Artificial Intelligence in Medicine 61 (2014) 21–34

Contents lists available at ScienceDirect

Artificial Intelligence in Medicine

journal homepage: www.elsevier.com/locate/aiim

Scalable approximate policies for Markov decision process models of hospital elective admissions

George Zhu, Dan Lizotte, Jesse Hoey*

School of Computer Science, University of Waterloo, 200 University Avenue W., Waterloo, Ontario, Canada N2L 1Z2

ARTICLE INFO

Article history:
Received 10 July 2013
Received in revised form 18 March 2014
Accepted 1 April 2014

Keywords:
Markov decision process
Monte-Carlo tree search
Elective admissions planning
Patient flow modelling

ABSTRACT

Objective: To demonstrate the feasibility of using stochastic simulation methods for the solution of a large-scale Markov decision process model of on-line patient admissions scheduling.
Methods: The problem of admissions scheduling is modeled as a Markov decision process in which the states represent numbers of patients using each of a number of resources. We investigate current state-of-the-art real time planning methods to compute solutions to this Markov decision process. Due to the complexity of the model, traditional model-based planners are limited in scalability since they require an explicit enumeration of the model dynamics. To overcome this challenge, we apply sample-based planners along with efficient simulation techniques that given an initial start state, generate an action on-demand while avoiding portions of the model that are irrelevant to the start state. We also propose a novel variant of a popular sample-based planner that is particularly well suited to the elective admissions problem.
Results: Results show that the stochastic simulation methods allow for the problem size to be scaled by a factor of almost 10 in the action space, and exponentially in the state space. We have demonstrated our approach on a problem with 81 actions, four specialities and four treatment patterns, and shown that we can generate solutions that are near-optimal in about 100 s.
Conclusion: Sample-based planners are a viable alternative to state-based planners for large Markov decision process models of elective admissions scheduling.

1. Introduction

Efficient management of patient admissions plays a critical role in reducing the overall cost of hospitalization at health care facilities. In this paper, we consider *elective patient admissions* for non-emergency care. A typical scenario is admissions for rehabilitation hospitals where patients stay for extended periods and are not expected to require emergency care.

Under this setting, hospital administrators are given increased flexibility to plan an efficient admission policy in order to maximize the amount of available resources at their disposal. Admission schedules can be planned at a tactical level in order to reach strategic targets of resource utilization as well as limit potential bottlenecks during a patient's course of stay.

Determining an optimal admission policy is a challenging problem. Several competing factors must be considered in order to balance the number of patients admitted at any given time with the current state of resource availability. These factors include:

1. Patient flow – a typical patient follows a treatment path during her course of stay that may require different resources (e.g. equipment, nurses, physicians) at various stages along the way.
2. Resource utilization – a resource that is overcapacity blocks the patient's treatment path (e.g. patients must wait until the resource becomes free) as well as limits other resources from being used to their full capacity.
3. Stochasticity – demand for resources may fluctuate at any given time. A patient's course of treatment and resource requirements are typically non-deterministic. For example, a change in the patient's condition may result in a different set of resource requirements than was originally planned.

As the number of resources and patients increases, the complexity of the problem increases exponentially. A systematic framework is needed to assist hospital administrators in making the most efficient admission policy decisions.

Numerous studies have been conducted on the topic of patient scheduling and resource allocation within the operations research and artificial intelligence communities. A popular approach is to model the problem as a Markov decision process (MDP) – a well established model for decision planning under

* Corresponding author. Tel.: +1 5198884567
E-mail address: jhoey@cs.uwaterloo.ca (J. Hoey).

http://dx.doi.org/10.1016/j.artmed.2014.04.001
0933-3657/

uncertainty. For our work, we adopt an MDP model from [1] that leverages similar resource consumption as *treatment patterns* in order to reduce the overall domain space. The model seeks to maximize resource utilization and can be easily extended to handle emergency admissions as well as minimizing patient wait-times.

We investigate several approaches to solving for the optimal or best available admission policy (given limited resources). Traditional MDP solvers that are applied in the research literature usually attempt to find an optimal solution. Although ideally an optimal solution is preferred, it is quite clear that alternative methods are required due to the well-known *curse of dimensionality* issue that arises as the problem size increases. Our goal is to find an efficient real-time solution that may be able to scale up to real-world settings. Our work includes applying established offline and online solvers against a real-time solving approach that has gained increasing popularity from the game playing world called *upper confidence bound applied to trees* (UCT) [2]. UCT is a stochastic simulation method (also known as a "Monte-Carlo" method), many of which have recently been proposed [3]. Our results show that as the problem size increases, the Monte-Carlo type solving approach that does not require an enumeration of state transition probabilities appears to be one viable option.

Our objective in this paper is to demonstrate that, for the established Markov decision process model for patient admissions presented by [1] there are existing stochastic simulation methods that can be used to allow for scalability into domains where optimal MDP solutions are intractable.

The following sections are organized as follows. Section 2 reviews the literature on patient scheduling and on real-time planning. Section 3 gives a background description of the general MDP model and planning approaches. Section 4 gives a formal description of the elective admissions MDP model and Section 5 offers a performance analysis over the various planning algorithms we investigated. We conclude in Section 6 with a discussion of future work.

2. Literature review

2.1. Patient scheduling

In recent years, health care planning and control has received considerable attention as providers face increasing pressure to reduce costs and improve operational efficiency. A large body of work exists in the operations research and management science literature on patient scheduling and resource allocation. Studies typically offer tailored solutions for specific settings such as outpatient clinics [4,5], emergency departments [6,7], bed demand [8], and various types of services (e.g. diagnostic, rehabilitation). Comprehensive surveys of patient scheduling can be found in [5,9–13].

Hans et al. [14] proposed a hierarchical framework that separates health care planning into the strategic, tactical, and operational levels. Strategic planning involves setting static long-term targets for resource capacity and staffing levels while tactical planning implement policies for reaching strategic targets under a dynamic setting. Operational planning is concerned with the day-to-day implementation of a tactical plan, usually involving real-time decisions and short-term planning.

Our work in elective admissions planning falls within the tactical and strategic stages, rather than the operational stage. Work on operational planning is vast, and we do not attempt to cover it in this short review. Some recent highlights include [15], who use a constraint solving system to generate patient-to-bed plans. This work was extended by [16] to non-deterministic (dynamic) domains using a meta-heuristic approach to obtain improved computational speeds. Bennett and Hauser recently proposed a patient-centric approach for clinical decision making to improve health outcomes [17].

As pointed out by [18], previous work on tactical/strategic admissions planning are either myopic (do not consider long time horizons), or, if they are not, fail to provide a scalable solution (as in the case of the model we use [1]). In this paper, we show that the full MDP model proposed in [1] may, in fact, provide a solution for realistically sized instances if a stochastic Monte-Carlo approach is taken. Further, our approach provides a bridge between tactical and operational planning by allowing real-time decision support in tactical planning.

The model we adopt for our work is a finite horizon version of the infinite horizon MDP proposed by Nunes et al. [1]. It is a general framework for multiple specialties/departments and resource types. The objective is to maximize resource utilization in accordance to pre-defined targets on multiple resource types. In the original study, they were able to generate an optimal policy for a test instance on a minimal setting but concluded that alternative planning approaches should be investigated for more realistic settings. One of the key features of the model is the use of *treatment patterns* to group multiple patients with similar demand dynamics. This allows health care providers to use their historical treatment data to potentially make drastic reductions in the model complexity.

The idea of treatment patterns was first proposed by Kapadia et al. [19,20]. Their objective was to model the patient's treatment over the course of an extended period (e.g. days, weeks) and defined a treatment pattern as "a quantified configuration of services delivered a unit of time". Instead of using location changes or health states, treatment patterns represent the amount of resources consumed during each time step to mark a patient's treatment path over the course of his/her stay. Thus, a patient's course of treatment can be described as a sequence of treatment patterns. The study demonstrated that distinct treatment patterns as well as shifts between them can be identified using historical data from a real rehabilitation hospital and cluster analysis.

Other recent MDP-based approaches include [21,22]. In [21], a finite-horizon MDP model was applied for multi-category patient scheduling in a diagnostic hospital. They used data from a real hospital to model patient arrival patterns and used net resource revenue in their objective function. As with [1], they found similar scalability issues with value iteration (VI) as the problem grew in size. Puterman and Queyranne [22] considered multi-priority patients and applied an approximate dynamic programming approach as their planner.

Mathematical programming is an alternative planning model that has been used extensively for general resource allocation problems. This approach was recently investigated in [18] in which they developed a mixed integer linear programming (MILP) framework with the objective of providing equitable distribution of resources and patient access times. A detailed comparison of MDP and MILP approaches would be an interesting avenue for future work.

The main focus of our work in this paper is on applying a Monte-Carlo tree search (MCTS) approach for generating an admission policy in real-time. Recent results from [23] show that this approach may be better suited for MDPs with very large state and action spaces. To our knowledge, there has not been a study that has investigated this approach for elective admissions planning.

2.2. Real time planning

Real time planning has received increasing focus as researchers look for ways to solve planning problems in an online fashion.