Artificial Intelligence in Medicine (2009) 47, 159—171

ARTIFICIAL INTELLIGENCE IN MEDICINE

http://www.intl.elsevierhealth.com/journals/aiim

Markov decision process applied to the control of hospital elective admissions

Luiz Guilherme Nadal Nunes [a,*], Solon Venâncio de Carvalho [b], Rita de Cássia Meneses Rodrigues [b]

[a] *Sarah Network of Rehabilitation Hospitals, SMHS Quadra 501 Conjunto A, Brasília, DF 70330-150, Brazil*
[b] *Brazilian National Institute for Space Research, Av. dos Astronautas, 1758, Jd. Granja, São José dos Campos, SP 12227-010, Brazil*

Received 24 April 2008; received in revised form 27 May 2009; accepted 28 July 2009

KEYWORDS
Markov decision processes;
Hospital admission control;
Patient flow modeling

Summary

Objective: To present a decision model for elective (non-emergency) patient admissions control for distinct specialties on a periodic basis. The purpose of controlling patient admissions is to promote a more efficient utilization of hospital resources, thereby preventing idleness or excessive use of these resources, while considering their relative importance.
Methods: The patient admission control is modeled as a Markov decision process. A hypothetical prototype is implemented, applying the value iteration algorithm.
Results: The model is able to generate an optimal admission control policy that maintains resource consumption close to the desired levels of utilization, while optimizing the established deviation costs.
Conclusion: This is a complex model due to its stochastic dynamic and dimensionality. The model has great potential for application, and requires the development of customized solution methods.

1. Introduction

There are several aspects involved in the need for controlling patient admissions. Admissions may be scheduled to satisfy different and sometimes contradictory goals, such as sustaining a high utilization of available hospital capacity (possibly resulting in some bottlenecks); or smooth throughput for a minimum length of patient stay (possibly resulting in some idleness). As mentioned by Kusters and Groot [1], controlling patients' admissions is the key activity which allows a hospital to balance the demand for patient facilities against the availability of these resources. Choosing the "right" patients from the

* Corresponding author at: Rede Sarah de Hospitais de Reabilitação, Centro Nacional de Controle de Qualidade, SMHS Quadra 501 Conjunto A, Brasília, DF 70330-150, Brazil.
Tel.: +55 61 3319 1743/9146 3509; fax: +55 61 3226 5280.
E-mail address: guilherme@sarah.br (L.G.N. Nunes).

0933-3657/$ — see front matter
doi:10.1016/j.artmed.2009.07.003

waiting list in order to reach this balance is not a simple task. It is unlikely that a single model capable of generating an optimal solution and dealing with the complexity of this problem exists. Since the final decision shall take into account a variety of factors, it can only be made by a human decision maker. However, it is possible to provide this decision maker with support. In this study we described and tested a model aimed at providing precisely this type of support.

Unlike Adan and Vissers [2], we do not consider the admission process as a daily occurrence. Rather, we wish to plan the number of admissions within fixed periods of time (e.g., one week or fifteen days). While controlling the number of patients admitted during these fixed periods of time, we aim at sustaining stabilized, desirable (target) levels of the average use of key hospital resources (such as exams, follow-up consultations, in-patient admissions, and surgical interventions), preventing idleness or excessive use of these resources. In other words, we developed a decision model that helps hospital managers plan at the strategic and tactical level, rather than at operational level. At the strategic level (input), top hospital managers must set the desirable levels of use, as well as the relative importance of each hospital resource. At the tactical level (output), the decision model sets, at the beginning of each planning period, an admission policy that, depending on the number of patients being served, their specialties and the pattern of resource consumption, informs the number of admissions to be achieved for each specialty during the next period. At the operational level, if the recommended number of admissions is always adhered to throughout each period, over the long-run, the hospital will be stabilized in relation to the target level for resource use and the relative importance of each resource.

According to Adan and Vissers [2], patients' admissions to a hospital can be divided into two types: non-scheduled and scheduled. Non-scheduled admissions, also called emergency admissions, concern patients that are immediately admitted as a consequence of medical conditions. Scheduled admissions, also called elective admissions, are selected from a waiting list for an admission date. In this study we focused on scheduled patient admissions. In particular, we analyzed the procedures for elective patient admission in a hospital that does not provide emergency care. Tertiary hospitals, such as rehabilitation services, fit this type of elective care. However, the model developed in this paper can also be applied, with modifications, to hospitals that provide emergency care.

Four studies represented an essential contribution to the development of our model, as follows: 1) a survey on theoretical models of admission planning in Gemmel and Van Dierdonck [3], which included, in detail, the studies of Groot [4], and Roth and Van Dierdonck [5]; 2) the original idea of *treatment patterns* combined with Markov models in order to estimate hospital resource utilization, from Kapadia et al. [6,7]; 3) Adan and Vissers [2]'s interesting cost function aimed at stabilizing hospital resource utilization at desirable levels—the authors translated the admission planning problem into a mathematical model in the form of a linear integer program; and 4) Markov processes in Puterman [8]—Markov models suit the stochastic characteristics behind patients' dynamic throughout hospital services.

There is rather extensive literature on Markov models applied to describe the stochastic dynamics of patients. We included a brief summary of the referred studies that, despite their distinct focus, applied the Markov theory. Smallwood et al. [9] and Kao [10] formed groups of patients admitted to the same specialty and with similar arrival rates at the hospital; next, a distinct and independent semi-Markov model was processed for each one of these groups. The authors developed the necessary formulation to estimate performance parameters, such as the average number of patients in a given department and the expected resource utilization. Hershey et al. [11] and Côté and Stein [12] also modeled semi-Markovian processes, but they considered the flow of one group of patients. The study of Hershey et al. [11] was focused on modeling a hospital as being formed by capacitated facility units, representing transitory states in a semi-Markov process. Côté and Stein [12] went further and introduced the Erlang distribution for governing the transition probabilities among states. Navarro [13] applied proportion calculations in order to estimate the transition probabilities among hospital facilities represented by discrete time Markov chains. Through this Markov chain, he obtained performance parameters, taking the hospital as a closed system comprised of recurrent states and a unique group of patients. Hincapié et al. [14] performed a longitudinal follow-up of groups of patients' arrivals at the hospital, and estimated the transition probabilities among hospital facilities; then, the authors developed a discrete time Markov chain to obtain performance parameters. Weiss et al. [15] presented a model similar to that of Smallwood et al. [9] and Kao [10], and proposed an iterative methodology for testing the validity of Markovian characteristic assumptions.

In the present study, we propose a new approach, modeling the control of patients' admissions as a

Markov decision process (MDP). MDPs have proven to be useful as models for sequential decision problems with stochastic characteristics and the Markovian property, i.e., future states and decisions are independent from past states and decisions, given the knowledge of the present situation of the system. At each decision instant in an MDP, the system's state is observed and one action is adopted. Based on this information [observed state; adopted action] we can calculate the probabilities of reaching any possible system state in the next decision instant, as well as the expected cost to be incurred until the next decision instant.

The decision process involved in admission planning is repeated over and over during hospital activities. At the tactical level, the decisions concerning the number of patients to be admitted can be made in equally spaced periods (e.g., one week or fifteen days). We assume that it is always possible to count the number of patients during the latest period and classify them into distinct treatment patterns. Kapadia et al. [6,7] showed us that it is possible to estimate the probabilities of admission of patients to a hospital in any treatment pattern, and that it is also possible, at the beginning of a period, to determine the transition probabilities among the treatment patterns for patients being served. This data allows us to estimate the consumption of each given resource throughout the next period. In this way we defined the elements for the stochastic dynamics related to the MDP.

The motivation for modeling the elective hospital admission system using an MDP was the identification of system characteristics geared towards the concepts of MDPs, namely: (1) a sequential decision process that operates under uncertainty, generating stochastic dynamics; (2) a possibility of observing the system's state at decision instants equally spaced over time (discrete time decisions); and finally, (3) Markovian characteristics can be assumed and modeled (the future is dependent only on the present state and the decision taken).

The proposed model is presented within this context in the following sections. Section 2: brief introduction of the MDPs background; Section 3: presentation of the hospital characteristics as they are considered in this study; Section 4: description of MDP elements—state space, action space, probabilistic dynamics, and cost function; Section 5: presentation of the implementation of a small-sized hypothetical model as an example; Section 6: discussion of practical considerations on the model. Our conclusions are presented in Section 7.

2. Markov decision process background

MDPs are deeply and comprehensively explored in Puterman [8]. In this section we briefly present the concepts applied in this paper. MDPs can generally be defined by the *n-tuple* (X, A, P, R), where: X is the set of states; A is the set of actions applicable depending on the states; P represents the transition probabilities among states, depending on an observed state and on an adopted action at a decision instant; and R determines the expected cost related to the observed state and the adopted action. At any decision instant t, we observe the system's state $x_t \in X$ and adopt an action $a_t \in A(x_t)$, which generates the expected cost $R(x_t, a_t)$ and the probability of the system being at any state $x_{t+1} \in X$ in the next decision instant $t+1$, represented by $P(x_{t+1}|(x_t, a_t))$. According to the characteristics of the problem at hand, as we modeled it, in this study we consider an MDP with denumerable finite state space and action space, discrete time evolution, and bounded costs.

A decision policy π is a sequence of functions $\pi = \{\pi_0, \pi_1, \ldots\}$, where $\pi_t: X \to A$ such that $\pi_t(x_t) = a_t \in A(x_t)$. If π_t is invariant with regards to t, then π is called a stationary decision policy. Let Π be the set of all possible decision policies. The problem intended to be solved is: given the initial state x_0 at decision instant $t = 0$, to find the decision policy $\pi^* \in \Pi$ that optimizes the cost function related to the problem.

The optimization criterion adopted here is the minimization of the *expected average cost* over infinite planning horizon. The expected average cost of a policy π, represented by $J_\infty^\pi(x)$, can be calculated through the following expression

$$J_\infty^\pi(x) = \lim_{H\to\infty} \frac{1}{H} E\left\{\sum_{t=0}^{H-1} R(x_t, \pi_t(x_t))|x_0 = x\right\},$$

where H represents the planning horizon, $\pi(x_t)$ is the action prescribed by the stationary policy for state π_t, and the expectation is taken according to the probabilities defined in P. Our objective is to find an optimal policy $\pi^* \in \prod$ that minimizes the expected average cost over all possible policies. As shown by Puterman [8], if X and A are finite, R is bounded, and the model is *unichain* (i.e., consists of a single class of recurrent states and a possibly empty set of transient states), there exists a stationary optimal policy.

Whenever a system modeled as a MDP is controlled by a stationary policy there is a Markov chain embedded in the process. This Markov chain has the state space of the original MDP, and the transition probabilities are given by $P(x_{t+1}|x_t,\ \pi(x_t)),\ \forall\ x_t,\ x_{t+1} \in X$. Let P^π represent the transition probability matrix of the embedded Markov chain, determined by the stationary decision policy π.

The steady state probability for one state is the probability of finding the system in this state after a large number of transitions (long-run). Let π^s represent the vector of steady state probabilities under the stationary decision policy π for all states. Assuming that under control of π the embedded Markov chain is *unichain*, π^s can be obtained by multiplying P^π by itself until it converges to the steady state probabilities. Alternatively, the row vector π^s can be obtained by solving the linear system $\pi^s = \pi^s P^\pi$, $\sum \pi^s = 1$. Through the application of the steady state probabilities we can determine performance parameters, such as the expected number of patients in each specialty and treatment pattern, as well as the expected utilization of each hospital resource.

In order to find an optimal stationary policy, we apply the well-known value iteration algorithm (VIA) [8]. Following, we introduce the main concepts for the VIA application. Let V_x^n represent the minimal expected total cost within the planning horizon of n periods when the initial state is $x \in X$. The VIA calculates recursively the value of V_x^n, $n = 1, 2,\ldots,$ where

$$V_x^n = \min_{a \in A(x)} \left\{ R(x,a) + \sum_{y \in X} P(y|(x,a)) V_y^{n-1} \right\},\ \forall\ x \in X,$$

and $V_x^0,\ \forall\ x \in X$, can be chosen arbitrarily. As n increases, the difference in one period $V_x^n - V_x^{n-1}$ gets closer to the minimum expected average cost.

When $n \to \infty$, the limits $m_n = \min_{x \in X}\{V_x^n - V_x^{n-1}\}$ and $M_n = \max_{x \in X}\{V_x^n - V_x^{n-1}\}$ both approximate the minimum expected average cost. By choosing V_x^0, such as $0 \le V_x^0 \le \min_{a \in A(x)}\{R(x,a)\}$, then $V_x^1 \ge V_x^0,\ \forall\ x \in X$, thus each term into the non-decreasing sequence $\{m_n,\ n \ge 1\}$ is non-negative. It follows, then, that

$$\frac{(M_n - m_n)}{m_n} \le \varepsilon \Rightarrow 0 \le \frac{(J_\infty^{\pi^n}(x) - J_\infty^{\pi^*}(x))}{J_\infty^{\pi^*}(x)} \le \varepsilon,$$

which means that, when $(M_n - m_n)/m_n \le \varepsilon$, the expected average cost of the policy π^n, set for the horizon n, cannot differ more than ε from the minimum expected average cost of the optimal policy π^*. Setting ε as a parameter works as the desired precision for the VIA approximation to the optimal solution value.

The VIA can be summarized as follows:

(1) Select $0 \le V_x^0 \le \min_{a \in A(x)}\{R(x,a)\}, \forall\, x \in X$; specify $\varepsilon > 0$; and set $n = 1$.
(2) Compute $V_x^n,\ \forall\ x \in X$, by $V_x^n = \min_{a \in A(x)} \left\{ R(x,a) + \sum_{y \in X} P(y|(x,a)) V_y^{n-1} \right\}$.
(3) Compute $m_n = \min_{x \in X}\{V_x^n - V_x^{n-1}\}$ and $M_n = \max_{x \in X}\{V_x^n - V_x^{n-1}\}$.
(4) If $(M_n - m_n)/m_n \le \varepsilon$, go to step 5; otherwise increment n by 1 and return to step 2.
(5) Set $\pi^n(x)$ as $\pi^n(x) = \arg max_{a \in A(x)}\{R(x,a) + \sum_{y \in X} P(y|(x,a)) V_y^{n-1}\},\ \forall\ x \in X$, and stop.

The VIA finds a stationary ε—*optimal* policy and an approximation to its average expected cost. With a sufficiently small ε, the policy π^n converges to an optimal policy.

3. Hospital elements

When designing the model for patients' admissions control, the hospital elements must be defined in a systematic framework. Here, we present the definitions for the three basic elements considered in the model: patients' demand, treatment patterns, and hospital resources.

3.1. Patients' demand characteristics

This is a model of an elective admission system, that is, no emergency cases are included. After registering for admission, patients are put on a waiting list to be admitted and begin their treatment program. The end of treatment date is the discharge date.

For the purposes of this study, the hospital is considered as having admissions in m medical specialties. We represent the specialties by index d, thus $d \in \{1,\ldots,m\}$. We also assume that there is a continuous flow of patients in every specialty waiting for treatment.

3.2. Treatment patterns

In order to attain the model's objective, that is, to stabilize resources utilization over the long-term, a methodology for measuring resources consumption is necessary. The approach proposed by Kapadia et al. [7] allows us to estimate hospital resources utilization. Kapadia et al. [7] described the patients' *courses of treatment* in the hospital through *treatment patterns*. Course of treatment bears some resemblance to the *path* of patients through the hospital, but instead of location change [16,17], state of health [10,18], or recovery state [19], the course of treatment uses services delivered to patients as the basic data to classify patients' progression during predefined treatment periods. Course of treatment comprises treatment (delivered services) patterns. A treatment pattern is a quantified configuration of services delivered during a predefined fixed period of time. This period of time must be chosen in accordance with the planning purposes; for example, fifteen days, one week, or even one day.

As shown by Kapadia [7,8], it is possible to reach a discrete number of characteristic treatment patterns for a hospital by selecting a sufficiently large and unbiased sample of patients, dividing the treatment time into constant period intervals, and applying cluster analysis to the periodic services configurations.

For the proposed model, considering one period interval, we assume there are n possible treatment patterns, E_i, $i = 1,\ldots,n$. Thus, for example, the course of treatment for one patient who spent a total of eight periods at the hospital can be described as a sequence of treatment patterns such as $E_2E_1E_1E_4E_2E_2E_3E_5E_n$. In this sequence, the patient's initial period in hospital was classified as treatment pattern E_2 (period 1), then he/she has spent periods 2 and 3 in treatment pattern E_1, period 4 in treatment pattern E_4, periods 5 and 6 in treatment pattern E_2, period 7 and 8 in treatment patterns E_3 and E_5, before discharge. We define the pattern E_n as representing discharge from hospital. This pattern must appear at the end of all treatment courses.

3.3. Hospital resources

We are concerned with the stabilization of the average resources utilization, while exerting admissions control on strategic/tactical levels over the long-term. Thus we consider average utilization of resources in fixed periods of time instead of exact amounts.

The treatment pattern E_i determines expected average amounts of hospital resources use in one period; for example, the average number of medical consultations, in-patient days, and magnetic resonance exams.

We consider k hospital resources denoted by L_j, $j \in \{1,\ldots,k\}$. We define L_{ij}, $i \in \{1,\ldots,n\}$ and $j \in \{1,\ldots,k\}$, as the expected average amount of resources of type L_j required in one period by one patient under a treatment pattern E_i. We also define $\max L_j$ as the available capacity of resource L_j in one period.

We assume that if a resource L_j is requested beyond its available capacity $\max L_j$ the hospital provides it, though at a higher cost. This elevated cost is imputed in the cost function (see Subsection 4.5).

4. Markov decision process elements

Considering the definition in the last section, the corresponding MDP elements are presented in this section.

4.1. State space

At every decision instant t, we assume that the hospital can be observed in a state x_t,

$$x_t = \{E^1_{1,t}, \ldots, E^1_{n,t}, \ldots, E^m_{1,t}, \ldots, E^m_{n,t}\},$$

where $E^d_{i,t}$ represents the number of patients from specialty d, $d \in \{1,\ldots,m\}$, under treatment pattern E_i, $i \in \{1,\ldots,n\}$, during the last period, between decision instants $t-1$ and t. Thus, the state space X comprises all possible states of type x_t.

4.2. Action space

At decision instants equally spaced over time, the number of patients in each specialty that will be admitted in the next period must be determined. Therefore, at each decision instant a type a action must be selected and adopted, such as

$$a = \{S^1, \ldots, S^m\},$$

where S^d represents the number of patients from specialty d, $d \in \{1,\ldots,m\}$, to be admitted in the next period. We consider that S^d is limited, ranging from 0 to $\max S^d$, where $\max S^d$ is the maximum admission capacity for patients from specialty d in one period. Thus, the action space A is defined as a finite set consisting of all possible type a actions.

4.3. Stochastic dynamics

We assume that each specialty during the course of treatment has distinct and independent stochastic dynamics, in relation to the transitions among treatment patterns.

The analysis of a sufficient number of treatment courses, counting the transitions from one treatment pattern to another, for each specialty, as done by Kapadia et al. [7,8], allows us to find the maximum likelihood estimation for the *one patient's transition probabilities among the treatment patterns*. The one patient's transition probabilities matrix for patients from specialty d, $d \in \{1,\ldots,m\}$, can be presented as

$$\begin{bmatrix} p^d_{11} & p^d_{12} & p^d_{13} & \cdots & p^d_{1n} \\ p^d_{21} & p^d_{22} & p^d_{23} & \cdots & p^d_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ p^d_{(n-1)1} & p^d_{(n-1)2} & p^d_{(n-1)3} & \cdots & p^d_{(n-1)n} \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix},$$

where p^d_{ij} represents the probability that one patient of the specialty d passes from treatment pattern E_i to treatment pattern E_j in one period. The one patient's transition probability matrix is a Markov chain where the states represent the treatment patterns, and there is one single absorbing state E_n representing the "hospital discharge".

Besides the transition probability matrixes, it is possible to determine, for every specialty, the probability of one patient being admitted to hospital in each one of the treatment patterns. The probability that one patient from specialty d spends his/her first period at hospital in treatment pattern E_i is represented by p^d_i. In order to take into account at least one medical appointment, we assume that the patients could not be directly classified into the discharge pattern (E_n) in the same period of admission, which means that $p^d_n = 0$ for $d = 1,\ldots,m$.

Prior to setting the transition probabilities among states in X we must define $P(\{E^d_{1,t+1},\ldots,E^d_{n,t+1}\}|(x_t,a))$, the probability of reaching the specified amount of patients from specialty d in each treatment pattern at the decision instant $t+1$, given that at decision instant t the hospital is in the state x_t and the action a is adopted. This probability can be calculated by adding up the probabilities of all possible combinations to reach $\{E^d_{1,t+1},\ldots,E^d_{n,t+1}\}$, starting from ($x_t$, a); and each combination probability can be obtained as a convolution of Multinomial distributions. The probability formulation is represented as follows:

$$P(\{E^d_{1,t+1},\ldots,E^d_{n,t+1}\}|(x_t,a)) =$$

$$\sum_{y_1=\max\left(0;E^d_{1,t+1}-\sum_{i=1}^{n-1}E^d_{i,t}\right)}^{\min(E^d_{1,t+1};S^d)} \quad \sum_{x_{11}=\max\left(0;E^d_{1,t+1}-\left(\sum_{i=2}^{n-1}E^d_{i,t}+y_1\right)\right)}^{\min(E^d_{1,t+1}-y_1;E^d_{1,t})} \cdots \sum_{x_{(n-2)1}=\max\left(0;E^d_{1,t+1}-\left(E^d_{n-1,t}+\sum_{i=1}^{n-3}x_{i1}+y_1\right)\right)}^{\min\left(E^d_{1,t+1}-\left(\sum_{i=1}^{n-3}x_{i1}+y_1\right);E^d_{n-2,t}\right)}$$

$$\sum_{y_2=\max\left(0;E^d_{2,t+1}-\left(\sum_{i=1}^{n-1}(E^d_{i,t}-x_{i1})\right)\right)}^{\min(E^d_{2,t+1};S^d-y_1)} \quad \sum_{x_{12}=\max\left(0;E^d_{2,t+1}-\left(\sum_{i=2}^{n-1}(E^d_{i,t}-x_{i1})+y_2\right)\right)}^{\min(E^d_{2,t+1}-y_2;E^d_{1,t}-x_{11})} \cdots$$

$$\sum_{x_{(n-2)2}=\max\left(0;E^d_{2,t+1}-\left(E^d_{n-1,t}-x_{(n-1)1}+\sum_{i=1}^{n-3}x_{i2}+y_2\right)\right)}^{\min\left(E^d_{2,t+1}-\left(\sum_{i=1}^{n-3}x_{i2}+y_2\right);E^d_{n-2,t}-x_{(n-2)1}\right)} \cdots$$

$$\sum_{y_{n-2}=\max\left(0;E^d_{n-2,t+1}-\left(\sum_{i=1}^{n-1}\left(E^d_{i,t}-\sum_{j=1}^{n-2}x_{ij}\right)\right)\right)}^{\min\left(E^d_{n-2,t+1};S^d-\sum_{j=1}^{n-3}y_j\right)} \quad \sum_{x_{1(n-2)}=\max\left(0;E^d_{n-2,t+1}-\left(\sum_{i=2}^{n-1}\left(E^d_{i,t}-\sum_{j=1}^{n-3}x_{ij}\right)+y_{n-2}\right)\right)}^{\min\left(E^d_{n-2,t+1}-y_{n-2};E^d_{1,t}-\sum_{j=1}^{n-3}x_{1j}\right)} \cdots$$

$$\sum_{x_{(n-2)(n-2)}=\max\left(0;E^d_{n-2,t+1}-\left(E^d_{n-1,t}-\sum_{j=1}^{n-3}x_{(n-1)j}+\sum_{i=1}^{n-3}x_{i(n-2)}+y_{n-2}\right)\right)}^{\min\left(E^d_{n-2,t+1}-\left(\sum_{i=1}^{n-3}x_{i(n-2)}+y_{n-2}\right);E^d_{n-2,t}-\sum_{j=1}^{n-3}x_{(n-2)j}\right)}$$

$$\sum_{x_{1(n-1)}=\max\left(0;E^d_{n-1,t+1}-\left(\sum_{i=2}^{n-1}\left(E^d_{i,t}-\sum_{j=1}^{n-2}x_{ij}\right)+\left(S^d-\sum_{i=1}^{n-2}y_i\right)\right)\right)}^{\min\left(E^d_{n-1,t+1}-\left(S^d-\sum_{i=1}^{n-2}y_i\right);E^d_{1,t}-\sum_{j=1}^{n-2}x_{1j}\right)} \cdots$$

$$\sum_{x_{(n-2)(n-1)}=\max\left(0;E^d_{n-1,t+1}-\left(E^d_{n-1,t}-\sum_{j=1}^{n-2}x_{(n-1)j}+\sum_{i=1}^{n-3}x_{i(n-1)}+\left(S^d-\sum_{i=1}^{n-2}y_i\right)\right)\right)}^{\min\left(E^d_{n-1,t+1}-\left(\sum_{i=1}^{n-3}x_{i(n-1)}+(S^d-\sum_{i=1}^{n-2}y_i)\right);E^d_{n-2,t}-\sum_{j=1}^{n-2}x_{(n-2)j}\right)}$$

$$\sum_{x_{1n}=\max\left(0;E^d_{n,t+1}-\left(\sum_{i=2}^{n-1}\left(E^d_{i,t}-\sum_{j=1}^{n-1}x_{ij}\right)\right)\right)}^{\min\left(E^d_{n,t+1};E^d_{1,t}-\sum_{j=1}^{n-1}x_{1j}\right)} \cdots \sum_{x_{(n-2)n}=\max\left(0;E^d_{n,t+1}-\left(E^d_{n-1,t}+\sum_{i=1}^{n-3}x_{in}\right)\right)}^{\min\left(E^d_{n,t+1}-\left(\sum_{i=1}^{n-3}x_{in}\right);E^d_{n-2,t}-\sum_{j=1}^{n-1}x_{(n-2)j}\right)}$$

$$\left\{\begin{array}{l} \dfrac{S^d!}{y_1!y_2!\ldots y_{n-2}!\left(S^d-\sum_{i=1}^{n-2}y_i\right)!}(p^d_1)^{y_1}(p^d_2)^{y_2}\ldots(p^d_{(n-2)})^{y_{n-2}}(p^d_{(n-1)})^{S^d-\sum_{i=1}^{n-2}y_i} \\ \dfrac{E^d_{1,t}!}{x_{11}!x_{12}!\ldots x_{1n}!}(p^d_{11})^{x_{11}}(p^d_{12})^{x_{12}}\ldots(p^d_{1n})^{x_{1n}} \\ \dfrac{E^d_{2,t}!}{x_{21}!x_{22}!\ldots x_{2n}!}(p^d_{21})^{x_{21}}(p^d_{22})^{x_{22}}\ldots(p^d_{2n})^{x_{2n}}\ldots \\ \dfrac{E^d_{n-1,t}!}{\left(E^d_{1,t+1}-\sum_{i=1}^{n-2}x_{i1}-y_1\right)!\ldots\left(E^d_{n,t+1}-\sum_{i=1}^{n-2}x_{in}\right)!}(p^d_{(n-1)1})^{E^d_{1,t+1}-\sum_{i=1}^{n-2}x_{i1}-y_1}\ldots(p^d_{(n-1)n})^{E^d_{n,t+1}-\sum_{i=1}^{n-2}x_{in}} \end{array}\right\}$$

where y_i represents the number of patients from specialty d admitted during the period between t and $t+1$, classified into treatment pattern E_i at decision instant $t+1$, and x_{ij} represents the number of patients from specialty d that pass from treatment pattern E_i at decision instant t to treatment pattern E_j at decision instant $t+1$. Two characteristics of the above complex formulation were drawn from the model: (1) once we had defined that it is not possible for one patient to be discharged in the same period of his/her admission, we excluded this possibility in the formulation (y_n was excluded); (2) as the treatment pattern E_n represents hospital discharge, and we had defined that there are no transitions from the discharge pattern, we excluded these possibilities (x_{ni} was excluded for $i=1\ldots n$). The formulation encompasses four restrictions:

(1) adding up y_i, $i\in\{1,\ldots,n-1\}$, we have the amount of patients from specialty d to be admitted during the period between t and $t+1$, as determined by action a, which means that $S^d=\sum_{i=1}^{n-1}y_i$;
(2) transitions of patients from specialty d, during the period between t and $t+1$, departing from treatment pattern E_i, must be equal to the amount of patients from specialty d in treatment pattern E_i at decision instant t, which means that $\sum_{j=1}^{n}x_{ij}=E^d_{i,t}$, for $i=1,\ldots,n-1$;
(3) the number of patients from specialty d at decision instant $t+1$ must be equal to the number of patients from specialty d at decision instant t added to the number of patients from specialty d admitted during the period between t and $t+1$, excluding the number of patients in the discharge pattern at decision instant t, which means that $\sum_{i=1}^{n}E^d_{i,t+1}=\sum_{i=1}^{n-1}E^d_{i,t}+S^d$;
(4) any combination that does not meet the above three restrictions has probability zero.

Considering that the stochastic dynamics is independent for each specialty, then $P(x_{t+1}|(x_t, a))$, the probability that the system moves from state $x_t\in X$ to state $x_{t+1}\in X$ given that action a is adopted in the decision instant t, can be obtained by the following product:

$$P(x_{t+1}|(x_t,a))=\prod_{d=1}^{m}P(\{E^d_{1,t+1},\ldots,E^d_{n,t+1}\}|(x_t,a)).$$

4.4. State space and action space limits

We assume that not all actions can be chosen for every state. In states where the expected average utilization in the next period for at least one of the k resources is over its available capacity we do not admit patients. This means that in these states the only action allowed is the one where $S^d=0$ for all $d\in\{1,\ldots,m\}$. Therefore, the possible actions are state-dependant. Formally, given the state $x_t\in X$, at any decision instant t, the set of possible actions $A(x_t)$ is composed by actions of type $a=\{S^1,\ldots,S^m\}$ so that, for all $d\in\{1,\ldots,m\}$,

$$S^d\in\{0,\ldots,\max S^d\} \text{ if } \sum_{i=1}^{n}L_{ij}\sum_{d=1}^{m}\sum_{l=1}^{n}E^d_{l,t}p^d_{li} \le \max L_j \text{ for all } j\in(1,\ldots,k),$$

$$\text{and } S^d=0 \text{ if } \sum_{i=1}^{n}L_{ij}\sum_{d=1}^{m}\sum_{l=1}^{n}E^d_{l,t}p^d_{li} > \max L_j \text{ for at least one } j\in(1,\ldots,k).$$

We can ensure that the state space X, related to these possible actions, is denumerable and finite, considering that: (1) admissions are ceased when at least one of the resources is expected to be overutilized in the next period; (2) the available capacity of all resources is limited; (3) there is an absorbing treatment pattern for all specialties that represents the hospital discharge; (4) we do not add up in the next state ($t + 1$ state) patients that are in the discharge pattern in t; and (5) in the initial state, we impose that $E^d_{i,0}$ is limited for all $d \in \{1,\ldots,m\}$ and $i \in \{1,\ldots,n\}$.

4.5. Cost function

Motivated by Adan and Vissers [2], we developed a cost function considering our model objective: determine the patients' number from each specialty to be admitted during fixed planning periods, aimed at stabilizing average hospital resources utilization at desirable levels, preventing idleness or excessive use while considering the relative importance of the resources.

We define the *target level* as the desired utilization of a resource in each planning period. The level attained should be as close as possible to the target. The target level for each hospital resource L_j, $j \in \{1,\ldots,k\}$, is the quantity designated by N_j, $j \in \{1,\ldots,k\}$.

In order to keep resources utilization close to the target levels we set costs for deviations. When the use of the resource L_j is below the target level by one unit during a period, we set the idleness cost at unit O_j. When the utilization of L_j is above the target level by one unit, we set the excess cost at unit B_j. We also impute a cost for the resource use above its available capacity. When the utilization of L_j is above the available capacity $\max L_j$ by one unit

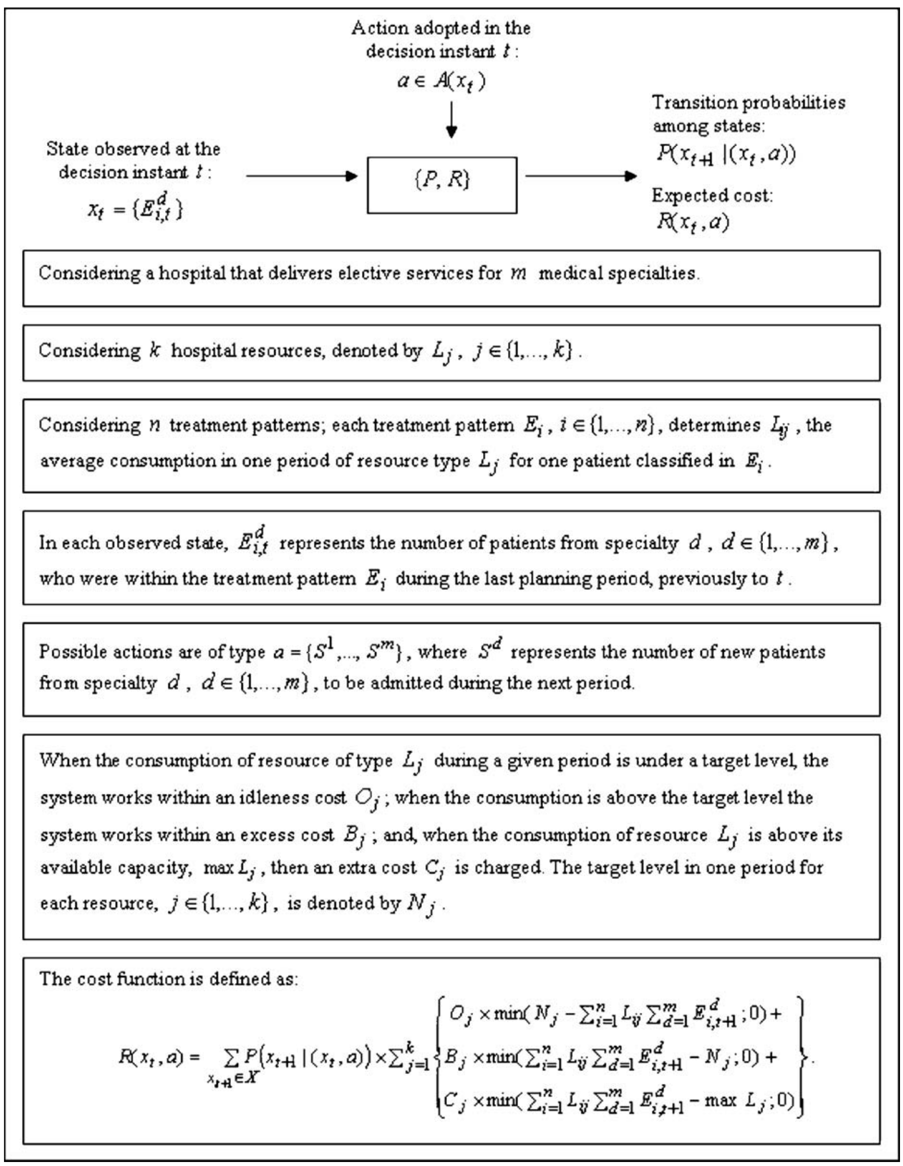

Figure 1 Markov decision process applied to the control of elective patients admissions.

we set the over cost unit C_j. The quantities for the parameters N_j, O_j, B_j, and C_j must be defined on strategic/tactical levels by hospital managers, taking into account the hospital objectives and the relative importance of each resource.

Considering that in the observed state x_t action a is adopted, in order to estimate the system performance during the next period, we define the expected cost function $R(x_t, a)$ as:

$$R(x_t,a) = \sum_{x_{t+1}\in X} P(x_{t+1}|(x_t,a)) \sum_{j=1}^{k} \left\{ \begin{array}{l} O_j \times \max\left(N_j - \sum_{i=1}^{n} L_{ij} \sum_{d=1}^{m} E^d_{i,t+1}; 0\right) + \\ B_j \times \max\left(\sum_{i=1}^{n} L_{ij} \sum_{d=1}^{m} E^d_{i,t+1} - N_j; 0\right) + \\ C_j \times \max\left(\sum_{i=1}^{n} L_{ij} \sum_{d=1}^{m} E^d_{i,t+1} - \max L_j; 0\right) \end{array} \right\},$$

where, as defined in Subsection 3.3, L_{ij} is the average amount of resources type L_j spent in one period by one patient who is under treatment pattern E_i.

4.6. Model summary

The model dynamics is summarized in Fig. 1: there are patients of m different specialties waiting for treatment; and there are k hospital resources whose utilization must be considered. At the beginning of each period, the system state is observed and a decision is made about the number of new patients that must be admitted from each specialty in the next period; from this decision, it is possible to determine the next state probabilities and the expected cost incurred in one period.

5. Model implementation

In this section we present the implementation of a small-sized hypothetical model. We describe the example design (Tables 1 and 2) and show some results to illustrate the correct working of the model (Tables 3 and 4).

Table 1 Model design.

Specialties	2 specialties	
Maximum admission capacity per period		
specialty 1	2 patients	
specialty 2	2 patients	
Resources	2 resources	
Available capacity per period		
L_1	5 units	
L_2	5 units	
Treatment patterns	3 patterns	
Average resource utilization per period		
	L_1	L_2
E_1	2.2 units	2.6 units
E_2	2.6 units	2.2 units
E_3 (discharge)	—	—
Desired utilization level		
L_1 target level	4 units	
L_2 target level	4 units	
Costs for deviation from desired utilization level		
	L_1	L_2
idleness cost	1.0	1.6
excess cost	1.5	1.0
over cost	1.0	1.0

Table 2 Probability configuration.

Transition probability among treatment patterns	E_1	E_2	E_3
specialty 1			
E_1	0.4	0.1	0.5
E_2	0.1	0.3	0.6
E_3	0.0	0.0	1.0
specialty 2			
E_1	0.2	0.1	0.7
E_2	0.1	0.2	0.7
E_3	0.0	0.0	1.0

Entering probability	specialty 1	specialty 2
E_1	0.5	0.4
E_2	0.5	0.6
E_3	0.0	0.0

We apply the value iteration algorithm (see Section 2) in order to obtain an optimal admission policy. We compare the optimal policy with two non-optimal policies: the "greedy" policy, and the "fixed" policy. Given an observed state, the greedy admission policy prescribes an action that minimizes the next period's expected cost without taking into account the possibilities for future periods. The fixed policy admits one patient from each specialty whenever the system is observed in a state that allows new admissions.

Considering the system under control of these three policies, we analyze performance parameters such as the average cost per period, the average number of patients from each specialty in the hospital per period, the average number of patients in each treatment pattern per period, the average utilization of each resource per period, and the average number of patients from each specialty admitted per period. As mentioned in Section 2, the performance parameters are computed through application of the steady state probabilities from the Markov chain embedded in the MDP under control of a stationary policy.

The calculations for the VIA, as well as for the other decision policies (greedy and fixed), were programmed in C language. The solutions were obtained using a common PC, with 1GB of RAM and 2.4 GHz Pentium IV processor.

The example model consists of two specialties, two hospital resources, and three treatment patterns (Table 1). We considered the maximum capacity of admission as being two patients per period for each specialty, and the available capacity of each resource as five units per period. A patient who spends one period in treatment pattern E_1 is expected to utilize in average 2.2 units of resource L_1 and 2.6 units of resource L_2; while in treatment pattern E_2 is expected the utilization of 2.6 unit of resource L_1 and 2.2 units of resource L_2. The treatment pattern E_3 represents the hospital discharge; there is no resource utilization in this pattern. We have established the desired stabilized level of utilization (target level) for each resource as being four units per period, corresponding to 80% of the available capacity.

Concerning the deviation costs from the target level, we have considered dimensionless quantities in order to control the relative importance for each resource. Each unit of deviation smaller than the target level for the resource L_1 has an idleness cost of 1.0; each unit of deviation larger than the target level has an excess cost of 1.5; and each unit of deviation larger than the available capacity has an over cost of 1.0. Considering the resource L_2, the idleness cost is 1.6, the excess cost is 1.0, and the over cost is 1.0.

The probability configuration considered for each specialty is shown in Table 2: the transitions probabilities of one patient among treatment patterns, and the probability of a new patient entering under each treatment pattern.

Table 3 Optimal policy, greedy policy, and fixed policy decisions for five selected states.

Selected states[a]: $\{E_1^1, E_2^1, E_3^1, E_1^2, E_2^2, E_3^2\}$	Number of patients assisted in the last period					Optimal policy decision[b] (S^1, S^2)	Greedy policy decision[b] (S^1, S^2)	Fixed policy decision[b] (S^1, S^2)
	in hospital	per specialty		per pattern				
		1	2	E_1	E_2			
{0,0,1,0,1,3}	1	0	1	0	1	(1,1)	(2,0)	(1,1)
{0,0,2,0,4,0}	4	0	4	0	4	(1,0)	(1,0)	(1,1)
{1,0,4,2,1,1}	4	1	3	3	1	(0,0)	(1,0)	(1,1)
{1,0,2,0,1,2}	2	1	1	1	1	(0,1)	(0,1)	(1,1)
{1,1,2,0,1,0}	3	2	1	1	2	(0,1)	(1,0)	(1,1)

[a] State: observed number of patients from each specialty per treatment pattern, in the last period.
[b] Policy: number of patients to be admitted per specialty, given the observed sate.

Table 4 Performance measures.

	Under the optimal policy control	Under the greedy policy control	Under the fixed policy control
Expected number of admissions per period			
specialty 1	0.28	0.85	0.98
specialty 2	0.99	0.31	0.98
admissions	1.27	1.16	1.95
Expected number of patients being served per period			
specialty 1	0.51	1.55	1.79
specialty 2	1.42	0.44	1.39
pattern E_1	0.87	1.01	1.54
pattern E_2	1.06	0.98	1.64
patients being served	1.93	1.99	3.18
discharge (pattern E_3)	1.27	1.16	1.95
Expected average resource utilization per period			
resource L_1	4.66	4.77	7.65
resource L_2	4.59	4.78	7.61
Expected cost per period (available capacity: $L_1 = 5$, $L_2 = 5$; target level: $L_1 = 4$, $L_2 = 4$)			
idleness cost	0.00	0.00	0.00
excess cost	1.58	1.94	9.09
over cost	0.00	0.00	5.27
total cost per period	1.58	1.94	14.36

The actions for five states selected from the model state space are shown in Table 3 to exemplify the decisions determined by the admission policies (optimal, greedy, and fixed). We also present for each selected state the corresponding number of patients in the hospital per specialty and per treatment pattern. For instance, in the fifth line of Table 3 (shaded line), this observed state indicates that, in the last period, two patients from specialty 1 were discharged, and three patients were being given hospital assistance: two patients in specialty 1 classified as being assisted in treatment patterns E_1 and E_2, and one patient in specialty 2 classified as being assisted in treatment pattern E_2. For this observed state the optimal policy determines the admission of one patient in specialty 2, the greedy policy determines the admission of one patient in specialty 1, and the fixed policy determines, as always, the admission of one patient in each specialty.

Hospital average performance measures over the long-run for the model under control of the optimal policy (OP), the greedy policy (GP), and the fixed policy (FP) are presented in Table 4. The expected number of admissions per period was greater under OP control (1.27 patients) than under GP control (1.16 patients). Under FP control the expected number of admissions was 1.95 patients per period. As the FP prescribes the admission of one patient in each specialty whenever the observed state allows new admissions, the expected number of admissions per period was the same for each specialty (.98 patients). The OP prescribed more admissions of patients in specialty 2 (.99 patients) than the GP (.31 patients), while the GP prescribed more admissions from specialty 1 (.85 patients) than the OP (.28 patients). In the long-run equilibrium, the number of discharges is equal to the number of admissions. In average there were 1.27 discharges per period under control of OP, 1.16 discharges under GP, and 1.95 discharges under FP.

The expected number of patients served per period under OP control (1.93 patients) was almost the same as under the GP control (1.99 patients). However, most patients under OP were treated in specialty 2 (1.42 patients) and were classified into treatment pattern E_2 (1.06 patients), while most patients under GP were treated in specialty 1 (1.55 patients) and were classified into treatment pattern E_1 (1.06 patients). The FP kept more patients being served (3.18 patients) than the OP and the GP, the majority of patients treated in specialty 1 (1.79 patients) and in treatment pattern E_2 (1.64 patients).

Considering the resource utilization and the operational costs, we observe in Table 4 that the average resource utilization per period was closer to the target level under OP control ($L_1 = 4.66$ and $L_2 = 4.59$) than under GP ($L_1 = 4.77$ and $L_2 = 4.78$), and consequently, the OP had lower expected average cost per period (1.58) than the GP (1.94). The cost incurred in both policy, OP and GP, was the excess cost in relation to the target level. The average resource utilization per period under

FP ($L_1 = 7.65$ and $L_2 = 7.61$) exceeded the target level and the available capacity of both resources, and, thus, the FP presented the largest average cost per period (14.36).

6. Discussion

Rather than the number of admission on a daily basis (operational level), we looked for the number of admissions within successive planning periods (strategic/tactical levels). In this context, with the aim of stabilizing average hospital resources utilization at desirable levels, preventing idleness or excessive use while considering the relative importance of the resources, we modeled the control of patients' admissions as a Markov decision process.

Based on the results described in the previous section, we can conclude that the model is able to generate an optimal admission control policy that keeps the resources utilization close to the desired levels while optimizing the established deviation costs. Furthermore, it enables performance parameters analysis of different admission control policies, as we did with the greedy policy and the fixed policy.

It should be noted that the model suits the admission process in elective hospitals, such as rehabilitation hospitals. However, with some modifications, the model developed in this paper can also be applied to emergency care hospitals.

A serious limitation of this model pertains to dimensionality. The small-sized example configuration presented in Section 5 comprises 5765 possible states, which generate, considering the system under control of the optimal policy, 482,504 possible transitions among states with no null probability. It is evident, as far as state and action spaces are concerned, that the model will assume very large dimensions when more realistic systems are evaluated. Fortunately, new approaches that apply concepts of reduced and sampled planning horizons, as well as evolutionary searches, are being developed for solving MDPs with large dimensions. Among them, we can mention the "parallel rollout" method for MDPs with large state space and small action space, and the "evolutionary random policy search" method for MDPs with small state space and large action space [20]. A time aggregation approach, as introduced in [21], can also be considered for solving MDPs with large state space and small action space. It is worth emphasizing that the authors of the present paper are developing a methodology that combines the aforementioned methods for solving MDPs with simultaneously large state and action spaces.

With regards to the model parameters, further studies are required for applying appropriate grouping methods [22] to real data in order to determine the treatment patterns and the related transition probability matrices. We are currently adjusting the model to real data from a rehabilitation hospital to include it in a future project.

7. Conclusion

Controlling the admission of patients into hospitals with limited resources is a traditional and common problem faced by health care systems. Modeling this control as an MDP is a new approach, which may lead to more effective decisions involving the balance between admission of elective patients and utilization of available hospital resources. Combined with an efficient solution method for large dimension MDPs, this method has great potential for application.

References

[1] Kusters R, Groot P. Modelling resource availability in general hospitals design and implementation of a decision support model. European Journal of Operational Research 1996;88: 428—45.
[2] Adan I, Vissers J. Patient mix optimization in hospital admission planning: a case study. International Journal of Operations and Production Management 2002;22:445—61.
[3] Gemmel P, Van Dierdonck R. Admission scheduling in acute care hospitals: does the practice fit with the theory. International Journal of Operations and Production Management 1999;19:863—78.
[4] Groot P. Decision support for admission planning under multiple resource constraints. Enschede: Febo; 1993.
[5] Roth A, Van Dierdonck R. Hospital resource planning: concepts, feasibility, and framework. Production and Operations Management 1995;4:2—29.
[6] Kapadia A, Chan W, Sachdeva R. Predicting duration of stay in a pediatric intensive care unit: a Markovian approach. European Journal of Operational Research 2000;124:353—9.
[7] Kapadia A, Vineberg S, Rossi D. Predicting course of treatment in a rehabilitation hospital: a Markovian model. Computers & Operations Research 1985;12:459—69.
[8] Puterman M. Markov decision processes: discrete stochastic dynamic programming. New York: John Wiley & Sons; 2005.
[9] Smallwood R, Murray G, Silva D, Sondik E, Klainer L. A medical service requirements model for health system design. Proceedings of the IEEE 1969;57:1880—7.
[10] Kao E. A semi-Markovian population model with application to hospital planning. IEEE Transactions on Systems Man and Cybernetics 1973;4:327—36.
[11] Hershey J, Weiss E, Cohen M. A stochastic service network model with application to hospital facilities. Operations Research 1981;29:1—22.
[12] Côté M, Stein W. An Erlang-based stochastic model for patient flow. Omega 2000;28:347—59.
[13] Navarro V. A systems approach to health planning. Health Services Research 1969;4:96—111.

[14] Hincapié D, Ospina J, Grisales H, Arroyave M, Valencia M, González G. Análisis markoviano de un proceso de estancia hospitalaria en un hospital de tercer nivel de complejidad. Revista Facultad Nacional de Salud Pública 2004;22:61—71.
[15] Weiss E, Cohen M, Hershey J. An iterative estimation and validation procedure for specification of semi-Markov models with application to hospital patient flow. Operations Research 1982;30:1082—104.
[16] Fetter R, Thompson J. A decision model for design and operation of a progressive patient care hospital. Medical Care 1969;7:450—62.
[17] Kao E. Modeling the movement of coronary patients within a hospital semi-Markov process. Operations Research 1981;22: 683—99.
[18] Kao E. A semi-Markov model to predict recovery of coronary patients. Health Services Research 1972;7: 191—208.
[19] Thomas W. A model for predicting recovery progress of coronary patients. Health Services Research 1968;3: 185—213.
[20] Chang H, Fu M, Marcus S. Simulation-based algorithms for Markov decision processes. New York: Springer; 2007.
[21] Cao X-R, Bhatnagar S, Fu M, Marcus S. A time aggregation approach to Markov decision processes. Automatica 2002;38:929—43.
[22] Kogan J, Nicholas C, Teboulle M, editors. Grouping multidimensional data: recent advances in clustering. New York: Springer; 2006.