The Knowledge Engineering Review, Vol. 5: 1, 1990, 15–33

Planning and decision theory

PETER HADDAWY[1] AND LARRY RENDELL

University of Illinois, Dept of Computer Science, 1304 W. Springfield, Urbana, IL 61801, USA

1 Introduction

Planning, the process of formulating and choosing a sequence of actions which when executed will likely achieve a desirable outcome, is a basic component of intelligent behaviour. An autonomous agent must be able to manipulate his environment in order to satisfy his needs and desires. Rather than simply reacting to immediate situations, an intelligent agent needs to plan for the future. In order to do this, he must have some knowledge about the state of the world, knowledge of possible actions he can perform, knowledge of how these actions affect the world, and a set of desires or preferences. The agent's objective is to achieve a world state that satisfies his desires. Planning then involves reasoning about the state of the world and the way in which one's actions affect the world.

2 Planning in artificial intelligence

Research on planning in AI can be separated into two major areas: plan generation and plan representation. Plan generation is primarily concerned with the composition of primitive actions to form plans, whereas plan representation is concerned with the representation of aspects of the world necessary for reasoning about plans. To generate plans, one needs a representation. Historically, the first formal representation for planning problems was the *situation calculus* (McCarthy & Hayes, 1969). Most AI planning systems to date have used representations based on situation calculus. This language is a variation of predicate calculus intended for representing time and change as a function of events by including explicit reference to situations in the language. A *situation* is a period of time over which no changes occur. Relations within a situation are states-of-affairs, which can be asserted to hold in a situation with a binary predicate T. For example, to say that block A is on block B in situation s_1 we might write $T(ON(A, B), s_1)$. Situations are described by finite sets of sentences. For example, the situation in which block A is on block B which is on block C might be described by $T(ON(A, B), s_k) \wedge T(ON(B, C), s_k)$.

Transitions from one situation to another are made only through the execution of an action or occurrence of an event. So an action takes us from one situation to another. The effect of an action is represented by a function that maps an action and situation to the situation that results from performing the action. An example of an action might be the UNSTACK operator which removes one block from on top of another and puts it on the table. The result of an action is described with the function *Do*. $Do(UNSTACK(x, y), s_3)$ denotes the situation that results from executing UNSTACK(x,y) in situation s_3. To say that if x is on y and x is clear then after the unstack operation, x is on the table and y is clear we might write:

$$T(ON(x, y), s) \wedge T(CLEAR(x), s) \rightarrow$$
$$T(ON-TABLE(x), Do(UNSTACK(x, y), s)) \wedge T(CLEAR(y), Do(UNSTACK(x, y), s))$$

2.1 Plan generation

The first planning system based on the situation calculus was developed by Cordell Green (Green,

[1] This work was supported in part by the first author's Cognitive Science/Artificial Intelligence fellowship and by NASA grant NAG 1-616.

1969). Green's method of planning was based on resolution theorem proving. Given a description of the initial situation, a goal description, and descriptions of available actions, the method attempts to produce a constructive proof of the existence of a state that satisfies the goal conditions. By cleverly using unification, the method returns the sequence of actions necessary to transform the initial situation into the goal situation. Unfortunately, the direct use of the situation calculus makes this approach extremely inefficient. Since the world description only specifies the initial situation, a large number of axioms is required to describe the properties unaffected by each possible action. For example, for a move operator we must specify that moving block A does not change the size of block A, the colour of block B, etc. These axioms are called *frame axioms* and the problem of having to specify them is called the *frame problem* (Hayes, 1973).

Fikes and Nilsson developed the STRIPS planning system partly in an attempt to solve the frame problem (Fikes & Nilsson, 1971). In the STRIPS planning system, the initial situation and the goal situation are described by finite conjunctions of literals. By focusing only on the current situation, STRIPS does not need to represent situations explicitly. Instead, transitions are represented in terms of changes to the current situation. Thus, actions are described by operators with three components: a precondition formula, a delete list, and an add list. The precondition formula must logically follow from the facts in the current situation description in order for the operator to be applicable. The delete list is a conjunction of literals that may no longer be true following the action and must be deleted from the current situation description upon application of the operator. The add list is a conjunction of literals that are true in the situation resulting from the action and must be added to the current situation description upon operator application. By describing actions in terms of their effects on the world and by assuming that the only changes to the world are those specified in the operators, STRIPS was able to eliminate frame axioms. Despite numerous limitations which will be discussed in the next section, this representation was highly successful. In fact, most planning systems to date have used variations of the STRIPS representation.

Given a database representing the initial situation, a goal description, and a set of operators, STRIPS constructs a sequence of operators that will transform the initial situation into one that satisfies the goal description. Conjunctive goals are decomposed and each of their components solved individually. Operators are selected by matching entries on the add list with the current goal. Planning proceeds by backward chaining and goal regression[2]: An operator that has an item on the add list matching a literal of the goal description is selected; the literal is removed from the goal description; the precondition formula is added to the goal description; and the operator is added to the list of operators in the partial plan. This process continues until the goal description matches the initial situation. When all the literals in a conjunctive goal are solved, STRIPS simulates the partial plan to determine whether the conjunctive goal itself is solved. The problem is that operators may interact such that operators in the plan undo the goals accomplished by previous operators. If such interactions are discovered, STRIPS must backtrack and attempt an alternative solution. The creation of a plan in which actions are simply ordered one after another is called *linear planning*. The process of backtracking in linear planning can be seriously inefficient.

At the time an action is chosen, constraints in the current partial plan may leave open several possible action orderings. By specifying the action ordering as little as necessary, the planning process can be made more efficient. This is the technique adopted by various *nonlinear planners* such as NOAH (Sacerdoti, 1975), NONLIN (Tate, 1977), DEVISER (Vere, 1983), SIPE (Wilkins, 1984).

Another way to make the planning process more efficient is to decompose a goal into several noninterfering subgoals. Each subgoal is solved individually and the seprate plans pieced together. A number of planners have been produced that first create component plans under the assumption that the individual subgoals do not interfere. The separate plans are then pieced together and modified if necessary to eliminate undesirable interactions (Sussman, 1975; Tate, 1974).

[2] *Backward chaining* is the process of searching for a plan by working from the goal state back to the initial state. *Goal regression* is the reduction of a goal to a subgoal on the basis of an action description. Execution of the described action in a situation satisfying the subgoal must result in a situation satisfying the goal description.

2.2 Plan representation

STRIPS and most subsequent planning systems can represent only a limited scope of planning problems. This is because of a number of inherent limitations in the STRIPS representation paradigm.

1 Since all changes in the world result from the planning agent's actions and the specification of the planning environment describes only the initial situation, events external to the agent cannot be represented.
2 Because actions are transformations from situation to situation, there is no concept of conditions holding or changing while an action takes place. Thus, only conditions that hold prior to execution can be used as necessary conditions for executing an action. For example, we cannot say that the wind must be blowing while sailing.
3 Only one event can occur between two successive situations, so one cannot represent plans with concurrent actions.
4 Goals are simply conditions that must hold in the final situation. There is no way to represent temporarily qualified goals such as "stay healthy the rest of your life".
5 The logic makes no attempt at representing causation; thus causal implication is not differentiated from material implication. Causation is important in planning since an agent's actions are intended to cause changes to the world.
6 The definition of a plan as a sequence of actions to achieve a particular goal is too restrictive. No facility is present for taking advantage of opportunities that may present themselves, such as seeing a car for sale at an excellent price and planning to buy it.
7 It is assumed that the planning agent has complete knowledge of all relevant aspects of the world and can completely predict the effects of his actions on the world. The representation language contains no facility for reasoning about the state of knowledge of the agent.
8 There is no way to represent quantitative goals such as "maximize wealth" or "Get to Tucson using as little fuel as possible".
9 There is no way of defining the quality of a plan. We have only a simple binary criterion—a plan either achieves the goal or it doesn't.

Much of the recent planning work in AI has concentrated on developing representations that overcome these limitations.

2.2.1 Time

Temporal logics and associated causal theories have been proposed as a means of overcoming limitations 1–5 (McDermott, 1982; Allen & Koomen, 1983; Allen, 1984; Pelavin & Allen, 1986; Shoham, 1987). These logics represent time explicitly rather than implicitly as states resulting from occurrences of events. Events take place over intervals of time, which means that we can reason about concurrent events and conditions during the occurrence of an event. Furthermore, goal descriptions can contain conditions that must hold during an interval.

McDermott (1982) develops a first order temporal logic for reasoning about processes and plans. He describes a semantic theory of future branching time and formalizes it using first-order axioms. So, for example, states are represented as terms in the logic. McDermott represents time as a partially ordered dense infinite set of points. A state is a snapshot of the universe at a point in time. Time chronicles (complete possible histories of the universe through time) are defined as totally ordered sets of states. His ontology distinguishes between facts and events. Both of these hold over time intervals but whereas facts also hold over their subintervals, events do not. Causality is represented with two special purpose predicates that describe an event causing another event and an event causing a fact.

Allen (1983) describes a linear time logic in which the primitive unit of time is the interval rather than the point. There are twelve possible relations between intervals and he provides axioms relating these. The ontology of facts and events is similar to McDermott's.

Pelavin & Allen (1986) is concerned with reasoning about plans that involve concurrent actions and external events. He extends Allen's temporal logic with two modal operators, INEV and IFTRIED, to reason about future branching time and action effects, respectively. IFTRIED is a counterfactual operator that associates the attempt of an action with the truth of a sentence. The semantics of the operator are based on Stalnaker's and Lewis's theories of counterfactuals. IFTRIED captures some aspects of causality, for example, an action cannot effect the state of the world at any time preceding its attempt. Causal effects of general events cannot be represented in the logic. Pelavin provides a formal semantic theory in terms of Kripke structures (Kripke, 1963) in which states are complete chronicles, as well as a sound axiomatization of the logic. Furthermore, he defines plans in terms of sequential and concurrent action composition.

Shoham (1987) presents two temporal logics, one similar to that of McDermott and an alternative modal logic of time intervals. Shoham argues that McDermott's point-based representation is more concise and intuitive than Allen's interval-based representation. He presents an ontology of temporal propositions that is more general than McDermott's or Allen's simple fact/event dichotomy. Within his McDermott-type temporal logic, Shoham develops a sophisticated computational account of causation.

2.2.2 Opportunism

Although opportunistic planning (item 6 above) represents an important type of reasoning, little work has been done in this area. Opportunistic planning involves reasoning about actions in response to observed situations in the environment that can be exploited to the planning agent's benefit. Hayes-Roth and Hayes-Roth (1979) present a cognitive model of opportunistic planning. They motivate and justify the model by analysis of subject protocols of an errand planning task. In this model, a planning agent has multiple goals, categorized by their degree of importance—primary, secondary, etc. The agent is not committed to achieving these goals; rather he looks at the opportunities present in the world to determine what the best plan is in terms of the goals he can achieve. The proposed planning model consists of multiple cooperating specialists working on the problem at various levels of abstraction and communicating through a blackboard. The planning process proceeds in both a top-down, goal-driven direction and a bottom-up, environment-driven direction. This contrasts with the purely top-down approach of traditional planners such as NOAH (Sacerdoti, 1975).

Hayes-Roth and Hayes-Roth view opportunistic planning as a largely procedural matter. Their discussion focuses on their characterization of opportunistic planning as being less structured than traditional planning and tending to proceed in a bottom-up fashion. Their model of opportunism does not distinguish between control issues and representation issues. Our view is that the key characteristic of opportunistic planning is not the way plans are created but rather the way goals are represented.

An opportunity is evaluated relative to an agent's desires. A planning agent that has only one goal to which he is committed and is otherwise indifferent about outcomes has only a very basic notion of opportunity—an opportunity is any situation that allows the achievement of that goal. So the set of possible world states is partitioned into only two types. The more complex the representation of the agent's desires, the more refined the notion of opportunity. An agent that has differential preferences among all possible outcomes of all possible plans will have the most refined notion of opportunity. Furthermore, the number of opportunities the agent can manage is inversely proportional to the number of constraints on his possible plans. The more the agent's goals or desires constrain the set of entertainable plans, the less the agent will be sensitive to factors in the environment.

Wilensky (1983) presents a theory of opportunistic planning that emphasizes representing and reasoning about goals. The planning agent may actively pursue numerous goals simultaneously. Goals may be both introduced and abandoned during the planning process. Goals are introduced by what Wilensky calls themes. A *theme* introduces a goal in response to internal states, e.g. hunger, or situations in the environment, e.g. opportunities or dangers. Themes implicitly encode the agent's desires, such as avoiding hunger or acquiring wealth. Goals may be abandoned when the agent detects interactions that prevent achieving all of his current goals. If two or more goals are incompatible some goals may be dropped. Alternatively, the agent may opt for partial fulfillment of his goals. Both of these mechanisms require a method for evaluating the desirability of various conjunctions of goals. Wilensky proposes assigning a number of each goal to represent its value and assuming that the goal values can be added to derive the value of a conjunction of goals. Reasoning about goals is governed by a set of four meta-themes: (1) do not waste resources, (2) achieve as many goals as possible, (3) maximize the value of the goals achieved, and (4) avoid impossible goals. Components of this planning theory have been implemented in the PANDORA planning system.

Although Wilensky's framework focuses on the right problem, his solution seems somewhat adhoc. It is not clear exactly how themes function and why they are preferable to simply having all the goals present all the time. Wilensky leaves open the question of where the numbers that represent the values of goals come from and what properties they possess. Finally, additivity of goal values is too restrictive an assumption—often the value of one goal will depend on the attainment of another. Wilensky's formulation of opportunistic planning, however, at least approaches the theoretically sound framework provided by decision theory.

2.2.3 Knowledge

In order to effectively plan with incomplete knowledge (item 7 above), an agent must have a planning representation that allows reasoning about his current knowledge state and about how this state may be altered through tests or observations. Ordinary first-order logic as well as situation calculus are insufficiently expressive for reasoning about knowledge. First-order logic allows partial knowledge in the sense that numerous models may satisfy the world description, but it is not possible within the logic to reason about the models. We need a mechanism within the representation language to reason about the state of knowledge explicitly. Researchers have taken two approaches to gain this expressive power: semantic and syntactic.

The semantic approach is most notably represented in the work of Moore (1985), Haas (1985), and Appelt (1985). The semantic approach makes use of modal logic to allow explicit reasoning about knowledge by introducing operators that quantify over ordinary first-order logic models. In modal logic such models are called *possible worlds.* A possible world is a complete description of one possible reality or state of affairs. Model logic is just first-order logic with the addition of modal operators, for example necessity $\Box$ and possibility $\Diamond$. The sentence $\Box\,\phi$ is true iff ϕ is true in all possible worlds and the sentence $\Diamond\phi$ is true iff ϕ is true in some possible world. Thus necessity and possibility amount to universal and existential quantification over the set of possible worlds, respectively. Knowledge is represented in modal logic by interpreting the necessity operator as "know" or "believe". This corresponds quite closely to our intuitive notions of what it means to know something. For example, if we entertain world views in which it is raining as well as ones in which it is not then we cannot be said to know what the weather conditions are. If we let the operator K stand for know then we can explicitly represent an agent's ignorance of the weather using the sentence

$$\neg K(\mathit{raining}) \wedge \neg K(\neg \mathit{raining}).$$

This sentence says that the agent does not know that it is raining and does not know that it is not raining.

Perhaps the best example of the use of modal logic in planning is the work of Moore (1985). Moore used modal logic to formulate a theory of the interaction between knowledge and action. He was concerned with reasoning about what knowledge an agent needs in order to perform an action and about what knowledge may be obtained by performing an action. Moore essentially extended the situation calculus by interpreting situations as possible worlds. The use of modal operators was then essentially equivalent to quantification over situations, something not previously done in planning. He represented actions using a result operator that maps a possible world and an event into the world that results from the occurrence of the event in the first world. An action is just an event caused by the agent.

In investigating the interaction between knowledge and action, Moore studied the problem of knowledge prerequisites for actions. His basic tenet was that these could be analyzed in terms of knowing what action to take and that this means knowing what action is referred to by an action description. An Agent can be said to know who or what something is if he has a rigid designator for that thing. A *rigid designator* is a term that denotes the same domain object in all possible worlds. For example, suppose that we have the action schema *Dial-combination* which takes as arguments a safe to be opened and its combination. We assume that the agent can open the safe once he knows which safe is to be opened as well as the combination, i.e., we assume that *Dial-combination* is a rigid function. Then the agent can carry out the action if he has rigid designators for the safe and the combination.

Moore's theory is capable of dealing with many aspects of dynamic knowledge in planning: The theory allows an agent to reason about his state of knowledge; it provides a general principle of what it means to know how to perform an action; and it provides a means of describing the effects of action on knowledge. Within this formalism, an agent can reason about what knowledge is required to perform an action, whether he possesses this knowledge, and what actions he may execute to gain the requisite knowledge.

Despite its elegance, Moore's modal logic approach has a number of important limitations. First, there is no way to distinguish between likely and unlikely states of the world. Thus Moore's theory must treat nearly certain outcomes and virtual impossibilities equally. In contrast, knowledge of the likelihoods of possibilities would allow an intelligent agent in many cases to focus its attention on just a few of the myriad different possibilities. Second, Morgenstern (1986) has pointed out that knowing how to perform an action is not the same thing as having a rigid designator for it. For example, calling Jane may involve pushing buttons or dialing depending on the world we find ourselves in. A rigid designator for dial would denote only one of these actions in all possible worlds. Third, the use of modal necessity to represent knowledge implies that an agent knows the deductive closure of all his beliefs. For example, this would mean that an agent that knows Peano's axioms would know whether Fermat's last theorem was true. Finally, Moore's approach does not allow quantification over propositions. For example, the approach cannot represent that John believed everything Mary told him.

Syntactic theories of belief (Haas, 1986; Morgenstern, 1986) were proposed as solutions to the last three problems mentioned above. The key idea behind this approach is the use of a quoting convention to allow a first order language to talk about sentences by referring to their names rather than having to use a higher order or modal logic. A believe predicate is introduced and appropriate axioms written to specify the desired properties of belief. For example, to say that John believes snow is white we would write (believe John "(white snow)"). Haas discusses in detail the use of his syntactic theory for reasoning about action and knowledge. Both Haas and Morgenstern note that the syntactic approach suffers from the problem of self-referential sentences. Thus the language allows construction of a sentence S such that S is true iff (know a (not S)) is true. Such a sentence is inconsistent since knowlege is taken as implying truth. Morgenstern proposes using a three-valued logic to solve this problem by assigning such sentences an indeterminate truth value.

2.2.4 Remaining issues

A number of the representation issues listed earlier remain open. Indeterministic causation is not handled by any of the existing temporal logics. Research has not addressed the representation of quantitative goals. A theoretically sound framework for opportunistic planning is lacking. Both modal and syntactic representations of knowledge cannot represent degrees of possibility. None of the representational frameworks defines plan quality in a satisfying way. Decision theory will be shown to address all of these issues, at least in part.

3 Decision theory

By viewing a decision as a commitment to an action, we may see planning as a form of sequential decision making, a subject that is addressed by a large body of literature in the area of Decision Theory (Berger, 1980; DeGroot, 1970; Raiffa, 1968; Raiffa & Schlaifer, 1961). Decision Theory emphasizes the representation of degrees of likelihood and of desires. The central concern of Decision Theory is the optimal choice of actions with respect to likelihoods and desires. Degrees of likelihood can be taken to represent the knowledge state of an agent. This knowledge state can be updated to adapt to a dynamic environment. Optimality of choice is defined in terms of maximum likelihood of maximally satisfying desires. Thus in formulating a plan, an agent is free to incorporate any opportunity likely to lead to a highly desirable outcome. The quality of a plan is defined in terms of its likelihood of achieving desirable and undesirable outcomes. These points and other strengths of the decision theoretic representation are discussed in detail in the following sections.

In the decision theoretic framework, desires are represented by *utilities.* Utility is an abstract scalar quantity that represents the decision maker's desires concerning possible outcomes of his actions. The decision maker may have multiple factors with respect to which he evaluates outcomes, each possibly without an obvious scale of value. For example, we might evaluate a car with respect to price, styling, reliability, comfort, etc. Even when there is a clear scale, the scale may not reflect the true value to the decision maker. The assignment of utilities to outcomes allows representation of relative desirability in such situations.

Decision Theory uses probability to represent likelihood or partial belief. Probability may be interpreted in various ways. Three main schools of thought have been predominant: logical (Keynes, 1921; Carnap, 1950), frequentist (Venn, 1986; von Mieses, 1957), and subjectivist (Ramsey, 1926; Savage, 1954; Kyburg et al., 1964). The logical theory takes probabilities to represent a logical relationship between a given hypothesis and given evidence. The frequentist view identifies probability with some suitably defined relative frequency. The subjectivist school takes probability to represent the degree of belief of a rational agent. Of the three views, the subjective interpretation is the most appropriate for representing the epistemic states of a planning agent. In a logic for reasoning about plans, we would like the quality of a plan to be defined in such a way that the behaviour recommended by the logic conforms to commonly accepted standards of rational behaviour. Subjective decision theory defines the optimality of decisions in terms of rationality of behaviour with respect to a set of beliefs and desires. Under the subjective interpretation, probabilities represent an ideally rational agent's degrees of belief, where belief is defined as a disposition toward action:

> We are driven therefore to the second supposition that the degree of a belief is a causal property of it, which we can express vaguely as the extent to which we are prepared to act on it. F.P. Ramsey (1926)

Concretely, subjective probabilities are interpreted in terms of willingness to bet. Thus to have a greater degree of belief in one proposition than another is to be willing to bet at higher odds on the first than on the second. Ramsey justifies this interpretation as follows:

> First, it is based fundamentally on betting, but this will not seem unreasonable when it is seen that all our lives we are in a sense betting. Whenever we go to the station we are betting that a train will really run, and if we had not a sufficient degree of belief in this we should decline the bet and stay home. F.P. Ramsey (1926)

This functional definition of belief is precisely what we want for building an intelligent planning agent because it gives us a clear mapping from the agent's internal state to the agent's behaviour. Furthermore, this definition makes belief a measurable quantity so that we can elicit people's degrees of belief.

Subjectivists define optimality of decisions by using the notion of *rationality.* It is difficult to find a clear definition of rationality in the literature and there is considerable debate as to what constitutes rational behaviour. It is clear, however, that just as categorical logic tells us what truth values we may assign to logical formulas given our assignments to other formulas, so a logic of partial belief should tell us what degrees of belief we may assign to events given our assignments to other events. So rationality concerns the consistency of beliefs with respect to other beliefs. In fact, the axioms of probability will show that the consistency requirements for probability subsume logical consistency. But there is more to rationality than this. Any decision theory is composed of three components: a representation of beliefs, a representation of desires, and a decision rule. Corresponding to these components, there are three basic aspects of rationality: (i) consistency of beliefs, (ii) consistency of desires, and (iii) consistency of choices with respect to beliefs and desires. In the next two sections we discuss *synchronic rationality*, the rationality of an agent at a fixed point in time. In section 3.2.1 we discuss *diachronic rationality*, the rationality of belief change. In section 3.3 we examine the relationship between causality and rationality.

3.1 Savage's decision theory

The discussion in this section follows loosely the development in (Kreps, 1988, Ch. 9). Savage's theory of choice under uncertainty (1954) is the classic and still most elegant formulation of subjective Bayesian Decision Theory. We begin by defining some terminology. Note that some of the terminology conflicts with that used in AI planning. Nevertheless, we use Savage's terminology to acquaint readers unfamiliar with Decision Theory with the standard terminology of the field. The *world* will be that object about which the decision making agent is concerned. A *state* of the world is a description of the world that leaves no relevant aspect undescribed. An *event* is a set of states of the world. A *consequence* (*or outcome*) is anything that may happen to the agent. An *act* is a function from the set of states to the set of consequences, so acts are deterministic in this framework. The components of the theory are: a set S of states, a set Z of consequences, a set F of acts, and a binary relation $\preceq$ that gives the agent's preference over the set of acts. The consequence of act f in state s will be denoted $f(s)$. Note that for conciseness we will use notation like $x \preceq f$ and $x \preceq y$, where x and y denote acts with constant consequences, independent of the state. Two important restrictions concerning the choice of states, acts, and consequences should be noted. The consequences must specify all factors relevant to the agent's decision and the states must be causally independent of the acts. More will be said about the second restriction in section 3.3.

Savage's theory is concerned with the consistent assignment of probabilities to states and utilities to consequences such that choosing the act that maximizes expected utility results in a choice that reflects the agent's preferences. The *expected utility* of an act is the sum of the utilities of its consequences, weighted by their probabilities. Savage provides seven axioms concerning the agent's preferences that guarantee the existence of such probability and utility functions. These axioms are intended to describe intuitive characteristics of preferences which people would agree should be possessed by a rational agent.

A1 $\preceq$ is a simple ordering.[3]

A2 There is at least one pair of consequences x and z such that $x \prec z$.

The first axiom says that the agent has preferences between any two consequences and that these are transitive. The second axiom simply guarantees that the agent has some actual preferences.

A3 Suppose f, g, f', and $g' \in F$ and $A \subseteq S$ are such that
(a) $f = f'$ on A, $g = g'$ on A,
(b) $f = g$ on $\bar{A}$, $f' = g'$ on $\bar{A}$,
(c) and $f \preceq g$,
then $f' \preceq g'$

[3] A relation $\preceq$ is a simple ordering iff for all x, y, z (1) either $x \preceq y$ or $y \preceq z$ and (2) $x \preceq y$ and $y \preceq z$ implies $x \preceq z$.

This axiom says that if f and g are two acts that agree with each other in $\bar{A}$ and $f \preceq g$; then if f and g are modified in $\bar{A}$ in any way such that the modified acts f' and g' still agree with each other in $\bar{A}$, it will also be the case that $f' \preceq g'$. The remaining axioms will make use of the notion of *conditional preference*. We say that $f \preceq g$ given A if $f' \preceq g'$, where $f' = f$ and $g' = g$ on A and $f' = g'$ off of A.

A4 If A is not null and $f(s) = x$ and $g(s) = y$ for all $s \in A$, then $f \preceq g$ given A iff $x \preceq y$.

An event A is *null* if for all f and g, $f \preceq g$ given A. The above axiom says that preference is not state dependent.

A5 Suppose x, y, x', y', f, g, f', g', a, and b are such that
(a) $y \prec x$ and $y' \prec x'$,
(b) $f(s) = x$ and $f'(s) = x'$ on A and $f(s) = y$ and $f'(s) = y'$ on $\bar{A}$, and
(c) $g(s) = x$ and $g'(s) = x'$ on B, and $g(s) = y$ and $g'(s) = y'$ on $\bar{B}$, and
(d) $f \preceq g$
then $f' \preceq g'$.

The axiom states the assumption that on which of two events, A and B, an agent chooses to stake a given prize does not depend on the prize itself.

A6 For all $A \subseteq S$,
(a) If $f \preceq g(s)$ given A for all $s \in A$ then $f \preceq g$ given A.
(b) If $g(s) \preceq f$ given A for all $s \in A$ then $g \preceq f$ given A.

The axiom says that if every possible consequence of the act g is at least as preferable to the agent as the act f considered as a whole, then $f \preceq g$, and similarly for the preferences reversed.

A7 If $f \prec g$ and x is any consequence; then there exists a finite partition of S such that, if f or g is so modified on any one element of the partition as to take the value x at every s there, other values being unchanged; then the modified f remains less than g, or the modified g remains less than the modified f, as the case may be.

This axiom does two important things. First, it restricts how good or bad a particular consequence can be. Second, it restricts how small a difference there can be between two acts f and g and still have $f \prec g$.

Together the seven axioms imply that

1 There exists a unique probability measure p on the algebra of all subsets of S, such that a is judged more likely than b by the agent iff $p(a) > p(b)$.
2 For p given above, there is a bounded utility function $u: Z \rightarrow R$ such that $f \preceq g$ iff $E[u(f(s)); p] \leqslant E[u(g(s)); p]$.[4] The utility function u is unique up to a positive affine transformation.

As a consequence, an agent is said to act rationally if he chooses the act a that maximizes his subjective expected utility.

$$SEU(a) = \sum_{s \in S} P(s)u(a(s)).$$

Savage has shown that the existence of a probability measure and a utility function that reflect the agent's preferences is a consequence of a set of axioms concerning the agent's qualitative preferences. Note that the fact that p is a probability measure implies that the agent's degrees of belief satisfy the axioms of probability. This will be discussed further in the next section. Savage's development constitutes what is called a *representation theorem*, in which the existence of probabilities and utilities are shown to follow from a set of intuitive properties. Other representation theorems are given by Ramsey (1926) and Jeffrey (1965). For a survey of representation theorems, see (Fishburn, 1981).

[4] $E[x; p]$ denotes the expected value of x w.r.t. p.

3.2 Probability theory

Mathematically speaking, a probability distribution P is a measure defined on a σ-algebra[5] $\mathfrak{F}$ over some set Ω, such that the following properties are satisfied

1 $0 \leqslant P(A) \leqslant 1$ for all $A \in \mathfrak{F}$
2 $P(\Omega) = 1$
3 If $A_1, A_2, \ldots \in \mathfrak{F}$ is a disjoint sequence of sets then $P(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k)$

One way to think of Ω is as the set of all possible outcomes of an experiment or observation. The probability of some event E is then the probability of the set of possible outcomes described by E.

Probability theory can be viewed as a generalization of modal logic (Gaifman, 1986; Haddawy & Frisch, 1990; Halpern, 1989). If we take the set over which probabilities are defined to be the set of possible worlds then we can see that probabilities are a way of assigning measures to sets of possible worlds. Thus probability, like modal logic, provides us with a way of talking about sets of possible models or world states. The difference is that with probabilities we can also talk about how likely these sets are.

The above three *axioms of probability* guarantee the consistency of a set of partial beliefs. A consistent set of beliefs is called *coherent*. The most widely accepted notion of coherence is that of *weak coherence* which requires that there be no set of odds determined by the agent's degrees of belief under which he is bound to lose, whatever the state of the world. This is intuitively appealing since an agent should not wish to lose objects of positive value under all possible conditions. It can be shown through the use of *Dutch Book* arguments (Shimony, 1955; Kemeny, 1955; Lehman, 1955) that conformity of beliefs to the axioms of probability is a necessary and sufficient condition for weak coherence. The Dutch Book argument represents an alternative to representation theorems as a means of justifying the axioms of probability. In more recent years it has fallen out of favour due to some of its restrictive assumptions, such as assuming a linear utility function. Nevertheless, its simplicity makes it appealing.

For example, let pS be the highest price you would pay for a lottery ticket which entitles the holder to a prize of S cents if A occurs and nothing otherwise. Then p is your *critical odds* on A at stake S. We define your degree of belief in A as your critical odds on A at a modest stake. We use a modest stake because people's utility functions over money are typically nonlinear, but are approximately linear over small ranges. One consequence of the axioms of probability is that $P(A) = 1 - P(\neg A)$. Suppose that you are playing bookie and that your beliefs violate this identity in the following sense: You are willing to offer odds of $\frac{1}{3}$ on A and odds of $\frac{1}{3}$ on $\neg A$. Now the better can fix the stakes so as to insure himself any positive net gain he desires. Suppose he fixes the stakes at \$10 on each of the two bets. Then he pays \$3.33 for each of the bets. Now no matter which result obtains he will receive a net profit of \$6.66 at your expense.

It should be emphasized that the above axioms of probability are consistency criteria and will not guarantee that an agent's beliefs will be "correct". This has a number of obvious consequences. First, two agents may have very different degrees of belief for the same event. Second, rationality does not guarantee good outcomes. If an agent's beliefs are not correlated with the frequency of events in the world then the agent may make rational decisions with poor outcomes, even in the long run. Rationality simply requires an agent's decisions to be consistent with his beliefs and desires. Furthermore, the rule of maximizing expected utility leads to decisions that roughly have a higher likelihood of leading to a good outcome than any other available decision. However, it is still perfectly possible for the outcome to be poor.

The advantage of using a decision theoretic framework with rationality as a performance criteria over say logical consistency is that it integrates an agent's beliefs and preferences into the planning process. This way plans can be ordered according to their desirability and the planning agent can

[5] A class $\mathfrak{F}$ of subsets of Ω is called a σ-algebra if it contains Ω and is closed under the formation of complements and countable unions.

focus his attention on more likely outcomes as well as outcomes with particularly extreme utility values.

3.2.1 Conditionalization

In the previous sections we discussed synchronic rationality. Now we move to the discussion of diachronic rationality, the rationality of belief change. In simple planning domains like the blocks world, it may be reasonable to assume that an agent can be initially endowed with all the knowledge he will ever need but in complex domains it is not possible to decide a priori what knowledge the agent will need to carry out all his possible tasks. Furthermore, if the world is dynamic then a world model which was accurate at one time may cease to be useful at a later time. An agent must be able to update his knowledge in order to keep his world model current. Thus, we are interested in the question of how a rational agent should incorporate information received from experimentation or observation.

The commonly accepted way of updating probabilities is through conditioning. The *probability of A conditional on or given B*, written $P(A \mid B)$, is typically defined as $P(AB)/P(B)$, when $P(B) > 0$ and undefined otherwise. If the effect of observing B is taken as raising the probability of B to 1 then conditioning can be used to assimilate this information by updating the probability of A from $P(A)$ to $P'(A) = P(AB)/P(B)$. This updating process is known as *Bayesian conditioning*. For example, suppose a die is tossed and we are told that it came up even. What is our probability now that it came up a six?

$$P(six \mid even) = P(six \,\&\, even)/P(even) = 1/3.$$

The question of whether conditioning is the only rational means of updating beliefs is, however, still open to debate.

Levi (1980; 1987) distinguishes between two distinct concepts that authors have typically lumped together under the term "conditionalization": confirmational conditionalization and temporal credal conditionalization. The principle of confirmational conditionalization says that an agent's confirmational commitments (conditional probabilities) represent the agent's current assessment of what his beliefs should be were his state of knowledge expanded in some way. Temporal credal conditionalization, on the other hand, says that if an agent's state of knowledge is expanded, he should update his remaining beliefs in accordance with his confirmational commitments. Temporal credal conditionalization is equivalent to confirmational conditionalization with the additional constraint of "confirmational tenacity". Confirmational tenacity says that an agent ought to hold his confirmational commitments fixed over the time that his knowledge is expanded. Closely related to confirmational tenacity is van Fraassen's *reflection principle* (1984). The principle states that $p(H \mid R_q) = q(H)$, where p is the agent's current probability function and R_q denotes that at some future time he will have probability function q. Reflection together with confirmational conditionalization imply temporal credal conditionalization.

de Finetti (1937) gives a Dutch Book argument for defining the conditional probability of A given B as $P(AB)/P(B)$ by using the notion of a conditional bet. A bet on A conditional on B is won or lost according as A occurs or not if B occurs and is called off if B does not occur. This argument pertains only to confirmational conditionalization and thus is a property of synchronic rationality.

Philosophers have attempted to give similar diachronic Dutch Book arguments in support of temporal credal conditionalization (van Fraassen, 1984; Goldstein, 1983). These arguments purport to show that if an agent's beliefs violate the *reflection principle* then he will accept a set of bets concerning his present and future beliefs which will lead to a certain loss. Both Levi (1987) and Maher (1990a) have shown the diachronic Dutch Book arguments to be fallacious. Indeed, Maher (Maher, 1990a; Maher, 1990b) provides some examples in which it may be rational for an agent to update his beliefs in ways other than by conditionalizing.

Then why prefer Bayesian conditioning over any other method of belief revision that results in coherent probabilities? There are two arguments that tend to recommend its use. Maher (1990a) provides the strongest defensible argument for the rationality of conditionalization. He first provides a result concerning the reflection principle, showing that "other things being equal, implementing a shift which violates reflection cannot have greater expected utility than implementing a shift which satisfies reflection". The "other things being equal" refers to assumptions in the proof, the most important being that the agent's acts not influence his utility other than via their influence on his subsequent decisions. This rules out situations in which beliefs influence utility other than via choice, as well as situations in which acts influence utility by providing evidence for states which they have no tendency to cause. Maher calls acts of the latter type *symptomatic*. More will be said about these in the next section. Finally, Maher's result concerning conditionalization is stated as follows

> The principle of conditionalization is a special case of the principle which says not to implement shifts that violate reflection. Like that more general principle, it is not a universal requirement of rationality; but it is a rationally acceptable principle in contexts where other things are equal...

The other, more classical argument for Bayesian conditioning has to do with convergence of beliefs. It was pointed out earlier that the subjective view of probability requires no agreement among the beliefs of different agents. If, however, two agents update their probabilities by conditioning on their observations and they share a common series of observations then their beliefs will tend to converge to a common distribution which attaches probability 1 to the true hypothesis. Savage (1954, section 3.6) shows this result for a finite set of agents who agree on which hypotheses have probability 0 and on the likelihoods for these hypotheses over an infinite sequence of observations which are independent and identically distributed given a hypothesis. Generalizations of this argument are given by Blackwell & Dubins (1962) and Schervish & Seidenfeld (1987). The latter develop an argument that does not require beliefs to be updated according to conditionalization. Thus Bayesian conditioning is a sufficient condition for convergence but not a necessary one.

If the means for acquiring information in a Bayesian framework is through conditioning, we must have some initial or prior probabilities to update. Just how these priors are determined is still the subject of much debate. Attempts at obtaining "objective" priors have lead to the use of principles such as invariance (Jaynes, 1968; Jeffreys, 1961; Villegas, 1977) and maximum entropy (Cheeseman, 1983; Jaynes, 1979; Konolige, 1979). Invariance says that the probabilities should be invariant with respect to certain transformations such as location or scale. The maximum entropy principle says that we should choose the prior distribution which satisfies any constraints we might have and which maximizes entropy. For a critique of the maximum entropy approach, see (Seidenfeld, 1986). Numerous other methods have been developed for eliciting subjective priors from human subjects (Hogarth, 1975; Savage, 1971). For a good discussion on the topic of priors see (Berger, 1980, Ch. 3).

It is often the case with tests that the conditional probabilities are most readily available as the probability of evidence given hypothesis, while we need the other direction for Bayesian conditioning. Bayes rule is useful in these situations. The rule says that

$$P(H \mid E) = \frac{P(E \mid H)P(H)}{P(E)}$$

and is derivable directly from the definition of conditional probability.

3.3 Causality and rationality

Thus far in our discussion of rationality we have focused on the representation of beliefs and desires. Now we examine the way in which causal factors influence problem representation. We outline the arguments for the progression from classical decision theory to causal decision theory and show that any decision theory that ignores causality will prescribe irrational behaviour in some situations.

The subsequent discussion follows the development in (Maher, 1990b). We mentioned in section 3.1 that within Savage's decision theory, states must be causally independent of acts. If this is not the case, the rule of maximizing subjective expected utility (SEU) may lead to irrational choices. This is easily shown with the following example of a smoker faced with the decision of whether to quit smoking (Maher, 1987). Suppose the available acts are "smoke" and "don't smoke", the states of the world are "cancer" and "no cancer", and the person's utilities are such that

u(smoke(cancer)) > u(don't smoke(cancer))
u(smoke(no cancer)) > u(don't smoke(no cancer)).

The rule of maximizing *SEU* then recommends that the best decision is to smoke. Of course this may not be rational because the act of smoking increases the chance of contracting cancer. In order to allow for situations in which acts influence states, Richard Jeffrey (1965) proposed a decision rule that looks at the probability of a state conditional on an act. His decision rule was to maximize conditional subjective expected utility

$$CSEU(a) = \sum_{s \in S} P(s \mid a)u(a(s)).$$

This rule recommends that the smoker refrain from smoking if smoking and cancer are correlated strongly enough. And the principle of maximizing *CSEU* agrees with the principle of maximizing *SEU* when the states are conditionally independent of the acts. This rule is fine if acts serve as evidence for states that they tend to cause but the principle of maximizing *CSEU* can lead to irrational choices if acts are symptomatic, i.e., they serve as evidence for states which they have no tendency to cause.

One of the most common examples of a symptomatic act is when an act and a state both tend to be brought about by a common cause. The following example, due to R.A. Fisher (1959) illustrates this case. Consider the situation of the smoker faced with the choice between smoking and not smoking. Suppose now that the tendency to smoke and the tendency to contract lung cancer are correlated not because smoking causes cancer but because they are both due to a common genetic cause. Suppose you do not know whether or not you possess the gene. If you decide to smoke, this will be good evidence that you possess the gene and, hence, will contract lung cancer. If you decide not to smoke, that will be good evidence that you do not possess the gene and, hence, will not contract lung cancer. Suppose that

P(cancer | smoke) = .9	P(cancer | don't smoke) = .1
u(smoke(no cancer)) = 10	u(don't smoke(no cancer)) = 9
u(smoke(cancer)) = 1	u(don't smoke(cancer)) = 0.

Then the CSEU of smoking is 1.9 while that of not smoking is 8.1. So the rule of maximizing CSEU recommends not smoking. But whether or not you smoke has no influence on whether you contract cancer in this case, and the rational choice is to smoke.

Various researchers have proposed Causal Decision Theories (CDTs) as a means of allowing states to be influenced by acts while properly handling symptomatic acts. For a unifying view of Causal Decision Theory see (Lewis, 1981). Two basic approaches have been taken. The first approach by Gibbard & Harper (1978) augments Savage's theory in such a way that states are forced to be causally independent of acts. This is done by representing states as conjunctions of counterfactuals. So for example in the smoking example presented earlier, one state might be: 'If I were to smoke then I would get cancer and if I were to abstain from smoking I would not get cancer." The causal dependence is expressed by the pattern of counterfactuals in each state. Applying this theory to the second smoking example and taking the states to be composed of counterfactuals of the form "If I were to smoke, I would contract cancer", yields the correct recommendation to smoke.

The other more general approach due to Skyrms (1980) gives up Savage's requirement that acts be functions from states to consequences and allows acts to be indeterministic. Skyrms represents indeterministic causality through objective probabilities. Causal factors are represented by *dependency hypotheses.* A proposition k is a dependency hypothesis if for all acts a and consequences c, kE specifies the objective probability of c given a, and no weaker hypothesis than k specifies all these probabilities. Thus for all acts a and consequences c, $p(c \mid ka)$ represents the objective probability of c given a. According to Skyrms's version of CDT, an agent is acting rationally just in case he chooses the act that maximizes the subjective expected objective expected utility

$$EOEU(a) = \sum_{k \in K} P(k) \sum_{c \in C} P(c \mid ka)u(c),$$

where C is the set of consequences, K is the set of dependency hypotheses, and the $P(k)$ represent the agent's subjective belief concerning which dependency hypothesis is correct. If we apply this CDT to the smoking problem and take the dependency hypotheses to be whether or not smoking causes cancer then we get the correct recommendation to smoke.

Until very recently it was thought that CDT was a general solution to the problem of symptomatic acts. Maher (1990b) has shown, however, that for sequential decision problems in which acts are symptomatic, CDT may recommend irrational choices. More specifically, suppose we have a symptomatic act that we choose to perform. Once we perform the act, we update our probabilities by conditioning on this knowledge. Now if we make a second decision, that decision will be influenced evidentially by the knowledge of performance of the first act. Maher provides scenarios in which CDT may recommend against cost-free observation.

For single stage decision problems, CDT recommends rational choices; it is only in the case of sequential decision problems that it has trouble. Thus CDT preserves synchronic rationality, which recommends it for decision problems where acts may be symptomatic. This result has an interesting relation to our earlier discussion of diachronic rationality. Remember that there is no strong rationality argument for conditioning as the only means of updating probabilities. These results discussed here suggest further cases (when acts are symptomatic) in which it may not be rational to update by conditioning.

3.4 Decision theoretic plan representation

Decision Theory can contribute to building robust planning systems in numerous ways. When planning under uncertainty, there is a particularly strong need to distinguish between evidential and causal correlation. The various Causal Decision Theories provide frameworks for representing causal factors that influence decision. Skyrms's theory is capable of representing indeterministic cause.

Decision theory addresses many issues dealt with in modal representation of knowledge. Probabilities, interpreted as degrees of belief represent the knowledge state of a rational agent. The agent can reason about the degree of likelihood of various possibilities. The agent's knowledge state can be updated by conditioning to adapt to a dynamic environment. Planning using the quantiative units of probability and utility provides a way of defining the value of information. If we express the cost of obtaining information in the same units as utility then we can weigh that cost against the projected benefits of the information in terms of reduction in uncertainty.

Traditional AI planners measure plan quality according to the heuristic metric of the number of actions involved in achieving the goal—the fewer the number of actions, the better the plan. A rational agent chooses plans that maximize expected utility, determined by its probabilities over states and its utilities over outcomes. So in Decision Theory, the quality of a plan is defined in terms of its expected utility and perhaps other factors such as variance and sensitivity to errors in information. This definition of plan quality is sensitive to factors like the probability of achieving both desirable and undesirable outcomes and the costs of executing the actions composing the plan.

A direct consequence of this representation of plan quality is the ability to define the notion of an approximate plan. The traditional planning paradigm provides no meaningful notion of an approximate plan—a plan either achieves the goal or it doesn't. Within the decision theoretic framework each plan has an associated expected utility. Instead of simply achieving a property we are using the more general notion of maximizing a property. An approximate plan is one which has a lower expected utility than the optimal plan. Approximate plans are valuable in situations where a planner has to respond in real time or within certain time constraints.

4 Combining AI and decision theoretic approaches to planning

4.1 Goals versus preferences

There is a basic conflict between the concept of a plan in AI and in Decision Theory. This conflict centers around the notion of a planning agent's goal. A goal is a symbolic description of the world, such as "show a profit this month" or "have block A on block B". In the traditional AI view of planning, a plan is a sequence of actions which when executed will cause the goal condition to be satisfied. So the planning agent is committed to achieving the goal. In contrast, decision theoretic planning is opportunity driven. A plan is simply any sequence of actions and the best plan is the one that maximizes expected utility. There is no one goal; rather the planning agent's desires are represented in terms of preferences over outcomes. This view of planning has the advantage that the planning agent is responsive to environmental factors but it also has two major drawbacks. First, the set of possible plans that must be considered is immense, since absolutely any sequence of actions is a possible plan. Second, if we wish to adhere strictly to Decision Theory, plans must project infinitely into the future. Since outcomes include all consequences of actions, we must consider all the ways an action might affect the future course of events. So there is no point at which we can stop planning. This concept of planning is, of course, unacceptable for realistic planning problems. As Savage put it

> Making an extreme idealization, which has in principle guided the whole argument of this book thus far, a person has only one decision to make in his whole life. He must, namely, decide how to live, and this he might in principle do once and for all. (Savage, 1954, p. 83)

It would be useful to incorporate the concept of a symbolic goal into the decision theoretic framework in order to focus the attention of the planner and to provide it with a criterion for determining how far into the future to plan.

There are two obvious ways of accomplishing this synthesis: one involving the notion of degree of commitment to a goal and the other involving restrictions on the form of the utility functions. We will say that an agent is committed to a goal to degree α if he only considers as admissible plans with probability at least α of achieving the goal, where the goal, as above, is symbolic description of the world. This definition of goal would focus a planner's attention by eliminating some plans from consideration and by providing symbolic information to guide search. Temporally qualified goals are particularly well suited for terminating plan elaboration. Since each action takes a positive amount of time, for many goals there will be a point after which further planning will not alter the probability of achieving the goal. Once it is shown that further actions along a plan path will not change the probability of achieving the goal, elaboration of the plan path can be stopped. If the elaborated plan has probability at least α of achieving the goal, it is kept for comparison with other plans. If not, it is discarded.Of course, by terminating plan elaboration we are open to a horizon effect that can cause us to ignore important potential consequences. For example, in an extreme case all actions following the last one explored could result in the planning agent's death. So by introducing goals into the decision theoretic framework, we are trading rationality in the ideal sense for practicality of the planning process.

What sort of behaviour would various degrees of commitment produce? A planner with minimal commitment to the goal would consider all plans with non-zero probability of achieving the goal. Since just about any plan has some minuscule probability of achieving the goal, this degree of commitment would not rule out many plans. Even this minimal commitment, however, would still be useful in determining when the planner can terminate plan elaboration.

A planner with complete commitment to a goal would only consider plans with probability one of achieving the goal. This would result in planning behaviour characteric of traditional AI planners such as STRIPS, with the added benefit that the planner could compare the quality of alternative plans that achieve the goal.

The problem with any non-zero degree of commitment is that we are imposing an arbitrary cutoff criterion. A plan with a probability of achieving the goal slightly lower than the degree of commitment would be deemed inadmissible. We would like some way of rewarding or penalizing plans based in a continuous way on their probability of achieving the goal. This suggests defining the goal in terms of the utility function. We could say that an agent has a goal G iff he prefers every outcome in which G is true to every outcome in which G is false. A utility function of this form is known as a lexicographic preference ordering (Keeney & Raiffa, 1976). In general, lexicographic ordering can have a sequence of attributes each of which takes precedence over those later in the sequence. Useful ways of restricting the forms of utility functions are discussed in the literature on Multiattribute Utility Theory (Keeney and Raiffa, 1976).

This notion of goal produces a continuous measure of plan quality based somewhat on the probability of achieving the goal but it does not serve the two purposes for which goals are most useful. First, the definition is too weak to provide any focusing effect. We must still consider plans with a low probability of producing an outcome that achieves the goal. Consider a plan with two outcomes: one in which the goal is true and one in which it is false. Suppose the outcome in which the goal is true has a very high utility and the other outcome has only slightly lower utility than the lowest utility outcome in which the goal is true. Even if this plan's probability of achieving the goal is small it could have a high expected utility. Furthermore, this definition provides no stopping criterion for planning.

4.2 Research

The first planning framework to incorporate both decision theoretic and AI planning techniques was described by Feldman and Sproull (1977). They augmented the basic STRIPS representation by introducing costs associated with actions, utilities of achieving various goals, and probabilities of states. By assuming that the utility of a plan could be calculated in terms of the utility of the goal achieved minus the costs of the actions involved, they could determine the expected utility of each plan and find the optimal plan. Furthermore, they introduced test actions that functioned by updating the probabilities of certain propositions. This work was pioneering in its intent but somewhat limited in its scope. Because it was fundamentally based on the STRIPS representation, it retained many of the STRIPS limitations. Many simplifying assumptions were impliticly incorporated into the representation of probabilities and utilities. Finally, the framework was developed specifically within the context of the simplistic monkey and bananas planning problem, making its generality unclear.

More recently, Wellman (1988) has investigated the use of decision theory in qualitatively reasoning about the utility of plans. His SUDO-PLANNER proves decision theoretically that certain classes of plans are dominated based on qualitative relations in the domain. He uses what he calls qualitative probabilistic networks (QPNs) to express constraints on the joint probability distribution over a set of variables. QPNs are constructed from a knowledge base describing the effects of actions and relations among events at multiple levels of abstraction. Dominance relations are established by manipulating the networks. He applies his system to a medical decision example.

Recently AI researchers have investigated the integration of temporal and decision theoretic representations. This work has focused on solutions to the persistence problem. Persistence is concerned with reasoning about how long a condition will remain true in the absence of explicit information concerning its truth or falsity in the future. It is closely related to the frame problem. Since plans are typically created to be executed some time in the future it is important for a planner to be able to reason about what the state of the world is likely to be at execution time. Hanks (1988, 1989) has developed a representation and system for computing an agent's strength of belief in a proposition at a point in time, based on observations about that proposition and information about the tendency of the proposition to persist over time. Time is represented as a set of totally ordered points and facts are believed with a particular strength to be true at a point in time. The probability of persistence of a condition between observations is calculated according to some general probabilistic frame axioms.

Dean and Kanazawa (1988a, 1988b) have also addressed the problem of persistence but have taken a somewhat different approach. They too model time as points and a fact is true at a point in time. Whereas Hanks uses general conditional probabilities to represent persistence rules, they use survivor functions from queuing theory. They consider both exponential and piecewise linear functions.

5 Summary

Research on planning in AI can be separated into the two major areas: plan generation and plan representation. Most AI planners to date have been based on the STRIPS planning representation. This representation has a number of limitations. Much recent work in plan representation has addressed these limitations. It was shown that Decision Theory can be used to remove a number of the limitations. Furthermore, the decision theoretic framework provides a precise definition of rational behaviour. There remain open questions within decision theory regarding belief revision and causality. It should be noted that these problems are not artifacts of the representation. Rather they arise because the rich representation allows their formulation. Some work integrating AI and decision theoretic approaches to planning has been done but this remains a largely untouched research area.

We see two main avenues for fruitful research. First, the straightforward decision theoretic formulation of planning is computationally impractical. Techniques need to be developed to do efficient decision theoretic planning. Work in AI plan generation has exploited information contained the structure of qualitative representations to guide efficient plan construction. These techniques should be applied to decision theoretic representations as well. Second, AI has developed many representations that allow useful structuring of knowledge about the world. Decision Theory has concentrated on representing beliefs and desires. Integration of AI and decision theoretic representations would yield powerful representation languages. Some of the benefits of such work can already be seen in the research combining temporal and decision theoretic representations.

Acknowledgements

The authors would like to thank Steve Hanks for sharing his insightful thoughts on modal logic and probability. We thank Patrick Maher for his thoughtful criticisms and for allowing us to see his prepublication manuscript. We would also like to thank Alan Frisch, Gregg Collins, George Monahan, Marianne Winslett, Rich Scherl, and an anonymous reviewer for helpful comments on an earlier draft of this paper.

References

Allen, JF, 1984, "Towards a general theory of action and time" *Artificial Intelligence* **23**(2) 123–154.

Allen, JF and Koomen, JA, 1983, "Planning using a temporal world model" In: *Proceedings of the eighth International Joint Conference on Artificial Intelligence* pp 741–747, Karlsruhe, W. Germany, August 1983.

Appelt, D, 1985, "Planning English referring expressions" *Artificial Intelligence* **26**(1) 1–33.

Berger, JO, 1980, *Statistical Decision Theory and Bayesian Analysis* Springer-Verlag.

Blackwell, D and Dubins, L, 1962, "Merging opinions with increasing information" *Annals of Mathematical Statistics* **33** 882–887.

Carnap, R, 1950, *Logic Foundations of Probability* Univ. of Chicago Press.

Cheeseman, P, 1983, "A method of computing generalized bayesian probability values for expert systems" In: *Proceedings of the Eighth International Joint Conference on Artificial Intelligence* pp 198–202, Karlsruhe, W. Germany, August 1983.

de Finetti, B, 1937, "La prevision: ses lois logiques, ses sources subjectives" *Ann. Inst. Henri Poincare* **7** 1–68 (English translation in Kyburg and Smolker, 1964).

Dean, T and Kanazawa, K, 1988, "Probabilistic causal reasoning" In: *Proceeding of the Fourth Workshop on Uncertainty in AI* pp 73–80, Minneapolis, MN, August 1988.

Dean, T and Kanazawa, Keiji, 1988, "Probabilistic temporal reasoning" In: *Proceedings of the Seventh National Conference on Artificial Intelligence* pp 524–528, St. Paul, MN, 1988.

DeGroot, M, 1970, *Optimal Statistical Decisions* McGraw-Hill.

Feldman, JA and Sproull, RF, 1977, "Decision theory and artificial intelligence ii: the hungry monkey" *Artificial Intelligence* **1** 158–192.

Fikes, RE and Nilsson, NJ, 1971, "Strips: a new approach to the application of theorem proving to problem solving" *Artificial Intelligence* **2** 189–208.

Fishburn, PC, 1981, "Subjective expected utility: a review of normative theories" *Theory and Decision* **13** 129–199.

Fisher, RA, 1959, *Smoking: The Cancer Controversy* Oliver and Boyd.

Gaifman, H, 1986, "A theory of higher order probabilities" In: *Proceedings of the Conference on Theoretical Aspects of Reasoning about Knowledge* pp 275–292, Monterey, California, 1986.

Gibbard, A and Harper, WL, 1978, "Counterfactuals and two kinds of expected utility" In: Hooker, Leach, and McClennen, eds., *Foundations and Applications of Decision Theory* pp 125–162, D. Reidel, Dordrecht, Holland, 1978.

Goldstein, M, 1983, "The prevision of a prevision" *J. American Statistical Association* **78**(384) 817–819.

Green, C, 1969, "Application of theorem proving to problem solving" In: *Proceedings of the First International Joint Conference on Artificial Intelligence* pp 219–239, Washington, DC (Also in Weber, BL and Nilsson, NJ, eds., *Readings in Artificial Intelligence* Morgan Kaufmann, Los Alto, CA.)

Haas, AR, 1985, "Possible events, actual events, and robots" *Computational Intelligence* **1** 59–70.

Hass, AR, 1986, "A syntactic theory of belief and action" *Artificial Intelligence* **28** 245–292.

Haddawy, P and Frisch, AM, 1990, "Modal logics of higher-order probability" In: Shachter, R, Levitt, TS, Lemmer, J and Kanal, LN, eds., *Uncertainty in Artificial Intelligence* Elsevier.

Halpern, YJ, 1989, "The relationship between knowledge, belief, and certainty" In: *Proceedings of the Fifth Workshop on Uncertainty in AI* pp 142–151, An expanded version appears as IBM Research Report RJ 6765.

Hanks, S, 1988, "Representing and computing temporally scoped beliefs" In: *Proceedings of the Seventh National Conference on Artificial Intelligence* pp. 501–505, St. Paul, Minn.

Hanks, S, 1989, "Projecting plans for uncertain worlds" PhD thesis, Yale University.

Hayes, PJ, 1973, "The frame problem and related problems in artificial intelligence" In: Elithorn, A and Jones, D, eds., *Artificial and Human Thinking* pp 45–59, Jossey-Bass.

Hayes-Roth, B and Hayes-Roth, F, 1979, "A cognitive model of planning" *Cognitive Science* **3**(4) 275–310.

Hogarth, RM, 1975, "Cognitive processes and the assessment of subjective probability distributions" *Journal of the American Association* **70** 271–294.

Jaynes, ET, 1968, "Prior probabilities" *IEEE Transactions on Systems, Science, and Cybernetics* **SSC-4** 227–241.

Jaynes, ET, 1979, "Where do we stand on maximum entropy" In: Levine and Tribus, eds., *The Maximum Entropy Formalism* MIT Press.

Jeffrey, RC, 1965, *The Logic of Decision* McGraw-Hill.

Jeffreys, H, 1961, *Theory of Probability 3rd edn.*, Oxford University Press.

Keeney, RL and Raiffa, H, 1976, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs* Wiley.

Kemeny, J, 1955, "Fair bets and inductive probabilities" *Journal of Symbolic Logic* **20** 263–273.

Keynes, JM, 1921, *A Treatise on Probability* MacMillan.

Konolige, K, 1979, "A computer based consultant for mineral exploration, Appendix D: Bayesian methods for updating probabilities" report, SRI.

Kreps, DM, 1988, *Notes on the Theory of Choice. Underground Classics in Economics* Westview Press.

Kripke, S, 1963, "Semantical considerations on modal logic" *Acta Philosophica Fennica* **16** 83–94 (*Proceedings of a Colloquium on Modal and Many-Valued Logics* Helsinki, 23–26 August, 1962).

Kyberg, HE, Jr and Smokler, HE, eds., 1969, *Studies in Subjective Probability* Wiley.

Lehman, RS, 1955, "On confirmation and rational betting" *Journal of Symbolic Logic* **20** 251–262.

Levi, I, 1980, *The Enterprise of Knowledge* MIT Press.

Levi, I, 1987, "The demons of decision" *The Monist* **70** 193–211.

Lewis, D, 1981, "Causal decision theory" *Australasian Journal of Philosophy* **56**(1) 5–30.

Maher, P, 1987, "Causality in the logic of decision" *Theory and Decision* **22** 155–172.

Maher, P, 1990, "Betting on theories (unpublished manuscript).

Maher, P, 1990, "Symptomatic acts and the value of evidence in causal decision theory" *Philosophy of Science* September.

McCarthy, J and Hayes, P, 1969, "Some philosophical problems from the standpoint of artificial intelligence" In: Meltzer, B and Michie, D, eds., *Machine Intelligence 4*, pp 463–502, Edinburgh University Press.

McDermott, DV, 1982, "A temporal logic for reasoning about processes and plans" *Cognitive Science* **6** 101–155.

Moore, RC, 1985, "A formal theory of knowledge and action" In: Hobbs, JR and Moore, RC, eds., *Formal Theories of the Commonsense World* Ablex.

Morgenstern, L, 1986, "A first order theory of planning, knowledge, and action" In: Vardi, M, ed., *Proceedings of the Conference on Theoretical Aspects of Reasoning About Knowledge*, pp 99–114, Monterey.

Pelavin, RN and Allen, JF, 1986, "A formal logic of plans in temporally rich domains" *Proceedings of the IEEE* **74**(10) 1364–1382.

Raiffa, H, 1968, *Decision Analysis* Addison-Wesley.

Raiffa, H and Schlaifer, R, 1961, *Applied Statistical Decision Theory* MIT Press.

Ramsey, FP, 1926, "Truth and probability" In: Mellor, DH, ed., *Foundations,* chapter 3, pp 58–100, Humanitities Press.

Sacerdoti, ED, 1975, "A structure for plans and behavior" Technical Note 109, SRI.

Savage, LJ, 1954, *The Foundations of Statistics* John Wiley & Sons (Second revised edition published 1972).

Savage, LJ, 1971, "Elicitation of personal probabilities and expectations" *Journal of the American Statistical Association* **66** 783–801.

Schervish, MJ and Seidenfeld, 1987, "An approach to consensus and certainty with increasing evidence" Technical Report 389, Carnegie-Mellon University, July 1987 (Forthcoming: *J. Statistical Inference and Planning*).

Seidenfeld, T, 1986, "Entropy and uncertainty" *Philosophy of Science* **53** 467–491.

Shimony, A, 1955, "Coherence and the axioms of confirmation" *Journal of Symbolic Logic* **20** 8–20.

Shoham, Y, 1987, *Reasoning About Change* MIT Press.

Skyrms, B, 1980, *Causal Necessity* Yale Univ. Press.

Sussman, GJ, 1975, *A Computer Model of Skill Acquisition* American Elsevier.

Tate, A, 1974, "INTERPLAN: A plan generation system which can deal with interactions between goals" memo MIP-R-109, Machine Intelligence Research Unit, Univ. of Edinburgh.

Tate, A, 1977, "Generating project networks" In; *Proceedings of the Fifth International Joint Conference for Artificial Intelligence* pp 888–893, Cambridge.

van Fraassen, B, 1984, "Belief and the will" *Journal of Philosophy* **81** 235–256.

Venn, J, 1966, *The Logic of Chance* MacMillan (new paperback edition, Chelsea, 1962).

Vere, S, 1983, "Planning in time: windows and durations for activities and goals" *IEEE Transactions on Pattern Analysis and Machine Intelligence* **5**(3) 246–267.

Villegas, C, 1977, "On the representation of ignorance" *J. American Statistical Association* **72** 651–654.

von Miesis, R, 1957, *Probability, Statistics and Truth* Allen and Unwin.

Wellman, MP, 1988, "Formulation of tradeoffs in planning under uncertainty" PhD thesis, MIT.

Wilensky, R, 1983, *Planning and Understanding.* Addison-Wesley.

Wilkins, DE, 1984, "Domain independent planning: representation and plan generation" *Artificial Intelligence* **22** 269–301.