Trade-offs in Decision-theoretic Planning*

Niels Peek and Peter Lucas
Department of Computer Science, Utrecht University
P.O. Box 80.089, 3508 TB Utrecht, The Netherlands
E-mail: {niels,lucas}@cs.ruu.nl

Abstract

Decision making under uncertainty can be viewed as a planning task, because it basically amounts to determining a sequence of actions that is optimal for a problem under study. Many different formalisms are used in decision making under uncertainty, with many different restrictions enforced. In this paper, a novel, general decision-theoretic planning framework is proposed, and used to analyse various representation formalisms for decision making. Several features of these formalisms are highlighted in terms of the framework, using examples from the domain of cancer-treatment planning in medicine.

Keywords: planning, decision theory, probabilistic networks.

1 Introduction

Decision theory provides a mathematical foundation for rational decision making under uncertainty by integrating notions of utility theory and probability theory. In brief, decision theory assumes that a *decision maker* faces a choice among various alternatives, where the effects of an alternative are described by a probability distribution over outcomes. Given the decision maker's preferences among outcomes, the theory recommends the alternative that optimises the yield related to the outcomes. The field of *decision analysis* is concerned with the practical application of decision-theoretic techniques. The theory has been successfully employed in many different problem areas, ranging from medical patient management to economic policy making. Representation formalisms for decision analysis include *decision trees*, *influence diagrams*, and *Markov decision processes*.

In the field of artificial intelligence (AI), the task of generating ordered sequences of actions in order to achieve some predefined objective is generally referred to as *planning*.

*This paper was published in: J. Hunt & R. Miles (eds.), *Research and Development in Expert Systems XIV (Proceedings of Expert Systems '97)*, SGES Publications, pp. 147–157, 1997. The investigations were (partly) supported by the Netherlands Computer Science Research Foundation with financial support from the Netherlands Organization for Scientific Research (NWO).

A planning system usually starts from a problem specification, the actions that may be performed, their expected outcomes or effects, and a description of objectives. The system's task is to yield a *plan*, i.e. a sequence of actions. Most classical AI planning systems assume that the effects of actions are deterministic (though possibly hard to predict). In many real-world situations, however, the effects of actions are essentially uncertain. The field of *planning under uncertainty* is concerned with planning tasks taking the inherent uncertainty in the effects of actions into account.

The term *decision-theoretic planning* [2, 1, 3] was recently introduced in AI to refer to approaches to planning under uncertainty incorporating decision-theoretic techniques. Compared to traditional applications of decision theory, these approaches do not so much differ in the type of problems addressed, but in the circumstance that notions like action, plan, and plan synthesis and execution are made explicit in the representation and problem-solving methods.

In this paper, several decision-theoretic representation formalisms will be analysed, in view of a novel framework of decision-theoretic planning. Features of the formalisms investigated are flexibility, expressiveness, and computational expenses. The purpose of this analysis is to obtain insight into the various trade-offs in building decision-theoretic expert systems. For decision support in uncertain domains, such systems are becoming increasingly popular as a mathematically well-founded alternative to traditional rule-based or hybrid rule-based and object-oriented systems. Planning of cancer treatment is used as an example domain throughout the paper.

2 Example: Cancer Treatment Planning

Modern medical management of cancer is a complicated process, guided by a large variety of clinical factors. Usually, more than one treatment modality is available to the clinician to treat a patient with a specific form of cancer; often a patient receives two or more types of treatment. Patient-specific information that is needed in the selection of appropriate treatment is only partly known with certainty; another part of the information is uncertain. Most of the information is gathered in the diagnostic process to assess the severity and extent of the disease. After treatment has been instilled, its success or failure is determined after a certain amount of time.

Rather than focusing on a single type of cancer, in this paper, a general scheme of cancer therapy is considered, which is depicted in Fig. 1. Preceding treatment, patient-specific information is collected, including the patient's age and the clinical stage of the disease. Assessment of the clinical stage of the disease is based on the patient's age, information obtained from taking biopsies at distant sites to show the presence of tumour cells, and the extent of the primary tumour. The patient's general health status is influenced by the clinical stage of the disease; when the clinical stage is unfavourable, the patient's health status is usually bad. In a few patients, appropriate biopsy material confirming the presence or absence of tumour cells cannot be obtained. Similarly, it is not always possible to determine the extent of the primary tumour. Nevertheless, even these patients

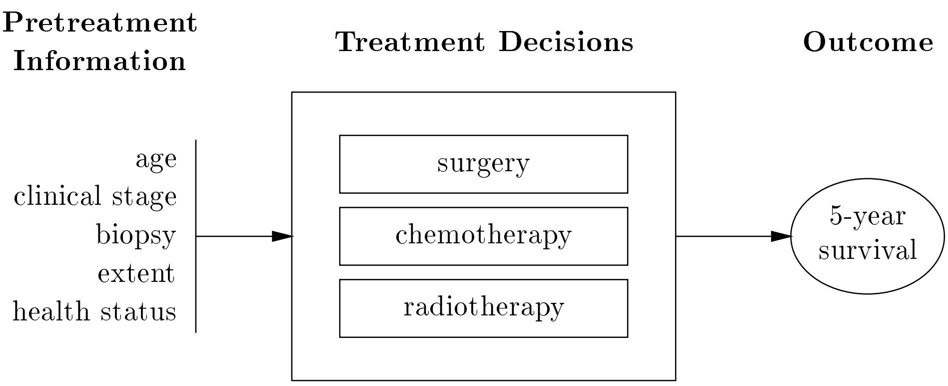


Fig. 1. Cancer treatment.

will have to receive treatment, taking into account the clinical picture and the uncertainty associated with the clinical evidence.

Generally spoken, the clinician can choose among the following treatment modalities: *surgery*, *chemotherapy* and *radiotherapy*. One or more of these treatment modalities can be applied to the patient in some order, depending on the patient's condition, and expected treatment outcome. The main factor used in optimising treatment outcome is 5-year survival, i.e. expected survival or death following treatment. Thus, cancer treatment can be essentially viewed as a form of planning under uncertainty, in which the following aspects are of importance: (i) a subcollection of treatment modalities is selected out of a collection of possibilities; (ii) treatment modalities are carried out in a particular order; (iii) during the therapeutic process, information may become available, influencing treatment decisions.

3 Decision-theoretic Planning

In this section, we develop a framework that combines notions from decision theory and AI planning systems. Subsequently, this framework will be used to analyse various decision-theoretic representation formalisms.

Consider the class of domains that can be represented by a set of discrete, finite *random variables* $X = \{x_1, \ldots, x_n\}$, $n \geq 1$, and a set of *actions* $A = \{a_1, \ldots, a_m\}$, $m \geq 1$. Random variables are taken to describe domain elements beyond the direct control of the decision maker. To express the joint assignment of values to random variables from a set $X' \subseteq X$, the notion of a *configuration* of X' is introduced, which is denoted by $c_{X'}$. The set of all possible configurations of X' is denoted by $C_{X'}$. The decision maker has the opportunity to perform actions from the set A in order to influence the actual configuration of the set of random variables X. It is assumed that actions are performed one by one, in a particular order. This sequencing of actions, for example a_1 followed by a_2, is denoted by $\langle a_1; a_2 \rangle$. The sequencing operator ';' is assumed to be associative; consequently, $\langle a_1; \ldots; a_k \rangle$, $k \geq 0$, is an action sequence of length k. We do not require that all actions in a sequence are distinct (i.e. actions may be repeated), but action sequences are required to be finite. The action sequence of length $k = 0$ is called the *null sequence*. The set of all action sequences over A is denoted by S_A.

Random variables and actions for a given problem domain are taken to interact in two ways. First, performing an action may affect the configuration of the random variables. Let $p(c'_X \mid c_X, s)$ denote the probability that we arrive at configuration $c'_X \in C_X$ after performing the action sequence $s \in S_A$, given the initial configuration $c_X \in C_X$. If $p(c'_X \mid c_X, s) > 0$, then we say that c'_X is a possible *outcome* of c_X and s. Since spontaneous changes are excluded, we have that $p(c'_X \mid c_X, \langle\,\rangle) = 0$ if $c'_X \neq c_X$. Second, performing an action may yield information on the actual configuration of random variables. This is laid down in an *observation function* $\omega : A \to 2^X$, where $\omega(a) \subseteq X$ indicates the subset of random variables whose actual configuration can be observed when action $a \in A$ is performed. The set of random variables that are observable at any time, independent of action choice, is denoted by X_0. The decision process starts with a given configuration c_{X_0} of X_0, called a *problem instance*.

A *plan* is a partially ordered sequence of actions, where different subsequences are selected on the basis of observations. More formally, the set of all plans over the random variables X and action variables A, consistent with an observation function ω, is the smallest set $\Pi(X, A, \omega)$, such that

1. for all $s \in S_A$, we have that $s \in \Pi(X, A, \omega)$,
2. if $a \in A$ and $X' \subseteq X_0 \cup \omega(a)$, then $\langle a; \sigma_{X'}\rangle \in \Pi(X, A, \omega)$, where $\sigma_{X'}$ is a function $\sigma_{X'} : C_{X'} \to \Pi(X, A, \omega)$, called a *switch*, and
3. if $\pi_1, \pi_2 \in \Pi(X, A, \omega)$ then $\langle \pi_1; \pi_2\rangle \in \Pi(X, A, \omega)$.

The first condition says that any action sequence in A is also a plan over X and A consistent with ω; for such plans, the sequence of actions to be performed will be the same in all situations. The second condition introduces additional flexibility: the possibility to choose a subplan on the basis of an observed configuration of random variables. The subplan resulting from applying the switch $\sigma_{X'}(c_{X'})$, $c_{X'} \in C_{X'}$, is called a *branch* of $\sigma_{X'}$; the configuration $c_{X'}$ is called the *case* corresponding to the branch. Actions occurring in some, but not all, branches of a switch are said to be *contingent*. Finally, the third condition expresses that plans can be concatenated; $\langle \pi_1; \pi_2\rangle$ denotes the plan in which subplan π_1 is (unconditionally) followed by π_2.

It is said that a plan $\pi \in \Pi(X, A, \omega)$ is *executed* if a branch is chosen recursively for each switch σ in π. The result is an action sequence, called a *ground instance* of π. The set of all possible ground instances of a plan π is denoted by $g(\pi)$; the cardinality of $g(\pi)$, i.e. the number of possible ground instances, is called the *complexity* of plan π. Given a plan π and a problem instance c_{X_0}, let $q(s \mid c_{X_0}, \pi)$ denote the probability of arriving at ground instance $s \in g(\pi)$ of π; this probability depends on the likelihoods of cases corresponding to the branches chosen which in turn depend on the problem instance c_{X_0} and actions performed in due course. Since we cannot arrive at an action sequence $s \in S_A$ that is not a ground instance of π, we will take $q(s \mid c_{X_0}, \pi) = 0$ if $s \notin g(\pi)$, for each $c_{X_0} \in C_{X_0}$.

In classical AI planning systems, the objective is to achieve some predefined propositional goal. When the effects of actions are uncertain, however, such a result is unattainable. In decision-theoretic planning, the objective is to optimise a *utility function* $u : C_X \times S_A \to \mathbb{R}$ that describes the decision maker's preferences for possible outcomes, i.e. outcome c_X is preferred to outcome c'_X after performing action sequence $s \in S_A$, if $u(c_X, s) \geq u(c'_X, s)$. In most applications, these preferences depend only on a part of the variables involved. The *expected utility* $eu(\pi \mid c_{X_0})$ of a plan π given problem instance c_{X_0} equals

$$eu(\pi \mid c_{X_0}) \;=\; \sum_{c_X \in C_X} \sum_{s \in g(\pi)} u(c_X, s) \cdot p(c_X \mid c_{X_0}, s) \cdot q(s \mid c_{X_0}, \pi) \qquad (1)$$

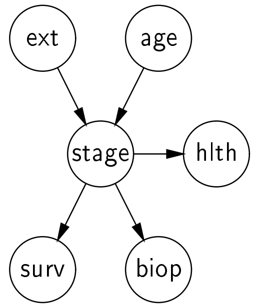
Many decision-theoretic representation formalisms provide a means to restrict the set of plans to consider for a given problem domain. Usually, restriction is accomplished by employing a grouping of actions in a partially ordered set $(D, \ll)$ of *decision variables*. Each decision variable $d \in D$ can take a value $a \in Dom(d)$, notation $d \leftarrow a$, where $Dom(d) \subseteq A$. The partial order $\ll$ on D will be used to express the temporal order in which decision variables should be given a value. Consequently, $\ll$ serves as a set of constraints on the order of actions in action sequences: we will say that an action sequence $\langle a_1; \ldots; a_k \rangle \in S_A$ *adheres* to $(D, \ll)$ if there is a chain $d_1, \ldots, d_k$ in $(D, \ll)$ such that for each i, $1 \leq i \leq k$, we have that $a_i \in Dom(d_i)$. Likewise, $\ll$ serves as a set of constraints on plans over A: we say that a plan $\pi \in \Pi(X, A, \omega)$ *adheres* to $(D, \ll)$ if any ground instance $s \in g(\pi)$ of π adheres to $(D, \ll)$.

Now, let $P_{\ll} \subseteq \Pi(X, A, \omega)$ be the subset of plans adhering to $(D, \ll)$, and let c_{X_0} be a problem instance. A plan $\pi \in P_{\ll}$ is called *optimal* with respect to $P_{\ll}$ given c_{X_0}, notation $\pi = \pi^*(P_{\ll}, c_{X_0})$, if we have that $eu(\pi \mid c_{X_0}) \geq eu(\pi' \mid c_{X_0})$ for any $\pi' \in P_{\ll}$. Finding an optimal plan for a given problem instance usually proceeds by employing dynamic programming methods, optimising the last action first. The computational complexity of finding an optimal plan depends on the number of plans to consider as well as on average plan complexity.

We now return to the cancer-treatment example described in the previous section to illustrate the above framework. In this example, we distinguish six random variables, representing age of the patient, health status, clinical stage, extent of the primary tumour, result of biopsy, and 5-year survival, respectively, and four actions representing the performance of surgery, the application of chemotherapy and radiotherapy, and the refrainment of treatment, respectively. These variables are summarised in Fig. 2a. Fig. 2b shows the graphical part of a *Bayesian belief network* [9] comprising the random variables distinguished; the network describes a joint probability distribution on these variables. There are two types of interaction between the random variables and the actions. First, performing any of the three therapeutic actions may affect the clinical stage and the extent of the primary tumour. Second, performing an action may yield new information on the configuration of some of the random variables. In our example, we assume that the age and health status of the patient are always known or observable (i.e. $X_0 = \{\mathsf{age}, \mathsf{hlth}\}$),

	Name	*Interpretation*	*Domain*
Random variables	age	age of the patient	$< 50, \geq 50$
	hlth	health status	*good, bad*
	stage	clinical stage	*fav, unfav*
	ext	extent of primary tumour	*small, large*
	biop	result of biopsy	*pos, neg*
	surv	5-year survival	*true, false*
Actions	surg	perform surgery	
	chem	perform chemotherapy	
	radio	perform radiotherapy	
	skip	do nothing	

(a)



(b)

Fig. 2. Variables and belief network for the example domain.

and that surgery yields biopsy material ($\omega(\mathsf{surg}) = \{\mathsf{biop}\}$), which is otherwise unavailable. The other random variables are unobservable.

We will use three binary decision variables, d_{surg}, d_{chem}, and d_{radio}, representing the separate decisions to apply each of the three treatment modalities, e.g. $Dom(d_{\mathsf{surg}}) = \{\mathsf{surg}, \mathsf{skip}\}$. Now suppose that we impose the constraint $d_{\mathsf{surg}} \ll d_{\mathsf{chem}}$. A simple treatment plan for a given patient is to perform surgery and, depending on the biopsy material found, to subsequently apply chemotherapy:

$$\begin{aligned} \pi &= \langle \mathsf{surg}\,;\, \sigma_{\{\mathsf{biop}\}} \rangle \\ & \quad \textit{where} \left\{ \begin{array}{lcl} \sigma_{\{\mathsf{biop}\}}(\mathsf{biop} = \textit{pos}) & = & \langle \mathsf{chem} \rangle \\ \sigma_{\{\mathsf{biop}\}}(\mathsf{biop} = \textit{neg}) & = & \langle \, \rangle \end{array} \right. \end{aligned} \tag{2}$$

The set of ground instances of the plan π equals $\{\langle \mathsf{surg}\,;\, \mathsf{chem} \rangle, \langle \mathsf{surg} \rangle\}$. Finally, a utility function is defined, based only on 5-year survival, where a maximum utility of 100 is assigned to survival, and a minimum utility of 0 to death.

4 Decision-theoretic Representations

The most popular representation formalisms for decision analysis are decision trees, influence diagrams and Markov decision processes. These formalisms will now be analysed in terms of the framework presented in the previous section, using examples from the cancer-treatment domain.

4.1 Decision Trees

A *decision tree* [11] is a rooted tree, where each internal node is either a *decision node*, denoting a point in time where the decision maker faces a choice among several actions, or a *chance node*, denoting a point in time where the decision maker can observe the actual value of a random variable. The root of the tree, which is usually a decision node, represents the start of the decision process. An external node, or leaf, of the tree represents the outcome of the scenario associated with the sequence of events and actions along the unique path to that node starting at the root. A numerical utility value is associated with each leaf in the tree. Fig. 3 shows a decision tree for a simplified version of the cancer-

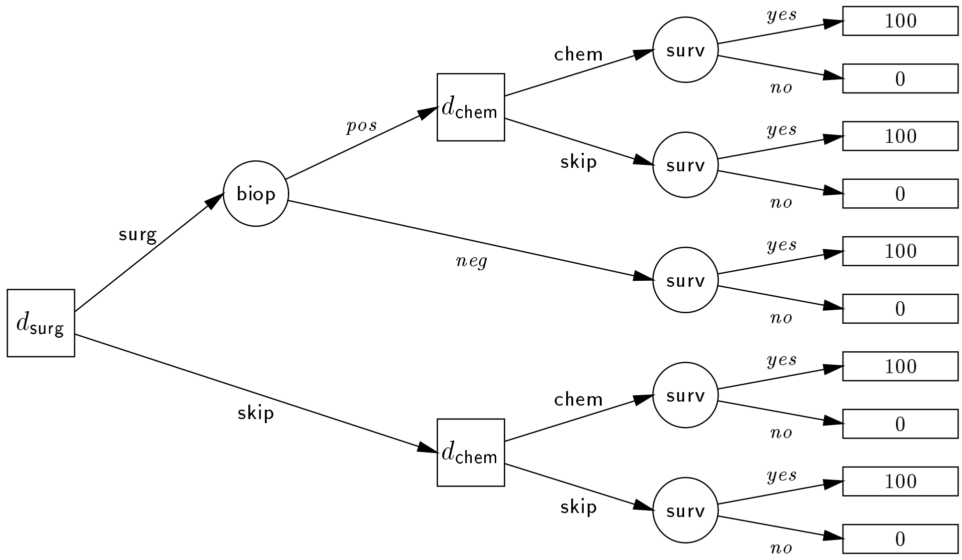


Fig. 3. Decision tree for cancer treatment.

treatment example, addressing only the decisions whether or not to perform surgery and to subsequently apply chemotherapy. Decision nodes, representing decision variables, are depicted by square boxes; chance nodes, representing random variables, are depicted by circles; outcomes are depicted by rectangular boxes labelled with the corresponding utility value. The branches emanating from an internal node correspond to the respective values that the associated variable may take. With each branch emanating from a chance node is associated a numerical value expressing the probability that the value will be observed in the given situation. The probabilities in a tree correspond to a single problem instance $c_{X_0} \in C_{X_0}$, and are typically obtained from an external probability model, for instance a belief network. Furthermore, note that there may be multiple nodes in a decision tree representing the same variable, if the variable occurs in multiple scenarios. The tree is asymmetrical, since finding biopsy material is contingent on the decision to operate.

In terms of the decision-theoretic planning framework, a decision tree provides an explicit, graphical enumeration of the collection of plans $P_{\ll} \subset \Pi(X, A, \omega)$ adhering to the partial order $\ll$ on decision variables provided by the tree. A plan $\pi \in P_{\ll}$ is selected from the collection by choosing a value for each decision node in the tree; chance nodes act as switches in the selected plan. For instance, by selecting the actions surg and chem for the decision nodes in Fig. 3, we select the plan of Eq. (2). The evaluation of a decision tree is the process of finding an optimal plan from the tree, consistent with Eq. (1). Since a decision tree is case-specific, it only provides an optimal plan for a single problem instance $c_{X_0} \in C_{X_0}$. The number of computational steps needed for evaluation is proportional to the number of nodes in the tree. Since a decision tree is an explicit enumeration of plans for a decision-theoretic problem, the size of the tree will typically grow exponentially in the number of actions and uncertain events modelled.

4.2 Influence Diagrams

An *influence diagram* [5] is a concise representation of a decision problem under uncertainty in a directed, acyclic graph. Fig. 4 shows an influence diagram graph for the cancer treatment example. Decision variables are represented by *decision nodes* in the graph, denoted by square boxes, and random variables by *chance nodes*, denoted by circles. Furthermore, there is a single, diamond-shaped node valuating the various outcomes of the decision process; this node, called the *value node*, may not have successors in the graph. We note that, in contrast with decision trees, any decision or random variable considered relevant for the decision problem at hand occurs only once in the representation. Arcs in the graph pointing to a decision node, called *informational arcs*, specify which information is available at the point in time when an action value is chosen for the node; it is assumed that a value is known for each of its parents in the graph at that point in time. Furthermore, it is required that there exists a directed path in the graph containing all the decision nodes, defining the order in which decisions are taken. The arcs pointing to a chance node, called *conditioning arcs*, determine which nodes have a direct influential or causal effect on the node. A chance node comes equipped with a *conditional probability table*, specifying a conditional probability distribution on its values given a value assignment to each of its parents in the graph. The value node is supplied with a *utility function*, defined over the configurations of the node's parents in the graph.

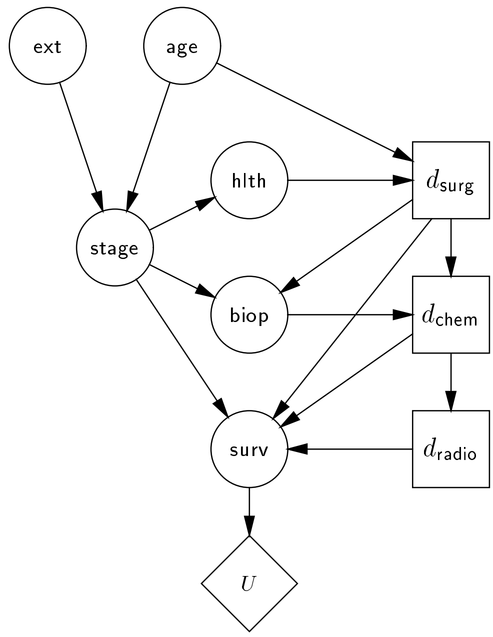


Fig. 4. Example influence diagram.

In terms of the decision-theoretic planning framework, an influence diagram provides a concise enumeration of a collection of plans $P_{\ll} \subset \Pi(X, A)$, where $\ll$ is the total order on decisions variables provided by the graph. A plan $\pi \in P_{\ll}$ is selected from the collection by providing a conditional action specification for each decision node given the values of the node's parents in the graph; chance nodes directly preceding decision nodes in the graph act as switches in the selected plan. The possibility to represent complex plans is limited by the total order on decision variables: influence diagrams cannot represent plans with contingent decisions. It is also not possible to represent contingent data gathering; for instance, in the example diagram, the value of the biopsy variable is assumed to be known before deciding on chemotherapy, even if the decision on surgery was negative.

In contrast with decision trees, an influence diagram models all the information needed for problem solving in a given decision-theoretic planning domain, including the relevant probabilities. Therefore, no external probability model is needed; a single influence diagram can be used to find optimal plans for all problem instances $c_{X_0} \in C_{X_0}$. Furthermore, the representation exploits conditional independences holding between random variables, and the fact that an action often affects only a small subset random variables directly, whereas other random variables are affected only indirectly, or not affected at all. Evaluating an influence diagram, i.e. finding an optimal plan, is possible by performing a sequence of transformations on the graph [12]. The general problem is combinatoric, but evaluation is feasible in polynomial time for sparse, bounded, graphs.

4.3 Markov Decision Processes

In a *Markov decision process* [10], or MDP for short, the dynamics of a planning problem under uncertainty are modelled as a discrete-time stochastic process over the set C_X of configurations of the random variables involved. The probability distributions governing transitions depend on action choice. That is, with each action $a \in A$ is associated a transition probability function $\tau_a : C_X \times C_X \to [0,1]$, where $\tau_a(c'_X \mid c_X)$ expresses the probability of arriving at configuration $c'_X \in C_X$ after performing action a in configuration $c_X \in C_X$. The process is Markovian in the sense that the current configuration of the random variables depends on the past only through the previous configuration and the action choice:

$$p(c'_X \mid c_X, \langle s; a \rangle) \;=\; \sum_{c''_X \in C_X} \tau_a(c'_X \mid c''_X) \cdot p(c''_X \mid c_X, s) \qquad (3)$$

At each step in the process, the decision maker receives an immediate reward reflecting the desirability of the actual configuration as compared to other configurations; the objective is to optimise some function of the overall reward sequence that expresses the decision maker's intertemporal trade-offs. Fig. 5 offers a schematic depiction of an MDP for the cancer-treatment example.

In MDPs, it is assumed that the decision maker can observe the actual configuration of all random variables at each point in time. *Partially observable Markov decision processes* (POMDPs) [8] are a generalisation of MDPs which permit uncertainty regarding the actual configurations and allow for observations depending on action choice. Formally, a POMDP model is an MDP model extended with an observation function $\omega : A \to 2^X$.

Using Eq. 3, we can compute the expected consequences of any possible action sequence $s \in S_A$ for a given (PO)MDP problem. Therefore, the formalism provides for a very powerful environment for decision-theoretic planning. A (PO)MDP model does not enumerate plans using partially ordered decision variables like decision trees and influence diagrams: any potential plan from the set $\Pi(X, A, \omega)$ is considered for evaluation. Generally speaking, the complexity of constructing an optimal plan for a given POMDP problem depends on the number of possible configurations, $|C_X|$, the number of available actions, $|A|$, and the ability to observe configurations in due course. For fully observable MDP

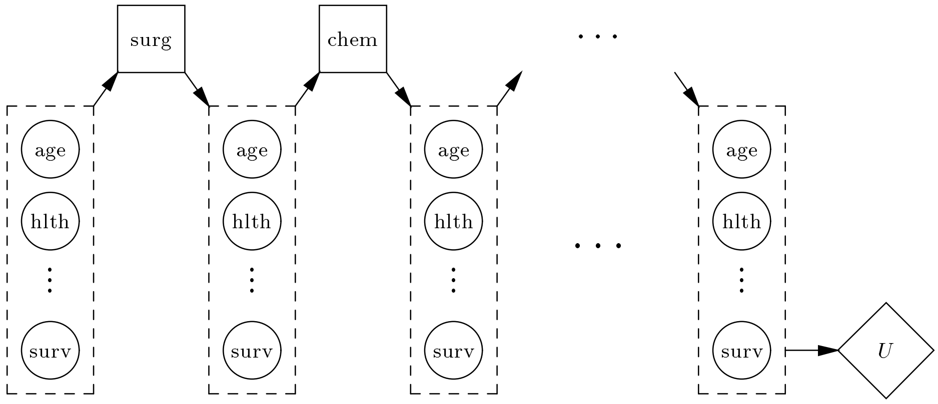


Fig. 5. Markov decision process for cancer treatment.

problems, efficient solution methods exist, based on the principle of dynamic programming [10]. Straightforward application of dynamic programming techniques to partially observable problems is not possible, and algorithms tend to be complicated and limited [7].

5 Comparison

In this section, we compare the three formalisms discussed in the previous section with respect to their flexibility and expressiveness as a knowledge-representation method. In particular, we will consider the possibilities to represent decision-theoretic plans, the ability to encode generic effects of actions, and the compactness of the representation.

In a decision tree, a decision-theoretic planning problem is represented by explicitly enumerating the possible plans to solve the problem. In the process of constructing a decision tree, a designer can decide on constraints on possible plans: any plan or subplan that is deemed unrealistic or irrelevant is simply left out. There are also no restrictions on the effects that actions may have on random variables. However, a decision tree is case-specific and does not encode such effects for a class of problems. Although the structure of a decision tree might be reused to deal with similar problems, for cases not represented in the tree, an external probability model is needed to compute the relevant probabilities. Hence, the decision-tree formalism yields a rather weak knowledge-representation formalism: although plans and constraints on plans can be represented explicitly, knowledge concerning probabilistic independences and influences among variables is always lacking.

An influence diagram is a Bayesian belief network augmented with decision nodes and a utility node. It provides a compact way to encode both decision-theoretic plans and probabilistic domain knowledge, by exploiting conditional independences between random variables and limitations in the effects of actions. However, all plans have the same basic structure with limited complexity, stemming from the fact that influence diagrams correspond to symmetrical decision trees. The order in which decision variables are considered is always fixed, and neither contingent decisions nor contingent data gathering can be represented.

Markov decision processes offer a very general formalism of decision-theoretic planning, based on a generic specification of the effects of actions on random variables. The only underlying assumption is that these effects are Markovian: the present depends on the past only through the last step. We note that this assumption is not made in decision trees and influence diagrams. The generality of the formalism may hamper straightforward application in practice. Problems of considerable size require the specification of a massive amount probabilities, and finding an optimal plan becomes intractable. Furthermore, the notion of plan is left implicit in the formalism, implying that knowledge about irrelevant, or even nonsensical plans cannot be represented explicitly.

6 Discussion

Decision theory is becoming increasingly popular as a mathematical foundation for building planning systems in uncertain domains. A knowledge-representation formalism for decision-theoretic planning, applicable to, for example, medical treatment planning, should provide sufficient expressive power to represent the most important features and nuances of a problem domain. In this paper, we have proposed a new, general framework for decision-theoretic planning based on two types of variables and a high-level description of plans. Existing representation formalisms from decision analysis have been analysed in terms of this framework, revealing their underlying degrees of expressiveness. It appears that each of these formalisms leaves some of the notions related to decision-theoretic planning implicit. In this sense, these formalisms are either too restrictive, or offer too much freedom, lacking the features of an appropriate knowledge-representation formalism.

In recent years, some research has been devoted to developing extensions to the representations discussed above, either to enhance their expressiveness as a knowledge-representation formalism (e.g. [13]), or to improve the computational efficiency of their evaluation (e.g. [3, 6]). Being the most general representation method, Markov decision processes are becoming increasingly popular in AI as a basis for decision-theoretic planning. For instance, Boutilier and colleagues [1] have developed an alternative framework for decision-theoretic planning, taking Markov decision processes as a starting-point. The generality of their framework ensures that many more restricted decision-theoretic formalisms can be represented, which they accomplish by switching to different representations. However, switching to different representations, although mathematically sound, yields little insight in terms of expressive power, whereas defining a formalism in terms of a number of framework parameters, as is done in our analysis, reveals underlying assumptions.

In the near future, the notions of time and external event will be added to our framework, in order to incorporate reasoning about time and spontaneous change in decision-theoretic planning.

Acknowledgement. The authors wish to thank Linda van der Gaag for valuable comments on an earlier draft of this paper.

References

[1] C. Boutilier, T. Dean, and S. Hanks. Planning under uncertainty: structural assumptions and computational leverage. In Ghallab and Milani [4], pages 157–171.

[2] T. Dean and M. Wellman. *Planning and Control.* Morgan Kaufmann, San Mateo, CA, 1991.

[3] A. Doan, P. Haddawy, and C.E. Kahn. Decision-theoretic refinement planning: a new method for clinical decision analysis. In *Proceedings of the 19th Annual Symposium on Computer Applications in Medical Care (SCAMC'95)*, 1995.

[4] M. Ghallab and A. Milani, editors. *New Directions in AI Planning.* IOS Press, Amsterdam, 1996.

[5] R.A. Howard and J.E. Matheson. Influence diagrams. In R.A. Howard and J.E. Matheson, editors, *Readings on the Principles and Applications of Decision Analysis.* Strategic Decisions Group, Menlo Park, CA, 1981.

[6] S.-H. Lin and T. Dean. Generating optimal policies for high-level plans with conditional branches and loops. In Ghallab and Milani [4], pages 187–200.

[7] W.S. Lovejoy. A survey of algorithmic methods for partially observed Markov decision processes. *Annals of Operations Research*, 28:47–66, 1991.

[8] G.E. Monahan. A survey of partially observed Markov decision processes: theory, models, and algorithms. *Management Science*, 28(1):1–16, 1982.

[9] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann Publishers, Palo Alto, 1988.

[10] M.L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming.* Wiley, New York, 1994.

[11] H. Raiffa. *Decision Analysis: Introductory Lectures on Choice under Uncertainty.* Addison-Wesley, Reading, MA, 1968.

[12] R.D. Shachter. Evaluating influence diagrams. *Operations Research*, 34(6):79–90, 1986.

[13] N. L. Zhang, R. Qi, and D. Poole. A computational theory of decision networks. *International Journal of Approximate Reasoning*, 11(2):83–158, 1994.