Rational Metareasoning and Compilation for Optimizing Decisions Under Bounded Resources

Eric J. Horvitz*
Medical Computer Science Group
Knowledge Systems Laboratory
Stanford University
Stanford, California 94305

Several years ago, our research group initiated a project to investigate the use of decision theory as a framework for reasoning about the design and operation of ideal agents under bounded resources. We have studied metareasoning, reasoning, and compilation within the framework of decision theory. Our model of rationality centers on the use of design-time and tractable run-time decision-theoretic analyses to control the detail and completeness of problem-level decision making. Unlike straightforward decision analyses, we apply the principles of decision theory to enriched models that include not only distinctions and outcomes in the world, but also distinctions and outcomes about cognition. In this paper, we shall review some earlier work on rational metareasoning and describe the benefits of integrating deliberative models of decision-theoretic reasoning and metareasoning with several classes of precomputed or *compiled* actions.

1 Introduction

Much of the work in artificial intelligence can be cast as the development of techniques on which agents with limited representational and inferential abilities can rely for success in the face of complex challenges. Over the last 4 years, investigators on the PROTOS[1] project have been developing techniques for generating and evaluating optimal computational behavior under limitations in reasoning resources. We are interested in decision-theoretic techniques for use in offline design, as well as for explicit application in the deliberative machinery of real-world reasoners.

Our earlier research elucidated the value of decision-theoretic metareasoning in the control of complex object-level problem solving. Here we shall highlight the promise for an integrated approach to rationality centering on the development of agents that make use of a variety

*This work was supported by a NASA Fellowship under Grant NCC-220-51, by the National Science Foundation under Grant IRI-8703710 and by the National Library of Medicine under Grant RO1LM0429. Computing facilities were provided by the SUMEX-AIM Resource under NIH Grant RR-00785.

[1] "Protos" is a partial acronym for *p*roject on computational *r*esources and *t*rade*offs*.

of classes of compiled knowledge in addition to apparatus for reasoning and metareasoning. After reviewing some highlights of our earlier work on partial computation and rational metareasoning, we shall discuss the promise of methods for integrating deliberative models of decision-theoretic reasoning and metareasoning with compiled knowledge. Finally, we shall discuss the role of learning in the context of partial compilation for optimizing an agent's behavior.

2 Rationality Under Resource Constraints

Our research has pursued the development of optimal reasoning strategies and behaviors for computational agents within their environments. That we are studying the optimization of reasoning does not necessarily mean that verifiable optimality is our end-goal. Rather, the endeavor can help us to make explicit the relative performance of alternative reasoning methodologies, and to gain additional insight about how we might go about enhancing the activity of our computational agents. More fundamentally, the pursuit of bounded optimality elicits a number of significant questions and research opportunities regarding ideal and resource-constrained *rationality.*

There are competing views on the nature of rationality. Indeed, rational reasoning and behavior has been a hotly debated topic for centuries. A familiar perspective, held by a great number of researchers in the behavioral and decision sciences, is that rational decisions are those that maximize a numerical measure of preference, termed *utility.* Utility is defined by the axioms of utility theory enumerated by von Neumann and Morgenstern over four decades ago [47]. The axioms of utility and probability comprise decision theory.[2]. From a decision-theoretic perspective, a reasoner that optimizes utility is termed *normative.*

2.1 Bounded Optimality

In 1955, Simon noted that we should consider constraints on cognitive resources in generating and evaluating the behavior of a decision maker in a complex situation. However, he and other early pioneers of symbolic-processing models of intelligence quickly moved away from decision theory. Citing the limited abilities of human decision makers and the forgiving nature of many problems in the world, Simon proposed that most intelligent behavior is oriented toward finding relatively simple solutions that are nonoptimal, yet are sufficient or *satisficing* [44,45]. The primary task was viewed as devising computational procedures that could forward the goals of an intelligent decision maker—and not necessarily remaining consistent with decision theory. This theme has stimulated broad artifical-intelligence research on relatively ill-characterized heuristic procedures in a wide array of domains. Beyond machine-intelligence research, a number of economists and psychologists have also investigated issues related to satisficing behavior under resource constraints [46,38,43]. Thus, bounded-rationality research has come to be associated with the centrality of identifying and recreating heuristic satisficing decision-making behaviors.

[2] A detailed review of past and current efforts to apply decision theory for tackling challenging artificial-intelligence problems is found in [27].

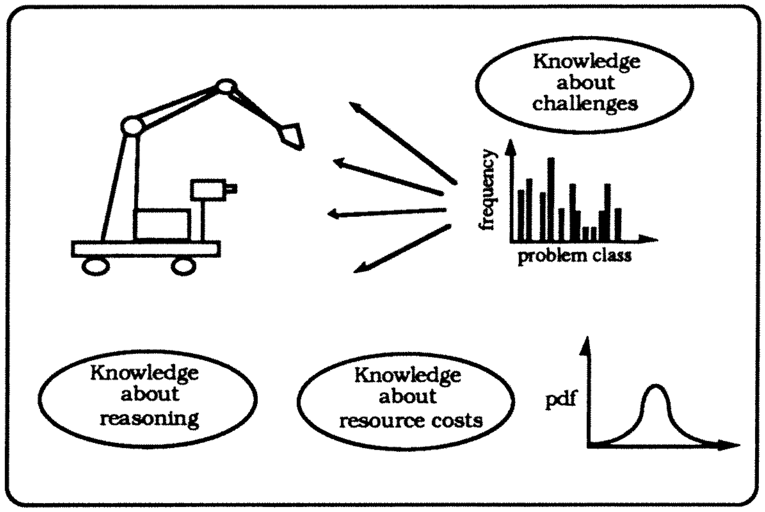

Figure 1: We desire our agents to guide their computation and to act with knowledge about the quality of results expected with increasing quantities of computation, knowledge about reasoning costs, and knowledge about challenges from the environment. Generally, all this knowledge is uncertain, as indicated by the prominence of the probability distribution.

Unfortunately, the nonnormative approaches to problem solving, characterizing the bulk of research under the label of *bounded rationality*, may stray far from the levels of utility that might be achieved through the pursuit of more sophisticated normative analyses. Losses may be especially significant in high-stakes decision making, given complex uncertainties about the world. Such potential losses—and opportunities for great gain—highlight the potential usefulness of decision theory for optimizing the value of behavior under resource constraints.

We use *bounded optimality* to distinguish our work on the optimal design of problem solvers and solution methodologies under bounded resources from traditional nonnormative approaches to reasoning under resource constraints [25]. Bounded-optimal systems perform ideally in some context or set of contexts under expected constraints in reasoning resources, such as in time and memory, given the expected utility structure of a problem. We define such optimality in the context of a decision-theoretic analysis of the costs and benefits of alternative hardware configurations and strategies for determining the best action. As portrayed in Figure 1, we must consider the knowledge about the solution methodologies available, the costs of reasoning resources, such as memory and time, and the expected challenges that will be faced in an environment over some period of time.

In general, we are uncertain about the nature of all of these classes of knowledge. Thus, we must generally reason under uncertainty and use probability distributions (or approximations of such distributions) about the costs and benefits of different reasoning strategies, and of expected challenges from the environment. Other factors include what an agent's uncertainty about future preferences is and how utility assigned to states over time will be weighted and combined in the future.

The determination of true bounded optimality requires proving tight lower-bounds on the solution of problems given the informational and computational constraints at hand. Our pursuit of bounded optimality does not mean we will always be able to prove optimality.

Indeed, bounded optimality may not always be welldefined. However, in the absence of theoretical limits, decision-theoretic meta-analyses can enable us to begin to reason about the relative optimality of agents of different constitutions in partially characterized environments [25,17]. We can also define different *classes* of bounded optimality under broad constraints in the constitution and abilities of reasoners. For example, we may seek and verify *strategic bounded optimality*—optimality given a limited repertoire of available reasoning strategies.

2.2 Epistemic and Autoepistemic Decision Models

The axioms of decision theory define rationality in terms of a *decision model.* Decision models capture knowledge about possible decisions and outcomes, and a state of information about the world in which a reasoner is immersed. Decision theory dictates that a decision-making system or agent should determine ideal actions by performing inference on its decision models. Given a decision model, the goal is to select an action that has the greatest utility.

It is important to note that traditional decision theory has nothing to say about the choice of a rational model. In practice, our agent either must use a predefined model of the world, or dynamically formulate and instantiate a decision model. After a model has been constructed, the agent can perform decision-theoretic inference based on the model.

2.2.1 Models of the Environment

Attempting to apply decision theory in a straightforward manner to detailed representations of the world can produce decisions that would be ideal only in a world of abundant time and memory. Such strategies typically are irrational in the real world of limited resources; the tasks of decision-theoretic model construction and inference are often computationally intensive.

The intuitions of early investigators in artificial intelligence on the inapplicability of implementations of decision theory are supported by recent complexity analyses. Research on the theoretical foundations of the complexity of exact probabilistic inference has demonstrated that automated probabilistic inference is intractable in the worst case [7]. Under constraints on time or on other resources, such as model construction and instantiation, approximations and more poorly characterized heuristic techniques often have a higher expected value than does complete normative reasoning. The potential for suboptimal procedures to dominate "optimal" analyses under resource constraints is easy to demonstrate with a formal analysis [26].

2.2.2 Models of Representation and Inference

Rather than reject the pursuit of a theoretical foundation for ideal belief and action, we have sought to extend coherently the principles of normative rationality to situations of uncertain, varying, and scarce reasoning resources. A central aspect of this research is the development of valuable approximation strategies and the enrichment of the models that are manipulated by decision-theoretic principles. We wish to extend naive decision-theoretic analyses that make use of models that represent only the challenge at hand with autoepistemic distinctions and relationships.

Our research to date has highlighted the usefulness of models that manipulate distinctions about problem solving in addition to knowledge about the world. We are studying rich multilevel models and associated approximation machinery that are promising in their ability to optimize the behavior of agents under static and varying constraints in reasoning resources. That is, we broaden the attention of decision-theoretic analyses beyond representations of the external world to include in our analyses distinctions about reasoning machinery and procedures in the cognitive world.

3 Utility of Agency

Let us first define some basic terminology necessary for delving deeper into rationality under bounded resources. We use *comprehensive value*, u_c, to refer to the utility attributed to the state of an agent in the world. This value is a function of the problem at hand, of the agent's best default action, and of the stakes of a decision problem. We call the net change expected in the comprehensive value, in return for an allocation of some computational resource to reasoning, the expected value of computation (EVC). The comprehensive utility, at any point in the reasoning process, can be viewed typically as a function of two components of utility: the *object-level* utility, u_o, and the *inference-related* cost function, u_i. The *object-level* utility is the expected utility associated with a computer result or state of the world without regard to the cost of reasoning that may be necessary to generate the state. The object-level utility is a function of a vector of attributes, $\vec{v}$. The *inference-related* component is the sum of the expected disutility associated with, or required by, the process of problem solving. This cost typically is the disutility of delaying an action while waiting for a reasoner to infer a recommendation.

In general, the inference-related cost is a function of a vector of resource attributes, $\vec{r}$, representing the quantity that has been expended of such commodities as time and memory. We have examined several classes of functions describing u_i, including the *urgency*, *deadline*, and *urgent-deadline* situations [26]. *Urgency* refers to the general class of inference-related utility functions that assign cost as some monotonically increasing function of delay. The *deadline* pattern refers to cases where $u_i(\vec{r})$ is 0 or insignificant until a certain level of resource is reached. At this point, computation must halt, and the maximum object-level utility attained before the halt must be reported immediately. Otherwise, the result is worthless. The *urgent-deadline* requires consideration of both the cost and availability of time. These cost functions are common in many real-world applications and are based in such universal interactions as lost opportunity and competition for limited resources.

3.1 Partial Computation

A key notion in rational metareasoning is that of partial results generated by partial computation. The notion of partial computation and its role in rationality has been analyzed by our group [25] and by Dean [8]. This work highlighted the notion that theoretical computer-science research has been limited in its focus on the analysis of the difficulty of achieving a well-defined final result [3,12]. Although such an assumption can bring useful simplification to analyses of computational complexity, it biases research toward policies that are indifferent to variation in the utility of a result or to the costs and availability of resources. Investigators working on computational complexity implicitly assume only one of two measures of utility to computational behavior: Either a final solution can be computed, which has maximum object-level utility, or a solution is not found in the time available, and the effort is worthless.

Under bounded resources, we do not generally have enough time to perform the computation required to generate an ideal answer. Approximation strategies or solutions of related problems can generate nonoptimal results for some fraction of the computation required for generating an ideal answer: wide variations in the value of a result to an agent, in the availability of resource, and in the cost of reasoning highlight the limitations of a focus on time complexity for termination on final results. We are interested in reasoning about optimizing quality of a solution in the time available. Thus, it is important to enumerate families of *partial results* for different ideal goals.

3.1.1 Partial Results

Partial results $\pi(I)$ result from applying a sequence of computation steps, or a *strategy*, $S_i(I, \vec{r})$, to a problem instance I, with the expenditure of resource $\vec{r}$. That is,

$$S_i(I, \vec{r}) = \pi(I)$$

The partial results are transformations of desired ideal results $\phi(I)$ along one or more dimensions of utility, where

$$0 \leq u_o[\pi(I)] \leq u_o[\phi(I)]$$

More specifically, the object-level utility function maps a real-valued object-level utility to a vector $\vec{v}$ of attributes in an approximation space $\mathcal{A}_o$ for $\phi(I)$ for each partial result. For the purposes of summarization, it can be useful to define a context-independent distance metric $D : \mathcal{A}_o \times \mathcal{A}_o \rightarrow \mathcal{R}$ between points in this space. We relate the difference in utility of $\phi(I)$ and $\pi(I)$ to a function of the context and this distance. An example of a widely-used, context-independent distance among results is the numerical approximation, where D is a simple unidimensional measure of precision (e.g., the result of a Taylor series carried to a particular term). In this case, $\phi(I)$ and $\pi(I)$ are separated by a distance in the space of reals.

Beyond the familiar numerical approximation, we can consider cases where D represents the divergence of $\pi(I)$ from $\phi(I)$ along higher-dimensional and more abstract properties of a computational result. Identifying poorly understood classes of partial results can highlight directions for research on approximation strategies and reasoning under resource constraints. Dimensions in $\mathcal{A}_o$ often are based on the use made of the result and are rooted in human preferences. Simulation methods partially characterize a probability of interest through probabilistically visiting portions of a problem, yielding a sequence of probability distributions over a set of states with additional computation. Randomized approximation algorithms generate results that take the form of inequalities on error bounds on an ideal result.

To highlight general principles of rational metareasoning and to demonstrate the richness possible in partial computation, we examined the problem of metareasoning about the task of sorting. In this work, multiattribute object-level utility functions were used to assign utility to incompletely sorted file of records, based on several different dimensions or attributes of *sortedness* [21,26]. Attributes explored include *disorder*, a measure of the average distance between current locations and expected final locations for records in a file, and *completeness*, the contiguous length of a file containing records sorted into their final positions. The multiattribute utility structure of alternative algorithms define problem-solving trajectories through multiattribute spaces. We examined how an agent could use knowledge about such trajectories to choose the best strategy to apply, given the cost of reasoning.

3.2 Graceful Degradation

Several properties of reasoning strategies and representations are desirable for reasoning under bounded resources [25]. We desire our strategies to provide us with partial results that have object-level value that increases monotonically with allocated resource. We additionally desire our strategies to be relatively insensitive to small reductions in resource fraction. Thus, we desire continuity in the refinement of partial results. Such *incremental-refinement* policies yield immediate object-level returns on small quantities of invested computation, reducing the risk of dramatic losses in situations of uncertain resource availability. This property is especially important for reasoning under uncertain resource constraints [25,8]. Finally, we desire the partial results to converge to an ideal answer with some finite allocation of resource. A class of reasoning strategies called spanning strategies are particularly useful for reasoning under varying resource constraints. *Spanning* strategies are approximation procedures that exhibit monotonicity, continuity, and convergence on an ideal object-level solution with the allocation of a finite quantity of resource.

In our work on metareasoning about fundamental computational tasks, we highlighted the utility of continuity by analyzing the example of sorting under resource constraints. In this study, we compared faster all-or-nothing $O(N \log N)$ sorting strategies [34] with several slower polynomial strategies. Although the latter strategies take longer to sort completely our test files, they have the ability to refine incrementally one or more object-level attributes of a partial sort. Given a variety of file sizes, prototypical utility models, and resource contexts, the slower polynomial sorting approaches were more valuable than the faster approaches. For example, mergesort, heapsort, and quicksort do not generate valuable intermediate results. If a deadline occurs at some time before completion, the result is useless or is associated with a cost. Under conditions of uncertain or poor resource availability, or of high cost of reasoning, we can generate a more valuable result by applying Shellsort, which is more conservative with a $O(N^{1.5})$ time complexity, yet is more graceful because it refines partial sorts continuously with time [26].

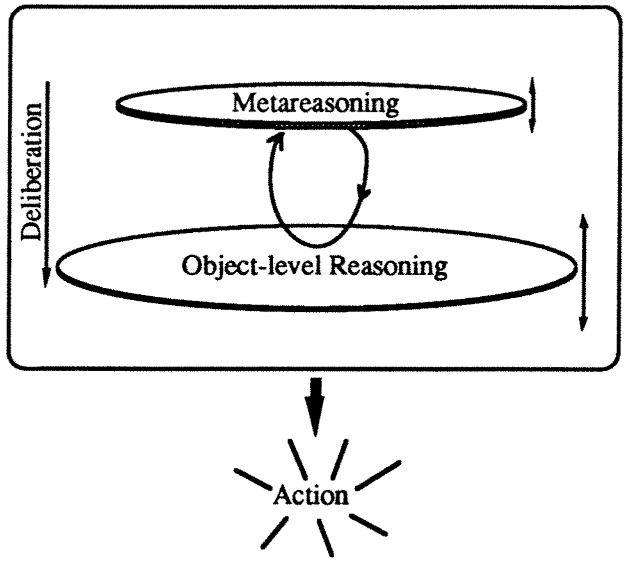

Figure 2: Optimization of the value of computation under uncertain resources and challenges, can necessitate the application of a portion of the available resource to metareasoning about optimal representation and reasoning.

4 Rational Metareasoning

Decision theory provides the foundations for a principled approach to metalevel decision making under the general condition of uncertainty. We shall first examine the general deliberative approach to decision-theoretic reasoning under resource constraints. That is, we shall look at the general notion of optimizing expected utility when we have available alternative or continuously tunable approximations strategies. Then we shall explore the integration of compiled knowledge for reasoning and metareasoning.

We are typically uncertain about the states that will be reached with future computation and about the utility of these states. We approach the problem of rationality under resource constraints by constructing one or more metareasoning problems to represent and control aspects of object-level representation and deliberation. Our goal is to decide which available strategy has the greatest value, given alternative available reasoning methods and possible object-level actions. This approach focuses our attention on the construction of rich multilevel models and on inference within these models.

4.1 Expected Value of Computation

Our goal is to optimize the utility of an agent's decision. To do this, we may have to apply a portion of the available reasoning resource to deliberate about *how* best to optimize utility. A chief component of metareasoning about decision making is determining the expected value of computation of alternative available strategies. We first review the calculation of the EVC.

The initial utility of an agent facing a challenge is $u_o(I)$, where I captures the current problem instance or *state* of the agent in the world. We are interested in the value of applying strategy S_i to the current instance to generate a better state. There is generally uncertainty in the object-level result that will folllow from the expenditure of computational resources on a strategy. Thus, we must consider probability distributions over changes in relevant attributes. These distributions capture our uncertain knowledge about how allocating costly resource will change $\vec{V}(I)$ to $\vec{V}[\pi(I)]$. For brevity, we shall consider a single distribution over partial results and a unidimensional measure of resource, r. The extension to multiple attributes is straightforward. We must apply knowledge of the form

$$S_i(I,r) = p_{S_i}[\pi^j(I)\,|\,r]$$

where p_{S_i} is a probability distribution over different possible partial results $\pi^j(I)$, conditioned on the allocation of resource r using strategy S_i.

The EVC is the change in comprehensive utility, u_c, associated with the generation of a result. If we assume that inference-related and object-level utilities can be decomposed, and are related through addition, the EVC is just the difference between the increase in object-level utility and the cost of the additional computation. Let us first consider the EVC without considering the cost of metareasoning itself. The EVC of applying strategy S_i with a quantity of resource r, is

$$\begin{aligned} EVC(S_i,I,r) &= u_c(S_i,I,r) - u_o(I) \\ &= \int_j u_o[\pi^j(I)] \times p_{S_i}[\pi^j(I)\,|\,r] - u_o(I) - u_i(r) \end{aligned} \tag{1}$$

In complex problems, we may be uncertain about the nature of the functions u_o and u_i used to map a object-level utility to attributes of partial results and disutility to resource expenditures. In such cases, we can extend this formula to sum over different weighted combination functions in addition to considering the uncertainty over different attributes of partial results.

So far, we have included the cost of reasoning about the solution of problem instance I, but we have not included the cost of metareasoning in our formula for computing the EVC. In practice, our goal is to minimize the cost of metalevel computation by elucidating tractable closed-form solutions to EVC estimation. We have found that it is possible to apply knowledge about properties of reasoning strategies and cost functions to develop a tractable economics of computation [28]. For example, consider the curves in Figure 4, which characterize the economics of partial computation. If we know the functions that describe different available spanning strategies, and know that the second derivative of the cost of computation is everywhere nonnegative, we know that EVC is equal to 0 and the comprehensive value of computation is maximized when the first derivative of the object-level utility function is equal to the first derivative of the cost function [23].

As tractable as such EVC analyses may be, however, metareasoning requires an additional allocation of resource; a metareasoner must diminish the resource available for object-level computation by the expected cost of metareasoning. To make explicit the inclusion of metareasoning costs, we denote by EVCM the expected value of computation including the costs of metareasoning. Assume that $r^{\mathcal{M}}$ is a fixed cost of metareasoning. The EVCM is

$$EVCM(S_i,I,\mathcal{M},r) = \int_j u_o[\pi^j(I)] \times p_{S_i}[\pi^j(I)\,|\,(r - r^{\mathcal{M}})] - u_o(I) - u_i(r) \tag{2}$$

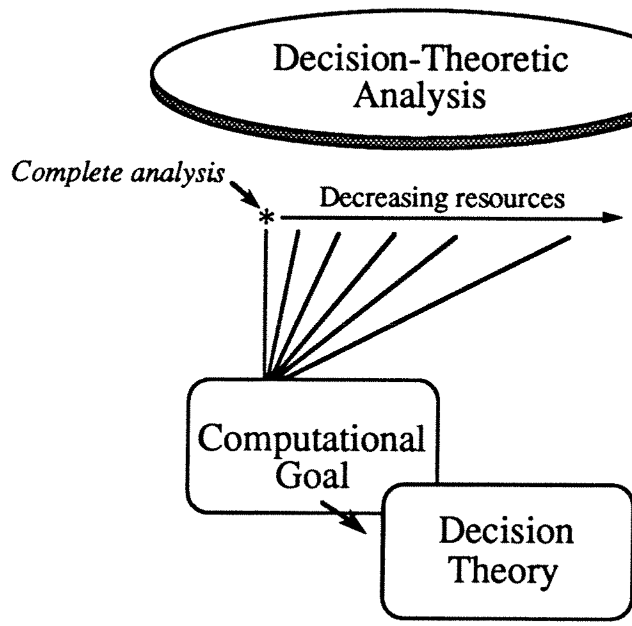

Figure 3: Rational metareasoning about partial computation can be applied to any computational goal, including such fundamental operations as sorting and searching, or to decision-theoretic inference itself. The metareasoning problem focuses on the value of alternative approximations of the object-level decision problem, as the solution ranges from the complete analysis (asterisk) to increasingly poorer approximations (direction of arrow). The rational control of decision-theoretic inference serves as a model of rationality under bounded resources.

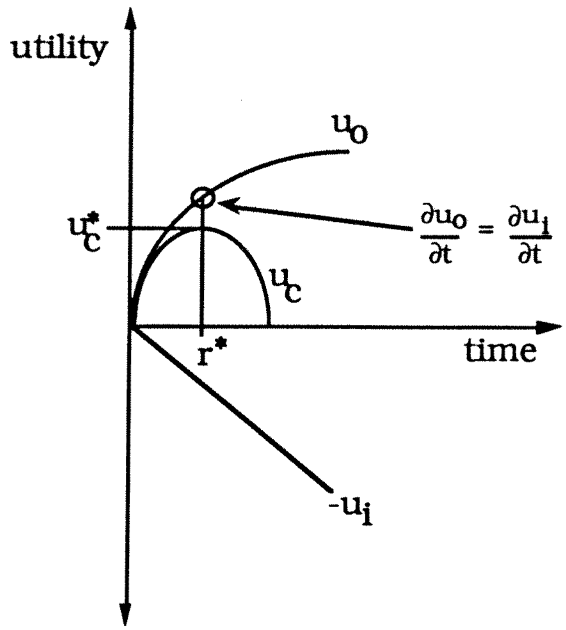

Figure 4: The expected refinement of a result by a spanning strategy, the costs (in this case, a linear cost function), and the comprehensive value of a result as additional resource is allocated to computation. The ideal comprehensive utility u_c^* is generated with the expenditure of an ideal quantity of resource r^*, where the derivatives of the object-level and inference-related utilities are equal.

where $r - r_m \geq 0$, and $\mathcal{M}$ is the metareasoning protocol. This EVCM formula is reflective in that it includes the costs associated with its own calculation. It is clear that the value of our metareasoning apparatus increases as the expected difference in $u_i(r)$ and $u_i(r^{\mathcal{M}})$ grows.

We could attempt to optimize the utility of an agent globally, considering alternative strategy sequences over a large range of resources from our initial state. Global analysis is feasible in some situations. Unfortunately, global analyses can introduce intolerable complexity to metalevel deliberation. Several simplifications can be used as approximate EVC analyses. As an exmaple, in many applications we have only one available reasoning strategy to address a particular challenge. In such cases, we need to consider only the optimality of alternative halting times. We can also reduce the costs of deliberation by performing greedy optimization over a prespecified small quantity of resource R, which increases the previously expended resource r to a new total expenditure r'. In such myopic analyses we continue to deliberate with additional packets of resource until the expected costs of the next allocated R will outweigh the benefits. Thus, a formulation of EVCM, emphasizing the marginal utility of expended additional resource R, is

$$EVCM(S_i, \pi^j(I), \mathcal{M}, R) = \int_{j'} u_o[\pi^{j'}(I)] \times p_{S_i}[\pi^{j'}(I) \mid (R - r^{\mathcal{M}})] - u_o[\pi^j(I)] - u_i(r) \quad (3)$$

where $R = r' - r$, $\pi^j(I)$ is our current result, and $\pi^{j'}(I)$ is the result expected after the next computation. That is, we continue to sum over the expected future utility, weighting each possible state by the probability of achieving that outcome, and to subtract the object-level value of our current state until the expected marginal gain is non-positive. At this point, we halt, and then act in the world. The total cost of myopic metareasoning includes the costs of metareasoning for each productive computation, in addition to a nonrefundable metareasoning penalty $u_i(r^{\mathcal{M}})$, associated with the last, nonproductive meta-analysis.

Rational metareasoning can be applied to the control of any computational or cognitive task. This point is highlighted by our exploration of the control of computation on tasks ranging from sorting a file of records to decision-theoretic inference. We believe that the optimal configuation of cognitive procedures for generating optimal behavior under bounded resources will frequently dictate some allocation of resource to metareasoning. A current research focus of work in our research group is addressing the value of metareasoning and control, given the architecture of our agents, and the costs of memory and inference. More speculatively, it is feasible that rational metareasoning may well play an important role in the cognitive systems of living creatures developed through long-term optimization under the pressures of competition.

Our approach to rational decision making under bounded resources is to construct a metalevel decision problem that represents distinctions about cognition, and to use this model to reason about and to control the nature and duration of alternative approximations to a complete decision-theoretic analysis. Our metareasoner depends on the existence of object-level approximation strategies that allow us to trade off the quality of an ideal analysis (asterisk in the figure) for more tractable, yet less precise, results. Elucidating and refining techniques for metareasoning and control are only useful if we have flexible problem-solving strategies to control. Thus, we stress that innovation at the object-level will typically be most crucial in constructing reasoners for performing under bounded resources.

4.2 Decision-Analytic Metaknowledge

Metareasoners make use of attributes of problem instances and of reasoning itself that can serve as evidence about the value of future computation. Such handles on the value of alternative reasoning pathways serve to characterize partially the nature of reasoning at lower metalevels and at the object level. As an example, the size and nature of a problem instance can serve as evidence in allowing a metareasoner to estimate the rate at which the quality of a result will increase over time. Another class of attributes that can serve as evidence of future computation is the nature of recent problem-solving behavior. There is rich research opportunity for the design of systems with the ability to inquire continually about and amass metaknowledge for calculating the value of future computation.

Decision-analytic metaknowledge includes knowledge about (1) model construction, (2) inference, (3) metareasoning, and (4) interactions among model construction, inference, and metareasoning. Model-construction metaknowledge captures attributes useful in reasoning about the value of continuing to employ strategies for generating and refining distinctions and relationships in a decision model. Inference metaknowledge includes distinctions useful in estimating the value of future inference. For example, we have been experimenting with inference metaknowledge that involves distinctions about the size and topology of a belief network. Such factors can serve as evidence in allowing a metareasoner to estimate the rate of convergence of upper and lower bounds on probabilities needed for a base-level decision problem. Metareasoning metaknowledge is information about distinctions used to characterize the expected value of increasing the fraction of time dedicated to metalevel deliberation, or moving to rely on results of a higher metalevel analysis. Finally, interaction metaknowledge captures knowledge about the interaction among model-building, inference, and metalevel deliberation. Such knowledge includes the relationship between models of higher quality and the growth of complexity of the inference task.

4.3 Decisions Under Resource Constraints

So far, we have described general principles of metareasoning without talking about a specific application. Let us now turn to work on the application of metareasoning techniques for controlling decision analysis. The process of decision analysis can be decomposed into the construction of a decision model followed by performance of decision-theoretic inference with the model. We thus must consider resource issues with both phases.

4.3.1 Partial Computation in Decision Making

In decision-theoretic inference, we are interested in partial computation for inference and model construction. Some partial-computation strategies for inference generate results of the form of bounds or distributions over belief by simulation or through the solution of specific portions of a reasoning problem [30]. We can also generate partial computation approaches to inference by incrementally increasing the abstraction of propositions or by decreasing the completeness of dependencies in a decision model [29]. Such completeness and abstraction moduation techniques have been successfully empolyed to control decision-theoretic inference within the Pathfinder reasoning system for pathology diagnosis [19]. Beyond the relatively straightforward partial results for inference and model construction, we can consider other classes of partial results for decision analysis under resource constraints. For example, it may be useful to develop a metric that represents a distance in a conceptual space describing properties of inference; in this regard, a set of partial results might be defined by the probability that a recommendation is inconsistent with one or more axioms of decision theory.

We dwell, in this paper, on inference from models that have been previously constructed. We note, however, that the principles of metareasoning presented apply also to processes for constructing and refining decision models with continuing computation. Decision models generated before the quiescence of the model-refinement process can be considered to be partial models. There are few examples of principled approaches to the construction of decision models. Model construction has been the domain of humans; the machine-intelligence community has proposed few normative approaches to the automation of this task. Exceptions include work on the automated selection of distinctions and relationships for use in a decision model based on the use of a decision-theoretic threshold analysis [18] and a qualitative analysis of important tradeoffs [50].

In metareasoning about belief and decision, we are typically interested in a probability ϕ of a state of the world given some evidence. The value of ϕ could be calculated precisely if an agent had sufficient computational resources. Under the general condition of insufficient time for a complete calulation, an agent can manipulate partial results, which are distributions

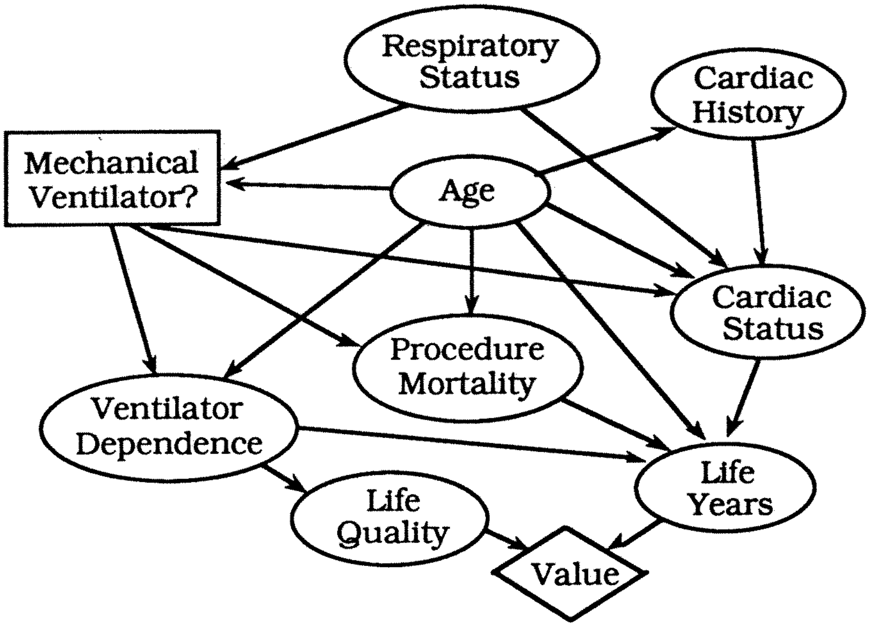

Figure 5: An influence diagram representing uncertain relationships among important distinctions (circular nodes) in intensive-care–unit medicine. Arcs between nodes indicate a dependency between the value of a node and belief assigned to the values of predecessor nodes. The possible actions are represented by a decision node (square). The best decision depends on several beliefs. The diamond represents the utility associated with different outcomes.

over ϕ, $p(\phi)$. Thus, our metareasoning problem is typically of the form

$$EVCM(S_i, p^j(\phi), \mathcal{M}, R) = \int_{j'} u_o[p^{j'}(\phi)] \times p_{S_i}[p^{j'}(\phi) \mid (R - r^{\mathcal{M}})] - u_o[p^j(\phi)] - u_i(r) \quad (4)$$

The EVCM for object-level decision-theoretic computation captures the value of continuing to refine probability distributions about events or predicates that are relevant for making a decision.

4.3.2 The Influence-Diagram Representation

Our decision-theoretic analyses have made use of *influence diagrams* [31] for representing and solving automated reasoning problems. The influence diagram is an acyclic directed graph containing nodes representing propositions and arcs representing interactions between the nodes. Nodes represent a set of mutually exclusive and exhaustive states; arcs capture probabilistic relationships between the nodes. Influence diagrams without preference or decision information are termed *belief networks.* A belief network defines a model for doing probabilistic inference in response to changes in information.

Algorithms for doing inference are expected to have a worst-case time complexity that is exponential in the size of the problem (e.g., the number of hypotheses and pieces of evidence) [7]. Several methods for inference in belief networks avoid intractability by exploiting independence relations to avoid the calculation of the joint-probability distribution. A variety

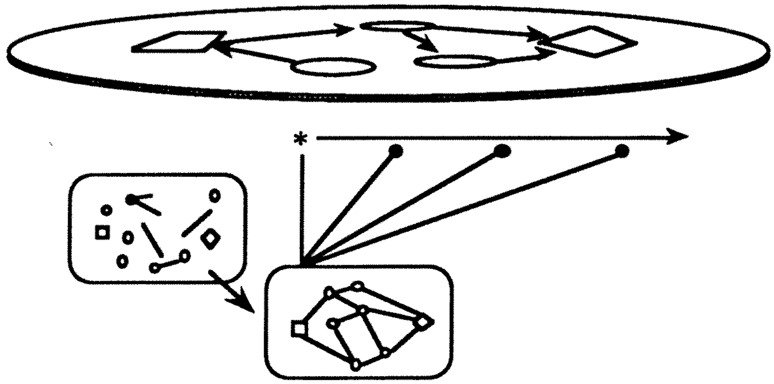

Figure 6: Under resource constraints, it can be valuable to enrich traditional decision models with metalevel decision problems. The metareasoning decision problem considers the value of expected partial results, as the quantity of resources ranges from that needed for a complete analysis to increasingly smaller allocations.

of exact methods has been developed, each designed to operate on particular topologies of belief networks [27].

The complexity of precise inference and the availability of alternative belief-network inference algorithms highlight the need for robust approximation strategies and intelligent control techniques. We have studied models of reasoning based on the application of influence diagrams, representing the metareasoning problem, to control more complicated object-level inference. The schematic portrayed in Figure 6 represents the general framework for rational action under bounded resources. We build a metalevel problem to reason about inference of varying degrees of approximation.

4.3.3 Empirical Study

We studies empirically the application of rational metareasoning to computation and action within alternative utility contexts in time-pressured medical decision making [28]. In some of this work, we endeavored to generate closed-form solutions to Equation 4 by making assumptions about the distributions over a probability of interest. After generating a tractable metareasoning model, we can apply knowledge about the behavior of alternative reasoning strategies. In particular, we have explored the behavior of recently developed graceful approximation methods for probabilistic inference. These strategies include a flexible variant of Pearl's method of conditioning [40], called *bounded conditioning* [30]. This approach satisfies the desirable properties of continuity, monotonicity, and convergence; the algorithm incrementally refines bounds on a probability of interest, continuing to tighten the upper and lower bounds on a probability of interest, with continuing computation, until reaching a point probability.

PROTOS work on the deliberative approach to ideal inference under constraints is captured Figure 7. A metareasoning decision problem is used to control the nature and extent of inference in a complex belief network. The structure of a sample belief network that we have analyzed is represented in the middle of Figure 7. This network represents uncertain

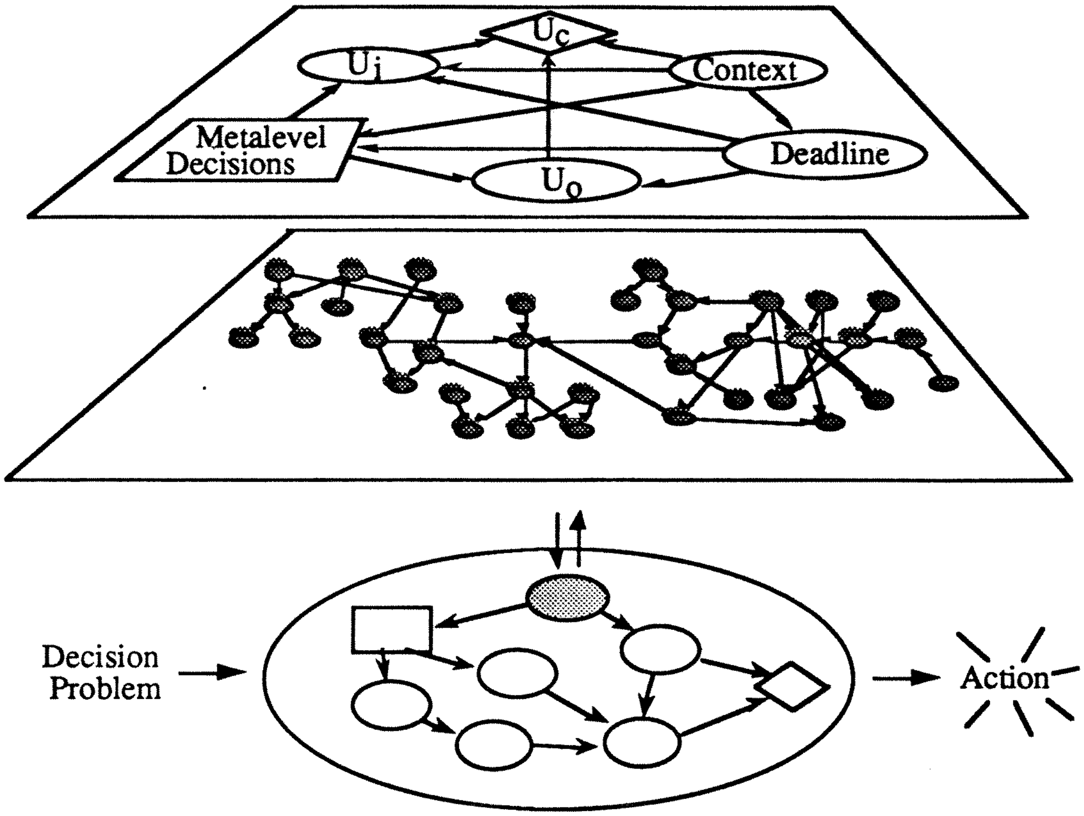

Figure 7: Research on deliberative metareasoning has explored the use of a metalevel decision model to control probabilistic inference in a complex object-level belief network. Object-level decision problems are passed to a metareasoner for evaluation.

relationships between observations and patient pathophysiology in intensive-care medicine. Object-level decision problems, requiring belief-network inference, are passed to the metareasoner, which determines the optimal dwell time.

We have pursued tractable solutions to the EVC by examining parameterized families of distributions. This work is described in more detail in [28]. For example, we have explored the use of rational metareasoning to control the application of probabilistic-bounding methods. In the work on intensive-care medicine, our reasoning system makes use of inference metaknowledge in the control of reflection and action. The inference metaknowledge is in the form a partial characterization of how the probability distribution over propositions of interest will change as reflection continues. This type of knowledge is central in reflection about the value of initiating or continuing decision-theoretic inference, as opposed to that of acting with the current best decision. We found that approximation strategies can typically deliver a large portion of the value of complete computation with a fraction of the resource [28].

4.4 History of Rational Metareasoning

Interest on the decision-theoretic control of computation has blossomed recently. To our knowledge, the general framework of rational metareasoning about partial computation was first explored and presented by our research group. In 1986, we introduced the notions of partial computation, enumerated desiderata of computation under bounded resources, and described the applicability of decision-theoretic control for optimizing the value of computation under resource constraints [21,22,25]. We proposed the decision-theoretic control of decision theory as a model of rational belief and action. Although PROTOS research has concentrated on metareasoning about decision-theoretic inference and knowledge representation, early investigations demonstrated that multiattribute decision-theoretic control of reasoning had promise for guiding the solution of a variety of tasks, including such fundamental problems as sorting a file of records or searching a large tree of possibilities [24]. To highlight the breadth of applicability of the ideas, we examined the use of the multiattribute decision theory in the control of sorting, within the PROTOS/ALGO system [26]. In 1988, Dean and Boddy independently introduced notions of partial computation [8]. Dean's group later moved to consider a decision-theoretic perspective on computation similar to ours, and has been exploring problems in temporal reasoning [33] and planning under resource constraints [4]. In 1988, Hansson and Mayer described independent work on evidential reasoning and value tradeoffs in search problems [15]. More recently, this team has been more recently examining the use of coherent evidential reasoning about the value of nodes to control search [16]. The same year, Fehling and Breese explored the application of desion-theory to the control of a robot planning problem [11], and Russell and Wefald examined the application of metareasoning to game-playing search [42]. This group has more recently explored single-agent search [49]. Agogino and colleagues have sought to apply principles of decision theory to select the best model to use in real-time reasoning in the control of milling machinery [1]. In related recent work, Doyle described a representation and analysis of a distributed approach to rationality [9], and Etzioni and Mitchell began to analyze the decision-theoretic control of learning [10].

Various aspects of the recent work on rational metareasoning can be viewed as a rediscovery of earlier speculation. The decision-theoretic approach to the control of reasoning was discussed by the statistician I.J. Good in the context of the direction of game-playing search over a decade ago [14]. Good had earlier discussed the explicit integration of the costs of inference within the framework of normative rationality, defining Type I rationality as inference that is consistent with the axioms of decision theory, regardless of the cost of inference, and Type II rationality as behavior that takes into consideration the costs of reasoning [13]. Related work in decision science has focused on the likely benefit of expending additional effort in performing decision analyses [39,48].

5 Compilation and Reflex

So far, we have discussed only deliberative approaches to reasoning and metareasoning. We have been working to move beyond reasoners that must deliberate explicitly about action. We wish to reduce complex deliberation in computer-based reasoners by developing decision-making techniques that rely to some extent on precomputed or *compiled* responses. Such knowledge can be generated at design time or learned by agents over their lifetimes.

There has been a trend in AI research to move away from deliberation toward compiled models of reasoning. For example, recent research on *reactive planning* has centered on the replacement of unwieldy solution mechanisms and detailed representations of knowledge with compiled *situation–action* rules [41,2,6,32,17]. Such rules enable agents to respond, in reflex fashion, to perceptual inputs. Investigators have sensed that, for many contexts, explicit representations and deliberation will not be necessary for satisficing performance.

5.1 Classes of Compiled Knowledge

We have been investigating several classes of compiled knowledge. These classes of compiled knowledge include (1) *situation–action rules*, (2) *platform rules*, and (3) *resource rules.* Each can play a distinct role at any level of analysis to increase the comprehensive value of decision making under resource constraints.

5.1.1 Situation–Action Rules

The simplest form of compiled knowledge is the *situation–action rule.* An observed situation is linked in reflex fashion to a final action. This form of rule is akin to the reflex response of a person to accidentally touching a hot stove. The state of VERY HOT is linked, without deliberation (besides the relatively low-overhead of retrieval), to MOVE AWAY.

5.1.2 Platform Rules

Platform rules are a generalization of situation–action rules used in conjunction with deliberative mechanisms *to make deliberation more efficient.* Such knowledge gives an agent's deliberative machinery a boost in solving a problem. Platform knowledge reduces the computational burden of reasoning. Compiled platform knowledge can be viewed as a mapping between a problem instance and a precompiled partial result, $\pi^{C}(I)$. If we require resource $r^{C}(I)$ to access the compiled knowledge, the availability of a compiled result allows us to substitute $u_o[\pi^{C}(I)] - u_i[r^{C}(I)]$ for $u_o(I)$ in our EVC formulae. The overall value of quantities of platform knowledge depends, in part, on the cost of memory and on the frequency of alternative challenges in an agent's environment. The expected gain associated with the storage and use of a particular piece of platform knowledge is the difference in optimal EVC achieved, in response to a challenge, in systems with and without the compiled knowledge. We must weight such gains by the frequency of a challenge over the agent's lifetime, and subtract the cost of purchasing and maintaining memory needed to encode the knowledge.

Platform rules include cached partial results for general classes of problem instances that can be refined with additional computation. Deliberative mechanisms can be designed with explicit reliance on the existence of platform knowledge. As an example, we can cache partial results for belief-network computation to make probabilistic inference more efficient. As another example, we can use platform rules about decision models to generate basic decision-model templates in reaction to salient features of a problem instance. Complementary deliberation can be used to custom-tailor the model to the specific situation at hand.

We can view precomputed, reflex knowledge of different degrees of completeness and levels of detail as lying at points along a compilation–deliberation continuum. The null case of zero platform knowledge defines the completely deliberative end of the spectrum. At the opposite end of the spectrum are situation–action rules. Situation–action rules can be viewed as a special case of platform knowledge where we do not perform additional deliberation.

5.1.3 Resource Rules

Situation–action rules and platform rules of different levels of compilation (capturing varying fractions of a problem-solving task) help an agent to make decisions in a more timely manner. That is, these classes of compiled knowledge provide a means for making decision making more efficient. Another class of compiled rules comprises reflex behaviors that serve to *generate additional resource*, by reducing the cost of delay, or pushing back a deadline. Resource rules are common in real-world domains that require difficult decisions and activity under time pressure. As an example of a resource rule in medicine, if a patient's blood pressure is falling dramatically, a common reaction is to add more fluids to the patient's circulatory system to raise the blood pressure. This action is typically only a temporary remedy. It serves to generate additional time for gathering information, for reasoning, and for acting to address the fundamental cause of the problem. In the case of our patient with falling blood pressure, the problem may involve a hidden hemorrhage or cardiac difficulties. Resource rules may be knowledge intensive and are typically custom-tailored to the domain at hand.

5.2 A Decision-Theoretic Perspective on Compilation

In most of the work on reactive planning, investigators have overlooked the principles of decision theory in the design and evaluation of compiled models. Thus, from the perspective of decision analysis, the compilation techniques to date typically are suboptimal for the general case of decisions under uncertainty. There have been several discussions of compilation, from the decision-theoretic perspective. Investigators have explored the derivation and consistency of compiled situation–action rules from underlying normative decision models [5], the relationship between deliberation and compiled knowledge of varying degrees of completeness [25], and the relationship of heuristically determined situation–action rules and decision-theoretic models [36]. Two recent analyses have focused on use of the principles of probability and utility theory for compiling knowledge.

A study by Heckerman and associates explores the application of decision theory to the problem of designing optimal sets of compiled situation–action rules. The analysis compares the value of completely deliberative reasoners to that of using only compiled models as a function of these factors. The analysis shows that the optimal compilation depends on (1) the nature and distribution of evidential relationships in a domain, (2) the utilities associated with alternative actions, (3) the costs of run-time delay, and (4) the costs of memory. The study describes optimal design-time rules, in the context of a binary decision problem, for choosing evidence sequences to optimize the value of compiled knowledge given the cost of memory, as well as the means to calculate the best decision, for selected subsets of observed features. A portion of the analysis examines the storage and access of arbitrary sequences of evidence within an efficient tree representation called a *trie*. A sample situation–action trie

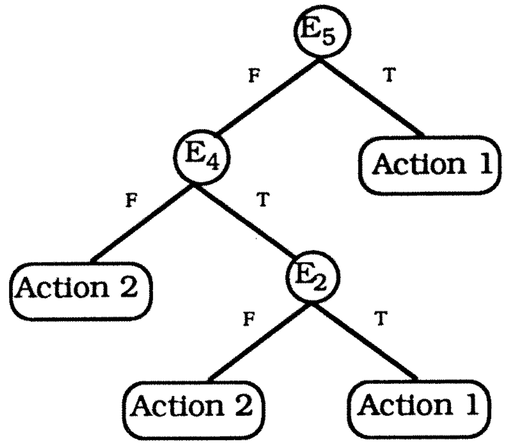

Figure 8: A sample trie containing compiled actions for a binary decision problem. In response to the observation of evidence, an action is immediately indicated.

is depicted in Figure 8. Here the compiled situation–action knowledge addresses the binary decision of whether an agent should take ACTION 1 or ACTION 2. From the example, if only evidence E_5 is observed, ACTION 1 has greatest expected value. If only E_4 is observed, then it is best to embark on ACTION 2. If, however, states E_4 and E_2 are observed, it is best to take ACTION 1.

Herskovits and Cooper have investigated the generation of platform knowledge for probabilistic inference. They compiled a set of conditional probabilities computed from a belief network in an offline manner [20]. The probabilities were selected through a utility-based simulation and also cached in a trie. Such precomputed probabilities can be used to increase the average response time of belief networks to probabilistic queries.

6 Integration of Compiled and Deliberative Reasoning

We are working to optimize the value of behavior under resource constraints by integrating compiled rules with deliberative reasoning. We can build systems that do complete analyses or different types of approximate analyses at different points along the deliberation–compilation continuum. We are interested in developing tools for designing and evaluating ideal configurations of metareasoning, reasoning, and reflex decision-making apparatus of agents, given a set of fundamental deliberative abilities, the expected problems encountered in an environment, and quantities or costs of memory and time. The ideal configuration depends on fundamental constraints on the constitution of the agent and on the nature of the challenges faced from the environment.

As depicted in Figure 9, in the context of our earlier work on deliberative models, we wish to add compiled knowledge to bolster the timeliness and accuracy of decisions made by our agents. In the general case, compiled models are incomplete. As pointed out in [17], if we store compiled knowledge solely as sets of arbitrary situation–action pairs, more than one cached situation may match the observed evidence. As an example, we may cache, as the

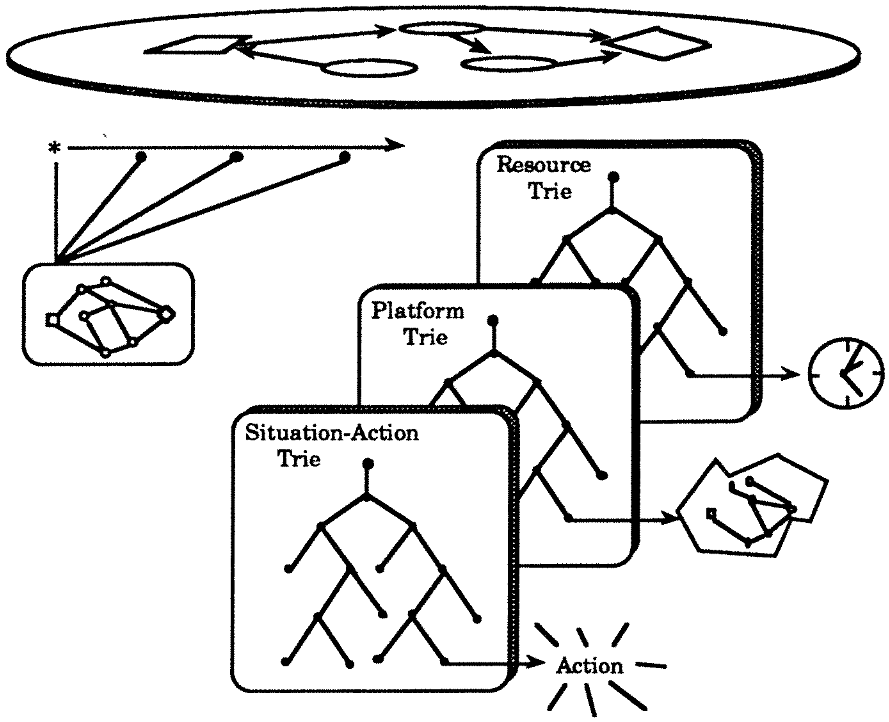

Figure 9: Several different classes of compiled knowledge can be used in an integrated fashion to make decision making more efficient. These classes include situation–action, platform, and resource knowledge. A metareasoner can make use of incomplete characterizations of the value of committing to a compiled situation–action rule versus deliberating about a challenge under expected resource constraints.

left-hand sides of a trie of situation–action rules, situations A, B and A, C and we may then observe the state A, B, C.

Incomplete compilations will typically ignore potentially important details of problems. Thus, the quality of compiled actions typically will be lower than the ultimate quality derived from computing. Under typical resource constraints, however, the compiled models might easily dominate the computed results.

In the case of the use of situation–action rules, the integrative task requires the encoding of knowledge about the relative expected value of deliberative and compiled strategies. This knowledge can be incomplete, and therefore can be represented as approximate or uncertain information. The generation and effective use of platform knowledge requires investigation of the relationships between compiled partial results and deliberation, and encoding of how the use of compiled platform knowledge is expected to enhance deliberation. Finally, we need to enrich our metareasoners such that they can quickly decide between the use of compiled situation–action rules and deliberation, determine the expected value of combining platform knowledge with deliberation under different situations, and consider the benefits and costs of firing resource rules.

6.1 Compilation of Metareasoning

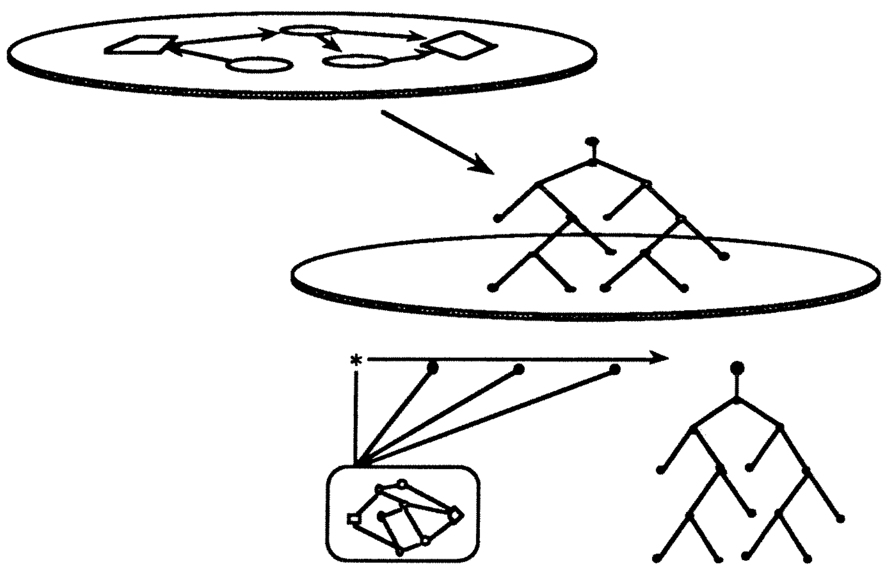

Figure 10: We can attempt to compile a portion or all of the metareasoning problem, depending on several factors, including the complexity of the model. Compilation of the metareasoning problem itself can make metalevel deliberation more efficient, allowing for more complex metareasoning models. Also, metareasoning compilation provides for quick access to object-level situation–action rules.

Metareasoning decision problems are choice candidates for compilation. We seek generally to formulate relatively tractable metareasoners for the control of difficult object-level reasoning so that we can spend a large fraction of available resource on object-level inference. Previous research in our group has explored deliberative models that have relatively tractable closed-form solutions to the value of computation problem. The discovery and application of a tractable closed-form solution to a metareasoning problem can be viewed as a form of metareasoning compilation.

In general, complex metareasoning problems may not allow for tractable closed-form solutions. These problems can pose unwieldy resource burdens on the solution of the problem. Attempting to compile some or all of the metareasoning problem into different forms of situation–*meta*-action rules can allow us to apply more complex metareasoning models.

The simplest compiled metareasoning strategies are default metareasoning policies. Such policies are commonplace in almost all simple object-level reasoners. That is, relatively inflexible object-level strategies, designed in the absence of metalevel control, implicitly impose a default metareasoning policy. Simple compiled metareasoning strategies can also be designed explicitly. These include such straightforward control strategies as the following: *(1) if a situation–action rule is available, take an action determined by the rule's consequent; (2) if it is not available, deliberate for a time dictated by the appropriate situation–meta-action rule.* More general compiled metareasoning models can make use of simple cached estimates of the expected value of deliberation versus the compiled response to a situation. Note that compiled metalevel control gives us instant access to object-level situation-action rules via situation–meta-action rules.

7 Multilevel Models

It may be valuable to custom-tailor metareasoning policies dynamically by adding one or more meta-metareasoners to control the nature and extent of metareasoning itself. The control of metareasoning can have great payoff for computation associated with difficult metareasoning analyses. For example, in the context of a difficult metareasoning problem, and the availability of several approximations to the meta-analysis, we may find great value in the use of a closed-form meta–meta-analysis that can limit metareasoning optimally.

7.1 Utility of Multilevel Metareasoning

There are costs and benefits of adding another level of analysis. The usefulness of a new metalevel for enhancing the object-level behavior of an agent is a function of the value of increased tractability and flexibility achieved through the control of preexisting levels of metareasoning, and of the memory and computation costs incurred by the additional analysis. Valuable meta-metareasoning depends clearly on the existence of flexible metareasoning strategies and on the availability of useful control knowledge at the new level of metareasoning. The optimization of the comprehensive EVC under the pressures of problems posed by an environment, and fundamental constraints on the reasoning capabilities of an agent, could dictate specific configurations of compiled knowledge and deliberative machinery at

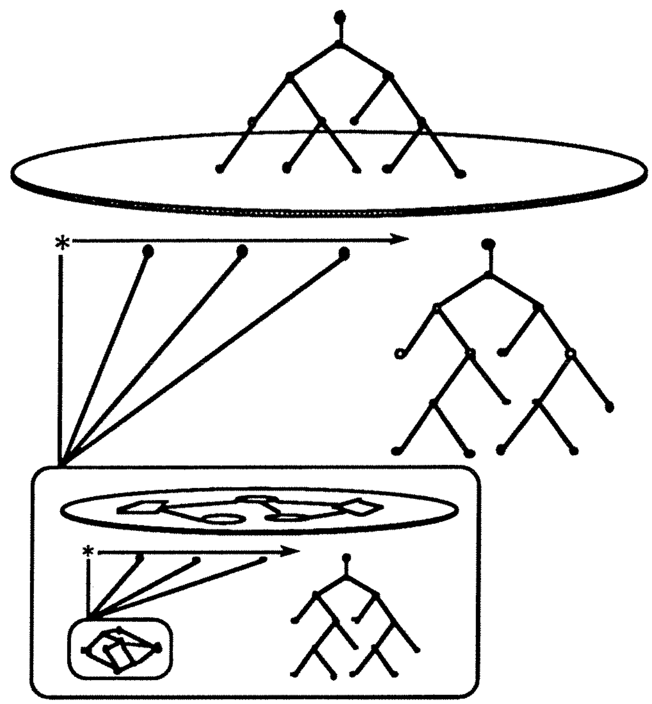

Figure 11: We can gain additional flexibility by adding new levels of metareasoning that apply knowledge about metareasoning to control a primary metareasoner. The model displayed here portrays the use of a compiled situation–action trie to control decisions about the value of explicit metalevel deliberation versus use of compiled rules to select the best metalevel strategy.

each level. Elucidating principles of optimal multiple-level metareasoning is a open problem area.

No matter how many metalevels we add, it is important to endow an agent with the ability to respond, in a reflex fashion, at the object level. In the general case, we wish to have a chain of situation–action rules from any metalevel to the object level, so as to provide efficient access to critical object-level reflexes. A more general integrated model comprises a *meta–*meta-analysis in the form of a tractable deliberative analysis or compiled situation–meta-meta-action rules. This meta-metareasoning problem controls decisions about the amount of metalevel deliberation versus object-level reasoning that should be performed, before an object-level action is taken. A sample compiled meta-metareasoning policy is the following: *(1) if an object-level situation–action rule is available, act; (2) if it is not available, deliberate for a time dictated by the metareasoning model.*

7.2 Addressing Problems of Analytic Regress

We note that questions of analytic regress tend to arise in discussions of metareasoning. If control is so valuable, why not control the controller itself, and why stop there? Will our agents have to grapple with infinite regress to optimize their decisions? Working to enhance the value of object-level computation by introducing metareasoning and control does not necessitate intractable metareasoning or infinite analytic regress. Concerns with analytic regress typically are based on assumptions about metareasoning problems that imply a need for the recursive application of meta-analysis.

One assumption, implicit in some discussions of analytic regress, is that the metareasoning problem will be at least as complex as the base problem and, therefore, will require the same kind of control as the base problem. We speculate that such an assumption of metareasoning complexity is based on a sense that a metareasoning problem must represent a great portion of the base problem that it is attempting to control. We may, indeed, be able to construct metaproblems with a complexity that rivals or exceeds that of the base problem. However, we can frequently build much simpler control problems that greatly enhance the comprehensive value of reasoners. These meta-analyses address particular aspects of object-level problem solving, and make use of specific classes of metaknowledge; they do not capture the full complexity of representation or inference at the base level.

Another assumption is that control decisions are necessarily sensitive to small changes in the accuracy of a metareasoning analysis. If this were so, optimizing the value of control decisions might invoke a large, or an infinite, number of metalevels. We believe, however, that this is the exception, rather than the rule. Within our models, control decisions appear to be insensitive to small changes in the accuracy of the analysis. When this not the case, we can seek to characterize a relevant spectrum of meta-metareasoning implications during the design phase; our experience leads us to be optimistic about our ability to dodge recursive real-time expansion of meta-analysis by analyzing the expected behavior of a metareasoner.

The theme of our metareasoning research is to extend object-level analyses by invigorating the base models with one or several levels of tractable metalevel decision problems and approximations, and to analyze the problem of metareasoning allocation at design time. We believe that work on tractable, closed-form meta-analyses hold the greatest opportunities for empirical and theoretical research. Nevertheless, there may be cases where the fundamental consitution of an agent and its environment dictates as optimal the application and management of a large or (in theory) infinite, tower of meta-analyses [26]. There has been some discussion of techniques for handling the analytic regress problem. Kripke has discussed a problem with infinite regress that arises in defining truth [35]. He applied a mathematical tool known as a transfinite hierarchy to grapple with the problem. Lipman has peformed analysis of an infinite-regress problem by applying similar techniques within a game-theoretic framework [37]. He demonstrates that, within his model, there is an equivalence between a base decision problem and an infinitely recursive analysis.

We have suggested that optimization problems that imply large hierarchies of metareasoners provide rich problems for convergence and stability analyses. For example, the EVC of an agent, based on a theoretical prescription for an infinite hierarchy of metareasoning analyses, might converge tractably to a positive value in a manner similar to the way the sum of an infinite series may closely approach a real number with a small number of terms. Also, perturbation and stability analyses developed within control theory may be useful in determining optimal configurations of large numbers of interrelated meta-analyses, especially for cases where control decisions are sensitive to minute perturbations of accuracy of analysis at a large number of metalevels. Great sensitivity of control reasoning to the completeness of metareasoning analyses can focus attention on a search for less sensitive, more stable levels of approximate meta-analyses. Convergence and stability analyses hold the promise of dictating finite concrete architectures and policies for an ideal bounded-optimal agent that rely on unwieldy recursive meta-analyses.

8 Learning, Compilation, and Rationality

I would like to end with some discussion on the importance of studying techniques for allowing the deliberative and compiled knowledge of our reasoners to be modified and extended based on an agent's expectations and experience within a specialized environment. From our perspective on rationality, short-term and long-term learning are viewed as local and more permanent compilation strategies, respectively. These dynamic compilation strategies allow an agent that has been endowed with an architecture that supports reflex, reasoning, and metareasoning, and with intrinsic knowledge for surviving in a broadly defined environment, to prosper when immersed in a specialized world. We view learning as the pursuit of optimal EVC and associated comprehensive value, in accordance with the distribution of problems seen in a specialized environment. There are several ways to combine the EVC, associated with an agent's responses to distinct challenges, into a more global value of agency over a period of time. We discuss the valuation of behavior over time in [25].

It may be optimal to allow a reasoner to allocate a quantity of memory for the dynamic generation and caching of platform knowledge. Depending on the constitution of an agent and the problems faced by that agent, it may be optimal to generate specific platform knowledge continually and, afterward, to destroy the knowledge to free up memory for the next local challenge. As an example, it can be useful to generate local expectation-driven partial solutions for probabilistic inference with the bounded-conditioning inference method

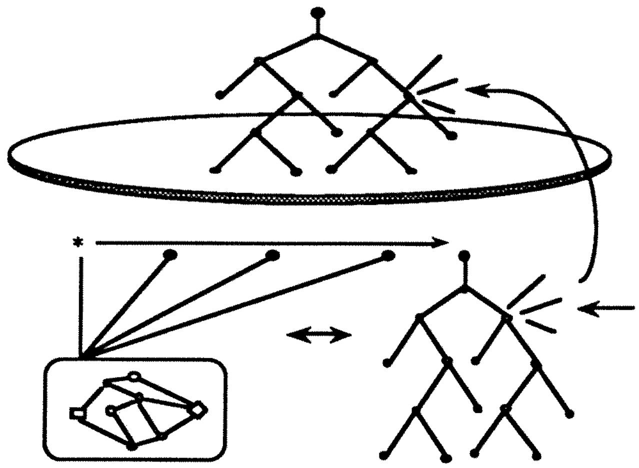

Figure 12: From the perspective of decision-theoretic compilation, learning can be viewed as enhancing or extending classes of compiled knowledge. Such learning can take place at the object level or within the metareasoning problem. In the case of refining platform knowledge, we wish reasoners to make complementary modifications in associated deliberative mechanisms and control policies.

during the idle time between observations in domains where observations impinge on an agent over time [30].

Knowledge about the interaction of situation–action, platform, and resource rules with deliberative mechanisms should be used in design-time selection of the best set of knowledge for a given quantity or cost of memory. Allowing for relatively permanent changes in compiled knowledge can enable agents designed for a range of environments to optimize their behaviors in specific environments. Utility-based rule-selection techniques can increase the effectiveness of reasoning at the object-level and at successive metalevels. Thus, we can view a portion of learning as the construction and storage of sets of reflex, platform, and resource knowledge that are tailored to particular problem tasks. Beyond the custom-tailoring and refinement of compiled knowledge, effective learning might also modify existing deliberative reasoning and metareasoning strategies. Such complementary refinement of deliberation could optimize the use of the compiled knowledge

9 Summary

We have reviewed issues surrounding the pursuit of rational behavior under resource constraints. We have endeavored to optimize the value of decisions by expanding traditional decision models into multilevel analyses that model cognitive processes, in addition to distinctions associated with problem challanges. We described the notion of partial computation, presented a general framework for optimizing the multiattribute utility of computation, and examined some of our work on the application of rational metareasoning to the control of fundamental computation tasks such as sorting, and for deliberation about rational belief and action. After presenting a deliberative approach to bounded-optimal decisions, we discussed more recent research on the integration of different classes of compiled knowledge with deliberative machinery. In particular, we described the use of *situation–action rules* to gain access to direct action or meta-actions, *platform rules* to enhance deliberation at any level of analysis, and *resource rules* to lower the costs associated with delay, or to fend off deadlines. After discussing the compilation of metareasoning, we presented opportunities for developing integrated multilevel reasoners. Finally, we touched on short-term and long-term learning as compilation strategies that raise the expected value of a agent's behavior. Our investigation continues to pursue rational belief and action through analysis of the rich relationships among reflex, reasoning, and metareasoning.

Acknowledgments

John Breese, Bruce Buchanan, Gregory Cooper, George Dantzig, Tom Dean, David Heckerman, Ronald Howard, Nils Nilsson, Edward Shortliffe, and Patrick Suppes have provided useful feedback on PROTOS research.

References

[1] A. Agogino and K. Ramamurthti. Real-time reasoning about time constraints and model precision in complex distributed mechanical systems. In *Working Notes: Spring Symposium Series on AI and Limited Rationality*, pages 1–5, Stanford University, March 1989. American Association for Artificial Intelligence.

[2] P. Agre and D. Chapman. Pengi: An implementation of a theory of activity. In *Proceedings of the Sixth National Conference on Artificial Intelligence*, pages 268–272, Seattle, WA, July 1987. AAAI-87.

[3] A.V. Aho, J.E. Hopcroft, and J.D. Ullman. *Data Structures and Algorithms.* Addison-Wesley, Menlo Park, CA, 1983.

[4] M. Boddy and T. Dean. Solving time-dependent planning problems. In *Proceedings of the Eleventh IJCAI.* AAAI/International Joint Conferences on Artificial Intelligence, August 1989.

[5] J.S. Breese. *Knowledge Representation and Inference in Intelligent Decision Systems.* PhD thesis, Department of Engineering-Economic Systems, Stanford University, 1987. Also available as Technical Report 2, Rockwell International Science Center, Palo Alto Laboratory.

[6] R. Brooks. Planning Is Just a Way of Avoiding Figuring Out What to Do Next. Technical Report Working Paper 303, M.I.T. Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 1987.

[7] G.F. Cooper. Probabilistic inference using belief networks is NP-hard. *Artificial Intelligence*, 2, 1989. In press.

[8] T. Dean and M. Boddy. An analysis of time-dependent planning. In *Proceedings AAAI-88 Seventh National Conference on Artificial Intelligence*, pages 49–54. American Association for Artificial Intelligence, August 1988.

[9] J. Doyle. Artificial intelligence and rational self-government. Technical Report CS-88-124, Carnegie-Mellon University, 1988.

[10] O. Etzioni and T. Mitchell. A comparative analysis of chunking and decision-analytic control. In *Working Notes: Spring Symposium Series on AI and Limited Rationality*, pages 42–45, Stanford University, March 1989. American Association for Artificial Intelligence.

[11] M.R. Fehling and J.S. Breese. A computational model for the decision-theoretic control of problem solving under uncertainty. Technical report, Rockwell International Science Center, April 1988. Rockwell Technical Report 837-88-5.

[12] M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W.H. Freeman and Co., New York, 1979.

[13] I.J. Good. Rational decisions. *J. R. Statist. Society B*, 14:107–114, 1952.

[14] I.J. Good. Dynamic probability, computer chess, and the measurement of knowledge. In E.W. Elcock and Michie D., editors, *Machine Intelligence 2*, pages 139–150. Wiley, New York, 1977.

[15] O. Hanson and A. Mayer. The optimality of satisficing solutions. In *Proceedings of Fourth Workshop on Uncertainty in Artificial Intelligence*, Minneapolis, MN, July 1988. American Association for Artificial Intelligence.

[16] O. Hansson and A. Mayer. Probabilistic heuristic estimates. In *Proceedings of the Second International Workshop on AI and Statistics*, Ft. Lauderdale, January 1989.

[17] D.E. Heckerman, J.S. Breese, and E.J. Horvitz. The compilation of decision models. In *Proceedings of Fifth Workshop on Uncertainty in Artificial Intelligence*, Windsor, Canada, August 1989. American Association for Artificial Intelligence.

[18] D.E. Heckerman and E.J. Horvitz. A principled approach to problem formulation. Technical Report KSL-89-67, Stanford University, Knowledge Systems Laboratory, Stanford, CA, June 1989.

[19] D.E. Heckerman, E.J. Horvitz, and B.N. Nathwani. Toward effective normative expert systems: The Pathfinder project. *Computers and Biomedical Research*, 1989. In press.

[20] E.H. Herskovits and G.F. Cooper. Algorithms for belief-network precomputation. Technical Report KSL-89-35, Stanford University, January 1989.

[21] E.J. Horvitz. Reasoning about inference tradeoffs in a world of bounded resources. Technical Report KSL-86-55, Stanford University, Knowledge Systems Laboratory, Stanford, CA, September 1986.

[22] E.J. Horvitz. Toward a science of expert systems. In Thomas J. Boardman, editor, *Proceedings of the 18th Symposium on the Interface of Computer Science and Statistics*, pages 45–52, Ft. Collins, Colorado, March 1986. American Statistical Association. Also available as Technical Report KSL-86-75: Knowledge Systems Laboratory: Stanford University, March 1986.

[23] E.J. Horvitz. The decision-theoretic control of problem solving under uncertain resources and challenges. Technical Report KSL-87-16, Stanford University, Knowledge Systems Laboratory, Stanford, CA, November 1987.

[24] E.J. Horvitz. Problem-solving design: Reasoning about computational value, tradeoffs, and resources. In *Proceedings of the NASA Artificial Intelligence Forum*, pages 26–43. National Aeronautics and Space Administration, November 1987. Also available as Technical Report KSL-87-64, Knowledge Systems Laboratory, Stanford University, October 1987.

[25] E.J. Horvitz. Reasoning about beliefs and actions under computational resource constraints. In *Proceedings of Third Workshop on Uncertainty in Artificial Intelligence*, Seattle, Washington, July 1987. American Association for Artificial Intelligence. Also in L. Kanal, T. Levitt, and J. Lemmer, ed., *Uncertainty in Artificial Intelligence 3,* Elsevier, pps. 301-324.

[26] E.J. Horvitz. Reasoning under varying and uncertain resource constraints. In *Proceedings AAAI-88 Seventh National Conference on Artificial Intelligence*, pages 111–116. American Association for Artificial Intelligence, August 1988.

[27] E.J. Horvitz, J.S. Breese, and M. Henrion. Decision theory in expert systems and artificial intelligence. *International Journal of Approximate Reasoning*, 2:247–302, 1988. Special Issue on Uncertain Reasoning.

[28] E.J. Horvitz, G.F. Cooper, and D.E. Heckerman. Reflection and action under scarce resources: Theoretical principles and empirical study. In *Proceedings of the Eleventh IJCAI*, pages 1121–1127. AAAI/International Joint Conferences on Artificial Intelligence, August 1989.

[29] E.J. Horvitz, D.E. Heckerman, K. Ng, and B.N. Nathwani. Heuristic abstraction in the decision-theoretic Pathfinder system. In *Proceedings of the 13th Symposium on Computer Applications in Medical Care.* IEEE Computer Society Press, October 1989.

[30] E.J. Horvitz, H.J. Suermondt, and G.F. Cooper. Bounded conditioning: Flexible inference for decisions under scarce resources. In *Proceedings of Fifth Workshop on Uncertainty in Artificial Intelligence*, Windsor, Canada, August 1989. American Association for Artificial Intelligence.

[31] R.A. Howard and J.E. Matheson. Influence diagrams. In R.A. Howard and J.E. Matheson, editors, *Readings on the Principles and Applications of Decision Analysis*, volume II, pages 721–762. Strategic Decisions Group, Menlo Park, CA, 1981.

[32] L. Kaelbling. An architecture for intelligent reactive systems. In A.L. Lansky and M.P. Georgeff, editors, *Reasoning About Actions and Plans: Proceedings of the 1986 Workshop*, pages 395–410. Morgan-Kaufmann, 1987.

[33] K. Kanazawa and T. Dean. A model for projection and action. In *Proceedings of the Eleventh IJCAI*. AAAI/International Joint Conferences on Artificial Intelligence, August 1989.

[34] D.E. Knuth. *Sorting and Searching.* Addision-Wesley, Reading, Massachusetts, 1973.

[35] S. Kripke. Outline of a theory of truth. *Journal of Philosophy*, 72:690–716, 1975.

[36] C. P. Langlotz, E. H. Shortliffe, and L. M. Fagan. Using decision theory to justify heuristics. In *Proceedings of the Sixth National Conference on Artificial Intelligence*, pages 215–219, Philadelphia, PA, 1986. AAAI-86.

[37] B.L. Lipman. How to decide how to decide hot to...: Limited rationality in decisions and games. Technical report, Carnegie Mellon University, Pittsburgh, February 1989.

[38] J.G. March. Bounded rationality, ambiguity, and the engineering of choice. *Bell Journal of Economics*, pages 587–608, 1978.

[39] J.E. Matheson. The value of analysis and computation. *IEEE Transactions on Systems Science, and Cybernetics*, 4:211–219, 1968.

[40] J. Pearl. A constraint-propogation approach to probabilistic reasoning. In L.N. Kanal and J.F. Lemmer, editors, *Uncertainty in Artificial Intelligence 1*, pages 357–369. Elsevier Science, Amsterdam, 1986.

[41] S. Rosenschein and L. Kaelbling. The synthesis of digital machines with provable epistemic properties. In *Proceedings of the Conference on the Theoretical Aspects of Reasoning About Knowledge*, pages 83–98, Asilomar, CA, 1986. AAAI.

[42] S.J. Russell and E. Wefald. Multi-level decision-theoretic search. In *Proceedings of the AAAI Spring Symposium on Game Playing*, Stanford, CA, March 1988. American Association for Artificial Intelligence.

[43] S. M. Shugan. The cost of thinking. *Journal of Consumer Research*, 7:99–111, 1980.

[44] H.A. Simon. A behavioral model of rational choice. *Quarterly Journal of Economics*, 69:99–118, 1955.

[45] H.A. Simon. The theory of problem solving. *Information Processing*, 71:261–277, 1972.

[46] H.A. Simon. Rationality as process and as product of thought. *J. American Economic Association*, 68(2):1–16, 1978.

[47] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior.* Princeton University Press, Princeton, NJ, 1947.

[48] S.R. Watson and R.V. Brown. The valuation of decision analysis. *J.R. Statist. Society A.*, 141(1):69–78, 1978.

[49] E. Wefald and S. Russell. Estimating the value of computation: The case of real-time search. In *Working Notes: Spring Symposium Series on AI and Limited Rationality*, pages 106–110, Stanford University, March 1989. American Association for Artificial Intelligence.

[50] M.P. Wellman. Formulation of tradeoffs in planning under uncertainty. Technical Report MIT/LCS/TR-427, Laboratory for Computer Science, Massachusetts Institute of Technology, August 1988.