Dynamic Decision Making in Stochastic Partially Observable Medical Domains: Ischemic Heart Disease Example.

Milos Hauskrecht
MIT Lab for Computer Science, NE43-421, 545 Technology Square, Cambridge, MA 02139
E-mail: *milos@medg.lcs.mit.edu*

Abstract

The focus of this paper is the framework of partially observable Markov decision processes (POMDPs) and its role in modeling and solving complex dynamic decision problems in stochastic and partially observable medical domains. The paper summarizes some of the basic features of the POMDP framework and explores its potential in solving the problem of the management of the patient with chronic ischemic heart disease.

Introduction

Dynamic decision problems in medicine usually deal with two sources of uncertainty. First is due to the action outcome uncertainty, the second is due to imperfect observability of the underlying process state. While most of the work on dynamic decision making addresses the issue of action outcome uncertainty, the feature of partial observability is often considered irrelevant or is abstracted out. Research work falling into this category includes the management of diabetes [5] or chronic heart disease [6]. The assumption of perfect observability may not work well for problems in which observations are imprecise indicators of the patient state and when investigative actions have significant cost (invasiveness, economic cost). In such cases careful evaluation of costs and benefits associated with both treatment and investigative actions with regard to the global objectives is necessary.

A framework that allows one to model both sources of uncertainty is the partially observable Markov decision process [1] [9] [7] [2] [3]. In the following I will summarize the POMDP framework and illustrate its potential on the problem of the management of patients with chronic ischemic heart disease.

Partially Observable Markov Decision Process

A *partially observable Markov decision process (POMDP)* describes the stochastic control process with partially observable states and formally corresponds to a 6-tuple (S, A, Θ, T, O, C) where S is a set of states, A is a set of actions, Θ is a set of observations, T is a set of transition probabilities between states that describe the dynamic behavior of the modeled environment under different actions, O stands for a set of observation probabilities that describe the relationship among observations, states and actions, and C denotes a cost model that assigns costs to state transitions and models payoffs associated with such transitions (alternative formulations can include rewards). Note that the basic model does not impose any restrictions on details of the actual representation and one can use a factored representation of states, observations and probabilistic relations with explicitly represented independencies.

The *decision (or planning) problem* in the context of POMDP requires one to find an action or a sequence of actions that minimizes the objective cost function. An *information state* represents all the information available to the agent at the decision time that is relevant to the selection of the optimal action. The information state consists of either a complete history of actions and observations or corresponding sufficient statistics ensuring the Markov property of the information process. An *objective function* represents (quantifies) control objectives by combining costs incurred over time using various kinds of models. Typically, the objective function is additive over time and based on expectations, e.g. the cost function often uses a finite horizon model $\min E(\sum_{t=0}^{n} c_t)$, that minimizes expected costs for the next n steps or an infinite discounted horizon $\min E(\sum_{t=0}^{\infty} \gamma^t c_t)$, with $0 \leq \gamma < 1$ being a discount factor. In the following we will focus on these two models.

Solving the POMDP Problem

For the n step-to-go problem the optimal value of the objective function, so called *value function*, and the optimal control can be computed using the Markov property of the information state process and Bellman's principle of optimality via standard recursive formulas [3]:

$$V_n^*(I_n) = \min_{a \in A} \rho(I_n, a) + \gamma \sum_{o \in \Theta_{next}} P(o|I_n, a) V_{n-1}^*(\tau(I_n, o, a))$$

$$\mu_n^*(I_n) = \mathrm{argmin}_{a \in A} \rho(I_n, a) + \gamma \sum_{o \in \Theta_{next}} P(o|I_n, a) V_{n-1}^*(\tau(I_n, o, a))$$

where $V^*(.)$ and $\mu^*(.)$ are optimal value and control functions, I_n denotes the current information state; $\rho(I_n, a)$ is the expected transition cost from state I_n under action a and can be computed as $\rho(I_n, a) = \sum_{s \in S} \sum_{s' \in S} P(s'|s, a) P(s|I_n) C(s, a, s')$; Θ_{next} is a set of observations that are available in the next step; τ is a transition function that maps the information state, new action and observation to the next step information state; and γ is the possible discount factor. Identical formulas (less the index denoting the number of steps to go) can be derived for the infinite discounted horizon problem.

The problem of finding optimal actions or policies can be computationally very expensive. It has been shown to be PSPACE-hard even for a single initial state and finite horizon cost function [8]. This is because the number of information states one potentially needs to visit grows exponentially with the number of steps to be explored. An even worse situation emerges when one is required to find the solution for all initial information states — the so called policy problem. Although exact methods for solving both decision and policy problems are available for standard POMDP models [9] [2] [3], the computational complexity of such methods leads to the exploration of various approximations that allow good solutions with less computation. Efficient approximation methods are mostly based on the idea of approximate dynamic programming (for the finite horizon case) and approximate value iteration (for the infinite discounted horizon). Methods can be based on sampling schemes combined with function approximation and fitting strategies, or based on reducing the complexity of the information vector space in various ways, e.g. through feature extraction mappings. Description of the exact and approximate solution methods is outside the scope of this paper (see [7] [2] [3]).

Management of Ischemic Heart Disease

The POMDP framework can be exploited in representing various complex problems of patient management, e.g. the management of chronic ischemic heart disease (IHD) [10] [6] [4]. The objective is to determine the optimal plan for managing the patient's chronic disease relative to cost criteria, including, e.g., invasiveness of the treatment, risk of death etc.

Basic components of the POMDP model include state, action and observation variables (see figure 1) describing the state of the patient, possible actions, and available observations.

State variables	Observation variables	Actions
status *dead* *alive* coronary artery disease ischemia level history of MI history of CABG acute MI	status *dead* *alive* history of MI history of CABG acute MI chest pain resting EKG stress test result angiogram result	no action medication treatment stress test angiogram investigation angioplasty (PTCA) coronary artery bypass surgery (CABG)

Fig. 1. Ischemic heart disease: basic model components

The dependencies between components of the model are captured in figure 2. The dynamics is represented by a controlled stochastic transition model with states and actions. States describe the underlying coronary artery disease (status of artery occlusion), severity of ischemia and other information influencing the transition, e.g. past MI. Actions (rectangle) decribe possible decision choices, e.g. no action (wait), medication treatment, angiogram investigation, or coronary artery bypass surgery.

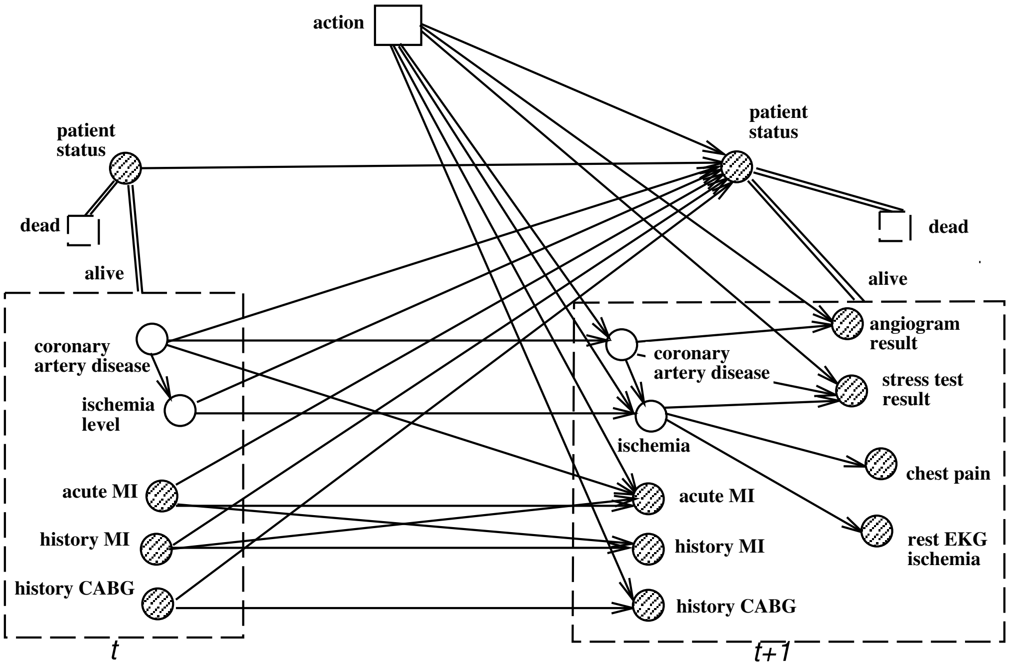

Fig. 2. Dynamic model of the ischemic heart disease

Observations (patterned circles) correspond to variables the decision-maker can observe and at any point in time. Observations can be conditioned on actions, e.g. angiogram result is conditioned on the choice of angiogram investigation. The conditioning is suitable when one wants to stress the presence of the test and its possible transition and cost effects. For example the angiogram investigation increases the incidence of MI or death and is also associated with higher invasiveness and economic cost. On the other hand many observations can be assumed "costless" and always available, e.g. history findings or EKG results.

The cost model describes payoffs associated with possible transitions, e.g. maximum cost is associated with the transition to the dead state, smaller but still substantial cost is associated with severe ischemia or occurrence of an MI. The decision criteria that try to reduce the expected cost then try to avoid these highly negative states.

Given the model, the objective of the problem solver is to find an action or a sequence of actions that minimizes the expected cost of management. Decision horizons applicable in the IHD case include both the finite horizon in which one optimizes the treatment for the next n time steps and infinite discounted horizon in which an infinite number of steps is considered, with more distant steps discounted.

Conclusion

Although the basic methodology for modeling dynamic decision processes via POMDPs has been available for some time, its potential has not been exploited in medical applications. In the paper I have described the basics of the POMDP framework and its solution methods. More detailed description of various exact and approximation methods can be found in [3]. The usefulness of the framework in medical settings is being examined with the example of the management of ischemic heart disease. Challenging problems in building and using the model include reliably estimating the many parameters and choosing appropriate approximation methods that yield timely and useful solutions.

This research was supported by the grant 1T15LM07092 from the National Library of Medicine. Peter Szolovits and William Long have provided valuable feedback on early versions of the paper and Hamish Fraser has helped with the ischemic heart disease example.

References

1. K.J. Astrom. Optimal control of Markov decision processes with incomplete state estimation. Journal of Mathematical Analysis and Applications, 10, pp. 174-205, 1965
2. A.R. Cassandra. Optimal policies for partially observable Markov decision processes. Brown University, Technical report CS-94-14, 1994.
3. M. Hauskrecht. Planning and control in stochastic domains with imperfect information. MIT EECS PhD thesis proposal, August 1996, 133 pages.
4. M. Hauskrecht. Dynamic decision making in stochastic partially observable medical domains. AAAI Spring symposium, pp. 69- 72, 1996.
5. R. Hovorka et.al. Causal probabilistic network modelling - An illustration of its role in the management of chronic diseases. IBM Systems Journal, 31:4, pp.635-648, 1992.
6. T.-Y. Leong. An integrated approach to dynamic decision making under uncertainty. MIT/LCS/TR-631, 1994.
7. W.S. Lovejoy. A survey of algorithmic methods for partially observed Markov decision processes. Annals of Operations Research, 28, pp. 47-66, 1991.
8. C.H. Papadimitriou, J.N. Tsitsiklis. The complexity of Markov decision processes. Mathematics of Operations Research, 12:3, pp. 441-450, 1987.
9. R.D. Smallwood, E.J. Sondik. The optimal control of partially observable processes over a finite horizon. Operations Research, 21, pp. 1071-1088.
10. J.B. Wong et.al. Myocardial revascularization for chronic stable angina. Annals of Internal Medicine, 113 (1), pp. 852-871, 1990.