Utility Models for Goal-Directed Decision-Theoretic Planners

Peter Haddawy[1], Steve Hanks

Department of Computer Science and Engineering
University of Washington
Seattle, WA 98195

Technical Report 93–06–04
June 15, 1993

[1] Department of EE & CS, University of Wisconsin–Milwaukee, Milwaukee WI 53201

Utility Models for Goal-Directed Decision-Theoretic Planners*

Peter Haddawy
Department of EE & CS
University of Wisconsin–Milwaukee
Milwaukee WI 53201
haddawy@cs.uwm.edu

Steve Hanks
Department of CS&E, FR–35
University of Washington
Seattle WA 98195
hanks@cs.washington.edu

June 22, 1993

University of Washington
Department of Computer Science & Engineering
Technical report 93-06-04

Abstract

AI planning agents are *goal-directed*: success is measured in terms of whether or not an input goal is satisfied, and the agent's computational processes are driven by those goals. A decision-theoretic agent, on the other hand, has no explicit goals—success is measured in terms of its preferences or a utility function that respects those preferences.

The two approaches have complementary strengths and weaknesses. Symbolic planning provides a computational theory of plan generation, but under unrealistic assumptions: perfect information about and control over the world and a restrictive model of actions and goals. Decision theory provides a normative model of choice under uncertainty, but offers no guidance as to how the planning options are to be generated. This paper unifies the two approaches to planning by describing utility models that support rational decision making while retaining the goal information needed to support plan generation.

We develop an extended model of goals that involves temporal deadlines and maintenance intervals, and allows partial satisfaction of the goal's temporal and atemporal

*Thanks to Tony Barrett, Denise Draper, Dan Weld, Mike Wellman, and Mike Williamson for many useful comments. Thanks to Meliani Suwandi for helping develop the refinement planning example. Haddawy was supported in part by NSF grant IRI–9207262 and in part by a grant from the graduate school of the University of Wisconsin-Milwaukee. Hanks is supported in part by NSF Grant IRI–9008670.

components. We then incorporate that goal information into a utility model for the agent. Goal information can be used to establish whether one plan has higher expected utility than another, but without computing the expected utility values directly; we present a variety of results showing how plans can be compared rationally, but using information only about the probability that they will satisfy goal-related propositions. We then demonstrate how this model can be exploited in the generation and refinement of plans.

Contents

1 Introduction **1**
1.1 Contributions 2
1.2 Outline 3

2 Classical Goal-Directed Agents **3**
2.1 Limitation of goal-directed behavior 4

3 Utility Models for Goal-Directed Agents **7**
3.1 Syntactic forms for goals 9
3.1.1 The language $\mathcal{L}_{tca}$ 9
3.1.2 Goal expressions 10
3.1.3 Utility function for individual goals 10

4 Partial Goal Satisfaction **11**
4.1 Partial satisfaction of the atemporal component 11
4.1.1 Goals with symbolic attributes 11
4.1.2 Goals with quantitative attributes 13
4.1.3 Conjunctive quantitative attributes 13
4.1.4 Summary 14
4.2 Partial satisfaction of the temporal component 14
4.2.1 Deadlines 14
4.2.2 Maintenance intervals 16
4.2.3 Summary 16

5 Utility Functions for Individual Goals **17**
5.1 Combining the temporal coefficient and the atemporal degree of satisfaction 17
5.1.1 Deadline goals 17
5.1.2 Maintenance goals 19
5.2 Summary 20

6 Using the Utility Functions to Rank Plans **20**
6.1 Deadline goals 21
6.1.1 Bounds for symbolic attributes 21
6.1.2 Bounds for quantitative attributes 23
6.1.3 Bounds for strictly ordered atemporal attributes 24
6.1.4 Example 25
6.2 Maintenance goals 26
6.2.1 Bounds for symbolic attributes 26
6.2.2 Bounds for quantitative attributes 28
6.2.3 Bounds for strictly ordered atemporal attributes 29
6.3 Summary 29

7 Multiple Goals, Residual Utility, and Computational Issues **29**
7.1 Plan generation and refinement 30
7.1.1 Partial-order planning and search control 31
7.1.2 Refinement planning 31

8 Summary and Related Work **41**
8.1 Related work 42
8.1.1 Goals and utility models 42
8.1.2 Decision-theoretic planning and control 43
8.1.3 Fuzzy decision theory 44
8.2 Future work 44

1 Introduction

Reasoning about and planning for an uncertain world raises both representational and algorithmic problems: we need to represent change, uncertainty, and value or utility, and to use those concepts to represent various plans of action. And given such a representation for the world and for plans that might be executed in the world, we further need an efficient way to generate plausible plans, anticipate their results, improve their performance, and choose the best option from among them.

Decision theory addresses the representational problem, providing a rational basis for choice under uncertainty. The framework starts with the agent's preferences over an abstract set of possible outcomes, and guarantees the existence of probability and utility functions such that acting to maximize expected utility is rational in the sense that it respects those preferences. (The planner need not explicitly perform the decision-theoretic analysis—it suffices that the planner act according to the recommendations that such an analysis would make.)

Decision theory does not, however, constitute a theory of reasoning about plans. The theory does not provide a vocabulary for describing planning problems, a method for generating options, or a computational model for choosing among plan alternatives. It merely dictates a rational choice *given* such a capability. Decision analysis—the study of applying decision-theoretic framework—addresses these limitations, but in a subjective and non-algorithmic fashion.

The capabilities provided by symbolic planning and decision theory are therefore complementary: the former provides methods for representing planning problems and generating alternative plans in response to externally supplied goals (but under restrictive conditions that don't include uncertainty); the latter provides a method for choosing among alternatives and a language that allows reasoning with uncertainty, but provides (1) no guidance in structuring planning knowledge, (2) no way of generating alternatives, and more generally, (3) no computational model. The first step toward integrating these two methodologies is to reconcile the two representations.

Planning problems are described in terms of

- a set of *operators* that effect change in the world,
- a description of some initial *world state* (usually expressed in some logical or quasi-logical language), and
- a *goal state description*.

Decision-theoretic problems, on the other hand, are stated in terms of

- a probability distribution over possible outcomes, and
- preferences over those outcomes.

The relationship between the probabilistic model of the world and the operators and initial state has been studied as a problem of probabilistic temporal inference, [Haddawy, 1991b], [Hanks and McDermott, 1993], [Hanks, 1993]. The relationship between the goal state description and the agent's utility model has received less attention, and is the topic of this paper.

1.1 Contributions

This paper presents a utility model for goal-directed agents that allows rational choice among planning alternatives but that also can be exploited by a plan-generation algorithm to guide the process of building effective plans. It therefore directly addresses the gap between AI and decision-theoretic planners, providing a richer representation language for the former and a computational framework for the latter.
The main contributions of the paper are:

1. A detailed analysis and taxonomy of goal forms. We break a goal into atemporal and temporal components: *what* is to be true and *when* it is to be true. We consider two main forms of temporal constraints: deadline goals and maintenance goals, and various forms of atemporal constraints, including symbolic and numeric goals.

2. A model of partial satisfaction that allows reasoning about partial satisfaction of a goal's atemporal component, temporal component or both. Partial satisfaction information for a goal is supplied in terms of

   - A function DSA measuring the extent to which the goal's atemporal component is satisfied at a point in time.
   - A function measuring the extent to which the goal's temporal component is satisfied:
     - For deadline goals, a function CT that measures the extent to which a time point meets the deadline.
     - For maintenance goals, a function CP that measures the extent to which a time interval satisfies the maintenance constraint.

   We develop a method for combining these two pieces of information into a coherent assessment of the goal's level of satisfaction.

3. A utility model for an agent that allows reasoning about trading off (partial) success in achieving one goal against (partial) success in achieving another, and trading off success in achieving goals with resource consumption.

4. Effective methods for comparing two plans: relationships that imply that one plan has higher expected utility than another, but that do not require computing the expected utilities directly.

5. An example of how the model can be exploited by a planning algorithm to prune the space of partial plans explicitly considered.

1.2 Outline

Section 2 begins by defining a goal-directed agent, and pointing out the limitations of a planning strategy guided only by conjunctive goal expressions. Section 3 defines an extended utility model for goal-directed agents, which includes a richer notion of goal than the one typically used by classical planners. Section 4 confronts the problem of specifying preferences over partially satisfied goals, and Section 5 defines a model that allows partial satisfaction of a goal's temporal and atemporal components simultaneously.

Sections 6 and 7 address the question of how the goal-based utility model can be exploited in comparing and building plans. Section 6 presents a variety of results describing circumstances under which deciding that one plan is preferable to another amounts to establishing a relationship between probabilities involving the respective goal expressions. Section 7 discusses how the model might be used in generating plans, exploring extensions to the classical partial-order generation algorithm and the idea of planning by refinement. Section 8 summarizes and discusses related work.

2 Classical Goal-Directed Agents

Before we proceed with a formal analysis of goals and utilities we need to define more precisely what role these concepts will play in a planning system. Goals play three roles in automated planning systems:

- Goals act as a device for communicating information about the planning problem. In particular they provide a concise definition of what constitutes a successful plan.
- Goals act as a means of limiting inference in the planning process by allowing the planner to backchain over goal propositions. In that sense they define exactly what is *relevant* to the planning problem.
- Goals limit the temporal scope of the planning problem, imposing a temporal "horizon," beyond which planning is irrelevant.

In the first case goals can be communicated more easily than utility functions, and in the second and third cases the goals' symbolic content can aid the search for good plans.

Figure 1 makes the relationship more clear: let us assume some *manager*, who has a utility function over outcome states. He is designing an *agent*, and has a model of the agent's capabilities. The manager wants to communicate information to the agent that will cause it to act so as to increase the manager's utility; he uses goal expressions to communicate that information to the agent. The agent uses this goal information—along with information like the anticipated state of the world at execution time and the cost of various resources —to produce a utility function that serves to guide its actions.

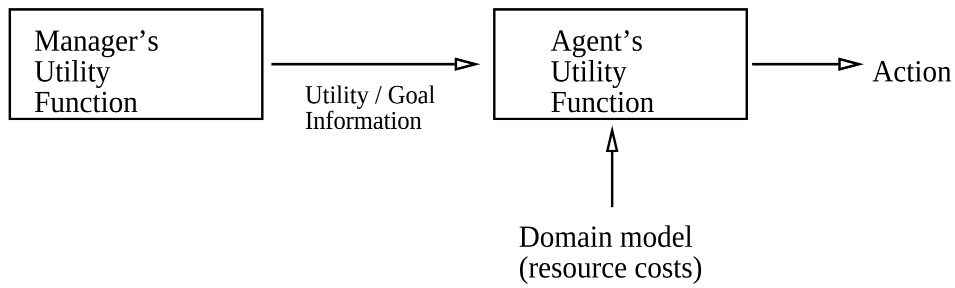


Figure 1: Goals as a means of communication

This model applies classical goal-directed agents as well as decision-theoretic agents. The question is what language the manager should use to convey the utility/goal information, and how that information restricts the possible behaviors available to the agent. The classical planning model restricts this information to a conjunction of goal propositions, which are supposed to hold at the end of plan execution. We will show that restricting the form of utility information to goal conjunctions of this form places severe restrictions on the agent's problem-solving abilities, then develop more expressive forms of goal expressions that still allow the agent to be an effective problem solver.

2.1 Limitation of goal-directed behavior

Suppose the manager communicates only symbolic goal expressions to the agent—he tells the agent to achieve some goal G, a conjunction $\mathsf{g}_1 \wedge \mathsf{g}_2 \ldots \mathsf{g}_n$, and the agent builds a plan that maximizes the probability that satisfying G will be true at the end of executing the plan. (This is the probabilistic analogue of logical goal-directed planning, in which the agent constructs a plan that *provably* achieves G.) What limits does this model place on the agent's preference struture? In other words, under what circumstances is planning to maximize expected utility equivalent to planning to maximize the probability of goal satisfaction?

We start by introducing some notation. We will talk about *time points* t, and can talk about a proposition ϕ being true over an interval of time $[t_1, t_2]$ by saying $HOLDS(\phi, t_1, t_2)$ (see Section 3.1.1 for more details).

A *plan* $\mathcal{P}$ can be viewed as defining a probability distribution over *chronicles*, where a chronicle is a set of time points representing one possible course of execution for the plan. We can therefore talk about the probability $\mathbf{P}(c|\mathcal{P})$.[1] We will generally be interested in the time point representing the moment the plan finishes executing; we will use $\mathrm{end}(c)$ to represent this point. The probability of success in the classical paradigm is the probability that the goal conjunction will hold when the plan finishes executing:

$$\mathbf{P}(\mathcal{P} \text{ succeeds}) \equiv \mathbf{P}(\mathsf{G}|\mathcal{P}) \equiv \sum_{c:HOLDS(\mathsf{G},\mathrm{end}(c),\mathrm{end}(c))} \mathbf{P}(c|\mathcal{P}) \tag{1}$$

[1][McDermott, 1982] defines chronicles in terms of a temporal logic, and [Hanks, 1990] and [Haddawy, 1991b] extend the notion to a probabilistic framework.

and the planner will try to find the plan maximizing that value.

A decision-theoretic planner has the same probabilistic model as the probabilistic planner, but it has an explicit utility function over chronicles, a function $\mathbf{U}(c)$ that maps chronicles into real values. The expected utility of a plan is defined to be

$$\mathbf{EU}(\mathcal{P}) \equiv \sum_{c} \mathbf{U}(c) \cdot \mathbf{P}(c|\mathcal{P}) \tag{2}$$

and the decision-theoretic planner will try to find the plan maximizing that value.

The question arises as to under what circumstances the two models are equivalent: for what utility models (definitions of $\mathbf{U}(c)$) is it the case that the plan that maximizes the probability of goal achievement (Equation (1)) will always be the plan that maximizes expected utility (Equation (2))?

The answer is that this relationship holds only for simple step utility functions, functions for which utility is a constant low value $\mathsf{U}\overline{\mathsf{G}}$ for chronicles in which the goal does not hold at the end of plan execution and a constant high value UG for chronicles in which it does. Figure 2 shows such a function—utility is represented along the vertical axis and the space of chronicles along the horizontal axis. G and $\overline{\mathsf{G}}$ designate the set of all chronicles in which the goal holds and does not hold, respectively.

Previous work, [Haddawy and Hanks, 1990], demonstrates the correspondence between these two policies, showing that the simple class of step functions pictured in Figure 2 is the *only* class of utility functions for which

$$\mathbf{P}(\mathsf{G}|\mathcal{P}_1) > \mathbf{P}(\mathsf{G}|\mathcal{P}_2) \Rightarrow \mathbf{EU}(\mathcal{P}_1) > \mathbf{EU}(\mathcal{P}_2) \tag{3}$$

holds for any plans $\mathcal{P}_1$ and $\mathcal{P}_2$. In other words, describing the desired state of the world in terms of a goal conjunction restricts a problem-solver's preference structure to those that can be characterized by a simple step function. ([Haddawy and Hanks, 1990] explores the relationship between goal satisfaction and probability maximization for other situations, e.g. cases in which goal expressions are combined with some preference information about resource consumption.)

The analysis points out four obvious limitations to the endeavor of planning to achieve a goal conjunction:

1. Temporal extent
   Defining plan success in terms of what is true at the end of execution—defining a successful chronicle only in terms of whether the goal holds at end(c))—rules out goals like deadlines (have g_1 true by noon and g_2 true by midnight), maintenance (keep g true continuously between noon and midnight), and prevention (make g_1 true, but without allowing g_2 to become true in the meantime). (The last is a combination of deadline and maintenance goals.) It is important not only *what* is accomplished, but also when or for how long.

2. Tradeoffs among the goals
   The classical model dictates that the satisfaction of all goal conjuncts is necessary and

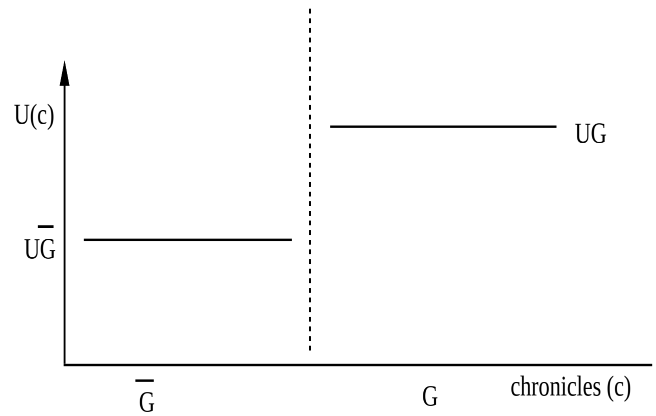


Figure 2: Step utility function

sufficient for success. More realistic is the view that satisfying some of the conjuncts is preferable to satisfying none, motivating the idea that success in achieving one conjunct could be traded off against success in achieving others.

3. Partial satisfaction
   Symbolic goals imply an all-or-nothing success criterion, either the goal form holds at the end of execution or it does not. This criterion is reflected in the step function: utility is either at a constant high value or at a constant low value. A realistic representation should allow the manager to communicate the fact that satisfying G is most preferred, but satisfying G partially is better than not satisfying it at all.

4. Incidental costs and benefits
   Symbolic goals imply that the goal attributes are the only aspects of an outcome that are relevant to assessing utility, which rules out the possibility that plan $\mathcal{P}_1$ is preferable to $\mathcal{P}_2$ because it is as effective at achieving the goal as $\mathcal{P}_2$, but does so more cheaply. Symbolic goals provide no way to specify the "more cheaply" part, nor do they provide the language to express the tradeoff between effectiveness in achieving the goal and the cost involved in doing so.

The analysis in this section demonstrates the steps necessary to unify goal-directed and decision-theoretic plan evaluation:

- The language of goals needs to be extended to represent partial goal satisfaction and resource-related utility.

- In order to effectively use these extended goal forms (more complicated utility functions) in planning we need to establish relationships like Equation 3: circumstances under which planning to maximize the probability of goal satisfaction ensures rational behavior in the decision-theoretic sense.

- We need to exploit these relationships as we build or refine plans.

This paper will have four concerns:

- Presenting a framework for analyzing a goal-directed agent's utility function, which includes goal and resource components.
- Defining a language for goals that provides for expressing preferences among situations involving partial satisfaction of the goals.
- Developing relationships that allow the agent to build plans based on the symbolic content of its goals, while simultaneously guaranteeing that following those plans will cause it to act so as to maximize expected utility.
- Demonstrating how those relationships can be exploited by a planning algorithm.

3 Utility Models for Goal-Directed Agents

This section defines a utility model for a goal-directed agent by describing the form of the function $\mathbf{U}(c)$ mentioned in the previous section. The task is simple for the simple goal model in the previous section:

$$\mathbf{U}(c) = \begin{cases} 1 & \text{if } \mathit{HOLDS}(\mathsf{G},\text{end}(\mathsf{c}),\text{end}(\mathsf{c})) \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

But we want our model to capture the fact that goal satisfaction can be measured at other time points and over intervals, that goals can be partially satisfied, that (partial) satisfaction of one goal can be traded off against (partial) satisfaction of other goals, and that resource costs affect the extent to which a plan succeeds as well.

In that case what can we say about a goal's role in the utility function? There are two main properties we want to capture:

1. That satisfying a goal to a greater extent is preferable to satisfying it to a lesser extent, all other things being equal[2].
2. That success in satisfying one goal component can be traded off against success in satisfying another goal, or against consuming resources.

We begin by defining a goal-oriented agent in terms of its top-level utility function:

Definition 1 *A* goal-directed agent *is a decision maker whose preferences correspond to the following utility function:*

$$\mathbf{U}(c) = \Sigma_{i=1}^{n} k_i \mathsf{UG}_i(c) + k_{\mathrm{r}} \mathsf{UR}(c) \tag{5}$$

[2]See [Wellman and Doyle, 1991] for a more general interpretation of this criterion.

The utility function is defined in terms of n subutility functions UG_i associated with the agent's n explicit goals. Each of these functions maps the chronicle into a real number between 0 and 1, where 0 represents the situation in which the goal is satisfied not at all in the chronicle and 1 indicates that the chronicle satisfies the goal fully.

The function UR (for "residual" or "resource" attributes) also maps a chronicle into $[0, 1]$, and measures the extent to which the chronicle produces or consumes non-goal attributes, e.g. time, fuel, wear and tear on equipment, money lost or gained. A value of 1 indicates the best possible use of non-goal attributes; a value of 0 indicates that resources were consumed at the theoretical maximum.

The model is also defined in terms of $n + 1$ numeric parameters: $k_1, k_2, \ldots k_n, k_r$. These parameters make explicit the tradeoffs among the component goals and between the goals and resource consumption. The ratio of any two of these numbers indicates the amount the agent is willing to sacrifice in satisfaction of one goal in order to satisfy another, or the amount of satisfaction in resource consumption he is willing to "spend" in order to accomplish an increase in satisfaction for a goal. We return to these tradeoffs in Section 7.

Equation (5) also implies that the agent's preferences over the component goals are *additive independent*: preferences over lotteries on the goals depend only on the marginal probability distributions for the goals and not on their joint probability distribution [Keeney and Raiffa, 1976, Sect 6.2]. In other words we assume that preference for a particular level of satisfaction for one goal does not depend on the extent to which the other goals are satisfied. The function furthermore implies that goal satisfaction is additive independent in resource consumption: the agent's preferences over patterns of resource consumption are the same regardless of the extent to which the top-level goals are satisfied.

The power of the independence assumption is that it allows us to reason easily about the tradeoff between levels of the various goals, and also about how much additional resource we would be willing to expend to improve the chances of satisfying a goal or to satisfy it more fully (see Section 7).

The independence restriction on top-level goals seems troubling on the surface, since AI planning research has focused mainly on interactions among the goals, and our assumption seems to rule out those interactions. In particular the assumption runs counter to the analysis of conjunctive-goal planning we presented in the previous section. Equation (5) means, for example, that we cannot represent conjunctive goals like "have the truck fueled and loaded by 7AM" as two separate top-level goals, since presumably satisfying either one without the other affords low utility but their conjunction affords high utility. This is indeed the case, and our argument is that these two statements do not involve two goals, but rather two components of a *single* goal. Section 4.1.1 addresses the problem of dealing with interactions among symbolic interacting components that comprise a goal—that section demonstrate how to represent traditional goals of the form "G is satisfied only if its components $\mathsf{g}_1 \ldots \mathsf{g}_n$ are all satisfied."

We should also note that the assumption of utility independence does not imply that that goals are *probabilistically* independent. One might object that two goals, say "have the truck at the depot by noon" and "have the truck clean," interact strongly if the only road to the depot is muddy and there's no way to wash the truck once at the depot. In particular there's no plan that might make them both true. But this is not a violation of *probabilistic* independence, not *utility* independence. The assumption of additive utility independence implies only that the *utility* derived from satisfying one top-level goal does not depend on the extent to which the other goals are achieved; it does not comment on the likelihood of achieving either in isolation or both simultaneously. In this example there might be no chronicle in which both propositions are true, in which case the likelihood of achieving the goal, and presumably the expected utility of any plan, will be low. But this interaction is properly reflected in the probabilistic model of the domain and the operators—it does not involve the agent's preferences.

The main implication of the additive independence is that the decision maker has to structure his preferences, identifying those that are utility independent and those that are not. The former are divided into separate goals and the latter become components of individual goals. Subsequent sections provide methods for describing interactions among goal components.

We now turn to the question of how the goal information—the $\mathsf{UG}_i(c)$ functions—is expressed. We will temporarily ignore the top-level utility function $\mathsf{U}(c)$ as well as the residual utility function $\mathsf{UR}(c)$, and concentrate on how to build utility functions for individual goals.

3.1 Syntactic forms for goals

Goals in classical planning algorithms consist of a symbolic expression to be achieved. We want to extend the idea in several directions:

- Goals should have a temporal extent—a time at which or an interval over which the proposition is to be achieved. The classical goal representation has the planner achieve the goal by the end of plan execution.

- Goals should be partially satisfiable—if the goal is to produce 10 widgets it might be better to produce 5 than none at all. Partial satisfaction can extend to the temporal component as well: if the deadline is noon, finishing by 12:05 might be better than finishing the next morning.

3.1.1 The language $\mathcal{L}_{tca}$

In order to talk about temporally qualified sentences in a probabilistic setting, we need a logic that can represent both time and probability. The logic of time, chance, and action $\mathcal{L}_{tca}$ is well suited to our purposes. A simplified version of the logic is described in [Haddawy, 1991a] and the full logic is described in [Haddawy, 1991b]. We describe here just that portion of the language relevant to this paper. We will use the single predicate *HOLDS* to associate facts with temporal intervals. The sentence $HOLDS(FA, t_1, t_2)$ is true if fact *FA* holds over the time interval t_1 to t_2. We impose the constraint that if a fact holds over an interval it holds over any subinterval.

Probability is treated as a sentential operator so it can be combined freely with other logical operators. We write $\mathbf{P}_t(\phi)$ to denote the probability of a formula ϕ at time t. The probability of a formula is defined as the probability of the set of chronicles in which the formula is satisfied. Although the language can represent the dynamics of probability over time by allowing the probability operator to be indexed by any time point, in this paper we will only index it by the current time (*now*) and sometimes will omit the index altogether: $\mathbf{P}(\phi)$ understood to mean $\mathbf{P}_{now}(\phi)$.

3.1.2 Goal expressions

We begin by breaking a goal into *atemporal* and *temporal* components. The former indicates *what* is to be achieved, the latter *when* it is to be achieved.

We define two types of temporal goals: deadline goals and maintenance goals. A deadline goal is a function of the deadline time point t_d, and its atemporal component is just a formula ϕ containing only the *HOLDS* predicate with only temporal parameter t_d. For shorhand we will notate a formula ϕ that contains only temporal parameters t, t' as $\phi(t, t')$. A deadline goal says only that ϕ should hold at the deadline point, but we will discuss below ways of expressing preferences over making ϕ "partially true" at t_d or making ϕ true "shortly after" t_d, or both. Two examples of deadline goals are:

- Have block A on block B on block C by noon.
  $\mathit{HOLDS}(\text{on(A,B)}, \mathit{noon}, \mathit{noon}) \wedge \mathit{HOLDS}(\text{on(B,C)}, \mathit{noon}, \mathit{noon})$

- Have two tons of rocks to the depot by 2:00 this afternoon.
  $\mathit{HOLDS}(=\!(\text{tons-delivered-to-depot,2}), \mathit{2PM}, \mathit{2PM})$

A maintenance goal represents a desire to keep a proposition true over an interval of time. It is a formula containing only the *HOLDS* predicate with temporal arguments t_b and t_e, the begin and end points of the maintenance interval. An example of a maintenance goal would be "keep the temperature between 65 and 75 degrees from 9am until 5pm:

$$\mathit{HOLDS}(\geq (\text{temp}, 65), \text{9am}, \text{5pm}) \wedge \mathit{HOLDS}(\leq (\text{temp}, 75), \text{9am}, \text{5pm})$$

3.1.3 Utility function for individual goals

The i^{th} individual goal appears in the top-level utility function as a function $\mathsf{UG}_i(\text{c})$ of a chronicle. Above we made the distinction between a goal's temporal and atemporal components, and that distinction is reflected in the UG_i function as well. We will explore this function in more detail below, but begin by dividing it into two components:

- A function $\mathsf{DSA}_i(t)$ which is a measure of the extent to which the *atemporal* component of the i^{th} goal is satisfied at time t (where t is implicitly a member of some chronicle c). DSA stands for "degree of satisfaction of the atemporal component."

- A function that defines the extent to which the goal's temporal component is satisfied, which depends on the form of the temporal component. For deadline goals we define a temporal coefficient $\mathsf{CT}_i(t)$, measuring the extent to which the deadline is satisfied at time t. For maintenance goals we define a persistence coefficient $\mathsf{CP}_i(t, t')$ measuring the extent to which the maintenance condition is satisfied over the interval t, t'.

Subsequent analysis addresses how the DSA, CT, and CP functions are specified and exactly how they are combined to form the utility function for the i^{th} goal, $\mathsf{UG}_i(c)$.

We next address the problems associated with specifying and reasoning with preferences over partial satisfaction of goal forms.

4 Partial Goal Satisfaction

So far we have defined the syntactic goal expressions in terms of an atemporal formula and a temporal parameter (time point or interval). We still need a language for expressing preferences over partial satisfaction of both. Partial satisfaction of the atemporal component might be possible by satisfying most but not all of the members of a conjunctive goal or by almost satisfying a numerical equality or inequality constraint. Partial satisfaction of the temporal component might be possible by achieving the atemporal component at a time soon after the deadline point, or keeping it true through most of the maintenance interval.

We start by considering in turn partial satisfaction of the atemporal component of the goal and partial satisfaction of the temporal (deadline or maintenance) component.

4.1 Partial satisfaction of the atemporal component

In the next sections we show how to specify partial satisfaction of two types of atemporal components:

- symbolic attributes—a conjunction of symbolic propositions like "a big red cylinder on the table," and
- quantitative attributes—the value of a real- or integer-valued quantity like the truck's fuel level or the number of items in a box.

In focusing on the atemporal goal component we are addressing the question of what form the goal's atemporal component, $\mathsf{DSA}_i(t)$, should take.

4.1.1 Goals with symbolic attributes

Here we consider how to define a function specifying partial satisfaction of symbolic-attribute goals. A symbolic attribute is any logical formula containing only the *HOLDS* predicate. For example, the (deadline) goal of having block A on top of a red block by noon would be represented as

$$\exists x\ HOLDS(On(A, x), noon, noon) \wedge HOLDS(Red(x), noon, noon)$$

We want to represent situations like one in which it is important to have A on an object, but perhaps less important that the object be red.

The degree of satisfaction (DSA) function for a symbolic-attribute goal is defined in terms of an application-supplied sequence S of mutually exclusive and exhaustive formulas $(\sigma_1, \sigma_2, ..., \sigma_n)$ such that

- σ_n is the actual atemporal component of the goal (thus representing complete satisfaction), and
- σ_i represents a greater degree of satisfaction than σ_j if $i > j$.

The application also provides a function $\mathsf{dsa}(\sigma)$ that associates a degree of satisfaction value with each σ_i. The function must have the property that $\mathsf{dsa}(\sigma_1) = 0$ and $\mathsf{dsa}(\sigma_n) = 1$. For the above example S might be

i	σ_i	$\mathsf{dsa}(\sigma_i)$
1	$\neg\exists x HOLDS(\mathrm{On}(A, x), t, t)$	0.0
2	$\exists x HOLDS(\mathrm{On}(A, x), t, t) \wedge \neg HOLDS(\mathrm{Red}(x), t, t)$	0.7
3	$\exists x HOLDS(\mathrm{On}(A, x), t, t) \wedge HOLDS(\mathrm{Red}(x), t, t)$	1.0

(Note that t is a free variable in these expressions; it is not the deadline point. We will discuss below the matter of what times the function should be evaluated at.)

The simplest such function would be one that admits no partial satisfaction of the goal. Recall the example from Section 3, to have the truck loaded and fueled by 7AM, where accomplishing one without the other yields no utility. The dsa function would be

i	σ_i	$\mathsf{dsa}(\sigma_i)$
1	$\neg(HOLDS(\text{truck-loaded}, t, t) \wedge HOLDS(\text{truck-fueled}, t, t))$	0.0
2	$HOLDS(\text{truck-loaded}, t, t) \wedge HOLDS(\text{truck-fueled}, t, t)$	1.0

Note that as we suggested, some partial satisfaction is accrued from making the *ON* relationship true even without the *RED*, but no satisfaction is accrued if *RED* is true without *ON*.

To be clear about the notation: the function DSA_i for the i^{th} goal is a function of a time point. It is defined in terms of a function dsa_i which is a function of the goal's atemporal component, e.g. a symbolic goal component. We integrate the two by taking $\mathsf{DSA}_i(t) = \mathsf{dsa}_i(\phi)$, where ϕ is the (unique) formula among the σ_i that is true at time t. There must be a unique such formula since the σ_i are mutually exclusive and exhaustive.

For some specialized types of symbolic goal structures the degree of satisfaction function can be defined more succinctly than by supplying the full table of σ_i and dsa values. For example, if the atemporal component is a conjunction of atomic formulas not sharing any variables, and if preferences over changes in the degree of satisfaction associated with each of the conjuncts are additive independent, then we can define the degree of satisfaction for each conjunct individually, and the utility associated with a time point is the sum of the utilities taken over all the conjuncts.

4.1.2 Goals with quantitative attributes

A quantitative attribute is a special kind of symbolic attribute: a logical formula containing only the *HOLDS* predicate in which the first argument expresses equality or inequality between a term and a numeric quantity. An example of a quantitative-attribute deadline goal is to have two tons of rocks at the depot by noon:

$$HOLDS(= (\text{tons-rocks-delivered-to-depot } 2), noon, noon)$$

The dsa function for such a goal could in principle be defined in the same way as we did for symbolic attributes above, but such a definition would be unmanageably large. Since the degree of satisfaction for quantitative attributes is simply a function of the quantity's magnitude, we can define the dsa function directly in terms of that magnitude. The degree of satisfaction for the goal to deliver a ton of rocks would simply be a function of the quantity of rocks delivered, e.g. $\mathsf{dsa}(r) = 1 - \frac{2000-r}{2000}$ where r is the weight of rocks delivered, measured in pounds. We again define $\mathsf{DSA}(t) = \mathsf{dsa}(r)$, where r is the number of rocks delivered at time t.

4.1.3 Conjunctive quantitative attributes

Things get a little more complicated when a goal's temporal component consists of a conjunction of quantitative attributes. Suppose we have the goal of delivering 12,000 pounds of polystyrene, 480 pounds of colorant, and 120 pounds of UV stabilizer to a plastics manufacturing plant by time t_2:

$$\begin{array}{l} HOLDS(= (\text{poly-delivered } 12{,}000), t_2, t_2) \ \wedge \\ HOLDS(= (\text{colorant-delivered } 480), t_2, t_2) \ \wedge \\ HOLDS(= (\text{UV-delivered } 120), t_2, t_2) \end{array}$$

These quantities were not chosen arbitrarily; they represent the quantities of materials needed to manufacture 2000 units of a particular product. They need to be present at the plant in a particular proportion in order to be useful. That is to say, only that quantity that is present in the right proportion can be used. This is a common characteristic of conjunctive quantitative attributes. Consequently, degree of satisfaction will be a function of the maximum amount of the materials that have been delivered in the required proportion. In this case, the necessary proportion is 100:4:1. So to derive the degree of satisfaction of a time in a chronicle, we let

$$q = min(x/100, y/4, z)$$

where x,y, and z are the quantities of polystyrene, colorant, and UV stabilizer, respectively. Then if we require 6 pounds of polystyrene to manufacture one unit, the degree of satisfaction would be some nondecreasing function of $\lfloor 100q/6 \rfloor$, normalized to range between 0 and 1.

4.1.4 Summary

We have defined the atemporal part of a goal's component in the utility function, for symbolic, quantitative, and conjunctive quantitative goals. This representation applies both to deadline and maintenance goals. The function $\mathsf{DSA}_i(t)$ measures the extent to which the atemporal component is satisfied at a particular time point (implicitly within a particular chronicle). In each case $\mathsf{DSA}(t)$ is defined in terms of an application-supplied function $\mathsf{dsa}(\phi)$, where ϕ is specific to the goal's atemporal form: for symbolic goals the programmer supplies a dsa function in the form of a table with entries for various combinations of the symbolic components, along with a partial-satisfaction number for each. For quantitative attributes he supplies a dsa function directly in terms of the quantity's value.

4.2 Partial satisfaction of the temporal component

Each goal's utility function also contains a weighting coefficient, CT_i for deadlines and CP_i for maintenance intervals, measuring the extent to which the deadline or maintenance interval was respected. CT_i is a function of a time point within a chronicle and measures the extent to which that time point meets the deadline. CP_i is a function of two time points, and measures the extent to which the maintenance interval was respected. These coefficients are specified independent of the goals' atemporal components. We now consider in turn partial satisfaction of deadline and maintenance constraints.

4.2.1 Deadlines

We first have to be precise about what is meant by a goal involving a deadline, for example "have the report on my desk by 10 tomorrow morning." We interpret the goal as a transfer of control over a resource, in this case the report. The deadline is the *earliest* time at which the transfer has value to the agent; in other words I am saying that I will be able to use the report at 10, but no earlier. So there is no utility associated *directly* with making the goal true before the deadline; there might or might not be utility associated with making it true after the deadline.

The word "directly" refers to the fact that there is obviously *some* advantage to delivering the report early. But that advantage is indirectly accrued: planning to have the report delivered an hour early makes it more likely that it will indeed be available at the deadline; in most cases the more slack built into the plan, the more likely the plan is to meet the deadline, even if things get behind schedule.

Of course there are also circumstances in which it is best to accomplish the goal as close to the deadline as possible. Perhaps the report is likely to be stolen if it is delivered early, or perhaps it has to be stored between the time it is delivered and the deadline, and storage incurs a cost. Both of these effects are indirect too, however, and should not be a part of the goal-related utility measure. Suppose that the probability of theft increases with the amount of time the report is delivered before the deadline. In that case projecting the plan will generate a chronicle in which the report is stolen and chronicles in which it is not. The

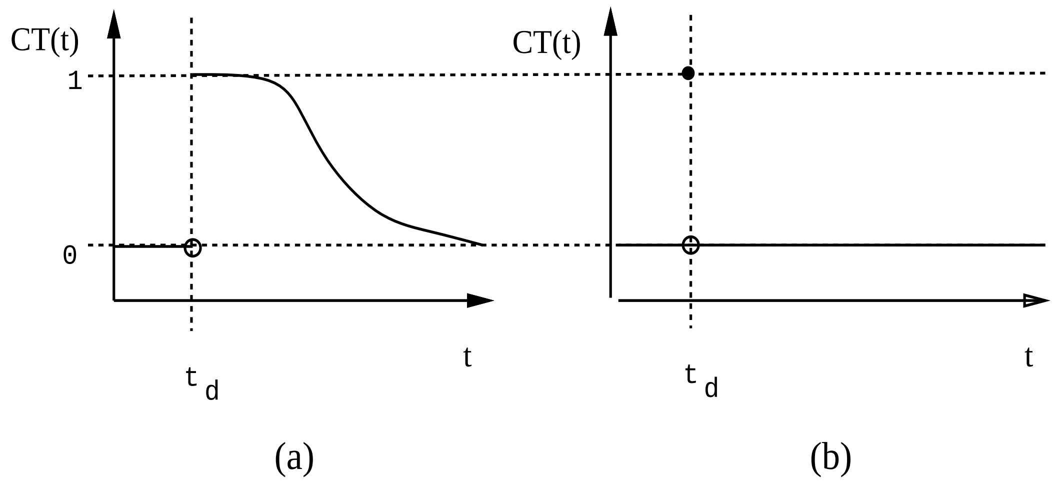


Figure 3: Two partial-satisfaction functions for deadlines

chronicle in which the report is stolen will not satisfy the goal at all, thus will have low utility. The chronicle in which it is not stolen will satisfy the goal fully (since the report will be available at the deadline) and will have high utility. Delivering the report earlier will increase the probability of the chronicle in which the report is stolen, thus will lower the expected utility of the plan, all else being equal.

In the second case the cost of storing the report is a resource, which is measured in the UR utility function. The earlier it is delivered the more storage cost is incurred, again lowering the utility of the plan.

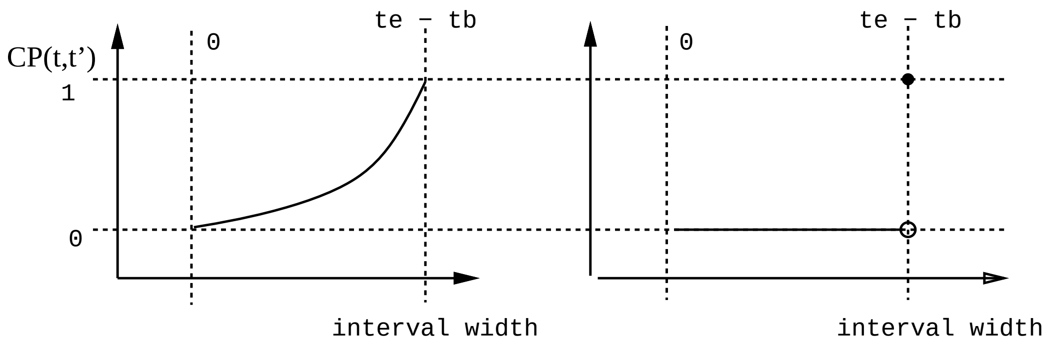
To summarize: the deadline represents the earliest point at which satisfying the goal has any direct value. At the deadline point itself the temporal component is fully satisfied. There might or might not be utility in satisfying the goal after the deadline—that will depend on the goal itself.

The coefficient $\mathsf{CT}_i(t)$ measure the degree to which the deadline is considered to be satisfied at time t. Our analysis requires that for a goal i with a deadline point t_d,

- $\mathsf{CT}_i(t) = 0$ for all $t < t_d$.
- $\mathsf{CT}_i(t_d) = 1$
- $\mathsf{CT}_i(t)$ is a nonincreasing function of t for all $t \geq t_d$.

If the deadline is absolute, then $\mathsf{CT}_i(t)$ will be 0 for all points $t \neq t_d$. CT_i need not be strictly decreasing. A flat region on the curve over some time interval after the deadline represents indifference about which time in that region the goal is satisfied.

Figure 3 shows two examples: the first (a) shows a situation in which full satisfaction is guaranteed for some period of time after the deadline, then decreases thereafter. The second (b) shows an absolute deadline: satisfying the goal is useful at the deadline point but at no others.



(a) (b)

Figure 4: Two partial-satisfaction functions for maintenance goals

4.2.2 Maintenance intervals

A maintenance goal specifies that a condition must hold throughout an interval, t_b, t_e. The analogue to the temporal weighting coefficient CT for a deadline goal (evaluated at a time point) is a function $\mathsf{CP}_i(t, t')$ over *subintervals of* $(t_b,\ t_e)$ that defines to what extent the interval satisfies the maintenance condition. We consider only the case in which CP_i is a function of the width of the interval, and require it to range between 0 and 1. Two analogous situations to Figure 3 would be (a) one in which satisfaction of the maintenance restriction declines gradually with the percentage of the interval covered, and (b) one in which *any* violation of the atemporal component renders the goal totally unsatisfied. Figure 4 shows these two cases.

4.2.3 Summary

At this point we have described the additional information the application must provide to define partial goal satisfaction for individual deadline and maintenance goals. For each goal i we require:

1. For the goal's atemporal component, a function $\mathsf{DSA}_i(t)$ that defines the extent to which the goal's symbolic component is satisfied at that point in time. $\mathsf{DSA}_i(t) = 0$ indicates no satisfaction of the goal formula at t; a value of 1 indicates complete satisfaction. DSA in turn is defined in terms of an application-supplied function dsa specific to the goal's atemporal content. For symbols this information will take the form of a table associating a number between 0 and 1 for formulas that represent partial satisfaction of the goal formula. For quantitative attributes the dsa function can be defined directly in terms of the attribute itself.

2. For the temporal component, either

- a function $\mathsf{CT}_i(t)$ that specifies the extent to which time t satisfies the deadline, or
- a function $\mathsf{CP}_i(t, t')$ that specifies the extent to which the interval (t, t') satisfies the maintenance requirement.

The utility function UG_i for an individual goal is a combination of these two functions, and the next section confronts the problem of how and when to combine them.

5 Utility Functions for Individual Goals

We have now supplied functional forms for describing satisfaction of a goal's atemporal and temporal component. The former is supplied in the form of a function $\mathsf{DSA}_i(t)$, defined in terms of the atemporal attribute. The DSA function varies between 0.0 and 1.0, representing respectively no satisfaction and complete satisfaction. The latter is supplied in terms of a weighting coefficient, either $\mathsf{CT}_i(t)$ or $\mathsf{CP}_i(t_1, t_2)$ depending on whether the goal is a deadline or a maintenance interval.

5.1 Combining the temporal coefficient and the atemporal degree of satisfaction

Both of the partial-satisfaction functions are functions of time points, whereas the goal's contribution to the utility function (the UG_i function) is defined in terms of an entire chronicle. To form the utility function for a deadline goal we need to evaluate and combine DSA and CT values at selected time points within the chronicle and to form the utility function for a maintenance goal we need to evaluate and combine the DSA_i and CP_i values over selected intervals within the chronicle.

The real problem therefore is *when* to evaluate the functions, and this is a difficult question only when partial satisfaction is allowed both in the atemporal and temporal components. Take a deadline goal, for example—if we were to allow partial satisfaction of the atemporal component but *not* the temporal component, we could just evaluate the atemporal goal expression DSA at the deadline point. Likewise, if we allow partial satisfaction of the temporal component but not the atemporal component then we could just evaluate the CT component at the earliest time point (at or after the deadline) at which the atemporal component is fully satisfied. But suppose that we allow partial satisfaction in both components—at what time point(s) should we evaluate and combine the DSA function and the CT coefficient?

5.1.1 Deadline goals

We motivate our solution with an example. Suppose we are employing a delivery agent whose task it is to deliver two tons of rocks to the depot by 2PM. Several trips might be required. Abstractly we can think about how to structure payments to the driver so that if he acts rationally he will act so as to maximize our utility.

A reasonable way to reward the driver is to pay him in proportion to each quantity of rocks he delivers on each trip, up to a total of two tons. The pay for each delivery would be discounted by an amount proportional to the amount of time by which each delivery misses the deadline. The amount the driver gets paid is then the sum of the rewards for all the deliveries he makes up to two tons. If the driver acts to maximize his expected pay, we will also be maximizing our utility relative to the goal.

Even though the deadline goal is stated in terms of the *level* of a quantity, it is important to note that it is proper to reward the driver for the quantities *delivered*, and not for the level attained. Otherwise the driver would be penalized if rocks were removed. But we also have to make sure that there is no incentive for the driver to remove rocks then immediately deliver them. So the proper reward structure is to pay for deliveries that increase the attribute's level of satisfaction. A delivery is analogous to an *increase* in the dsa value of the atemporal component, not directly to a dsa value.

We also make the following assumptions about changes in atemporal utility:

- All preferences for lotteries over changes in dsa at any time t are the same as preferences for lotteries at any other time t'.

- There are a countable number of changes in the dsa over the course of a chronicle.

Under these assumptions the appropriate reward structure for the agent—and by analogy the proper expression for the goal's utility—is an additive utility function:

$$\begin{aligned}\mathsf{UG}_i(c) = {} & \mathsf{DSA}_i(t_d) + \\ & \sum_{\{t>t_D:\ \neg\exists t'\ (t_D\leq t'<t)\wedge \mathsf{DSA}_i(t')\geq \mathsf{DSA}_i(t)\}} (\mathsf{DSA}_i(t) - max_{t_D\leq \hat{t}<t}\mathsf{DSA}_i(\hat{t}))\cdot \mathsf{CT}_i(t)\end{aligned} \tag{6}$$

where t_d is the time of the deadline. The intuition is that we accrue utility for the level of satisfaction that is true *at* the deadline point (thus rewarding early satisfaction), and also for every time the DSA function *increases* over a previous value. The amount of utility accrued for a change is the increase in atemporal utility over the previous maximum, weighted by the CT function that discounts the change according to how well it satisfies the deadline.

The basic form of this utility function is similar to that presented by Meyer [Keeney and Raiffa, 1976, Sect. 9.3.2] for the utility of a consumption stream. The only difference is that Meyer's formulation defines utility directly in terms of consumption—which is the change in the agent's wealth—while our DSA functions specify this change in utility indirectly. Another minor difference is that Meyer's formulation allows consumption to be negative, indicating a loss of wealth, whereas we are only interested in positive changes in degree of satisfaction.

Consider, for example, the goal to have two tons of rocks at the depot by noon. We can use one of two trucks. The big truck carries two tons in one load but is slow: it will get all two tons to the depot by 2pm. The small truck carries only one ton but is fast: it will get one ton there by 1pm and two tons there by 2pm. Which plan affords higher utility? The answer depends, of course, on the utility decay associated with missing the deadline compared to the utility decay associated with missing some rocks. Suppose that the degree of satisfaction of zero tons is 0.0, the degree of satisfaction of one ton is 0.5 and the degree of satisfaction of two tons is 1.0. Suppose further that the temporal coefficient is a linearly decreasing function that goes from 1.0 at noon to 0.0 at 3pm. So $\mathsf{CT}(t) = 1 - t/3$, where t is the number of hours past noon. Based on these values, the utility of the chronicle that results from using the big truck is $(1)(1/3) = 1/3$ and the utility of the chronicle that results from using the small truck is $(1/2)(2/3)+(1/2)(1/3) = 1/2$. Hence we prefer to use the small truck over the big truck. This preference fits our intuition since based on the specifications of the atemporal utility function and the temporal coefficient, we associate some benefit with having some portion of the two tons of rocks at the depot earlier.

This example is an instance of a general probabilistic dominance relation that holds for deadline goal utility functions as defined above. If

$$\forall i, t, n.\, \mathbf{P}(\{c : \mathsf{DSA}_i(t) \leq n\}|\mathcal{P}_1) < \mathbf{P}(\{c : \mathsf{DSA}_i(t) \leq n\}|\mathcal{P}_2),$$

where i ranges over all goals and $\{c : \mathsf{DSA}_i(t) \leq n\}$ denotes the set of chronicles in which $\mathsf{DSA}_i(t) \leq n$, then $\mathcal{P}_1$ has higher expected utility than $\mathcal{P}_2$ and is thus preferred.

5.1.2 Maintenance goals

A maintenance goal states that a condition should be maintained over an interval of time. Partial *atemporal* satisfaction of a maintenance goal is expressed (as in deadline goals) in terms of the degree of satisfaction of the atemporal component. Partial temporal satisfaction is expressed in terms of the length of the interval or intervals over which the atemporal component is partially satisfied. Partial temporal satisfaction can be a non-linear function of interval length. For example, we may wish to keep a machine running from 9:00 until 5:00 and if it has a production cycle time of 30 minutes, we may not accrue any reward for having it running for intervals shorter than 30 minutes. So we define a persistence coefficient function CP that maps a temporal interval into a number between zero and one.

The question again is how to combine the atemporal and temporal satisfaction into a utility function for the maintenance goal. We need to measure how long the atemporal component persists at each level of partial satisfaction. Since we are measuring the amount of time over which a quantity persists, we need to fix the quantity and find the intervals over which the quantity is at that level. If we have a DSA function that is only either 0 or 1, we simply sum the CP values of the intervals over which the atemporal component is satisfied. If the DSA can have intermediate value the procedure is roughly to integrate over all possible DSA values and find intervals over which DSA is maintained at a given level.

Notice that if the satisfaction level is low over some interval and higher over a second interval then the atemporal component was at least at that low level of satisfaction over both intervals. For example, suppose that the DSA is above some high level, say 0.8, throughout the interval of interest but that it fluctuates above that level at some high frequency. Suppose further that our CP function assigns zero satisfaction to any interval shorter than half of the interval of interest. If we measure the utility in terms of the intervals over which the DSA is at a given level, we would wind up assigning zero utility to this chronicle even though the

DSA is above 0.8 for the entire interval. So the following expression assigns utility in terms of the intervals over which the DSA is maintained *above* each possible value.

$$\mathsf{UG}_i(c) = \int_0^1 \sum_{\{I:\, \forall t \in I\, \mathsf{DSA}_i(t) \geq x \,\wedge\, \neg \exists I' \supset I\, \forall t \in I'\, \mathsf{DSA}_i(t) \geq x\}} \mathsf{CP}_i(I)\, dx \tag{7}$$

where I and I' denote temporal intervals. Notice the symmetry between the above expression and that for deadline goals. To compute the utility of a deadline goal, we sum changes in DSA over time. For maintenance goals, we sum time over DSA i.e., we sum the CP of intervals over which DSA exceeds a given value over all values of DSA.

5.2 Summary

This section completes the definition of utility functions for individual goals, providing a definition for the UG_i functions. The main problem we solved is how to combine partial satisfaction of a goal's temporal component with partial satisfaction of a goals' atemporal component.

For deadline goals the problem amounts to deciding *when* to evaluate the DSA function—the degree of satisfaction of the goal's atemporal component. The main idea is to evaluate the DSA function once at the deadline point, and then at every time point at which the DSA function increases. At every such point the *increase* in atemporal satisfaction is weighted by the temporal coefficient CT_i.

For maintenance goals combining the temporal and atemporal components involves identifying the intervals of DSA over which to evaluate CP. The main idea is to evaluate the CP of the longest continuous intervals over which the DSA exceeds each possible value.

Now that we have the basic form of the agent's top-level utility function we turn to the problem of how to use that function to establish qualitative differences in the quality of alternative plans.

6 Using the Utility Functions to Rank Plans

One of the main goals of the present work is to use information in the utility function's symbolic structure to guide the building of good plans, which generally will involve demonstrating that one plan is preferable to another. At worst establishing this relationship involves computing the expected utility of both plans, a process that requires generating and evaluating all possible chronicles for each.

By exploiting the structure of the utility functions we can establish the same relationships without performing the full expected-utility calculation. We do so by establishing relationships among the individual goals' symbolic components that indicate corresponding relationships among the corresponding utility functions. Here is an abstract characterization of the sorts of relationships we will establish:

Suppose that one of the agent's goals is g and that there are two formulas ϕ and ψ that bear some relation to g. More particularly, the truth of ϕ indicates a *high* goal-related utility (DSA) whereas the truth of ψ indicates a *low* DSA value. Further suppose that there are two alternative plans, $\mathcal{P}_1$ and $\mathcal{P}_2$. If the probability that ϕ is true by some time t_1 given $\mathcal{P}_1$ is at least α, and if the probability that ψ is true at all times before t_2 given $\mathcal{P}_2$ is at least β, and if α exceeds some function of β, $\mathsf{dsa}(\phi)$, $\mathsf{dsa}(\psi)$, t_1, and t_2, then $\mathcal{P}_1$'s expected utility is guaranteed to be greater than $\mathcal{P}_2$'s.

Having established a relationship of that form the planner need only establish two probability bounds associated with propositions ϕ and ψ in order to eliminate $\mathcal{P}_2$ from further consideration.

Two advantages accrue from this technique:

1. it reduces the general problem of expected-utility calculations to the more specific task of deciding whether a particular probability exceeds a particular threshold, and

2. it allows us to perform the expected-utility analysis incrementally—at each stage of the planning process we can eliminate some plans from consideration, again limiting the amount of inference needed to choose a good course of action.

We will demonstrate these advantages in Section 7, but first establish relationships for the goal types we defined in the previous section.

6.1 Deadline goals

We first derive results for deadline goals, attacking symbolic atemporal attributes, then quantitative atemporal attributes.

6.1.1 Bounds for symbolic attributes

We are comparing two plans on the basis of their performance on a goal whose temporal component is a deadline t_d and whose atemporal component is symbolic. We have a formula ϕ that indicates a high level of goal satisfaction (dsa value) and a formula ψ with a low dsa value. Consider the case in which $\mathcal{P}_1$ is likely to achieve ϕ at some time at or close to the deadline, and $\mathcal{P}_2$ is likely to achieve at best ψ only some time well after the deadline. Under what circumstances can we say that $\mathcal{P}_1$ has higher expected utility than $\mathcal{P}_2$?

Suppose that the goal's atemporal component is

$$HOLDS(\text{On(A,B)}, t_d, t_d) \wedge HOLDS(\text{On(B,C)}, t_d, t_d)$$

and we define the dsa function as follows:

i	σ_i	$\mathsf{dsa}(\sigma_i)$
1	$\neg\text{On}(\text{B},\text{C})$	0.0
2	$\neg\text{On}(\text{A},\text{B}) \wedge \text{On}(\text{B},\text{C})$	0.5
3	$\text{On}(\text{A},\text{B}) \wedge \text{On}(\text{B},\text{C})$	1.0

In other words we accrue partial satisfaction by achieving On(B,C) alone, but no partial satisfaction by achieving On(A,B) alone. In that case the two formulas might be

$$\begin{aligned}\phi &= \mathit{HOLDS}(\mathrm{On(B,C)}, t, t') \\ \psi &= \neg \mathit{HOLDS}(\mathrm{On(B,C)}, t, t').\end{aligned}$$

Now suppose that for plan $\mathcal{P}_1$ we can find some time point $t_1 \geq t_d$ such that

$$\mathbf{P}(\phi(t_1, t_1)|\mathcal{P}_1) \geq \alpha$$

and for plan $\mathcal{P}_2$ we can find some time point $t_2 \geq t_d$ such that

$$\mathbf{P}(\forall t, t' \, (t_d \leq t \leq t' < t_2) \rightarrow \psi(t, t')|\mathcal{P}_2) \geq \beta$$

.

Under what conditions can we say that plan $\mathcal{P}_1$ is preferable to plan $\mathcal{P}_2$? We must determine the lowest value of $\mathsf{EU}(\mathcal{P}_1)$ and the highest value of $\mathsf{EU}(\mathcal{P}_2)$ consistent with these two constraints. Let σ_ϕ be the formula of lowest dsa consistent with ϕ (in the example $\sigma_\phi = \sigma_3$). We are guaranteed the existence of such a formula since the σ_i are exhaustive. The expected utility of plan $\mathcal{P}_1$ is minimized if

1. with probability α, σ_ϕ is achieved at time t_1 and the goal is not partially achieved at any time earlier than t_1, and

2. with probability $1 - \alpha$ the goal is completely unsatisfied.

So by Equation 6 we have

$$\mathsf{EU}(\mathcal{P}_1) \geq \alpha \cdot \mathsf{dsa}(\sigma_\phi) \cdot \mathsf{CT}(t_1) + (1 - \alpha) \cdot 0$$

Now let σ_ψ be the formula of *highest* dsa consistent with conjunct ψ ($\sigma_\psi = \sigma_2$). The expected utility of plan $\mathcal{P}_2$ is highest if

1. with probability β, σ_ψ is achieved by the deadline and the goal is completely achieved immediately after time t_2, and

2. with probability $1 - \beta$ the goal is completely satisfied at the deadline.

Again by Equation 6:

$$\mathsf{EU}(\mathcal{P}_2) \leq \beta[\mathsf{dsa}(\sigma_\psi) + (1 - \mathsf{dsa}(\sigma_\psi))\mathsf{CT}(t_2)] + (1 - \beta) \cdot 1$$

And we therefore know that $\mathcal{P}_1$'s expected utility is higher than $\mathcal{P}_2$'s if

$$\alpha > \frac{\beta[\mathsf{dsa}(\sigma_\psi) + (1 - \mathsf{dsa}(\sigma_\psi))\mathsf{CT}(t_2)] + (1 - \beta)}{\mathsf{dsa}(\sigma_\phi) \cdot \mathsf{CT}(t_1)} \tag{8}$$

The values for σ_ϕ and σ_ψ and the times t_1 and t_2 will determine how useful our probability bounds are. A t_1 close to the deadline and a ϕ that is inconsistent with low-utility σ_i values will give us a high lower bound on $\mathsf{EU}(\mathcal{P}_1)$. Similarly, a t_2 well after the deadline and a ψ inconsistent with high utility σ_i values will give us a low upper bound on $\mathsf{EU}(\mathcal{P}_2)$.

6.1.2 Bounds for quantitative attributes

Now suppose that our deadline goal is stated in terms of some quantity Q. Assume that the dsa is a monotonically increasing function of the quantitative attribute[3]. Suppose we can establish that $\mathcal{P}_1$ will achieve some high level of Q by some time close to the deadline, and that $\mathcal{P}_2$ can not establish more than some low level of Q until well after the deadline. What do the levels and times have to be in order to conclude that $\mathcal{P}_1$ dominates $\mathcal{P}_2$?

Suppose we know that for plan $\mathcal{P}_1$ we find a time point $t_1 \geq t_d$ and some attribute value k_1 for which

$$\mathbf{P}(\mathit{HOLDS}(\geq (Q,\ k_1), t_1, t_1)|\mathcal{P}_1) \geq \alpha$$

and for plan $\mathcal{P}_2$ we find a time point $t_2 \geq t_d$ and a value k_2 such that

$$\mathbf{P}(\forall t, t'\, (t_d \leq t \leq t' < t_2) \rightarrow \mathit{HOLDS}(\leq (Q,\ k_2), t, t')|\mathcal{P}_2) \geq \beta$$

Under what conditions can we say that $\mathcal{P}_1$ dominates plan $\mathcal{P}_2$? The expected utility of $\mathcal{P}_1$ is lowest if

1. with probability α, we achieve a level k_1 for Q at time t_1 and Q has value zero at all times prior to t_1, and

2. with probability $(1-\alpha)$, Q has its minimum value at all times.

Under those circumstances we know that

$$\mathsf{EU}(\mathcal{P}_1) \geq \alpha \cdot \mathsf{dsa}(k_1) \cdot \mathsf{CT}(t_1)$$

The expected utility of plan $\mathcal{P}_2$ is highest if

1. with probability β, Q attains level k_2 at the deadline and the maximum possible value of Q is attained immediately after t_2, and

2. with probability $1-\beta$ the maximum possible value of Q is attained by the deadline.

Then we know that

$$\mathsf{EU}(\mathcal{P}_2) \leq \beta \cdot [\mathsf{dsa}(k_2) + (1 - \mathsf{dsa}(k_2)) \cdot \mathsf{CT}(t_2)] + (1-\beta),$$

and $\mathcal{P}_1$ is preferred to $\mathcal{P}_2$ if

$$\alpha > \frac{\beta \cdot [\mathsf{dsa}(k_2) + (1 - \mathsf{dsa}(k_2)) \cdot \mathsf{CT}(t_2)] + (1-\beta)}{\mathsf{dsa}(k_1) \cdot \mathsf{CT}(t_1)} \tag{9}$$

[3] A symmetric analysis can be performed for monotonically decreasing atemporal utility.

6.1.3 Bounds for strictly ordered atemporal attributes

Additional structure in the atemporal component can make dominance relationships easier to come by. In this section we will consider a common atemporal component that has an additional structural feature: a ordered conjunction of expressions in which each conjunct dominates subsequent conjunct—any chronicle that satisfies g_1 is preferable to *every* chronicle that does not satisfy g_1, even it satisfies all the others. For example, if our conjuncts are g_1, g_2, and g_3 our strictly ordered atemporal utility function would satisfy the constraint that satisfying g_1 is preferable to not satisfying it, regardless of whether g_2 or g_3 are satisfied:

$$\begin{aligned}
\mathsf{dsa}(g_1 \wedge g_2 \wedge g_3) &> \mathsf{dsa}(\neg g_1 \wedge g_2 \wedge g_3) \\
\mathsf{dsa}(g_1 \wedge \neg g_2 \wedge g_3) &> \mathsf{dsa}(\neg g_1 \wedge g_2 \wedge g_3) \\
\mathsf{dsa}(g_1 \wedge g_2 \wedge \neg g_3) &> \mathsf{dsa}(\neg g_1 \wedge g_2 \wedge g_3) \\
\mathsf{dsa}(g_1 \wedge \neg g_2 \wedge \neg g_3) &> \mathsf{dsa}(\neg g_1 \wedge g_2 \wedge g_3) \\
\\
\mathsf{dsa}(g_1 \wedge g_2 \wedge g_3) &> \mathsf{dsa}(\neg g_1 \wedge \neg g_2 \wedge g_3) \\
\mathsf{dsa}(g_1 \wedge \neg g_2 \wedge g_3) &> \mathsf{dsa}(\neg g_1 \wedge \neg g_2 \wedge g_3) \\
&\vdots \\
\mathsf{dsa}(g_1 \wedge \neg g_2 \wedge \neg g_3) &> \mathsf{dsa}(\neg g_1 \wedge \neg g_2 \wedge \neg g_3)
\end{aligned}$$

and likewise, given that g_1 is true, it's always preferable to satisfy g_2 regardless of g_3's state:

$$\begin{aligned}
\mathsf{dsa}(g_1 \wedge g_2 \wedge g_3) &> \mathsf{dsa}(g_1 \wedge \neg g_2 \wedge g_3) \\
\mathsf{dsa}(g_1 \wedge g_2 \wedge \neg g_3) &> \mathsf{dsa}(g_1 \wedge \neg g_2 \wedge g_3) \\
\mathsf{dsa}(g_1 \wedge g_2 \wedge g_3) &> \mathsf{dsa}(g_1 \wedge \neg g_2 \wedge \neg g_3) \\
\mathsf{dsa}(g_1 \wedge g_2 \wedge \neg g_3) &> \mathsf{dsa}(g_1 \wedge \neg g_2 \wedge \neg g_3)
\end{aligned}$$

Finally we note that if g_1 and g_2 are both true then g_3 is preferable to its negation:

$$\mathsf{dsa}(g_1 \wedge g_2 \wedge g_3) > \mathsf{dsa}(g_1 \wedge g_2 \wedge \neg g_3)$$

Notice that we did not specify that g_2 dominates g_3 in cases where g_1 is false. The reason is that we will often assign zero utility to all states in which the most important goal is not satisfied. For example, given the goal to have the truck fueled and clean by noon, it might do us no good to have it clean if it is not fueled.

The structure of a utility function of this form allows us to prune away suboptimal plans by considering each conjunct individually in the priority order dictated by the dsa values.

As above suppose that we have a time t_1 at which $\mathcal{P}_1$ is likely to achieve g_1, but for $\mathcal{P}_2$ g_1 is liable to be false until at least t_2:

$$\begin{aligned}
\mathbf{P}(g_1(t_1, t_1)|\mathcal{P}_1) &\geq \alpha \\
\mathbf{P}(\forall t, t' \, (t_d \leq t \leq t' < t_2) \rightarrow \neg g_1(t, t')|\mathcal{P}_2) &\geq \beta
\end{aligned}$$

Under what conditions can we say that plan $\mathcal{P}_1$ is preferable to plan $\mathcal{P}_2$? We have no information about the probabilities of the atemporal component's other conjuncts, so a lower bound on the expected utility of $\mathcal{P}_1$ is

$$\mathsf{EU}(\mathcal{P}_1) \geq \alpha \cdot \mathsf{dsa}(g_1 \wedge_i \neg g_i) \cdot \mathsf{CT}(t_1)$$

and an upper bound on the expected utility of plan $\mathcal{P}_2$ is

$$\mathsf{EU}(\mathcal{P}_2) \leq \beta \cdot [\mathsf{dsa}(\neg g_1 \wedge_i g_i) + (1 - \mathsf{dsa}(\neg g_1 \wedge_i g_i))\mathsf{CT}(t_2)] + (1 - \beta).$$

So plan $\mathcal{P}_1$ is preferred to plan $\mathcal{P}_2$ if

$$\alpha > \frac{\beta[\mathsf{dsa}(\neg g_1 \wedge_i g_i) + (1 - \mathsf{dsa}(\neg g_1 \wedge_i g_i))\mathsf{CT}(t_2)] + (1 - \beta)}{\mathsf{dsa}(g_1 \wedge_i \neg g_i)\mathsf{CT}(t_1)} \tag{10}$$

In this case we can eliminate some suboptimal plans using based only on their ability to satisfy the first goal conjunct. Having done so we can then compare the probability of satisfying the first and second conjuncts to the probability of satisfying the first conjunct but not the second, and so forth. We continue until we have incorporated all the conjuncts, at which point we can compute the expected utility of any remaining plans.

This ability to consider the conjuncts sequentially has important consequences for the projection process. [Hanks, 1993] presents a probabilistic projection algorithm demonstrating that:

- it can be much cheaper to project a plan with respect to a single proposition (like one of our goal conjuncts) than it is to reason about all of the plan's effects,
- it can be much cheaper to determine whether the probability of a proposition exceeds a specific threshold than it is to compute the exact value of that probability.

6.1.4 Example

Let's consider a specific example of how these relationships can be used to compare plans. Suppose our goal is to be home by 6:00 with some Thai food and some beer:

$HOLDS$(Loc(me,home), 6:00, 6:00) $\wedge$
$HOLDS$(Possess(me,thai-food), 6:00, 6:00) $\wedge$
$HOLDS$(Possess(me,beer), 6:00, 6:00)

Assume that these propositions persist: any of the goals I achieve before the deadline will be true at the deadline.

Suppose we have the dsa function appearing in Figure 5, and that the temporal coefficient function falls off linearly from 1.0 at 6:00pm to 0.0 at 10:00pm, so $\mathsf{CT}(t) = 1 - t/4$, where t is measured in hours past 6:00pm. Now consider two plans $\mathcal{P}_1$ and $\mathcal{P}_2$ such that for plan $\mathcal{P}_1$ we have established a time t_1 prior to which the formula $\phi =$ Loc(me,home) is likely to be true,

i	σ_i	$\mathsf{dsa}(\sigma_i)$
1	¬Loc(me,home)	0.0
2	Loc(me,home) ∧ ¬Possess(me,thai-food) ∧ ¬Possess(me,beer)	0.4
3	Loc(me,home) ∧ ¬Possess(me,thai-food) ∧ Possess(me,beer)	0.5
4	Loc(me,home) ∧ Possess(me,thai-food) ∧ ¬Possess(me,beer)	0.8
5	Loc(me,home) ∧ Possess(me,thai-food) ∧ Possess(me,beer)	1.0

Figure 5: Example atemporal utility function

and for plan $\mathcal{P}_2$ we have established that will likely not get home ($\psi = \neg$Loc(me,home)) before $t_2 = $ 9:00. The precise formulas are:

$$\begin{aligned} \mathbf{P}(\mathit{HOLDS}(\text{Loc(me,home)}, t_1, t_1)|\mathcal{P}_1) &\geq \alpha \\ \mathbf{P}(\forall t, t' \, (6{:}00 \leq t \leq t' < 9{:}00) \; \rightarrow \neg \mathit{HOLDS}(\text{Loc(me,home)}, t, t')|\mathcal{P}_2) &\geq \beta \end{aligned}$$

In this case we can apply Equation (8) directly to establish that $\mathcal{P}_1$ dominates $\mathcal{P}_2$ if

$$\alpha > \frac{1 - .75\beta}{.4}$$

So if $\beta = .9$ then for any α greater than .8125 plan $\mathcal{P}_1$ will dominate $\mathcal{P}_2$, and plan $\mathcal{P}_2$ can be eliminated from consideration.

6.2 Maintenance goals

Now we consider a similar analysis for maintenance goals—those in which the temporal component requires that the atemporal component hold over an interval. The general relationship we will establish is between a plan $\mathcal{P}_1$ that is likely to keep a proposition ϕ true over a long subinterval of the maintenance interval, and ϕ has a high dsa value. About $\mathcal{P}_2$ we have established that some low-dsa formula ψ is likely to hold over all subintervals (i.e. ψ is the best that $\mathcal{P}_2$ is likely to achieve). In that case, if ϕ is sufficiently good compared to ψ, and if the interval over which ϕ is maintained is sufficiently long, and if $\mathcal{P}_1$ is sufficiently likely to actually achieve ϕ, then we can prove that $\mathcal{P}_1$ dominates $\mathcal{P}_2$.

6.2.1 Bounds for symbolic attributes

For symbolic attributes we once again measure degree of satisfaction on the basis of formulas related to the goal, and their associated dsa values. For maintenance goals we are supposed to make the formulas true over an interval $[t_b, t_e]$. Suppose that for plan $\mathcal{P}_1$ we have constants t_1, t'_1 such that ($t_b \leq t_1 \leq t'_1 \leq t_e$) for which we can establish that

$$\mathbf{P}(\phi(t_1, t'_1)|\mathcal{P}_1) \geq \alpha$$

($\mathcal{P}_1$ is likely to make ϕ true over a subinterval $[t_1, t_1']$ of the maintenance interval $[t_b, t_e]$), and for plan $\mathcal{P}_2$ we have constants t_2, t_2' such that $(t_B \leq t_2 \leq t_2' \leq t_E)$ for which we can establish that

$$\mathbf{P}(\psi(t_2, t_2')|\mathcal{P}_2) \geq \beta$$

($\mathcal{P}_2$ is likely to make ψ hold over subinterval $[t_2, t_2']$ of the maintenance interval $[t_b, t_e]$.)

Under what conditions can we say that $\mathcal{P}_1$ dominates $\mathcal{P}_2$? Again we must determine the lowest value of $\mathsf{EU}(\mathcal{P}_1)$ and the highest value of $\mathsf{EU}(\mathcal{P}_2)$ consistent with these two constraints. Let σ_ϕ be the formula of lowest dsa consistent with ϕ. The expected utility of plan $\mathcal{P}_1$ is lowest if

1. with probability α, σ_ϕ holds throughout t_1, t_1' and the atemporal component is not partially satisfied at any time outside of $[t_1, t_1']$, and
2. with probability $1-\alpha$ the atemporal component is completely unsatisfied throughout $[t_b, t_e]$.

So by Equation (7) we have

$$\begin{aligned}
\mathsf{EU}(\mathcal{P}_1) &\geq \alpha \cdot \int_0^{\mathsf{dsa}(\sigma_\phi)} \mathsf{CP}(t_1, t_1')\, dx \\
&= \alpha \cdot \mathsf{CP}(t_1, t_1') \cdot \mathsf{dsa}(\sigma_\phi)
\end{aligned}$$

Now let σ_ψ be the formula of *highest* dsa consistent with ψ. The expected utility of plan $\mathcal{P}_2$ is highest if

1. with probability β, σ_ψ holds throughout $[t_2, t_2']$ and the atemporal component is completely satisfied at all time outside $[t_2, t_2']$, and
2. with probability $1-\beta$ the atemporal component is completely satisfied throughout the interval $[t_b, t_e]$.

Again by Equation (7):

$$\begin{aligned}
\mathsf{EU}(\mathcal{P}_2) &\leq \beta \cdot \left[\int_0^{\mathsf{dsa}(\sigma_\psi)} \mathsf{CP}(t_B, t_E)\, dx \;+\; \int_{\mathsf{dsa}(\sigma_\psi)}^1 \mathsf{CP}(t_B, t_2) + \mathsf{CP}(t_2', t_E)\, dx\right] \\
&\quad + (1-\beta)(1) \\
&= \beta[\mathsf{dsa}(\sigma_\psi) + (\mathsf{CP}(t_B, t_2) + \mathsf{CP}(t_2', t_E)) \cdot (1 - \mathsf{dsa}(\sigma_\psi)) - 1] + 1
\end{aligned}$$

and then $\mathcal{P}_1$ dominates $\mathcal{P}_2$ if

$$\alpha > \frac{\beta[\mathsf{dsa}(\sigma_\psi) + (\mathsf{CP}(t_B, t_2) + \mathsf{CP}(t_2', t_E)) \cdot (1 - \mathsf{dsa}(\sigma_\psi)) - 1] + 1}{\mathsf{CP}(t_1, t_1') \cdot \mathsf{dsa}(\sigma_\phi)} \tag{11}$$

6.2.2 Bounds for quantitative attributes

In this case our atemporal goal is stated in terms of some quantity Q, and the dsa function associated with Q is monotonically increasing. We discover that $\mathcal{P}_1$ is likely to achieve a level of Q at least equal to k_1 throughout some subinterval $[t_1, t_1']$ of the maintenance interval $[t_b, t_e]$; $\mathcal{P}_2$ is likely to achieve a level of Q at *most* k_2 throughout some subinterval $[t_2, t_2']$ of $[t_b, t_e]$. The two equations are:

$$\begin{aligned} \mathbf{P}(\mathit{HOLDS}(\geq (Q,\ k_1), t_1, t_1')|\mathcal{P}_1) &\geq \alpha \\ \mathbf{P}(\mathit{HOLDS}(\leq (Q,\ k_2), t_2, t_2')|\mathcal{P}_2) &\geq \beta \end{aligned}$$

Under what conditions can we say that $\mathcal{P}_1$'s expected utility is greater than $\mathcal{P}_2$'s? The expected utility of $\mathcal{P}_1$ is lowest if

1. with probability α, a level k_1 for Q is maintained throughout the interval $[t_1, t_1']$ and Q has the value corresponding to a dsa of zero at all times outside of $[t_1, t_1']$, and

2. with probability $(1-\alpha)$ Q has the value corresponding to a dsa of zero at all times during $[t_b, t_e]$.

Under those circumstances we know that

$$\mathsf{EU}(\mathcal{P}_1) \geq \alpha \cdot \mathsf{CP}(t_1, t_1') \cdot \mathsf{dsa}(k_1)$$

The expected utility of plan $\mathcal{P}_2$ is highest if

1. with probability β, Q maintains a level of k_2 throughout the interval t_2, t_2' and Q maintains the value corresponding to complete satisfaction at all times outside t_2, t_2', and

2. with probability $1-\beta$ Q maintains the value corresponding to complete satisfaction throughout the interval t_B, t_E.

Then we know that

$$\mathsf{EU}(\mathcal{P}_2) \leq \beta[\mathsf{dsa}(k_2) + (\mathsf{CP}(t_B, t_2) + \mathsf{CP}(t_2', t_E)) \cdot (1 - \mathsf{dsa}(k_2)) - 1] + 1$$

and then $\mathcal{P}_1$ is preferred to $\mathcal{P}_2$ if

$$\alpha > \frac{\beta[\mathsf{dsa}(k_2) + (\mathsf{CP}(t_B, t_2) + \mathsf{CP}(t_2', t_E)) \cdot (1 - \mathsf{dsa}(k_2)) - 1] + 1}{\mathsf{CP}(t_1, t_1') \cdot \mathsf{dsa}(k_1)} \tag{12}$$

6.2.3 Bounds for strictly ordered atemporal attributes

Finally we reconsider the case of Section 6.1.3, where the atemporal component consists of a conjunction of formulas that can be ordered such that each conjunct dominates those later in the sequence. This time we establish that $\mathcal{P}_1$ is likely to make the dominant attribute g_1 true over some subinterval $[t_1, t_1']$, and $\mathcal{P}_2$ is likely to make that same attribute false over the subinterval $[t_2, t_2']$:

$$\begin{aligned} \mathbf{P}(g_1(t_1, t_1')|\mathcal{P}_1) &\geq \alpha \\ \mathbf{P}(\neg g_1(t_2, t_2')|\mathcal{P}_2) &\geq \beta \end{aligned}$$

Under what conditions can we say that plan $\mathcal{P}_1$ is preferable to plan $\mathcal{P}_2$? This is just a special case of an atemporal attribute, where the formula of lowest dsa consistent with g_1 is $\sigma_\phi \equiv g_1 \wedge_i \neg g_i$ and the formula of highest dsa consistent with $\neg g_1$ is $\sigma_\psi \equiv \neg g_1 \wedge_i g_i$. So by the results of section 6.2.1 plan $\mathcal{P}_1$ is preferred to plan $\mathcal{P}_2$ if

$$\alpha > \frac{\beta[\mathsf{dsa}(\neg g_1 \wedge_i g_i) + (\mathsf{CP}(t_B, t_2) + \mathsf{CP}(t_2', t_E)) \cdot (1 - \mathsf{dsa}(\neg g_1 \wedge_i g_i)) - 1] + 1}{\mathsf{CP}(t_1, t_1') \cdot \mathsf{dsa}(g_1 \wedge_i \neg g_i)} \tag{13}$$

6.3 Summary

The main point of this section was to establish circumstances under which the question of whether one plan is preferable to another (in the sense of having higher expected utility) could be answered by determining bounds on the probabilities of goal-related propositions. The general procedure for determining that a plan $\mathcal{P}_1$ dominates a plan $\mathcal{P}_2$ is to find a proposition ϕ associated with high (atemporal) satisfaction that $\mathcal{P}_1$ is likely to achieve and a proposition ψ associated with low atemporal satisfaction that $\mathcal{P}_2$ is likely to achieve. We then compare the worst case for $\mathcal{P}_1$ (that it establishes ϕ with low probability, and otherwise provides no atemporal goal satisfaction) agains the best case for $\mathcal{P}_2$ (that it establishes ψ with low probability, and the rest of the time provides complete goal satisfaction). This analysis leads to an inequality involving the probabilities of ϕ and ψ, indicating circumstances under which $\mathcal{P}_1$ dominates $\mathcal{P}_2$. We provided equations for symbolic, numeric, and ordered conjunctive atemporal attributes, first for deadline goals (Equations (8), (9), and (10)) and analogously for maintenance goals (Equations (11), (12), and (13)).

The analysis in this section studies a single goal in isolation; we next consider a multi-goal utility model, and how it can be exploited in the process of generating plans.

7 Multiple Goals, Residual Utility, and Computational Issues

The previous sections established bounds on the probabilities of outcomes that could be used to determine whether one plan is preferable to another. These analyses were performed on the utility functions for individual goals, implicitly assuming that utility for other goals and for resource consumption were the same.

The assumption that global utility is linear additive in the goal utilities (Section 3) makes explicit the tradeoff between satisfying the top-level goals, and between satisfying a goal and consuming resources. The assumption that resource consumption (counted in the residual utility term) is independent of the goal utilities means that we can regard resource consumption as a goal as well. Take the two-goal case, for example, for which we can write the equation for the expected utility of a plan $\mathcal{P}$ as follows:

$$\mathbf{EU}(\mathcal{P}) = k_1\mathbf{EU}_1(\mathcal{P}) + k_2\mathbf{EU}_2(\mathcal{P}) + k_r\mathbf{EU}_r(\mathcal{P}) \tag{14}$$

where

$$\mathbf{EU}_i(\mathcal{P}) = \Sigma_c \mathsf{UG}_i(c)\mathbf{P}(c|\mathcal{P}).$$

Now suppose that we have already generated a plan $\mathcal{P}_1$ and we know its expected utility $\mathbf{EU}(\mathcal{P}_1) = u_1$.

Further suppose that we have begun generating an alternative plan $\mathcal{P}_2$, in particular we have generated a subplan that achieves the first goal. From this subplan we can compute

1. an upper bound on the utility associated with first goal, $\mathbf{EU}_1(\mathcal{P}_2) \leq u_{12}$, and

2. a minimum level of resource consumption which provides us with an upper bound on residual utility, $\mathbf{EU}_r(\mathcal{P}_2) \leq u_{r2}$.

We can then calculate that for $\mathcal{P}_2$ to be preferred to $\mathcal{P}_1$ it must *at least* satisfy goal two to the degree

$$\mathbf{EU}_2(\mathcal{P}_2) \geq \frac{u_1 - k_1 u_{12} - k_r u_{r2}}{k_2}$$

which represents the required utility level for the second goal assuming no additional resource consumption. Examining the symbolic structure of the second goal may indicate exactly what propositions must be made true or what quantity level must be attained, and by when, in order to attain that level of utility. The ratio between k_r and k_2 along with the residual utility function indicates how efficiently the second goal must be satisfied as well.

7.1 Plan generation and refinement

The results in previous sections all involved comparing two complete plans; the relationships we provided reduced the question of which was preferable in the sense of maximizing utility to one of establishing a relationship between probabilities over the symbolic attributes that comprise one of the goal expressions. The algorithm in [Hanks, 1993] exploits both the symbolic content of the relationship and the numeric threshold to limit inference in establishing whether or not this relationship holds. Therefore deciding which of two partial plans is preferable might be considerably cheaper than computing the expected utility of each.

The result earlier in this section demonstrates that given one complete plan and a partial plan we can characterize the interesting possible completions of the partial plan—the ones that satisfy the remaining goals with a certain probability and with a certain effectiveness if the second plan is to be preferred to the first. Therefore deciding if a partial plan is worth pursuing can be considerably cheaper than examining all of its completions.

The question arises in general, however, as to what extent these relationships can aid the planning process. What can we say about the relationship between two partial plans? The next two sections discuss how our model might be applied to two plan-generation paradigms: partial-order planning and refinement planning.

7.1.1 Partial-order planning and search control

The BURIDAN planner [Kushmerick *et al.*, 1993] is an extension to classical nonlinear planners that allows uncertainty in the initial world state and in the effects of operators. The BURIDAN algorithm takes as input a problem description (a goal formula and a probability distribution over initial world states) and a probability threshold, and produces a plan that satisfies the goal with probability at least equal to the threshold. The analysis in this paper is complementary: it is a theory of how to generate the threshold values, but does not suggest an algorithm for exploiting them.

Two problems (at least) complicate the problem of applying our theory of utility to a probabilistic planner like BURIDAN. The first is that the BURIDAN algorithm generates a *lower bound* on the probability that the current plan or its completions satisfy the goal, and generally the lower the threshold the less work BURIDAN need do to generate an appropriate plan. A lower bound is of limited use in and of itself, however: you generally can't prove one plan superior to another without a lower bound on the performance of one and an *upper bound* on the other. The problem with a generative or transformational planner is that it is can be impossible to establish the point at which the plan cannot be further improved.

The second limitation in applying our theory to BURIDAN is that its notion of a goal is limited to propositional formulas, so concepts of partial satisfaction, deadlines, and resources are difficult to represent. Ongoing work is directed toward enhancing BURIDAN's representation language so it can naturally represent the concepts developed in this paper, and also to explore how this framework can be used to provide the planner with search-control information.

7.1.2 Refinement planning

One special case of partial planning that is amenable to our analysis is that in which the planner's only operation is to *refine* its current plan, which amounts to replacing an abstract action in the plan with a more specific version of that action. More precisely a refinement operator can never increase the set of possible outcomes consistent with the plan's execution, and thus tends to resolve uncertainty about its quality. An abstract plan's outcomes are sets of outcomes of more concrete plans. Since different probability and utility values may be associated with each specific outcome, in general a probability range and a utility range will be associated with each abstract outcome. Thus the expected utility of an abstract plan is represented by an interval, which includes the expected utilities of all possible instantiations of that abstract plan. Refining the plan, i.e. choosing an instantiation, tends to narrow the interval. Comparing two abstract plans can stop as soon as their expected-utility intervals no longer overlap[4].

The type of abstraction we use has been formalized by Tenenberg [1991] within the framework of the STRIPS representation and termed inheritance abstraction. He contrasts this to the type of abstraction used in ABSTRIPS [Sacerdoti, 1974], which Tenenberg calls relaxed model abstraction. For examples of the use of inheritance abstraction in plan generation in a deterministic setting see [Nau, 1987, Anderson and Farley, 1988]. The present work generalizes the notion of inheritance abstraction by introducing time, probability, and utility into the representation.

We illustrate the planning technique with an example. Consider the following problem of planning a delivery task. We wish to generate the best plan for delivering two tons of tomatoes from a farm to a warehouse within 85 minutes. The utility of a plan outcome is a function of the deadline goal and the residual utility, which is determined by the amount of fuel consumed. The components of the utility function are shown in Figure 7. The overall utility is

$$\mathbf{U}(c) = \mathsf{UG}(c) + (.02)\mathsf{UR}(c)$$

The delivery plan will consist of driving a truck from the depot to the farm, loading the truck, and driving the loaded truck to the warehouse. Planning is the process of generating this plan as well as choosing among the various ways this plan can be realized in such a way that expected utility is maximized. The descriptions of the available actions are shown in Figure 6. The action descriptions are similar to those in [Hanks, 1990]. Actions have conditional effects, and are represented by a tree with conditions labeling the branches. The leaves of the tree are labeled with the outcomes of the action. Deterministic actions are labeled with a single outcome. The probability of an outcome is the probability conditioned on the action and the conjunction of all conditions on all branches leading to the outcome. Outcomes are described in terms of a duration, as well as any changes in attribute values. Attributes are represented as functions of time. The variable t represents the beginning time of the action, making the action descriptions temporally indexical.

There are two possible routes we can take from the depot to the farm: road A and road B. Road A is longer but has no delays, while travel along road B might be delayed due to construction. The probability that construction is taking place is 0.2. These options are represented by the first two action descriptions in the Figure 6.

Once at the farm we must load the truck. We have two trucks at the depot to choose from: an open truck and a closed, cushioned truck. The open truck is easy to load, while the closed truck can be loaded easily if a special quick-loading device is available. There is

[4]Or when the intervals narrow to the point where the distinction between the two does not warrant further attention, as [Russell and Wefald, 1991b] point out.

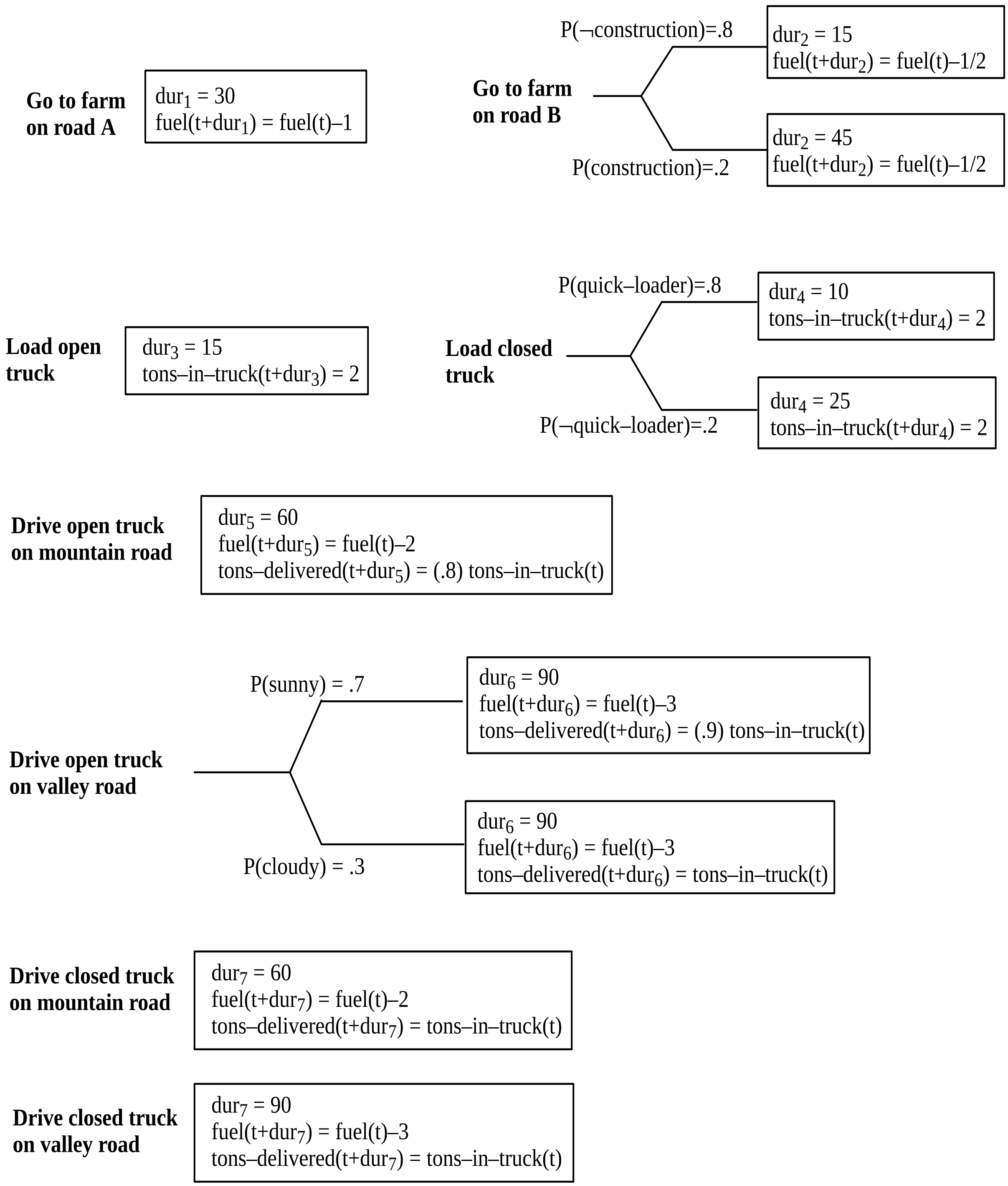


Figure 6: Action descriptions

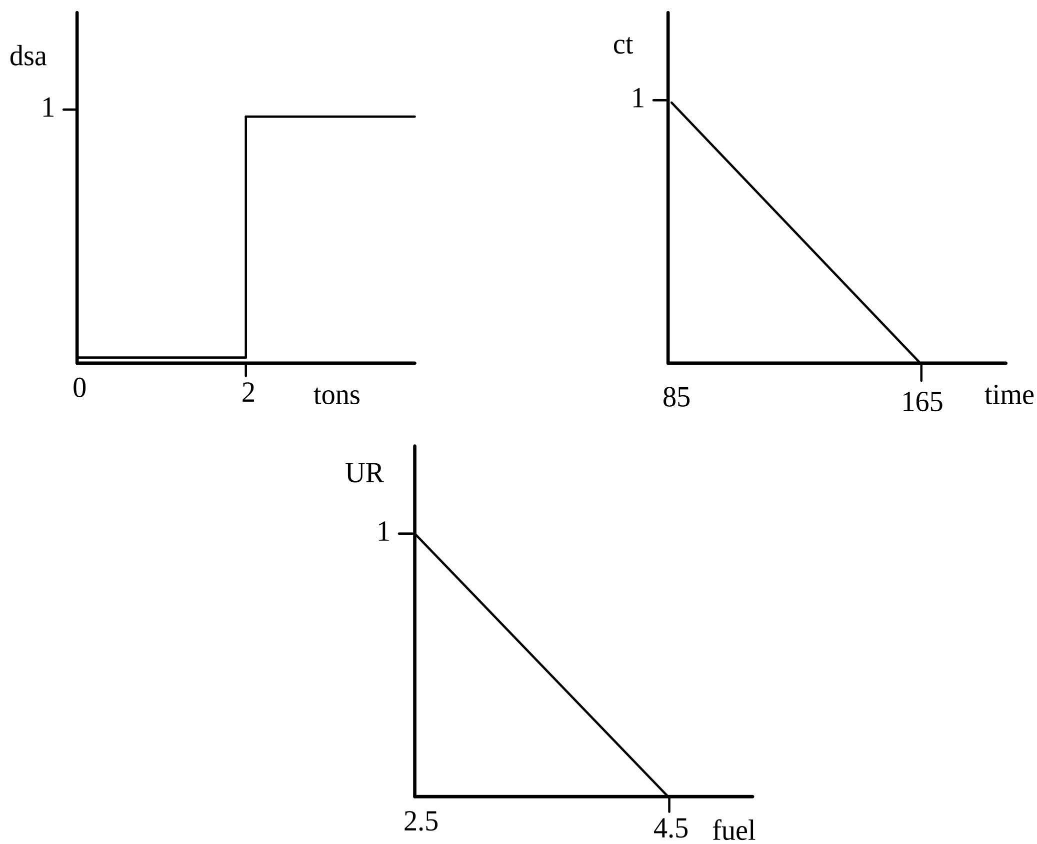


Figure 7: Specification of delivery utility function.

an 80% chance this device will be available. The next two diagrams in the Figure 6 depict these two actions.

Once the truck is loaded we must drive it to the warehouse. We have two routes to choose from: the mountain road and the valley road. The mountain road is shorter but bumpy. If we drive the open truck on the mountain road, the bottom 20% of the tomatoes will be crushed. If we drive the open truck on the valley road and the sun is shining, the top 10% of the tomatoes will be spoiled by the sun. This combination of options results in the last four action descriptions in the Figure 6.

In order to perform the refinement type planning described, we organize the planning knowledge in an abstraction/decomposition network, shown in Figure 8. Solid links show decompositions [Charniak and McDermott, 1985, Ch9], while dashed links show possible refinements. For example the task "deliver tomatoes" is decomposed into the sequence of the two actions "go to farm" and "load & drive truck". The arrow between the actions signifies temporal succession. The abstract action "go to farm" can be realized by driving on road A or road B. Since both elements of a decomposition must be realized, decomposition links are AND links and since either element of an instantiation may be used, instantiation links are OR links. So this network forms an AND/OR tree.

Descriptions of the abstract actions are shown in Figure 9. A set of actions is abstracted by grouping together their outcomes into abstract outcomes which represent the range of possible outcomes of the instantiations. Care must be taken to group together similar outcomes. For example, the "drive open truck" action is an abstraction of "drive open truck on mountain road" and "drive open truck on valley road." Since the single outcome of "drive open truck on mountain road" is more similar to the outcome of the upper branch of "drive open truck on valley road" than to the outcome of the lower branch, it is grouped with the former. While similarity is fairly clear in this case, determining similarity of outcomes with multiple attributes may not be straightforward in general. (In fact, one may wish to produce more than one abstraction based on different similarity measures and try each one on a given problem. One would use the hierarchy that resulted in the most pruning for the given problem.) The outcomes of the abstract action are now specified simply as the range of outcomes grouped together. Once the outcomes have been abstracted, probabilities must be assigned to them. This is done by taking the range of probabilities of the outcomes in the group. So for the "drive open truck" action, the range on the upper branch is [min($\mathbf{P}$(sunny), 1) max($\mathbf{P}$(sunny), 1)], which is just [$\mathbf{P}$(sunny) 1].

Given this representation of the planning problem, we evaluate plans at the abstract level and prune suboptimal plans before refining candidate plans further. There are eight possible plans implicitly encoded in the abstraction/decomposition network, and we want to choose the one that maximizes expected utility.

According to the network, the task of delivering tomatoes is first decomposed into going to the farm and loading, then driving the truck. We can first choose either which road to take to the farm or which truck to load and drive. Suppose we make the latter choice first. Because the utility function is a function over chronicles, the value of a particular action in a plan depends on when in the plan it occurs, so options can only be evaluated in the context

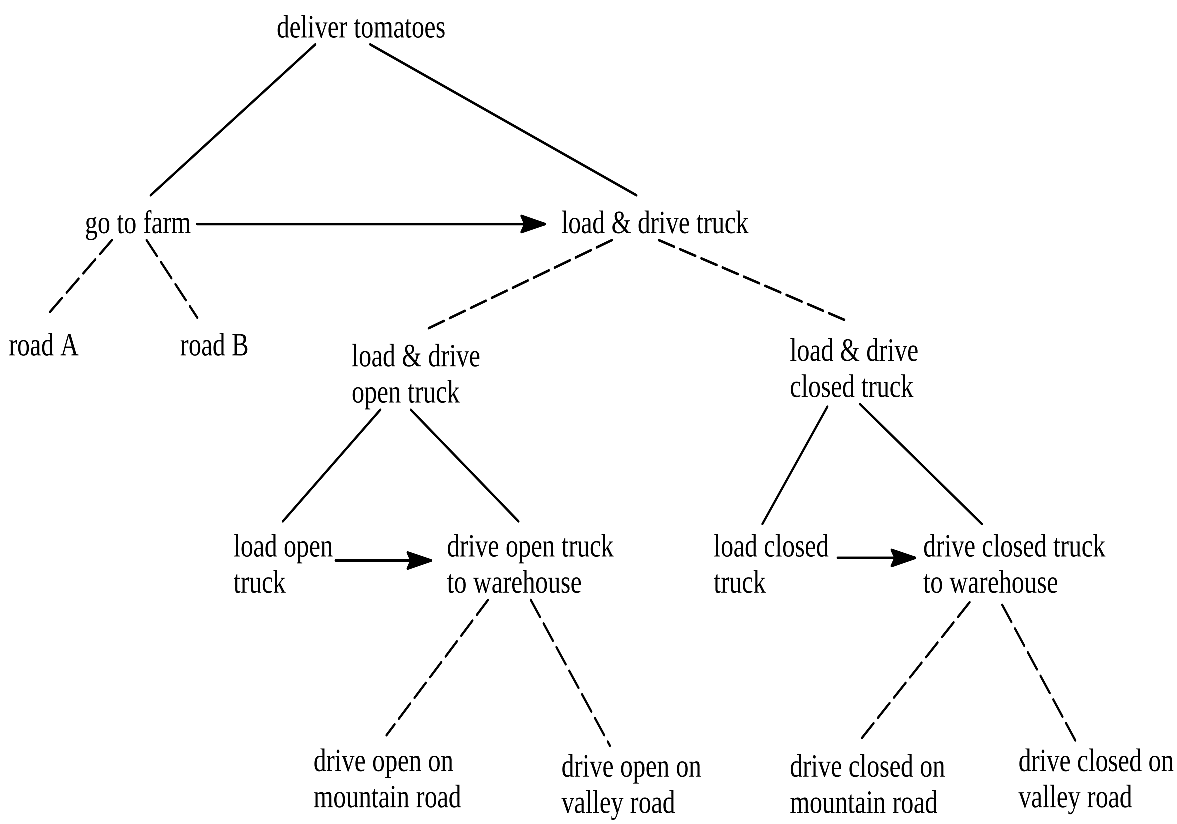


Figure 8: Abstraction/decomposition network

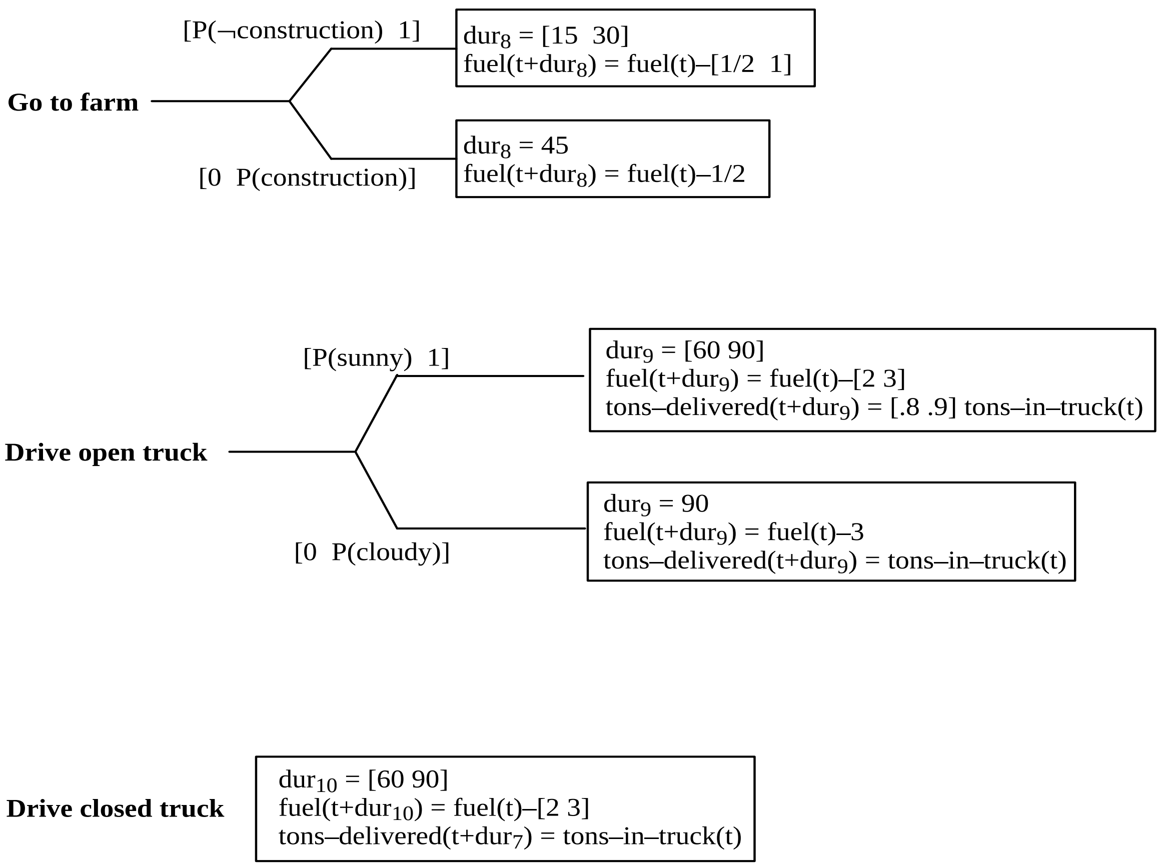


Figure 9: Abstract action descriptions

$\mathcal{P}_1$: Go to farm $\longrightarrow$ Load & drive open truck					
chronicle	time	fuel	tons	U(chronicle)	probability
c_1	[90 135]	[2.5 4]	[1.6 1.8]	[.005 .02]	[.56 1]
c_2	[120 135]	[3.5 4]	2	[.380 .5725]	[0 .3]
c_3	[120 150]	[2.5 3.5]	[1.6 1.8]	[.01 .02]	[0 .2]
c_4	150	3.5	2	.1975	[0 .06]
$\mathcal{P}_2$: Go to farm $\longrightarrow$ Load & drive closed truck					
chronicle	time	fuel	tons	U(chronicle)	probability
c_1	[85 130]	[2.5 4]	2	[.4425 1.02]	[.64 .8]
c_2	[100 145]	[2.5 4]	2	[.255 .8325]	[.16 .2]
c_3	[115 145]	[2.5 3.5]	2	[.255 .635]	[0 .16]
c_4	[130 160]	[2.5 3.5]	2	[.0675 .4475]	[0 .04]

Table 1: Highest level abstract plan outcomes.

of an overall plan. Consequently, we combine the "load & drive open truck" and the "load & drive closed truck" actions with the "go to farm" action to obtain a complete abstract plan that can be evaluated.

Each abstract plan results in four chronicles (Table 1). For example, for plan $\mathcal{P}_1$ we obtain these chronicles by concatenating the "go to farm," "load open truck," and "drive open truck" action descriptions. Doing so results in four chronicles, since "go to farm" and "drive open truck" each have two branches. Rather than showing the complete chronicles, the relevant attributes of the resulting chronicles are summarized in Table 1, which shows for each chronicle the range of times, fuel consumption, and tons of tomatos delivered. The time refers to the time that the delivery is made. We assume that the plan begins execution at time zero and we take the beginning time of an action to be the beginning time of the previous action plus its duration. So for example the fuel consumption for chronicle c_1 is computed according to

$$fuel(dur_8) \; = \; fuel(0) - [.5 \; 1]$$

$$fuel(dur_8 + dur_9) \; = \; fuel(dur_8) - [2 \; 3]$$

so

$$fuel(dur_8 + dur_9) \; = \; fuel(0) - [2.5 \; 4]$$

We illustrate how the utility range is computed by showing the computation for chronicle c_2 of plan $\mathcal{P}_1$; the computation for the other chronicles is similar. Let UG_{min} and UG_{max} be the lower and upper bounds on the goal utility, respectively, and let UR_{min} and UR_{max} be the lower and upper bounds on residual utility, respectively. Since we are computing the utility of a deadline goal, we can minimize the utility of the abstract chronicle by choosing the latest delivery times and the smallest delivery amounts, we can maximize the utility by choosing the earliest delivery times and the largest amounts. So

$$\mathsf{UG}_{max}(c_2) = dsa(2) \cdot ct(120) = 0.5625$$

$$\mathsf{UG}_{min}(c_2) = dsa(2) \cdot ct(135) = 0.375$$

$$\mathsf{UR}_{min}(c_2) = 0.25$$

$$\mathsf{UR}_{max}(c_2) = 0.5$$

Since our global utility function is $\mathbf{U}(c) = \mathsf{UG}(c) + (0.02)\mathsf{UR}(c)$, we have $\mathbf{U}(c_2) = [0.380\ 0.5725]$.

Using the utility and probability information in the table we can compute the expected utility bounds for each plan. Computing a bound for a plan involves choosing the probabilities within the given ranges that minimize and maximize the expected utility. We can do so by solving a small linear programming problem in which the objective function is the expression for the expected utility and the constraints are the probability bounds. For example, the upper bound for plan $\mathcal{P}_1$ can be computed by maximizing the objective function

$$.02p_1 + .5725p_2 + .02p_3 + .1975p_4$$

subject to the constraints

$$.56 \leq p_1 \leq 1$$

$$0 \leq p_2 \leq .3$$

$$0 \leq p_3 \leq .2$$

$$0 \leq p_4 \leq .06$$

$$p_1 + p_2 + p_3 + p_4 \ = \ 1$$

The maximizing probabilities are

$$\begin{aligned} p_1 &= .56 \\ p_2 &= .3 \\ p_3 &= .08 \\ p_4 &= .06 \end{aligned}$$

So the upper bound on expected utility is

$$(.56)(.02) + (.3)(.5725) + (.08)(.02) + (.06)(.1975) = .1964$$

So for plan $\mathcal{P}_1$ and $\mathcal{P}_2$ we obtain the expected utility ranges

$\mathcal{P}_{2.1}$: Go to farm $\longrightarrow$ Load closed truck $\longrightarrow$ Drive closed on mountain road					
chronicle	time	fuel	tons	U(chronicle)	probability
c_1	[85 100]	[2.5 3]	2	[.8275 1.02]	[.64 .8]
c_2	[100 115]	[2.5 3]	2	[.64 .8325]	[.16 .2]
c_3	115	2.5	2	.645	[0 .16]
c_4	130	2.5	2	.4575	[0 .04]
$\mathcal{P}_{2.2}$: Go to farm $\longrightarrow$ Load closed truck $\longrightarrow$ Drive closed on valley road					
chronicle	time	fuel	tons	U(chronicle)	probability
c_1	[115 130]	[3.5 4]	2	[.4425 .635]	[.64 .8]
c_2	[130 145]	[3.5 4]	2	[.255 .4475]	[.16 .2]
c_3	145	3.5	2	.26	[0 .16]
c_4	160	3.5	2	.0725	[0 .04]

Table 2: Intermediate level abstract plan outcomes.

$\mathbf{EU}(\mathcal{P}_1) = [.005\ .1964]$
$\mathbf{EU}(\mathcal{P}_2) = [.3673\ .9825]$.

Since the lower bound for $\mathcal{P}_2$ is greater than the upper bound for $\mathcal{P}_1$, we can eliminate from consideration all possible refinements of $\mathcal{P}_1$ and concentrate on refining $\mathcal{P}_2$. So at this point we have chosen the option of using the closed truck. By making this choice, we have pruned away the left-hand subnetwork underneath the "load & drive truck" node in Figure 8, resulting in pruning half the space of possible plans from consideration.

We are left with two more actions to refine: "go to farm" and "drive closed truck to warehouse." Suppose we next choose to refine the "drive" action. Again the instantiations involving the mountain road and the valley road must be evaluated in the context of a complete plan. So we compose the descriptions of the concrete actions "drive closed on mountain road" and "drive closed on valley road" with the descriptions of the concrete action "load closed truck" and the abstract action "go to farm." Table 2 summarizes the outcomes for the two alternative plans. We use this information to compute expected-utility bounds for the two alternatives:

$\mathbf{EU}(\mathcal{P}_{2.1}) = [.7533\ .9825]$
$\mathbf{EU}(\mathcal{P}_{2.2}) = [.3683\ .5975]$.

Notice that the EU intervals for the two plans are contained in the EU interval for the abstract plan of which they are a refinement. Since the lower bound for plan $\mathcal{P}_{2.1}$ is greater than the upper bound for plan $\mathcal{P}_{2.2}$, we can eliminate $\mathcal{P}_{2.2}$ from consideration, pruning away two more possible concrete plans. By eliminating plan $\mathcal{P}_{2.2}$, we have chosen to take the mountain road.

Finally we refine plan $\mathcal{P}_{2.1}$. Our two options are taking either road A or road B to the farm. The outcomes of the plans incorporating these options are summarized in table 3.

$\mathcal{P}_{2.1.1}$: Go to farm on road A $\longrightarrow$ Load closed truck $\longrightarrow$ Drive closed on mountain road					
chronicle	time	fuel	tons	U(chronicle)	probability
c_1	100	3	2	.8275	.8
c_2	115	3	2	.64	.2
$\mathcal{P}_{2.1.2}$: Go to farm on road B $\longrightarrow$ Load closed truck $\longrightarrow$ Drive closed on valley road					
chronicle	time	fuel	tons	U(chronicle)	probability
c_1	85	2.5	2	1.02	.64
c_2	100	2.5	2	.8325	.16
c_3	115	2.5	2	.645	.16
c_4	130	2.5	2	.4575	.04

Table 3: Concrete plan outcomes.

Since the plans now include only concrete actions, the attribute, probability, and utility values are now point values. The expected utilites of the two remaining plans are

$$\mathbf{EU}(\mathcal{P}_{2.1.1}) = .79$$
$$\mathbf{EU}(\mathcal{P}_{2.1.2}) = .9075$$

so we choose plan $\mathcal{P}_{2.1.2}$. Since this is a complete concrete plan, we have generated the plan that maximizes expected utility and are finished.

Although it is probably unrealistic to expect any planner to employ only refinement operators, the idea of modeling an execution system using a hierarchy of abstract actions is consistent with Firby's [1989] RAP system, and the architecture for planning and execution proposed in [Hanks and Firby, 1990].

8 Summary and Related Work

Our goal in this work was to take the concept of goals as they have been used in symbolic planning systems and simultaneously extend their functionality and recast the intuitions in a form that can be exploited by a decision-theoretic planning algorithm.

Our framework involves building a utility model first by identifying the agent's top-level goals plus the residual attributes that measure resource consumption and production in service of those goals. The assumption of utility independence among these attributes means that their interactions can be summarized by $n+1$ numeric parameters representing the relative weights for the n goals and residual attributes.

We then extended the notion of a goal to one that involves both a temporal component (deadline or maintenance interval) and an atemporal component (a formula to be achieved subject to the temporal constraint). We discussed various forms for the atemporal goal component: symbolic and numeric attributes, conjunctions of numeric attributes, and ordered conjunctions.

For each of the goals the user supplies two components that describe preferences over partial-satisfaction scenarios: the atemporal degree of satisfaction function and the temporal weighting coefficient. The former defines what it means to satisfy the goal's atemporal component, either partially or fully. The latter indicates how utility declines as a function of missing the deadline or violating the maintenance interval. We described how to combine those functions to produce a utility function for the entire goal. The problem was how to combine partial satisfaction of both the atemporal and the temporal components simultaneously, and we developed model in which utility is accrued each time the value of the atemporal component *increases* over the interval defined by the temporal coefficient.

We then showed how the model's information, both numeric and symbolic, could be exploited to compare plans: to decide whether one plan's expected utility was greater than another and to generate bounds on the quality of a partial plan in order that it should be chosen over an alternative. The general form of these relationships was to consider a plan $\mathcal{P}_1$ that was likely to achieve at worst a high level of satisfaction, and a plan $\mathcal{P}_2$ that was likely to achieve at best a low level of satisfaction. The result was a function of the respective formulas, their likelihoods, and their times, ensuring that $\mathcal{P}_1$'s expected utility was greater than $\mathcal{P}_2$'s, regardless of the two plans' other effects.

Finally we demonstrated how these relationships could be exploited by two existing planning techniques, particularly in the area of refinement planning.

8.1 Related work

A discussion of related work should begin with a mention of multiattribute decision theory, especially [Keeney and Raiffa, 1976]. What we have done is built a multiattribute utility model for goal-oriented planning problems that feature partial goal satisfaction and deadlines.

Our discussion of strictly ordered goals was motivated by the work in *goal programming* [Schniederjans, 1984] a mathematical optimization technique that deals with ordered conflicting goals.

8.1.1 Goals and utility models

In the AI literature the work closest to our own is by Wellman and Doyle [1991], [Wellman and Doyle, 1992], which also analyzes the relationship between goals and preference structures. Their work confronts the question of what it means to say that an agent has some goal. The most fundamental difference between their work and ours is that they begin by examining an agent's preference structure directly and produce a definition of what it means to say that an agent has a goal γ. In contrast, we adopt various intuitive notions about goals at the outset (*e.g.* that they are utility independent at the top level), and structure the agent's utility function (and therefore his preferences) to accommodate those ideas. Our work is mainly oriented toward using the resulting structure to build and exploit representations for concepts like partial satisfaction and temporal deadlines in the process of building and comparing plans.

A few efforts have been made in the classical planning literature to extend the form of goal expressions: [Drummond, 1989] introduces a crudeform of maintenance goals, allowing the constraint that a proposition remain true throughout the execution of a plan. [Vere, 1983] implements a concept related to deadline goals: a temporal *window* or interval within which an action must be executed, and [Dean *et al.*, 1988] handles deadlines and actions with duration. None of these efforts incorporate uncertainty or partial satisfaction into the representation, nor do they consider partial satisfaction of the goal's atemporal component.

8.1.2 Decision-theoretic planning and control

Decision-theoretic techniques have been applied both to the planning problem [Koenig, 1992], [Feldman and Sproull, 1975], [Dean and Kanazawa, 1989] and to the problem of controlling reasoning—choosing among computational actions as well as those that make physical changes to the world [Russell and Wefald, 1991a], [Boddy, 1991], [Etzioni, 1991], [Horvitz *et al.*, 1989].

Most of these works use some variant of the utility model common to Markov decision processes [Howard, 1960]: there is a reward function associated with certain states and a cost function associated with actions. The value of a plan is the value associated with the end state less the cost of the actions (e.g. the time they consume) that comprise the plan.

This model implies a number of assumptions about the agent's preference structure. First it assumes that the value and cost attributes are measured in units that are directly comparable. Second the idea that reward is accrued as a result of arriving at a "goal" state means that any value or cost associated with achieving the goal must be captured in the cost attribute. There is no obvious way to capture maintenance and deadline goals, and no model of partial goal satisfaction, either temporal or atemporal.

Etzioni's [1991] model is more similar to ours: admitting both partial satisfaction of the goals and also the idea that the value of achieving the goal will tend to change over time. Both of these elements are supplied directly to the model, in the form of three functions:

- a function $i(g)$ measuring the "intrinsic value" of goal g,
- a function $d(s, g)$ measuring the extent to which goal g is satisfied in state s, and
- a function $F(i(g)d(s, g), s)$ measuring the extent to which the benefit of goal g should be realized in state s.

The first two functions correspond roughly to our atemporal component, the third to our temporal weighting coefficient. There is no analogue in his model to our discussion of maintenance goals. He makes the same assumption we do about the utility independence of top-level goals.

These efforts are basically complementary to ours: they make simplifying assumptions about the utility model and concentrate on algorithms to solve the planning problem; we develop a richer utility model but provide no algorithm, only relationships implied by the model that might guide a planning algorithm. The challenge will be to integrate our model and the relationships it implies into these algorithms.

8.1.3 Fuzzy decision theory

The notion of partially satisfied goals and their role in the decision-making process appears prominently in the literature on fuzzy mathematics and decision analysis. In particular our notion of a degree of satisfaction function bears close resemblance to a fuzzy-set membership function. The seminal paper in this area is [Bellman and Zadeh, 1980]; also see the papers in [Zimmerman *et al.*, 1984], of which the most relevant to this paper is [Dubois and Prade, 1984]. They discuss the role of *aggregation operators* in the decision-making process. In the language of fuzzy-set theory a goal may be expressed as a fuzzy set, a plan's membership function with respect to that set indicates the extent to which the plan satisfies that goal. An aggregation operator combines membership functions for individual goals into an aggregate membership function which is an indicator of global success—this is called the *decision set*. A decision maker then selects an alternative that is "strongly" a member of the decision set. Dubois and Prade categorize and analyze various aggregation functions.

So our analysis is similar to the efforts in fuzzy decision making in that it emphasizes the representation problems associated with expressing partial satisfaction of goals. Fuzzy sets may be a more appropriate representation than degree of satisfaction when the latter (a numeric function) cannot reasonably be assessed. If we can only assess vague satisfaction measures like "reasonably well satisfied," "utter failure," and "complete success," the fuzzy-set methodology provides a way to incorporate these measures into a precise analysis. As such it is essentially complementary to our analysis.

8.2 Future work

The formal model can be extended to cover more types of goals. The representation in this paper covers only goals that mention facts but goals can refer to events as well. An example might be "flip the switch at noon." Goals mentioning events would be restricted to deadline goals since it does not make sense to maintain an event over an interval of time. Deadline goals involving events could be represented in a way similar the representation of goals with a symbolic atemporal component. But the current definition could not be used unchanged since we have defined DSA in terms of formulas that hold at time points, while events occur over time intervals.

Deadline goals will often have a maintenance component to them: we want to achieve a given state by a time and once achieved we want it to persist for a given period of time. The expressions for deadline and maintenance goals could be combined to represent such goals. We would perform a maintenance goal computation at each time that a term for deadline goal utility is calculated, i.e., at each time that DSA changes to a value higher than any previous value. At each of these times, we would perform a maintenance goal computation.

The most important area of future work is to incorporate our model into decision-theoretic planning algorithms. We have mentioned several candidates above: the state-space planners based on Markov decision processes, and two symbolic algorithms—probabilistic nonlinear planning and refinment planning. We are currently working on an implementation of the refinement planning algorithm. But to handle a realistic range of problems, the algorithm needs to be extended in several ways. In the example, actions have only discrete outcomes but more realistically some actions would be described in terms of a continuous distribution over outcomes, e.g. a normal distribution over durations. Representing such distributions is not difficult but we need to find ways of grouping such distributions to describe the outcomes of abstract actions. Our network decomposed actions into totally ordered sequences of actions. We need to incorporate other decompositions such as partial orders. Our example included no interacting actions. We need to develop mechanisms to handle interactions such as the effects of one action depending on the choice of a previous action. This particular interaction can be handled by an application of abstraction. It is probably unrealistic to expect that a planner will explicitly store all possible action decompositions, so we need to integrate the refinement algorithm with a nonlinear planner to generate the decompositions on the fly. Finally we need to test the effectiveness of both the representation and the planning algorithm on some large problems.

References

[Anderson and Farley, 1988] J. Anderson and A. Farley. Plan abstraction based on operator generalization. In *Proceedings AAAI*, pages 100–104, 1988.

[Bellman and Zadeh, 1980] R.E. Bellman and L.A. Zadeh. Decision-making in a Fuzzy Environment. *Management Science*, 17:B141–B164, 1980.

[Boddy, 1991] Mark Boddy. Solving Time-Dependent Problems: A Decision-Theoretic ApproachPlanning in Dynamic Environments. Technical Report CS–91–06, Brown University, Department of Computer Science, May 1991.

[Charniak and McDermott, 1985] E. Charniak and D. McDermott. *Introduction to Artificial Intelligence*. Addison-Wesley, Reading, MA, 1985.

[Dean and Kanazawa, 1989] Tom Dean and Keiji Kanazawa. A Model for Projection and Action. In *Proceedings IJCAI*, pages 985–990, 1989.

[Dean *et al.*, 1988] T. Dean, J. Firby, and D. Miller. Hierarchical Planning involving deadlines, travel times, and resources. *Computational Intelligence*, 4(4):381–398, 1988.

[Drummond, 1989] M. Drummond. Situated Control Rules. In *Proceedings of the First International Conference on Knowledge Representation and Reasoning*, May 1989.

[Dubois and Prade, 1984] Didier Dubois and Henri Prade. Criteria Aggregation and Ranking of Alternatives in the Framework of Fuzzy Set Theory. In H.J Zimmerman, L.A. Zadeh, and B.R Gaines, editors, *Fuzzy Sets and Decision Analysis*, pages 209–240. North Holland, 1984.

[Etzioni, 1991] Oren Etzioni. Embedding Decision-analytic Control in a Learning Architecture. *Artificial Intelligence*, 1–3(49):129–160, 1991.

[Feldman and Sproull, 1975] J.R. Feldman and R.F. Sproull. Decision Theory and Artificial Intelligence II: The Hungry Monkey. *Cognitive Science*, 1:158–192, 1975.

[Firby, 1989] R. James Firby. Adaptive Execution in Complex Dynamic Worlds. Technical Report 672, Yale University, Department of Computer Science, January 1989.

[Haddawy and Hanks, 1990] Peter Haddawy and Steve Hanks. Issues in Decision-Theoretic Planning: Symbolic Goals and Numeric Utilities. In *DARPA Workshop on Innovative Approaches to Planning, Scheduling, and Control*. Morgan Kaufmann, November 1990.

[Haddawy, 1991a] P. Haddawy. A Temporal Probability Logic for Representing Actions. In Richard Fikes James Allen and Erik Sandewall, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference*, pages 196–207. Morgan Kaufman, San Mateo, CA, 1991.

[Haddawy, 1991b] P. Haddawy. *Representing Plans Under Uncertainty: A Logic of Time, Chance, and Action*. PhD thesis, University of Illinois, 1991. (Report no. UIUCDCS-R-91-1719).

[Hanks and Firby, 1990] Steve Hanks and R. James Firby. Issues and Architectures for Planning and Execution. In *DARPA Workshop on Innovative Approaches to Planning, Scheduling, and Control*. Morgan Kaufmann, November 1990.

[Hanks and McDermott, 1993] Steve Hanks and Drew McDermott. Modeling a Dynamic and Uncertain World I: Symbolic and Probabilistic Reasoning about Change. *Artificial Intelligence*, 1993. To appear.

[Hanks, 1990] Steven Hanks. Projecting Plans for Uncertain Worlds. Technical Report 756, Yale University, Department of Computer Science, January 1990.

[Hanks, 1993] Steven Hanks. Modelling a Dynamic and Uncertain World II: Action Representation and Plan Evaluation. *Journal of Logic and Computation*, 1993. Submitted.

[Horvitz *et al.*, 1989] Eric J. Horvitz, Gregory F. Cooper, and David E. Heckerman. Reflection and Action Under Scarce Resources: Theoretical Principles and Empirical Study. In *Proceedings IJCAI*, pages 1121–1127, 1989.

[Howard, 1960] Ronald A. Howard. *Dynamic Programming and Markov Processes*. MIT Press, 1960.

[Keeney and Raiffa, 1976] Ralph L. Keeney and Howard Raiffa. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. John Wiley & Sons, 1976.

[Koenig, 1992] Sven Koenig. Optimal Probabilistic and Decision-Theoretic Planning using Markovian Decision Theory. Technical Report UCB/CSD 92/685, Computer Science Division (EECS), UC Berkeley, May 1992.

[Kushmerick *et al.*, 1993] Nick Kushmerick, Steve Hanks, and Dan Weld. An Algorithm for Probabilistic Planning. *Artificial Intelligence*, 1993. Submitted.

[McDermott, 1982] Drew McDermott. A Temporal Logic for Reasoning About Processes and Plans. *Cognitive Science*, 6:101–155, 1982.

[Nau, 1987] D. Nau. Hierarchical abstraction for process planning. In *Proc. of the Int's Conference on Applications of Artificial Intelligence in Engineering*, 1987.

[Russell and Wefald, 1991a] Stuart J. Russell and Eric H. Wefald. *Do the Right Thing: Studies in Limited Rationality.* MIT Press, 1991.

[Russell and Wefald, 1991b] Stuart J. Russell and Eric H. Wefald. Principles of Metareasoning. *Artificial Intelligence*, 1–3(49):361–396, 1991.

[Sacerdoti, 1974] E.D. Sacerdoti. Planning in a hierarchy of abstraction spaces. *Artificial Intelligence*, 5(2):115–135, 1974.

[Schniederjans, 1984] Marc J. Schniederjans. *Linear Goal Programming.* Petrocelli Books, 1984.

[Tenenberg, 1991] J.D. Tenenberg. *Reasoning About Plans*, chapter Abstraction in Planning, pages 213–283. Morgan Kaufmann, San Mateo, CA, 1991.

[Vere, 1983] S. Vere. Planning in Time: Windows and Durations for Activities and Goals. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 5, 1983.

[Wellman and Doyle, 1991] Michael P. Wellman and Jon Doyle. Preferential Semantics for Goals. In *Proceedings AAAI*, 1991.

[Wellman and Doyle, 1992] Michael P. Wellman and Jon Doyle. Modular Utility Representation for Decision-Theoretic Planning. In *Proceedings, First International Conference on AI Planning Systems*, 1992.

[Zimmerman *et al.*, 1984] H.J Zimmerman, L.A. Zadeh, and B.R Gaines, editors. *Fuzzy Sets and Decision Analysis.* North Holland, 1984. TIMS Studies in the Management Sciences, Volume 20.