Proceedings of the 2020 Winter Simulation Conference
K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, eds.

OCCUPANT-ADAPTIVE INDOOR ENVIRONMENTAL CONTROLLER USING DQN

Seongkwon Cho
Han Sol Shin
Youngsik Choi
Yun Dam Ko
Cheol-Soo Park

Department of Architecture and Architectural Engineering
College of Engineering, Seoul National University
1 Gwanak-ro, Gwanak-gu
Seoul, 08826, SOUTH KOREA

ABSTRACT

Heating, ventilation, and air conditioning (HVAC) controller should be energy-efficient and responsive to occupant's personal preference. Recently, Deep Reinforcement Learning (DRL) has received increasing attention due to its capability of learning and adapting to building dynamics and occupant behavior. In this study, the authors use deep Q-network to control two air conditioners of a residential building. The agent of DQN is trained to learn thermal preference of virtual occupants and to take appropriate control actions that can achieve energy savings and occupant thermal satisfaction. The simulation results show that the DQN controller adapts itself to personal thermal preference while successfully reducing energy consumption of building systems.

1 INTRODUCTION

An ideal heating, ventilation, and air conditioning (HVAC) controller should be self-learning and then self-adaptive according to occupant's personal thermal preference (Park 2018). In this regard, DRL has advantage of learning dynamic relationships between HVAC systems, occupant behavior, and indoor built environment. In addition, DRL can be used for indoor environmental control without resorting to high-fidelity simulation models compared to Model Predictive Control (Wang 2020). For this purpose, deep Q-network (DQN) is applied in this study to control air conditioners of residential buildings. This study shows that the agent of DQN can learn occupant behavior and determine personalized room air set-point temperature to achieve both energy savings and occupant thermal satisfaction.

2 IMPLEMENTATION OF DQN

An apartment, in Seoul, South Korea, with two air conditioners (denoted by AC in Figure 1) was selected as a target building. A simulation model of the building was realized by EnergyPlus, a building energy simulator developed by US DOE. Occupancy schedule was determined based on a survey data. With regard to the occupant behavior (OB), the following was assumed: If the indoor temperature is greater than 26°C, the air conditioner is turned on with set-point 23°C. If the indoor temperature is less than 23°C, it is turned off. Please note that the box OB [23 26] in Figure 1 represents occupant's personal thermal comfort range.

Artificial agents of deep Q-network (DQN) learn how to map indoor environment to actions that can maximize reward based on previous experience data consisting of four tuples (state, action, reward, next state). The DQN controller can take seven actions (on/off, six setpoint temperatures from 22°C to 27°C). Because the apartment has two air conditioners, the number of possible actions is 49 at each time-step. In other words, the control option space becomes 49^{96} because the artificial agent should make optimal decision at the sampling time of 15 minute per day. In this study, the DQN agent was trained to learn occupant's thermal preference based on feedback data (Figure 1). When a control action does not meet the occupant's thermal preference, it is assumed that the occupant's override is triggered. Please note that a penalty of 100 was added to the reward function to keep the indoor air temperature within the comfort range.

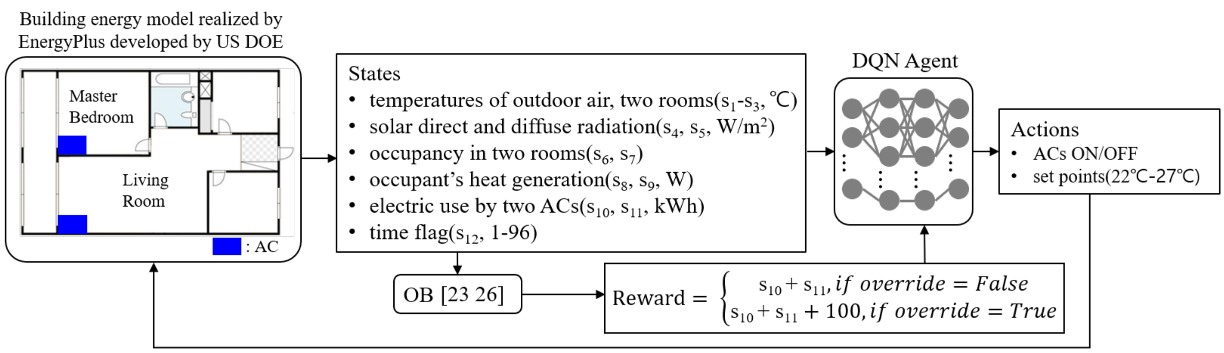

Figure 1: Implementation process of the DQN controller with states, actions, reward, and OB model.

3 RESULTS AND CONCLUSIONS

The DQN is designed with two hidden layers that have 30 nodes each. Simulation was conducted for 10 days in cooling season (July 30–August 8) per episode and 200 episodes were iterated to explore optimal control policy. Temperatures of outdoor air and two rooms for five days (August 4-8) are shown in Figure 2 (left). While occupied, the temperatures of two rooms are controlled between 23°C and 26°C that meet the occupant's thermal preference range. Energy use per episode is shown in Figure 2 (right). After 50 episodes, the energy consumption had reduced to approximately 382 kWh. This signifies that the DQN learns occupant's thermal preference, leading to occupant-adaptive optimal control policy.

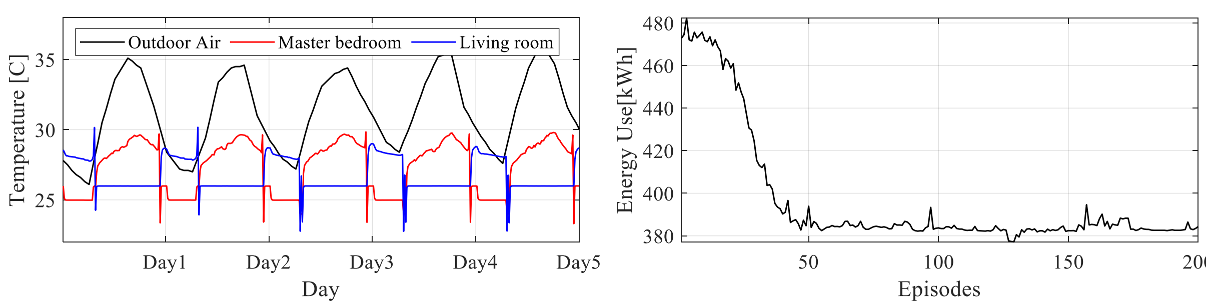

Figure 2: Temperatures of outdoor air and two rooms (left). Energy use of the two ACs per episode (right).

ACKNOWLEDGMENTS

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry & Energy(MOTIE) of the Republic of Korea (No. 20192010107290).

REFERENCES

Park, J.Y., T. Dougherty, H. Fritz, and Z. Nagy. 2018. "LightLearn: An adaptive and occupant centered controller for lighting based on reinforcement learning". *Building and Environment* 147:397-414

Wang, Z., and T. Hong. 2020. "Reinforcement learning for building controls: The opportunities and challenges". *Applied Energy* 269:115036