Explicit temporal models for decision-theoretic planning of clinical management*

Niels Peek
Department of Computer Science, Utrecht University
P.O. Box 80.089, 3508 TB Utrecht, The Netherlands
e-mail: niels@cs.uu.nl

Abstract

The management of patients over a prolonged period of time is a complicated task involving both diagnostic and prognostic reasoning with incomplete and often uncertain knowledge. Various formalisations of this type of task exist, but these often conceal one or more essential ingredients of the problem. This article explores the suitability of partially observable Markov decision processes to formalising the planning of clinical management. These processes allow for explicit representation of clinical states of the patient, the management strategy employed, the objectives of treatment, and the role of time and change in reasoning. However, practical application is hampered by their coarse representational granularity and complex formulation. It is discussed how probabilistic network representations can be used to alleviate these obstacles. The resulting method is illustrated with a real-world example from the domain of paediatric cardiology.

Keywords: Decision-theoretic planning; Markov decision processes; Probabilistic networks; Paediatric cardiology

1 Introduction

The planning of therapy over a prolonged period of time often requires the ability to predict the interplay between the natural history of disease and effects of clinical actions. Controversies about therapy are frequently rooted in prognosis, as most predictions cannot be made with certainty, and trade-offs have to be made between the expected outcomes of current decisions and future decision-making opportunities. A typical example of this situation is found in the field of paediatric cardiology. In the management of patients with a congenital cardiac anomaly, there is always a trade-off between the benefits gained by waiting before surgical intervention in the hope that the patient's condition will improve, and the risks caused by the natural history of these disorders [19]. Furthermore, over the last decade, improved methods for non-invasive diagnostic testing have questioned the self-evidence of conducting invasive tests.

Although the need for decision support in this domain is recognised by paediatric cardiologists, no system currently exists to support the treatment-planning process. A few systems exist to help the clinician reaching a diagnosis for children with congenital heart disease [32, 30, 28]. A larger number of systems have been built to support the management of

*Appeared in: *Artificial Intelligence in Medicine* **15**(2):135–154,1999.

patients with acquired cardiovascular disorders, most notably ischemic heart disease. In this application field, the emphasis has also been on data interpretation and diagnosis, but recently some systems (e.g., the Heart Disease Program [17]) have a component for reasoning about the expected effects of therapy.

In this article, we present a formalisation of the problem of patient management over a prolonged period of time. This is the problem of taking appropriate clinical action at appropriate points in time, where the effects of actions cannot be predicted with certainty. Furthermore, knowledge concerning the disease state will generally be limited, and diagnostic tests will have to be considered repeatedly. In general terms, this problem may be characterised as prognostic assessment and time-critical action planning under uncertainty. As a case study we have selected the congenital cardiac anomaly of ventricular septal defect. The formalisation serves as the basic representation for a system under construction that is aimed to provide the cardiologist with interactive and transparent decision support. This goal requires the formalisation to reveal the structure of the management problem as apprehended by clinicians in the field; essential ingredients of the problem will need explicit representation.

Throughout the history of AI in medicine, a large amount of research work has been devoted to formalising decision problems under uncertainty. The focus of work in this area has gradually shifted from models with static features to ones that emphasise the dynamic aspect of clinical decision-making problems. The integrated approach to time-critical decision making under uncertainty has been termed *decision-theoretic planning* [8]. Although various formalisations of this type of planning have been proposed, it is now widely acknowledged that the theory of *Markov decision processes* (MDPs) [11] provides a suitable unifying framework in this field. Markov decision processes are models for sequential decision making under uncertainty, taking into account both the outcomes of current decisions and future decision-making opportunities. The generalisation to *partially observable Markov decision processes* (POMDPs) [2] allows for the expression of many different decision-making scenarios, including reasoning with incomplete information, and planning of both information-gathering and intervening actions.

The present work explores the applicability of POMDPs in the formalisation and solution of time-critical management problems in medicine. Although the expressive power of POMDPs permits the explicit formulation of most ingredients of complex decision-theoretic planning problems, the framework is seldomly applied in practice. This is due to two reasons. First, for some types of knowledge involved in these problems, little organisational support is provided by POMDP models. And second, the complexity of the approach, both in representational and computational respects, hampers efficient application. Our formalisation incorporates a solution to these deficiencies, as recently proposed in the literature, based on probabilistic network representations [22, 7].

This article is organised as follows. In Section 2, we discuss the problem of clinical treatment of children with a ventricular septal defect. This problem is further analysed in Section 3. Section 4 reviews Markov decision process models and associated solution techniques. In Section 5, we discuss the representation of Markov decision processes in probabilistic networks. The resulting representation method is then compared to related approaches to sequential decision making under uncertainty in Section 6. The paper is completed with a discussion in Section 7.

2 Treatment planning for VSD

Ventricular septum defect (VSD) is a relatively well-understood disorder with many clinical features that are characteristic for congenital heart disease in general. Anatomically speaking, VSD is a hole in the ventricular septum, the fibromuscular wall that separates the left and the right ventricle. The main pathophysiological consequence of the defect is blood flow ("shunt") from the left to the right ventricle due to ventricular pressure differences. The shunt size, i.e., the amount of blood flowing through the defect, depends primarily on the size of the defect and the pulmonary vascular resistance. During the fetal stage, the muscular pulmonary arteries are small in diameter with a thick smooth muscular wall, thus preventing massive shunting by their high resistance. In the first weeks following birth, the arteries change to thin-walled structures with increased internal diameter. These changes are accompanied by a decline in pulmonary vascular resistance, resulting in an increased shunt size.

Left-to-right shunting causes oxygenous blood to be pumped into the lungs again. As a result, the pulmonary vascular pressure will rise, and systemic cardiac output will decrease. The latter effect usually causes the patient to look pale and to be easily sweating. With large defects, the high pulmonary arterial pressure (*pulmonary hypertension*) may lead to heart failure. Heart failure accounts for most of the typical symptoms associated with VSDs, such as shortness of breath, feeding problems, oedema, and growth arrearage. Severe heart failure may result in cardiomegaly (enlarged heart), hepatomegaly (enlarged liver), and oedema.

About 70% of all VSDs close spontaneously due to normal growth of tissue [14]. The majority (54%) closes in the first two years of life, but spontaneous closures have been reported to occur up to the age of 31 years; the rate of closure appears to follow an exponential decay rate. With small defects, the clinical course is favourable throughout infancy and childhood [13]. Patients with moderate-sized defects may develop large left-to-right shunts and associated complications in infancy, but the majority of this group can be managed medically without surgical intervention. Patients with large defects are more difficult to manage, because of the risks of mortality in the first year of life due to heart failure and associated pulmonary infections. Furthermore, elevated pulmonary vascular resistance may develop over time as a response to continuous pulmonary overflow and hypertension [9]; this is termed *Eisenmenger's complex*. As a result, the shunt size will decrease and the symptoms of heart failure will vanish. However, Eisenmenger's complex eventually leads to severe, irreversible damage to the pulmonary arteries (*arteriopathy*); this is accompanied by a reversal of the shunt direction, which can be recognised by cyanosis (a bluish tint to the skin). Early surgical intervention is therefore strongly recommended for these patients, but once the pulmonary arteries are damaged by the Eisenmenger reaction, surgical closure of the VSD will only worsen the condition of the patient. The majority of patients with repair of uncomplicated VSD in infancy or early childhood have an excellent result with no clinical signs or symptoms and apparently normal life-expectancy [21].

For the clinician, the main problem is to decide if and when to submit a patient to surgery. Usually, the patient's condition is monitored without surgical intervention during the first year of life. During this period, non-invasive diagnostic tests such as auscultation and echocardiography are conducted repeatedly, and when necessary, medical treatment is given to reduce the effects of shunting and improve the overall condition of the patient. Sometimes, X-ray images of the chest are made to inspect the size of the heart and pulmonary arteries. After the first year of life, the risks associated with surgical intervention have dropped, and a decision whether surgery is necessary has to be made. In cases of doubt concerning shunt size and the state of the pulmonary arteries, cardiac catheterisation or pulmonary biopsy may be performed prior to that decision to obtain more information. Therapy is considered completed after closure of the defect, either spontaneously or by surgical intervention.

In general terms, this problem may characterised as prognostic assessment and action planning under uncertainty, where the timing of actions is essential. A trade-off is required between the benefits gained by waiting before intervention in the hope that the patient's condition will improve, and the risks caused by natural history. Careful timing of clinical investigations can improve the ability to predict the future course of disease. Furthermore, risks and costs[1] of invasive testing have to evaluated against their potential information gain.

3 Formalising the treatment-planning problem

In this article, we develop a formalisation of the prognosis and treatment planning problem for VSD patients. Our aim is to provide automated support for the paediatric cardiologist who has to decide upon therapy for individual patients. Interviews with field experts have led us to the conviction that this is best accomplished by a "white-box" system, in which the user can perceive what is going on, and can interact by proposing alternatives or adjust admissible plans. This implies that we will have to use a formalisation that reveals the structure of the problem as apprehended by clinicians in the field.

3.1 Basic ingredients

There are three key concepts as revealed in the description of the VSD domain that need distinction: (1) the clinical state of the patient and its development over time, (2) the management strategy employed by the clinician, and (3) the overall objective of management. Parts of the *clinical state* of the patient pertain to symptoms and signs associated with the disease, and are readily observable by the clinician. Other parts pertain to pathophysiological parameters hidden from direct observation, though some of these parameters may be measured by clinical investigations. Furthermore, clinical states exhibit internal structure relating the various parts, and subsequent states in time are related by natural development of the disease. An analysis of the second concept, the *management strategy*, makes clear that this can be considered as composed of series of decisions. Here, each decision to conduct some clinical action has *test* and *treatment effects*. Test effects pertain to observations regarding the state of the patient yielded by performing the action, whereas treatment effects pertain to expected changes to that state caused by the action. The third concept, the *objective* of the management strategy, addresses the trade-off between benefits and costs associated with conducting clinical actions on the one hand, and the risks caused by natural history of the disease on the other. The notion of time is essential for each of these ingredients and therefore needs to be made explicit in a formalisation as well.

3.2 Clinical state

To describe the clinical state of the patient we have chosen to use a set of 62 stochastic variables. These variables and their respective value ranges were assessed in co-operation

[1]Throughout this article, we will use the term "costs" for both financial and non-financial consequences (e.g., pain, stress) of clinical actions.

Variable	*Interpretation*	*Domain*
size	defect (VSD) size	*null*, *small*, *moderate*, *large*
shunt	shunt size	*none*, *small*, *large*, *reversed*
res	relative pulmonary vascular resistance	*normal*, *increased*, *high*, *very_high*
fail	heart failure	*absent*, *mild*, *moderate*, *severe*
symp	heart failure symptoms	*absent*, *mild*, *moderate*, *severe*
pmhyp	pulmonary hypertension	*absent*, *mild*, *moderate*, *severe*
pmart	pulmonary arteriopathy	*absent*, *mild*, *moderate*, *severe*
cyan	central cyanosis	*false*, *true*
death	death	*false*, *true*

Table 1: Stochastic state variables for the VSD domain.

with a domain expert and by studying a collection of medical records of VSD patients. They can be divided in five groups, according to their role in clinical reasoning:

1. Primary pathophysiology (size and location of the VSD, pressure and resistance gradients in heart and great vessels, shunt direction and size);
2. Complications (heart failure, pulmonary infections, cardiomegaly);
3. Signs and symptoms (shortness of breath, feeding problems, oedema, growth arrearage);
4. Clinical findings (results of cardiac ultrasound imaging, catheterisation, X-ray imaging and taking pulmonary biopsies); and
5. Risks (mortality and morbidity).

We will use a subset of nine variables from this set in the illustrations throughout this article; these variables, and their possible value sets, are listed in Table 1. The majority of the selected variables actually pertain to continuous entities. For instance, the shunt direction and size, usually expressed as the ratio of pulmonary and systemic blood volumes, may range from $1:1$ (the normal situation, no shunt) to $2:1$ or greater with large left-to-right shunts; furthermore, a ratio of slightly less than $1:1$ is found with right-to-left shunts. Nevertheless, discretised value ranges are generally believed to provide sufficient level of detail for decision-making tasks in this field. Our choice was therefore to take over commonly used discretisations, which facilitated communication with domain experts and using results from clinical trials reported in the literature. Returning to the example, we thus only distinguish absence of shunting ($1:1$), small left-to-right shunting ($1:1 - 2:1$), large left-to-right shunting ($> 2:1$), and reversed (i.e., right-to-left) shunting ($< 1:1$).

3.3 Actions

The available clinical actions for the VSD domain, their effects to state variables, and their associated observations are summarised in Table 2. Medical treatment may be used to control heart failure, and surgery may be used to close the defect. Information concerning the clinical state of the patient may be obtained by making echocardiographic images, chest X-ray images, cardiac catheterisation, and taking pulmonary biopsies. Unfortunately, the latter three of these test actions carry adverse effects and/or expose the patient to mortality risks; their

Action	*Interpretation*	*Effects*	*Observations*
echo	echocardiography	–	size
med	medical treatment	fail	–
xray	chest X-ray imaging	–	pmhyp
cath	cardiac catheterisation	death	shunt
biop	pulmonary biopsy	–	pmart
surg	perform surgery	size, death	size

Table 2: Available treatment actions for the VSD domain.

deployment is therefore subject to a trade-off.

The next step in the formalisation is now to make an integrated, formal description of the potential temporal developments of clinical states, the interaction between actions and clinical states, and the long-term objective of action planning. A suitable framework to provide such a description is found in the theory of Markov decision processes. We elaborate on this theory in the next section.

4 Markov decision processes

Markov decision processes (MDPs) [11, 26] are models for sequential decision making under uncertainty, which take into account both immediate and long-term consequences of decisions. Basically, the theory assumes that a person, called the *decision maker*, is charged with the responsibility of choosing a sequence of actions in order to influence a stochastic process. The immediate result of each action choice is that the process under consideration evolves to a new state according to a probability distribution determined by the action choice. Furthermore, the decision maker receives an immediate reward reflecting the desirability of the new state compared to other possible states. The goal is to optimise some function of the overall reward sequence that expresses the decision maker's intertemporal trade-offs. *Partially observable Markov decision processes* (POMDPs) [2] are a generalisation of Markov decision processes where the decision maker has limited knowledge concerning the process state, and action choice determines the acquisition of state information. Consequently, in POMDP problems the trade-off between actions does not only concern their immediate and long-term effects, but also their information-gathering properties. The generalisation to POMDPs is significant in problem settings where state uncertainty is a central issue that cannot be discarded; we note that this holds for most domains in clinical medicine, including for the VSD domain.

4.1 POMDP model

Formally, a POMDP model is a tuple (T, X, A, P, ω, R), where

- T is a set of *decision moments*,
- X is a set of *stochastic state variables*, jointly defining the set of *states*,
- A is a set of available *actions*,
- P is a set of *transition probability functions*,

- ω is an *observation function*, and
- R is a set of *reward functions.*

The basic form of the model is quite general and the literature on POMDPs provides a plethora of possible choices for each of the elements. The qualifier "Markov" refers to the fact that the transition probability, observation and reward functions depend on the past only through the current process state and the most recent action selected by the decision maker. Below, we discuss the appropriate choices for each of the model elements from the perspective of clinical VSD management.

The set T of decision moments denotes the points in time where the decision maker is expected to choose an action. As the duration of patient management in our problem domain is typically bounded to the first years of life, we will confine ourselves to the case where T is finite; one then speaks of a *finite-horizon process.* Standard POMDP theory requires us to fix the timing of decisions in advance for a given problem domain. However, in the clinical practice of VSD treatment, decision moments are established by the cardiologist in due course. Therefore, we use the set T as a *grid* of decision moments, where each time point $t \in T$ is a potential decision moment. We take $T = \{0, 1, 2, \ldots, N\} \subset \mathbb{N}$, where $t \in T$ denotes the age of the patient expressed as number of life months. We adopt the convention that no action is chosen at the last potential decision moment $t = N$: this moment is included for evaluation of the final state only.

The state space of a POMDP model is characterised by a finite set $X = \{X_1, \ldots, X_m\}$ of discrete, stochastic variables. For the VSD domain example in this article, we have that $m = 9$; the variables are listed in Table 1. A proposition of the form $X_i = x$ (i.e., variable $X_i \in X$ has value x) will be called a *configuration* of variable X_i; the set of all configurations of variable X_i is denoted by C_{X_i}. As the configuration of state variables may change over time, we will use superscripts to distinguish specific configurations at different time points; e.g., $c^t_{X_i}$ denotes the configuration of variable X_i at time point $t \in T$. The set of all configurations of a subset $X' = \{X_{i_1}, \ldots X_{i_k}\} \subseteq X$ of state variables now is defined as $C_{X'} = C_{X_{i_1}} \times \ldots \times C_{X_{i_k}}$. Consequently, the state space of a POMDP (i.e., the set of potential clinical states) equals $C_X = C_{X_1} \times \ldots \times C_{X_m}$. POMDP theory does not prohibit us from using a single state variable (capturing all the relevant information on the patient in condensed form), but a state space characterised by multiple stochastic variables offers several advantages from a representational perspective; we return to this subject in Section 5.

For the set A of available actions in the VSD domain we refer to Table 2. Our approach differs from the standard model presented in the literature in that we allow for multiple, simultaneous actions at each potential decision moment $t \in T$. That is, the clinician is expected to choose a subset $\alpha \subseteq A$ of actions to be performed, where choosing the empty set $\varnothing$ is interpreted as a *skip*-action, i.e., deciding to refrain from action at that point in time. Using this action, the clinician can make decisions at irregular time points (by skipping intermediate points) and 'fill up' the remaining time when some satisfactory state is reached. The dynamics of the decision process, given a sequence of action choices, are schematically depicted in Figure 1.

Treatment effects of actions are described by the set $P = \{p^\alpha_t : C_X \times C_X \to [0, 1] \mid t \in T, \alpha \subseteq A\}$ of time- and action-dependent transition probability functions, where $p^\alpha_t(c^{t+1}_X \mid c^t_X)$ denotes the probability of arriving at state c^{t+1}_X after performing action set $\alpha \subseteq A$ in state c^t_X at moment t. Test effects are modelled by the observation function $\omega : A \to \wp(X)$, where $\omega(a)$

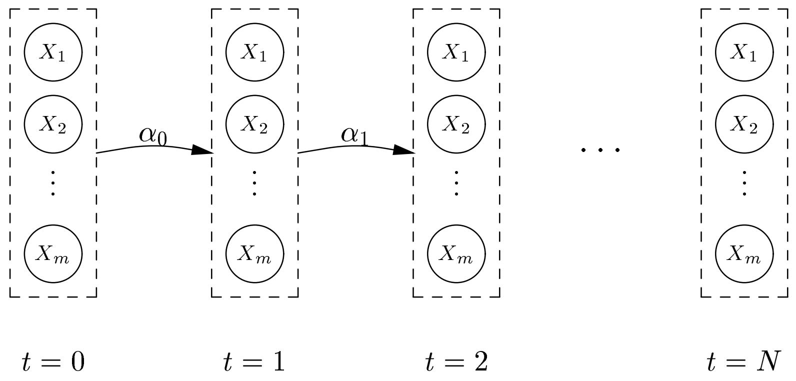


Figure 1: Dynamics of Markov decision processes.

denotes the set of state variables whose current values can be observed by the decision maker when action a has just been performed. For an action set $\alpha \subseteq A$, we define $\overline{\omega}(\alpha) = \bigcup_{a \in \alpha} \omega(a)$ to be the observable set. The sets of observable state variables for the VSD domain are listed along with the available actions in Table 2.

Finally, the immediate costs and benefits associated with clinical states and action choices are described by the reward functions $r_t : C_X \times \wp(A) \to \mathbb{R}$ for time points $t = 0, \dots, N-1$, and $r_N : C_X \to \mathbb{R}$ for time point $t = N$. Here, $r_t(c_X^t, \alpha)$ denotes the reward received when, at time point $0 \leq t \leq N-1$, the current state is $c_X^t \in C_X$ and action set $\alpha \subseteq A$ is chosen by the decision maker. This reward value reflects the relative (un)desirability of that state and action choice only; potential future developments are disregarded. The value $r_N(c_X^N)$ denotes the final reward received when the process ends in state $c_X^N \in C_X$ (recall that there is no action choice at the last point in time $t = N$); with this reward, we do take into account potential future developments and life-expectancy of the patient. For the VSD domain, we can apply several reductions to the general form of the reward functions, stemming from the following observations. Given a VSD patient, we are primarily concerned with the patient's life-expectancy; our secondary interest is to minimise the accumulated costs over the course of therapy. As in our domain the patient's clinical state after therapy includes a description of survival so far, and gives the best indication for future life-expectancy, the reward model for $t = 0, \dots, N-1$ can be simplified to $r_t : \wp(A) \to \mathbb{R}$. At time point $t = N$, the relevant variables are those describing the current survival of the patient and the state of the pulmonary arterioles. The reward function for the final moment therefore equals $r_N : C_{\{\mathsf{death},\mathsf{pmart}\}} \to \mathbb{R}$. We note that these reductions crucially depend on the fact that the normal life-expectancy (i.e., that of healthy persons) greatly exceeds the planning horizon in our domain.

4.2 Optimisation criterion

To complete the specification of the decision problem at hand, a *utility function* $u : \mathbb{R}^{N+1} \to \mathbb{R}$ is specified, reflecting the decision maker's intertemporal trade-offs. We say that $u(r_0, \dots, r_N)$ is the *utility* associated with the reward sequence $r_0, \dots, r_N$. Solving a sequential decision problem involves repeatedly applying maximisation and expectation to the utility function. These operators have special properties when applied to utility functions that are *time-separable.* Time-separable utility functions allow maximisations and expectations to be performed over separate rewards in the sequence, thus requiring only a subspace of the utility

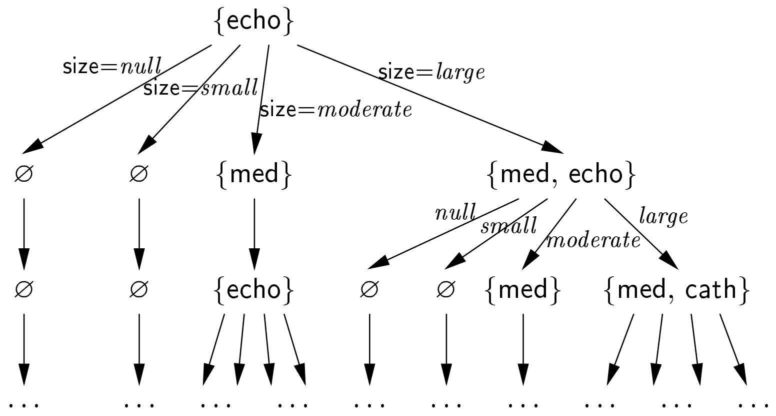


Figure 2: Example treatment plan for the VSD domain.

function to be examined when solving the decision problem. This significantly reduces the dimensionality of the solution operations. The most popular choice for the function u is *linear additive utility*, i.e., $u(r_0, \ldots, r_N) = \sum_{i=0}^{N} r_i$; it reflects the preferences of a risk-neutral decision maker. It is the form of utility that is convenient for our purposes in the VSD domain, given that the final reward takes into account the life-expectancy of the patient.

4.3 Solving a POMDP

A POMDP problem specification now consists of a POMDP model, an initial probability distribution on the state space, and a utility function. The *sequential decision problem* is to choose, prior to the first decision, a *decision-theoretic plan* to maximise expected utility. Such a plan provides the decision maker with a prescription for choosing an action set at each point in time, given the history of past actions and observations. The majority of POMDP research (e.g., [18, 16]) focuses on finding *stationary Markovian* plans that maximise expected utility; this type of plan neglects the timing of actions and considers the most recent observation for the action choice only. We believe, however, that these limited plans are not suited for medical decision problems that require the clinician to carefully trace out a course of action over a longer period of time; non-Markovian plans are preferred for such problems. A natural representation of non-Markovian plans is found in rooted trees, where the nodes represent action set choices, and the links stand for possible observations. Figure 2 depicts an example treatment plan for the VSD domain that covers the first three decision moments.

Generally speaking, the complexity of finding utility-maximising (non-Markovian) plans depends on the size of the state space, the number of available actions, and the horizon length. For finite-state (fully observable) MDP problems, efficient solution methods exist, based on the principle of dynamic programming [3, 26]. Unfortunately, this does not hold for problems involving partial observability. Solving a POMDP problem directly necessitates keeping track of entire process histories, of which the sizes grow exponentially in the size of the state space and action sets. A more promising approach is based on transforming the POMDP model to an equivalent MDP model (called the *belief MDP*) in which process states are probability distributions on the state space of the original POMDP model [2]; transition probabilities for the belief MDP are derived through Bayes' rule. Dynamic programming techniques can then be applied on the belief MDP to solve problems for the original POMDP, but because of the continuity of the belief state space, algorithms are complicated and limited. Solving POMDPs with a short, finite planning horizon is nevertheless feasible [18].

We conclude that POMDPs provide a powerful framework for sequential decision-making problems where both uncertainty in action outcome and imperfect observability are essential. When we reconsider the ingredients of the VSD treatment-planning problem as distinguished in Section 3, we find that most of them are formally described in the POMDP model. The notion of clinical development of the patient over time is modelled by successive configurations of the state variables, clinical actions with their test and treatment effects are described by the transition probability and observation functions, and the objective of therapy is captured by the reward and utility functions. However, the internal structure of clinical states and the relations between individual state variables cannot be captured by POMDP models. In other words, the granularity of the POMDP representation is too coarse to allow knowledge to be expressed at the level of individual variables. Furthermore, the POMDP formulation becomes impractical as the number of potential clinical states increases, because the number of transition probabilities that must be assessed grows exponentially in the size of the state space. To provide for these deficiencies, it has been suggested to augment these models with a factorised representation of the transition probability functions in probabilistic networks [5]. In the next section, we explore this possibility in more detail.

5 Graphical representations of POMDPs

In the characterisation of the state space of a POMDP, we choose to use multiple variables instead of a single variable. Multiple variables provide more structure, and thus help to simplify the state descriptions; such a state space characterisation is therefore called *structured* [5]. The main advantage of structured state space is that they offer the possibility to factorise the transition probability distributions describing the effects of actions. Factorisation of a joint probability distribution is based on conditional independence relations induced by the distribution, and allows for a reduction in the required number of model parameters, and for more efficient probabilistic inferences. *Bayesian belief networks* [22] provide for a concise, graphical representation of factorised joint probability distributions. More recently, the belief network framework was extended to cope with dynamic stochastic systems [7, 6], where the joint probability distribution on the variables in the network evolves over time. These networks are generally referred to as *temporal* belief networks. Temporal belief networks have been suggested as a suitable way to express the dynamics of POMDP models [5], equally facilitating the solution methods by exploiting independencies between state variables, and by making explicit persistence of states and ramification of action effects [7].

5.1 Two-stage temporal belief networks

Now, let $X(t)$ denote the set of state variables at time point t. A *two-stage temporal belief network* (2TBN) [5] is a belief network with two sets of variables $X(t)$ and $X(t+1)$, where each arc is drawn either a variable from $X(t)$ to a variable from $X(t+1)$, between two variables both from either $X(t)$ or $X(t+1)$. The former type of arc is called a *temporal* dependency; the latter is called an *atemporal*, or *synchronic*, dependency. 2TBNs allow for a compact representation of probabilistic state transitions, where temporal dependencies model the direct effects of actions, and synchronic dependencies model intra-stage correlations. We

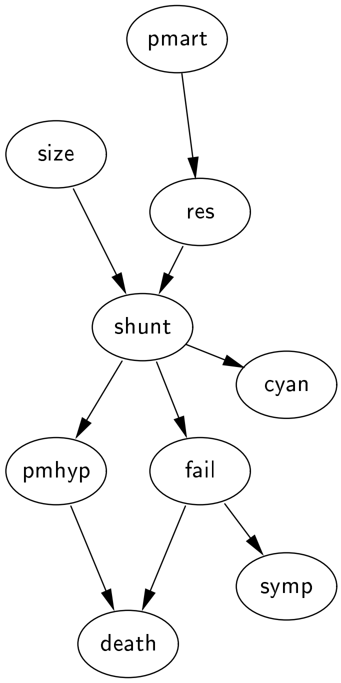


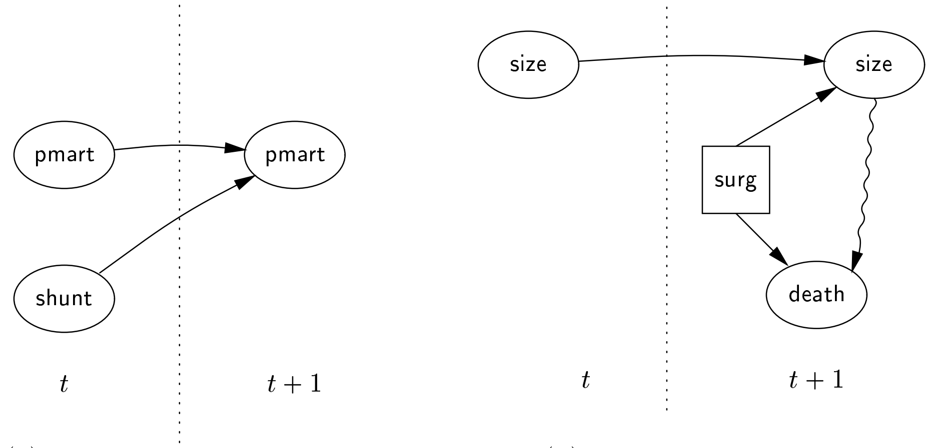
Figure 3: Synchronic belief network for the VSD domain.

assume that both inter-stage transition rates and intra-stage correlations may vary over time, but that probabilistic independence relations do not. In other words, the structure of the 2TBN is similar for all time points.

5.2 Belief networks for the VSD domain

A belief network for the VSD domain has recently been constructed with aid of a domain expert [24]; a simplified version of the atemporal part of this network is shown in Figure 3. For $X = \{X_1, \ldots, X_m\}$, let $\Gamma_t = \{\gamma_{t,i} : C_{\sigma(X_i)} \times C_{X_i} \to [0,1] \mid i = 1, \ldots, m\}$ be the set of synchronic probability assessment functions associated with the belief network at time point $t \in T$, where $\sigma(X_i)$ denotes the parent set of variable X_i in the synchronic dependency graph; $\gamma_{t,i}(c^t_{X_i} \mid c^t_{\sigma(X_i)})$ denotes the probability of variable X_i having configuration $c^t_{X_i}$ at time point $t \in T$, given the configuration $c^t_{\sigma(X_i)}$ of its parents. The synchronic probability assessment functions for time point $t = 0$ are taken to represent the initial probability distribution on the state space in the POMDP problem specification.

We distinguish two kinds of temporal dependency: those induced by action choices (*exogenous change*) and those stemming from persistence or action-independent change of state variables (*endogenous change*). Both kinds of dependency can be modelled using 2TBNs, comprising only those state variables that are directly relevant. Figure 4a shows the 2TBN modelling endogenous change due to Eisenmenger's complex: pulmonary arteriopathy may result from left-to-right shunts. This process is progressive, and therefore we also take into account the former state of the pulmonary arteries. Figure 4b shows the 2TBN modelling exogenous change due to surgery. The surgery action itself is depicted by a square box, called an *action node*. Successful surgery results in a closed defect (i.e., size = *null*). However, the chances of success equally depend on the size of the existing defect: large defects are more



(a) Eisenmenger's complex (b) surgical intervention

Figure 4: 2TBN models.

difficult to patch. Furthermore, a small risk is associated with the operation. The twisted arc here indicates that there is also a path in the synchronic dependency graph from size to death. If we merge the atemporal belief network and each of the 2TBNs modelling endogenous and exogenous change, a complete graphical model for transitions and observations is obtained. This model is shown in Figure 5.

Formally, let $\tau(X_i)$ be the set of temporal predecessors of node $X_i \in X$ in the complete model, that is, there is an arc from $X_j(t)$ to $X_i(t+1)$ for every $X_j \in \tau(X_i)$, and let $A_i \subseteq A$ be the set of actions directly affecting X_i, i.e. there is an arc from action node a to X_i for every $a \in A_i$. Furthermore, let $\Delta_t = \{\delta_{t,i} : C_{\sigma(X_i)} \times C_{\tau(X_i)} \times \wp(A_i) \times C_{X_i} \rightarrow [0,1] \mid i = 1, \ldots, m\}$ be the set of transition probability assessment functions associated with the complete graphical model at time t; $\delta_{t,i}(c^{t+1}_{X_i} \mid c^{t+1}_{\sigma(X_i)}, c^{t}_{\tau(X_i)}, \alpha_i)$ denotes the probability of X_i having configuration $c^{t+1}_{X_i}$ at time $t+1$ given that its synchronic predecessors then have configuration $c^{t+1}_{\sigma(X_i)}$, its temporal predecessors had configuration $c^{t}_{\tau(X_i)}$ at the previous moment t, and subsequently action set choice $\alpha_i \subseteq A_i$ was made. We assume that $\delta_{t,i}(c^{t+1}_{X_i} \mid c^{t+1}_{\sigma(X_i)}, c^{t}_{\varnothing}, \varnothing) = \gamma t, i(c^{t+1}_{X_i} \mid c^{t+1}_{\sigma(X_i)})$, i.e., synchronic probability assessment functions are used when there are no temporal predecessors and no action is selected that directly affects variable X_i. In the VSD example, this holds for five or six variables, depending on the action set choice. Two variables, size and pmart are subject to endogenous change connected to the natural course of disease.

The transition probability distributions associated with the POMDP model can now be factorised according to the independency relations portrayed by the graphical part of the complete 2TBN (i.e., with all action nodes added). For each potential decision moment $t \in T$ and each action set $\alpha \subseteq A$, we have that the transition probability function p^{α}_{t} can be written as

$$p^{\alpha}_{t}(c^{t+1}_{X} \mid c^{t}_{X}) \quad = \quad \prod_{i=1,\ldots,m} \delta_{t,i}(c^{t+1}_{X}(\{X_i\}) \mid c^{t+1}_{X}(\sigma(X_i)), c^{t}_{X}(\tau(X_i)), \alpha \cap A_i)$$

for each $c^{t}_{X}, c^{t+1}_{X} \in C_X$, where $c^{t}_{X}(X')$, $c^{t+1}_{X}(X')$ denote the configurations of $X' \subseteq X$ within c^{t}_{X} and c^{t+1}_{X}, respectively.

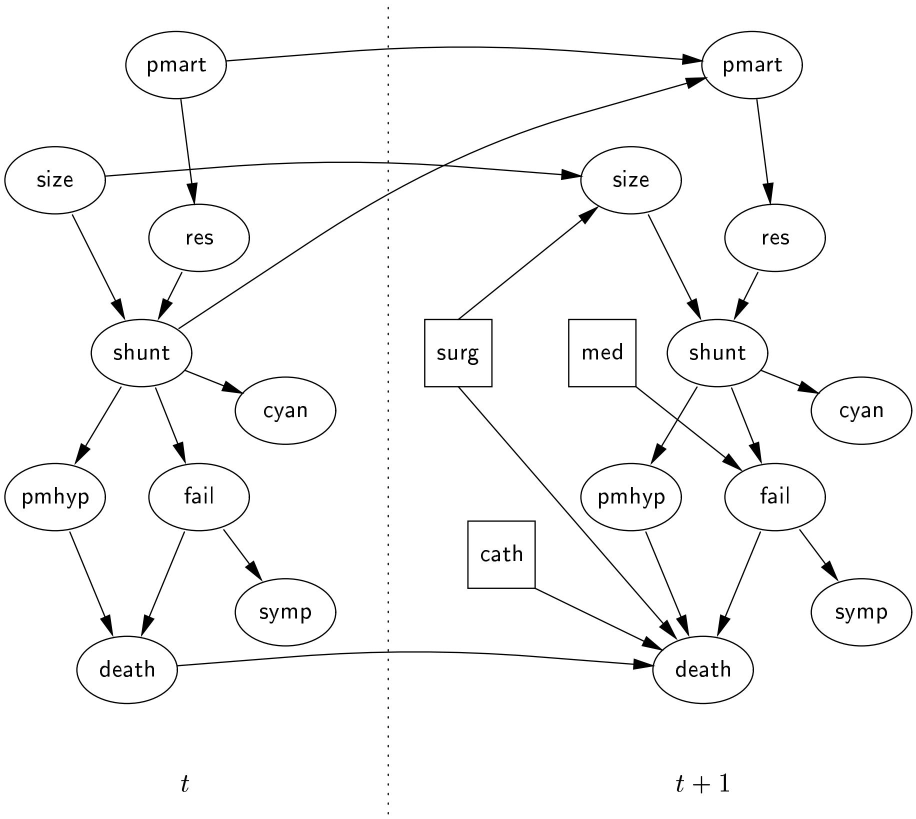


Figure 5: Full 2TBN with action nodes for the VSD domain.

5.3 Representational complexity reduction

We will now provide theoretical bounds on the reduction in the number of probability estimates that have to be assessed for the model. Let d be an upper bound on the number of values that each state variable can take, *i.e.*, $|C_{X_i}| \leq d$ for each $X_i \in X$. So, for $X = \{X_1, X_2, \ldots, X_m\}$, we have that $|C_X| \leq d^m$. Furthermore, let $|A| = k$, so that there are 2^k different action sets. Then, for each $t \in T$ and each $\alpha \subseteq A$, we need at most $d^m \cdot d^m = d^{2m}$ transition probabilities, so for the POMDP model, the upper bound on the total number of parameters needed per time point is $d^{2m} \cdot 2^k$.

In the graphical model, there are m state nodes, and with each node X_i is associated a transition probability assessment function $\delta_{t,i}$ for which at most $d \cdot d^{|\sigma(X_i)|} \cdot d^{|\tau(X_i)|} \cdot 2^{|A_i|}$ probabilities have to be specified. If the numbers of synchronic and temporal predecessors are bounded by l_s and l_t, respectively, and the number of actions directly affecting X_i is bounded by q, for each $X_i \in X$, then the total number of parameters for the complete model is at most $m \cdot d \cdot d^{l_s} \cdot d^{l_t} \cdot 2^q = d^{1+l_s+l_t} \cdot 2^q$. Note that in the most extreme case, where $l_s = m - 1$, $l_t = m$ and $q = k$, we obtain the number $m \cdot d^{2m} \cdot 2^k$. Fortunately, in most practical applications, we have that $l_s + l_t \ll 2m - 1$ and $q \ll k$. The reasons for this are as follows. Reconsidering the division of state variables in Section 3.2, we see that temporal dependencies are most likely to be found among variables in the first and last groups, i.e., those pertaining to pathophysiological conditions and risks. Sometimes complications of disease may also carry over time, but symptoms and clinical findings will typically depend on the current pathophysiology and complications only; they have no temporal predecessors in the graphical model. Conversely, the number of synchronic predecessors of pathophysiological variables will be small as these variables represent the causes of variables in other groups. Similar observations hold for the number of variables affected by actions. Actions typically affect risk variables, and one or more variables within another group.

In the example model for the VSD domain, we have 9 state variables ($m = 9$) with each at most 4 configurations ($d = 4$), and 6 available actions ($k = 6$); the numbers of synchronic and temporal predecessors in the 2TBN, however, are bounded by 2 ($l_s = 2$ and $l_t = 2$), and each action affects at most 2 state variables ($q = 2$). The reduction in the number of probabilities required per time point is therefore in the order of 10^9.

Summarising, 2TBNs with action nodes provide for a concise, graphical representation of the dynamics of POMDP models. Concision is arrived at by exploiting both intra-stage and inter-stage independencies between state variables, and explicitly representing the limitations in the effects of actions. This allows for a representation of deterministic and probabilistic relations at the level of individual state variables, and a reduction in the number of model parameters that have to be assessed.

6 Comparison with related work

Decision theory is becoming increasingly popular as a mathematical foundation for building planning systems in uncertain domains. This section briefly reviews representation formalisms that have been applied to time-critical decision-theoretic planning problems in medicine. The fundamental trade-off in this field is between representational expressiveness and model transparency on the one hand, and efficiency of solution techniques on the other [15, 23]. Most formalisms impose a number of restrictive assumptions on the type of problem that may be addressed, to enhance computational efficiency. Unfortunately, these assumptions are not always made clear. Other formalisms are more general in nature, at the penalty of high computational cost or even intractability.

The idea of using a temporal belief network to model the evolution of a process over time has been applied in several medical treatment-planning systems. Andreassen et al. [1] implemented a differential equation model of carbohydrate metabolism in the form of a belief network with (Markovian) time slices for predictions of 24-hour blood glucose profiles. Berzuini et al. [4] have proposed a general, temporal belief network model for monitoring and controlling biomedical processes; the model was applied to the monitoring of cancer patients receiving post-operative cytotoxic chemotherapy. In these applications, the emphasis is on predicting future states of the patient. Actions may be selected to control one of the variables in the network, but the problem of simultaneously planning investigative and intervening actions is not addressed. The support provided by these systems is therefore restricted to a single facet of patient management.

Another approach is based on extending the *influence diagram* representation [12, 29]. An influence diagram is a 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. The graphical structure allows representation of, and reasoning with, conditional independences between random variables and locality of action effects. Quaglini et al. [27] developed a system for managing anaemic patients in which a *Markov node* was added to the influence diagram representation for prognostic purposes. In a more general fashion, influence diagrams can be extended with repeating network structure and a time-separable utility function [31, 25]. A limitation of this approach is found in the fact that influence diagrams correspond to the restricted class of symmetrical decision trees; this imposes various restrictions on plan structure and flexibility.

A few applications of Markov decision processes to medical decision problems have been developed. Magni and Bellazzi [20] apply MDPs with structured state spaces to the problem of therapy planning for patients with hereditary spherocytosis. The optimal plan for their model was compared with the recommendations yielded by a Markov decision tree, and found to yield a slight improvement in life expectancy. Like in our method, the effects of actions in their model are described in probabilistic networks; an important difference is that they assume perfect observability of the system state at all times. Although this assumption yields a significant reduction in the computational complexity of solution methods, it is not tenable in most medical domains: the aspect of incomplete and noisy observation often plays an essential in clinical decision making.

Hauskrecht [10] developed a POMDP model with a structured state space for the management of patients with ischemic heart disease. The solution (plan) computed for this model was not experimentally tested or compared to other models. It seemed, however, in most cases to make reasonable therapeutic recommendations that are in concordance with clinical practice. In some cases the recommendations were doubtful; this was attributed to simplifications in the model and errors in the (subjectively estimated) model parameters.

In recent years, attempts have been made to develop a common vocabulary for comparing the major approaches to decision-theoretic planning. These attempts generally take some form of Markov decision process as the unifying mathematical framework, comparable to the role of first-order predicate logic in analysing traditional knowledge-representation formalisms. For instance, Leong [15] developed a formalism for time-critical decision-theoretic planning that allows for multiple graphical perspectives (visualising the same information in different ways) and incremental language extension (gradually expanding the scope of the problem addressed). The underlying mathematical representation is a semi-Markov decision process; in this type of process, state transitions are stochastic not only with respect to the resulting state, but also regarding their time lapse. The formalism was applied to the problem of managing patients with atrial fibrillation.

7 Discussion

The planning of clinical management requires the ability to predict the interplay between the natural history of disease and effects of intervening actions over time. Often, such predictions cannot be made with certainty, and trade-offs have to be made between the expected benefit of current and future decisions. We have shown how partially observable Markov decision processes can be used to formalise this type of problem, providing an explicit representation of clinical states of the patient, the management strategy employed, and the objectives of treatment. Augmented with structured state spaces and graphical representations of the transition probability functions, the POMDP framework provides a powerful knowledge-representation formalism for the problem of choosing an optimal course of action for a patient whose physiological conditions may vary over time. We conclude the article with a discussion of some of the potential weaknesses of this formalism, and how they can be coped with.

The most prominent assumption underlying the POMDP framework is that the transition probability, observation and reward functions obey the Markov property. That is, they depend on the past only through the current process state and the most recent action choice. Although this is a restriction on the temporal expressiveness of the formalism, one can often "work around" it by adding state variables that repeat or abstract parts of the process history (e.g., trend variables). A further restriction stems from the fact that POMDPs model continuous time through discretisation, and transitions are assumed to have fixed time duration. This can be quite unnatural, especially when there are multiple pathophysiological developments that may occur in parallel, but with different durations; this situation is found in the VSD domain. A possible solution for this problem is the extension to semi-Markov decision processes, which allow for transition times with variable duration.

Structured state spaces help us to simplify the description of possible states, and allow for a concise, graphical representation of the transition probability functions. However, they also tend to increase the size of the state space, as the space defined by multiple variables often includes configurations that cannot occur in practice (e.g., contradictory value combinations) or are redundant (when some variables become irrelevant given a specific configuration of other variables). This is unadvantageous as the complexity of POMDP solution methods critically depends on the size of the state space. A solution to this problem, recently proposed by Hauskrecht [10], is to use state descriptions at different levels of detail. A variable at a given level then represents one or more variables at lower levels. The structuring thus allows one to characterise the state space using descriptions of different complexity and size.

Acknowledgement. The investigations were (partly) supported by the Netherlands Computer Science Research Foundation with financial support from the Netherlands Organisation for Scientific Research. The author wishes to thank Jaap Ottenkamp for his help with constructing the domain model, and Peter Lucas, Silja Renooij, and three anonymous reviewers for their comments on the manuscript.

References

[1] S. Andreassen, R. Hovorka, J. Benn, K.G. Olesen, and E.R. Carson, A model-based approach to insulin adjustment, in: M. Stefanelli, A. Hasman, M. Fieschi, and J. Talmon, eds., *Proc. Third Conf. Artificial Intelligence in Medicine* (Springer Verlag, Berlin, 1991) 239–248.

[2] K.J. Åström, Optimal control of Markov decision processes with incomplete state estimation, *J. Math. Anal. Appl. 10* (1965) 174–205.

[3] R.E. Bellman, *Dynamic Programming* (Princeton University Press, Princeton, NJ, 1957).

[4] C. Berzuini, R. Bellazzi, and D. Spiegelhalter, Bayesian networks applied to therapy monitoring, in: *Proc. Seventh Annu. Conf. Uncertainty in Artificial Intelligence (UAI–91)* (1991) 35–43.

[5] C. Boutilier, T. Dean, and S. Hanks, Planning under uncertainty: structural assumptions and computational leverage, in: M. Ghallab and A. Milani, eds., *New Directions in AI Planning* (IOS Press, Amsterdam, 1996) 157–171.

[6] P. Dagum and A. Galper, Forecasting sleep apnea with dynamic network models, in: *Proc. Ninth Annu. Conf. Uncertainty in Artificial Intelligence (UAI–93)* (1993) 64–71.

[7] T. Dean and K. Kanazawa, A model for reasoning about persistence and causation, *Comput. Intell. 5* (1989) 142–150.

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

[9] T.P. Graham and H.P. Gutgesell, Ventricular septal defects, in: A.J. Moss and F.H. Adams, eds., *Heart Disease in Infants, Children, and Adolescents*, (Williams & Wilkins, Baltimore, 1995) 724–746.

[10] M. Hauskrecht, *Planning and control in stochastic domains with imperfect information*, Ph.D. thesis, Dept. of Electrical Engineering and Computer Science, MIT, 1997.

[11] R.A. Howard, *Dynamic Programming and Markov Processes* (MIT Press, 1960).

[12] R.A. Howard and J.E. Matheson, Influence diagrams, in: R.A. Howard and J.E. Matheson, eds., *The Principles and Applications of Decision Analysis II*, (Strategic Decisions Group, Menlo Park, CA, 1984) 719–762.

[13] L. Kidd, D. Driscoll, and W. Gersony, Second natural history study of congenital heart defects: results of treatment of patients with ventricular septal defects, *Circulation 87* (1993) 138–151.

[14] L.J. Krovetz, Spontaneous closure of ventricular septal defect, *Am. J. Cardiol. 81* (1998) 100–101.

[15] T.-Y. Leong, *An Integrated Approach to Dynamic Decision Making under Uncertainty*, Ph.D. thesis, Dept. of Electrical Engineering and Computer Science, MIT, 1994.

[16] M.L. Littman, *Algorithms for Sequential Decision Making*, Ph.D. thesis, Dept. of Computer Science, Brown University, 1996.

[17] W.J. Long, H. Fraser, and S. Naimi, Reasoning requirements for diagnosis of heart diease, *Artif. Intell. Med. 10* (1997) 5–24.

[18] W.S. Lovejoy, A survey of algorithmic methods for partially observed Markov decision processes, *Ann. Oper. Res. 28* (1991) 47–66.

[19] F.J. Macartney, D.J. Spiegelhalter, and M.L. Rigby, Medical management, in: R.H. Anderson and F.J. Macartney and E.A. Shinebourne and M. Tynan, eds., *Paediatric Cardiology I* (Churchill Livingstone, Edinburgh, 1987) 421–442.

[20] P. Magni and R. Bellazzi, The optimal dynamic therapy: a decision-theoretic approach, in: P. Borne, M. Ksouri, and A. El Kamel, eds., *Computational Engineering in Systems Applications (CESA–98)* (1998).

[21] J.H. Moller, C. Patton, R.L. Varco, and C.W. Lillehei, Late results (30 to 35 years) after operative closure of isolated ventricular septal defect from 1954 to 1960, *Am. J. Cardiol. 68* (1991) 1491–1497.

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

[23] N.B. Peek and P.J.F. Lucas, Trade-offs in decision-theoretic planning, in: J. Hunt and R. Miles, eds., *Research and Development in Expert Systems XIV (Proceedings of Expert Systems '97)* (SGES Publications, 1997), 147–157.

[24] N.B. Peek and J. Ottenkamp, Developing a decision-theoretic network for a congenital heart disease, in: E. Keravnou, C. Garbay, R. Baud, and J. Wyatt, eds., *AIME '97: Proc. Sixth Conf. Artificial Intelligence in Medicine Europe* (Springer Verlag, Berlin, 1997) 157–168,

[25] G.M. Provan and J.R. Clarke, Dynamic network construction and updating techniques for the diagnosis of acute abdominal pain, *IEEE Trans. Pattern Anal. Mach. Intell. 15(3)* (1993) 299–307.

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

[27] S. Quaglini, C. Berzuini, R. Bellazzi, M. Stefanelli, and G. Barosi, Therapy planning by combining AI and decision theoretic techniques, in: D. Hunter, ed., *Proc. Second Conf. Artificial Intelligence in Medicine* (1989) 125–134.

[28] N.E. Reed, M. Gini, P.E. Johnson, and J.H. Moller, Diagnosing congenital heart defects using the Fallot computational model, *Artif. Intell. Med. 10* (1997) 25–40.

[29] R.D. Shachter, Evaluating influence diagrams, *Oper. Res. 34(6)* (1986) 79–90.

[30] D.J. Spiegelhalter, A.P. Dawid, S.L. Lauritzen, and R.G. Cowell, Bayesian analysis in expert systems, *Stat. Sci. 8(3)* (1993) 219–283.

[31] J.A. Tatman and R.D. Shachter, Dynamic programming and influence diagrams, *IEEE Trans. Syst. Man Cybern. 20(2)* (1990) 365–379.

[32] W.B. Thompson, P.E. Johnson, and J.B. Moen, Recognition-based diagnostic reasoning, in: *Proc. Eighth Int. Joint Conf. on Artificial Intelligence (IJCAI–83)* (1983) 236–238.