary debriefing of experts or informants, as practiced by decision analysts. Less familiar but potentially rewarding techniques seek to infer such information from plans. Specifically, what information about the beliefs and preferences of the planner can one obtain from the subgoal relationships of a plan, from temporal ordering, or from other structural information? Can one infer when the satisfaction of one subgoal is more important than the satisfaction of another? How much do such inferences depend on the background knowledge used in constructing the plan?

Sympathetic Planning

Successful collaboration depends on identifying and adopting preferences and other information adequate to make decisions on behalf of one's collaborators or to further their own or mutual objectives. One of the most common collaborative settings is that of advisory systems, which collaborate with their users to help solve the user's problems.

For example, in domains involving elements of uncertainty and risk, a user can seek the advice of an automated planning expert, not about alternative outcomes, which the user might understand full well, but about alternative courses of action or plans for obtaining desired outcomes. The user can hope to benefit from the computer's enhanced ability to search large plan spaces and perhaps from its greater domain knowledge. It goes without saying that the advice will not be appropriate unless it is based on an accurate model of the user's preferences. The task of the system then is one of decision-theoretic planning, in which the system holds beliefs (including probabilistic beliefs) constant and adjusts the utilities to accommodate the user. Thomason (1998) has called this task *sympathetic* planning, drawing on the common connotation of the word sympathy as feeling what another person feels, if only in a diluted way. As mentioned earlier, even if one uses a fully quantitative version of decision theory to evaluate the expected utility of plans, the need to elicit and explain models of preferences in terms familiar and convenient to the advisee will require a qualitative approach to representing utility. Common-sense reasoning and communication about decisions seem to treat risk as an independent ingredient ("I would like to spend less money on a trip, but I think travel by car is too unsafe this time of year") rather than integrate it into the alternatives by treating lotteries as the objects of preferences. Savage's representation of quantitative utilities in terms of preferences over lotteries is, of course, a fundamental theoretical idea but is not well suited to eliciting user preferences in task-oriented interviews.

Sympathetic planning tasks, as illustrated by the challenging example presented early in this article, exercise many of the representational and reasoning abilities desired of a qualitative approach to decision making, including formation of representations of the preferences of another, generation and evaluation of alternative actions, and revision of preferences. Sympathetic planning seems a natural and potentially useful extension of current planning systems, offering the challenge of integrating the planning with a version of decision theory that is able to cope with domains that involve uncertainty, substantial variation in the utilities of outcomes, and users who might differ substantially in their needs and preferences. Recent research trends indicate growing interest in the problem of sympathetic planning (see Ha and Haddawy [1998b, 1997b]; Boutiler et al. [1997]; and Boutilier, Brafman, and Geib [1997]), although to our knowledge, almost nobody develops advisory systems that use decision theory as the foundation of their advice.

Conclusion

Formal studies of decision making have their origins in the seventeenth century, reaching a point of maturity in the mid and late 1940s with reasonably solid mathematical foundations and reasonably practical quantitative methods. Just as it attained this maturity, however, Herbert Simon and others led an exodus from the new theory, charging it with requiring excessive information, memory, and reasoning power. The critique hardly budged the main part of decision theory but led to exploration and development of the notion of qualitative goals (generalizing Simon's utility aspiration levels), formalization of goal-based problem solving, and the modern field of AI. Work on reasoned deliberation then proceeded largely oblivious to the notions of economic decision theory, facing instead large hurdles posed by fundamental representational issues and inadequate computational tools. The new areas eventually developed to the point where researchers realized the need to reconnect the methods of AI with the qualitative foundations and quantitative methods of economics.

Although the field of qualitative decision theory benefits from the substantial contributions of economic decision analysis and AI, many central questions of interest remain unanswered by the relatively small literature produced by recent studies, to the point where this brief survey constitutes more an indication of directions for future work than a presentation of answered questions. Although the existing literature reveals an emerging field with some promising ideas and potentially important applications, a large gap still separates the ideas from workable applications. Development of foundations for qualitative decision theory might require as much effort as the earlier development of foundations for quantitative decision theory. One might identify the most important trends as grappling in one way or another with the challenge of reworking a field that provides powerful theoretical arguments for representations of preferences that are not at all commonsensical and that can be difficult to elicit. How to emerge from this foundational stage with an apparatus that will integrate with problem-solving applications remains unclear at present. Lack of a single dominant approach also complicates matters; different people have quite different ideas about how to proceed. It might take a while for a few dominant paradigms to emerge, but we fully expect further studies to yield expansions of decision theory that serve applications well.

Note

1. The statement and proof of the soundness theorem in Section 7 of Thomason and Horty (1996) are flawed. See www.eecs.umich.edu/~rthomaso/dominance.html for a corrected version.

References

Armendt, B. 1988. Conditional Preference and Causal Expected Utility. In *Causation in Decision, Belief Change, and Statistics, Volume 2,* eds. W. L. Harper and B. Skyrms, 3–24. Dordrecht, The Netherlands: Kluwer Academic.

Arrow, K. J. 1959. Rational Choice Functions and Orderings. *Econometrica* 26:121–127.

Asher, N., and Morreau, M. 1995. What Some Generic Sentences Mean. In *The Generic Book,* eds. G. Carlson and F. J. Pelletier, 300–338. Chicago: Chicago University Press.

Bacchus, F., and Grove, A. 1997. Independence in Qualitative Decision Theory. Paper Presented at the AAAI Spring Symposium on Qualitative Preferences in Deliberation and Practical Reasoning, 24–26 March, Stanford, California.

Bacchus, F., and Grove, A. 1996. Utility Independence in Qualitative Decision Theory. In *KR'96: Principles of Knowledge Representation and Reasoning,* eds. L. C. Aiello, J. Doyle, and S. Shapiro, 542–552. San Francisco, Calif.: Morgan Kaufmann.

Bacchus, F., and Grove, A. 1995. Graphical Models for Preference and Utility. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence,* 3–10. San Francisco, Calif.: Morgan Kaufmann.

Bacchus, F.; Grove, A. J.; Halpern, J. Y.; and Koller, D. 1996. From Statistical Knowledge Bases to Degrees of Belief. *Artificial Intelligence* 87(1–2): 75–143.

Bell, J., and Huang, Z. 1997. Dynamic Goal Hierarchies. Paper Presented at the AAAI Spring Symposium on Qualitative Preferences in Deliberation and Practical Reasoning, 24–26 March, Stanford, California.

Berleant, D., and Kuipers, B. J. 1997. Qualitative and Quantitative Simulation: Bridging the Gap. *Artificial Intelligence* 95(2): 215–255.

Although the existing literature reveals an emerging field with some promising ideas and potentially important applications, a large gap still separates the ideas from workable applications.

Blythe, J. 1999. Decision-Theoretic Planning. *AI Magazine* 20(2). Forthcoming.

Boutilier, C. 1994. Defeasible Preferences and Goal Derivation, Department of Computer Science, University of British Columbia. Available at www.cs.ubc.ca/spider/cebly/craig.html.

Boutilier, C.; Brafman, R.; and Geib, C. 1997. Prioritized Goal Decomposition of Markov Decision Processes: Towards a Synthesis of Classical and Decision-Theoretic Planning. In Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence, ed. M. Pollack, 1156–1163. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.

Boutilier, C.; Brafman, R.; Geib, C.; and Poole, D. 1997. A Constraint-Based Approach to Preference Elicitation and Decision Making. Paper presented at the AAAI Spring Symposium on Qualitative Preferences in Deliberation and Practical Reasoning, 24–26 March, Stanford, California.

Boutilier, C.; Dean, T.; and Hanks, S. 1999. Decision-Theoretic Planning: Structural Assumptions and Computational Leverage. *Journal of Artificial Intelligence Research* 10. Forthcoming.

Brafman, R. I., and Tennenholtz, M. 1996. On the Axiomatization of Qualitative Decision Criteria. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence and the Ninth Innovative Applications of Artificial Intelligence Conference, Volume 1,* eds. H. Shrobe and T. Senator, 76–81. Menlo Park, Calif.: AAAI Press.

Chajewska, U.; Getoor, L.; Norman, J.; and Shahar, Y. 1998. Utility Elicitation as a Classification Problem. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence,* 79–88. San Francisco, Calif.: Morgan Kaufmann.

D'Ambrosio, B. 1999. Inference in Bayesian Networks. *AI Magazine* 20(2). Forthcoming.

Dean, T., and Wellman, M. P. 1989. On the Value of Goals. In Proceedings of the Rochester Planning Workshop: From Formal Systems to Practical Systems, eds. J. Tenenberg, J. Weber, and J. Allen, 129–140. Technical report, 284, Computer Science Department, University of Rochester.

Doyle, J. 1995. Contextual Equivalence and Ceteris Paribus Comparatives. Unpublished manuscript, Laboratory for Computer Science, Massachusetts Institute of Technology. Available at www.medg.lcs.mit.edu/doyle.

Doyle, J. 1994. Reasoned Assumptions and Rational Psychology. *Fundamenta Informaticae* 20(1–3): 35–73.

Doyle, J. 1992. Rationality and Its Roles in Reasoning. *Computational Intelligence* 8(2): 376–409.

Doyle, J. 1991. Rational Belief Revision. In *KR'91: Principles of Knowledge Representation and Reasoning,* eds. J. F. Allen, R. Fikes, and E. Sandewall, 163–174. San Francisco, Calif.: Morgan Kaufmann.

Doyle, J. 1990. Rational Belief Revision. Paper presented at the Third International Workshop on Nonmonotonic Reasoning, 31 May – 3 June, Stanford, California. Available at www.medg.lcs.mit.edu/doyle/ftp/rbr90.ps.

Doyle, J. 1988. On Rationality and Learning. Techni-

cal report, CS-88-122, Computer Science Department, Carnegie-Mellon University.

Doyle, J. 1983. Some Theories of Reasoned Assumptions. Technical report, CMU-CS-83-125, Department of Computer Science, Carnegie Mellon University.

Doyle, J. 1980. A Model for Deliberation, Action, and Introspection. Technical report, AI TR 581, Artificial Intelligence Laboratory, Massachusetts Institute of Technology.

Doyle, J., and Wellman, M. P. 1994. Representing Preferences as Ceteris Paribus Comparatives. In Decision-Theoretic Planning, 69–75. Technical report SS-94-06, Menlo Park, Calif.: American Association for Artificial Intelligence.

Doyle, J., and Wellman, M. P. 1991. Impediments to Universal Preference-Based Default Theories. *Artificial Intelligence* 49(1–3): 97.

Doyle, J.; Shoham, Y.; and Wellman, M. P. 1991. A Logic of Relative Desire (Preliminary report). In *Proceedings of the Sixth International Symposium on Methodologies for Intelligent Systems,* ed. Z. Ras, 16–31. Lecture Notes in Computer Science. Berlin: Springer-Verlag.

Draper, D.; Hanks, S.; and Weld, D. 1994. Probabilistic Planning with Information Gathering and Contingent Execution. In *Proceedings of the Second International Conference on AI Planning Systems,* ed. K. Hammond, 31–36. Menlo Park, Calif.: AAAI Press.

Dubois, D., and Prade, H. 1986. *Possibility Theory: An Approach to Computerized Processing of Uncertainty.* New York: Plenum.

Dubois, D.; Prade, H.; and Sabbadin, R. 1997. A Possibilistic Logic Machinery for Qualitative Decision. Paper presented at the AAAI Spring Symposium on Qualitative Preferences in Deliberation and Practical Reasoning, 24–26 March, Stanford, California.

Dubois, D.; Godo, L.; Prade, H.; and Zapico, A. 1998. Making Decision in a Qualitative Setting: From Decision under Uncertainty to Case-Based Decision. In *KR'98: Principles of Knowledge Representation and Reasoning,* eds. A. G. Cohn, L. Schubert, and S. C. Shapiro, 594–605. San Francisco, Calif.: Morgan Kaufmann.

Dubois, D.; Le Berre, D.; Prade, H.; and Sabbadin, R. 1998. Logical Representation and Computation of Optimal Decisions in a Qualitative Setting. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence,* 588–593. Menlo Park, Calif.: AAAI Press.

Eels, E. 1982. *Rational Decision and Causality.* Cambridge, U.K.: Cambridge University Press.

Fishburn, P. C. 1977. Multiattribute Utilities in Expected Utility Theory. In *Conflicting Objectives in Decisions,* eds. D. E. Bell, R. L. Keeney, and H. Raiffa, 172–196. New York: Wiley.

Fox, J., and Parsons, S. 1997. On Using Arguments for Reasoning about Actions and Values. Paper presented at the AAAI Spring Symposium on Qualitative Preferences in Deliberation and Practical Reasoning, 24–26 March, Stanford, California.

Georgeff, M., and Rao, A. 1995. The Semantics of Intention Maintenance for Rational Agents. In Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, ed. C. Mellish, 704–710. Menlo Park, Calif.: International Joint

Conferences on Artificial Intelligence.

Gibbard, A. F., and Harper, W. L. 1978. Counterfactuals and Two Kinds of Expected Utility. In *Foundations and Applications of Decision Theory,* eds. C. Hooker, J. Leach, and E. McClennen, 125–162. Dordrecht, The Netherlands: D. Reidel.

Gratch, J., and DeJong, G. 1996. A Statistical Approach to Adaptive Problem Solving. *Artificial Intelligence* 88(1–2): 101–142.

Ha, V., and Haddawy, P. 1998a. Geometric Foundations for Interval-Based Probabilities. In *KR'98: Principles of Knowledge Representation and Reasoning,* eds. A. G. Cohn, L. Schubert, and S. C. Shapiro, 582–593. San Francisco, Calif.: Morgan Kaufmann.

Ha, V., and Haddawy, P. 1998b. Toward Case-Based Preference Elicitation: Similarity Measures on Preference Structures. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence,* 193–201. San Francisco, Calif.: Morgan Kaufmann.

Ha, V., and Haddawy, P. 1997. Problem-Focused Incremental Elicitation of Multi-Attribute Utility Models. In *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence (UAI-97),* 215–222. San Francisco, Calif.: Morgan Kaufmann.

Haddawy, P. 1991. A Temporal Probability Logic for Representing Actions. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference,* 313–324. San Francisco, Calif.: Morgan Kaufmann.

Haddawy, P., and Hanks, S. 1998. Utility Models for Goal-Directed Decision-Theoretic Planners. *Computational Intelligence* 14(3): 392–429.

Haddawy, P., and Hanks, S. 1992. Representations of Decision-Theoretic Planning: Utility Functions for Deadline Goals. In *KR'92. Principles of Knowledge Representation and Reasoning: Proceedings of the Third International Conference,* eds. B. Nebel, C. Rich, and W. Swartout, 71–82. San Francisco, Calif.: Morgan Kaufmann.

Haddawy, P., and Hanks, S. 1990. Issues in Decision-Theoretic Planning: Symbolic Goals and Numeric Utilities. In *DARPA Workshop on Innovative Applications to Planning, Scheduling, and Control,* 48–58. San Francisco, Calif.: Morgan Kaufmann.

Halpern, J. Y. 1990. An Analysis of First-Order Logics of Probability. *Artificial Intelligence* 46(3): 311–350.

Hank, S., and Weld, D. 1992. Systematic Adaptation for Case-Based Planning. In *Proceedings of the First International Conference on Artificial Intelligence Planning Systems,* 96–105. San Francisco, Calif.: Morgan Kaufmann.

Hansson, B. 1990. Preference-Based Deontic Logic. *Journal of Philosophical Logic* 19(1): 75–93.

Hansson, S. O. 1996. What Is Ceteris Paribus Preference? *Journal of Philosophical Logic* 25(3): 307–332.

Hare, R. M. 1963. *Freedom and Reason.* Oxford, U.K.: Oxford University Press.

Horty, J. F. 1994. Moral Dilemmas and Nonmonotonic Logic. *Journal of Philosophical Logic* 23(1): 35–65.

Horty, J. F. 1993. Deontic Logic as Founded on Nonmonotonic Logic. *Annals of Mathematics and Artificial Intelligence* 9:69–91.

Howard, R. A. 1966. Information Value Theory. *IEEE Transactions on Systems Science and Cybernetics* SSC-2:22–26.

Howard, R. A., and Matheson, J. E. 1984. Influence Diagrams. In *The Principles and Applications of Decision Analysis, Volume 2,* eds. R. A. Howard and J. E. Matheson, 719–762. Menlo Park, Calif.: Strategic Decisions Group.

Jeffrey, R. C. 1977. A Note on the Kinematics of Preference. *Erkenntnis* 11:135–141.

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

Koehler, J. 1996. Planning from Second Principles. *Artificial Intelligence* 87(1–2): 145–146.

Lewis, D. K. 1981. Causal Decision Theory. *Australasian Journal of Philosophy* 59(1): 5–30.

Linden, G.; Hanks, S.; and Lesh, N. 1997. Interactive Assessment of User Preference Models: The Automated Travel Assistant. In Proceedings of User Modeling '97. Available at www.cs.washington.edu/homes/hanks/Papers/index.html.

Mantha, S. 1994. Towards a Logical Foundation for Qualitative Decision Theory. Paper presented at the Symposium on Decision-Theoretic Planning, 21–23 March, Stanford, California.

Minton, S. 1990. Quantitative Results Concerning the Utility of Explanation-Based Learning. *Artificial Intelligence* 42(2–3): 363–391.

Moorhouse, J., and McKee, G. 1997. Modeling Evaluative Judgments. Paper presented at the AAAI Spring Symposium on Qualitative Preferences in Deliberation and Practical Reasoning, 24–26 March, Stanford, California.

Packard, D. J. 1975. A Preference Logic Minimally Complete for Expected Utility Maximization. *Journal of Philosophical Logic* 4(2): 223–235.

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

Russell, S. J., and Wefald, E. 1991. *Do the Right Thing: Studies in Limited Rationality.* Cambridge, Mass.: MIT Press.

Sacks, E. P., and Doyle, J. 1992. Prolegomena to Any Future Qualitative Physics. *Computational Intelligence* 8(2): 187–209.

Savage, L. 1972. *The Foundations of Statistics.* 2d ed. New York: Dover.

Shafer, G. 1997. *The Art of Causal Conjecture.* Cambridge, Mass.: MIT Press.

Shapiro, S.; Lespérance, Y.; and Levesque, H. J. 1995. Goals and Rational Action in the Situation Calculus—A Preliminary Report. Paper presented at the AAAI Fall Symposium on Rational Agency: Concepts, Theories, Models, and Applications, Nov. 10-12, Massachusetts Institute of Technology, Cambridge, Mass.

Shoham, Y. 1988. *Reasoning about Change: Time and Causation from the Standpoint of Artificial Intelligence.* Cambridge, Mass.: MIT Press.

Shoham, Y. 1997a. A Mechanism for Reasoning about Utilities (and Probabilities): Preliminary report. Paper presented at the AAAI Spring Sympo-

sium on Qualitative Preferences in Deliberation and Practical Reasoning, 24–26 March, Stanford, California.

Shoham, Y. 1997b. A Symmetric View of Probabilities and Utilities. In Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence, ed. M. Pollack, 1324–1329. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.

Shoham, Y. 1997c. Two Senses of Conditional Utility. In *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence (UAI-97),* 429–436. San Francisco, Calif.: Morgan Kaufmann.

Tan, S.-W., and Pearl, J. 1994a. Qualitative Decision Theory. In *Proceedings of the Twelfth National Conference on Artificial Intelligence,* eds. B. Hayes-Roth and R. Korf, 928–933. Menlo Park, Calif.: AAAI Press.

Tan, S.-W., and Pearl, J. 1994b. Specification and Evaluation of Preferences under Uncertainty. In *KR'94: Principles of Knowledge Representation and Reasoning,* eds. J. Doyle, E. Sandewall, and P. Torasso, 530–539. San Francisco, Calif.: Morgan Kaufmann.

Thomason, R. H. 1998. Qualitative Decision Theory and Interactive Problem Solving (Extended Abstract). In Interactive and Mixed-Initiative Decision-Theoretic Systems, eds. P. Haddawy and S. Hanks, 107–113, Technical report SS-98-03. Menlo Park, Calif.: American Association for Artificial Intelligence.

Thomason, R. H. 1993. Towards a Logical Theory of Practical Reasoning. In Reasoning about Mental States: Formal Theories and Applications, ed. John Horty and Yoav shoham, 133–142. Technical report SS-93-05. Menlo Park, Calif.: American Association for Artificial Intelligence.

Thomason, R. H., and Horty, J. F. 1996. Nondeterministic Action and Dominance: Foundations for Planning and Qualitative Decision. In *Theoretical Aspects of Rationality and Knowledge: Proceedings of the Sixth Conference (TARK 1996),* ed. Y. Shoham, 229–250. San Francisco, Calif.: Morgan Kaufmann.

van der Torre, L. W. N. 1997. *Reasoning about Obligations: Defeasibility in Preference-Based Deontic Logic.* Tinbergen Institute Research Series 140. Amsterdam: Thesis.

van der Torre, L. W. N., and Tan, Y.-H. 1997. Prohairetic Deontic Logic and Qualitative Decision Theory. Paper presented at the AAAI Spring Symposium on Qualitative Preferences in Deliberation and Practical Reasoning, 24–26 March, Stanford, California.

von Wright, G. H. 1963. *Norm and Action.* London: Routledge and Keegan Paul.

Voorbraak, F. 1997. Decision Analysis Using Partial Probability Theory. Paper presented at the AAAI Spring Symposium on Qualitative Preferences in Deliberation and Practical Reasoning, 24–26 March, Stanford, California.

Wellman, M. P. 1996. Market-Oriented Programming: Some Early Lessons. In *Market-Based Control: A Paradigm for Distributed Resource Allocation,* ed. S. Clearwater. River Edge, N.J.: World Scientific.

Wellman, M. P. 1985. Reasoning about Preference Models. Technical report, 340, Laboratory for Computer Science, Massachusetts Institute of Technology.

Wellman, M. P. 1990a. *Formulation of Trade-Offs in Planning under Uncertainty.* San Francisco, Calif.: Morgan Kaufmann.

Wellman, M. P. 1990b. Fundamental Concepts of Qualitative Probabilistic Networks. *Artificial Intelligence* 44:257–304.

Wellman, M. P., and Doyle, J. 1992. Modular Utility Representation for Decision-Theoretic Planning. In *Proceedings of the First International Conference on AI Planning Systems,* 236–242. San Francisco, Calif.: Morgan Kaufmann.

Wellman, M. P., and Doyle, J. 1991. Preferential Semantics for Goals. In *Proceedings of the Ninth National Conference on Artificial Intelligence,* eds. T. Dean and K. McKeown, 698–703. Menlo Park, Calif.: AAAI Press.

Wellman, M. P.; Breese, J. S.; and Goldman, R. P. 1992. From Knowledge Bases to Decision Models. *Knowledge Engineering Review* 7(1): 35–53.

Young, R. M. 1997. Generating Concise Descriptions of Complex Activities. Ph.D. dissertation, Intelligent Systems Program, University of Pittsburgh.

Jon Doyle, a fellow of the American Association for Artificial Intelligence, serves as principal research scientist in the Laboratory for Computer Science at the Massachusetts Institute of Technology (MIT). He received his Ph.D. in AI from MIT in 1980. His current research interests include mathematical foundations for theories of limited rationality, qualitative decision theory, task-oriented information markets, and the history of AI. His e-mail address is doyle@mit.edu.
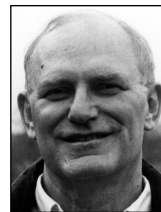
Richmond H. Thomason is a professor of computer science, linguistics, and philosophy at the University of Michigan at Ann Arbor. He received a Ph.D. in philosophy from Yale University in 1965. His current interests in AI include logic and AI, computational models of discourse, and qualitative decision theory. His e-mail address is rich@thomason.org.