Balancing comfort and energy consumption of a heat pump using batch reinforcement learning with fitted Q-iteration

José Vázquez-Canteli\textsuperscript{a}, Jérôme Kämpf\textsuperscript{b}, Zoltán Nagy\textsuperscript{a,*}

\textsuperscript{a} Intelligent Environments Laboratory, Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, Austin, TX, USA
\textsuperscript{b} ENERGY Institute, Haute Ecole d’Ingénierie et d’Architecture de Fribourg (HEIA-FR), Fribourg, Switzerland

Abstract

In this study, a heat pump satisfies the heating and cooling needs of a building, and two water tanks store heat and cold respectively. Reinforcement learning (RL) is a model-free control approach that can learn from the behaviour of the occupants, weather conditions, and the thermal behaviour of the building in order to make near-optimal decisions. In this work we use of a specific RL technique called batch Q-learning, and integrate it into the urban building energy simulator CitySim. The goal of the controller is to reduce the energy consumption while maintaining adequate comfort temperatures.

© 2017 The Authors. Published by Elsevier Ltd.
Peer-review under responsibility of the scientific committee of the CISBAT 2017 International Conference – Future Buildings & Districts – Energy Efficiency from Nano to Urban Scale

Keywords: artificial intelligence, energy management, thermal comfort, building simulation, energy storage

* Corresponding author. Tel.: +1-512-897-0841; E-mail address: nagy@utexas.edu
1. Introduction

Space heating is a large portion of the energy that buildings consume. In the residential sector, buildings account for approximately 30% of the global energy consumption, which grew at an average annual rate of 1.8% a year between 1971 and 2010 [1]. Some advanced control approaches such as Model-Predictive Control (MPC) are often too costly to implement in small residential settings because they require identifying the dynamic model of the system to be controlled. On the other hand, model-free control approaches allow a simpler implementation without the need for a thermal model of the building. Reinforcement learning is a type of model-free controller that is able to adapt to the changing environmental conditions, including not only weather conditions but also human preferences and behavior.

Reinforcement learning has been proved to be an effective algorithm for optimal energy control in low exergy buildings [2]. In this paper, we propose a batch reinforcement learning (BRL) controller with fitted Q-iteration to minimize the energy consumption of a heat pump that supplies heat and cooling to two water tanks. These tanks store and supply the energy to a building. The BRL controller learns from the outdoor temperatures, indoor temperature of the building, and the temperature of the water tanks in order to estimate the best times to provide additional heating or cooling energy. In order to test the suitability of the proposed controller, we implement it in CitySim, a building energy simulator developed at EPFL that computes an estimation of the energy demand for heating, cooling, and lighting of every building [3]. This approach allows evaluating the performance of the controller under different weather scenarios and in buildings of different characteristics.

1.1. Reinforcement learning

Reinforcement learning can be formalized using a Markov Decision Process (MDP). An MDP contains four elements: a set of states \( S \), a set of actions \( A \), a reward function \( R: S \times A \) and transition probabilities between the states \( P: S \times A \times S \in [0,1] \). The policy \( \pi \) then maps states to actions as \( \pi: S \rightarrow A \), and the value function \( V^\pi(s) \) of a state \( s \) is the expected return for the agent when starting in the state \( s \) and following the policy \( \pi \), i.e.

\[
V^\pi(s) = r(s, \pi(s)) + \gamma \sum P(s, \pi(s), s') V^\pi(s')
\]  

(1)

where \( r \) is the reward received for taking the action \( a = \pi(s_k) \), and \( \gamma \in [0,1] \) is a discount factor for future rewards. An agent that uses \( \gamma = 1 \) will give greater importance to seeking long term rewards, whereas an agent using \( \gamma = 0 \) will assign a greater value to states that lead to high immediate rewards. Reinforcement learning is particularly useful when the model dynamics (\( P \) and \( R \)) are not known, and have to be determined or estimated through interaction of the agent with the environment as depicted in Fig. 1 (a). Two approaches can be used to determine the values \( V^\pi \) of every state. In the model-based approach, the rewards and transition probabilities of the model are first learned, and then used to find the values by solving the system of equations represented by eq. (1). In the model-free approach, the agent learns the values associated to every (s, a) pair without explicitly calculating the transition probabilities or the expected rewards [4].
دریافت فوری متن کامل مقاله

امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات