

Engineers help artificial intelligence to learn more safely in the real world

December 19 2022, by Mary Fetzer



Credit: Pixabay/CC0 Public Domain

Penn State researchers are looking for a safer and more efficient way to use machine learning in the real world. Using a simulated high-rise office building, they developed and tested a new reinforcement learning algorithm aimed at improving energy consumption and occupant comfort in a real-world setting.

Greg Pavlak, assistant professor of architectural engineering at Penn State, presented the results from the paper he co-authored, "Constrained Differentiable Cross-Entropy Method for Safe Model-Based Reinforcement Learning," at the Association for Computing Machinery International Conference on Systems for Energy-Efficient Built Environments (BuildSys) Conference, which was held Nov. 9–10 in Boston.

"Reinforcement learning agents explore their environments to learn optimal actions through trial and error," Pavlak said. "Due to challenges in simulating the complexities of the real world, there is a growing trend to train [reinforcement learning](#) agents directly in the real world instead of in simulation."

However, deploying reinforcement learning in real environments presents its own challenges, according to the researchers.

"Two critical requirements for real-world reinforcement learning are efficient learning and safety considerations," said paper co-author Sam Mottahedi, who was a Penn State doctoral student of architectural engineering when the study was conducted. "Some reinforcement

learning systems require millions of interactions and multiple years to learn the optimal policy, which is not practical in real-world scenarios. Additionally, there is the potential for them to make bad decisions that generate undesirable results or lead to unsafe outcomes."

This concern led the researchers to ask the question: How do we develop algorithms that enable these types of reinforcement learning agents to learn safely in the [real world](#) without making very bad decisions that cause things to break or people to get hurt?

The researchers used an existing model-based reinforcement learning approach to train their model to make decisions. This artificial intelligence agent—the [control algorithm](#)—employs trial and error to interact with the environment, which for their project was a building.

"The safety critical factor of our research was, at a minimum, to not break anything in the building and ensure that occupants are always comfortable," Pavlak said. "While we don't have to worry about someone getting hit by a car, which is a concern for reinforcement learning in [self-driving cars](#), we do have to worry about building equipment operating constraints."

The researchers wanted to minimize [energy use](#) without violating thermal comfort, which ranges from -3, too cold, to +3, too warm. If the control algorithm completed an action that resulted in comfort being outside the -0.5/+0.5 range, it would be penalized. The control algorithm was able to maintain a -0.5/+0.5, which is an acceptable standard in the building industry.

"If the controller is set up to find the best [energy consumption](#), for example, it will be rewarded for achieving this good behavior," Pavlak said. "Alternatively, if it does something that increases energy consumption, it will be penalized for bad behavior. This trial-and-error

approach reinforces learning by gathering information so the controller can decide what to do next."

For this project, the researchers simulated a large office building in a Chicago climate zone. An equipment concern in a real 30-story building might include anything with a large motor, such as the chillers that are used to cool the building.

"Large motors don't like to move quickly," Pavlak said. "For example, a large chiller might be turned on once a day and turned off once a day—a total of two events—to avoid damaging the equipment. If our agent's actions resulted in more than two chiller events in a single day, it would be penalized."

The researchers compared their model-based approach to other common approaches for reinforcement learning, including the use of a model-free algorithm. A model-based agent can plan its action because it's able to predict the reward for it. A model-free agent actually needs to carry out the action to learn from it.

"The model-free algorithm tends to work well but violates some of the safety constraints," Pavlak said. "It also takes a lot more time to learn good behavior, sometimes years or tens of years."

The researchers' model learned about 50 times faster than a traditional model-free method, accomplishing in a month what the other approach needs years to do. And because of the way the researchers incorporated the safety factors, their model had fewer—sometimes zero—violations of the safety critical aspects.

According to Pavlak, adding safety constraints makes reinforcement learning a game of balancing trade-offs. The reinforcement agent could maximize energy consumption, which is a good behavior, by turning the

power completely off. However, doing so would negatively impact occupant comfort, which is bad behavior.

Moving forward, the researchers want to continue to work on the learning speed and reduce overall learning time.

"When a controller starts from scratch, it has to learn everything," Pavlak said. "But once you've trained that controller for one building, you can try it out on a similar [building](#) or reuse parts of it on the next project. Not starting from scratch could potentially lead to faster learning."

The paper is also published as part of the *Proceedings of the 9th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*.

More information: Sam Mottahedi et al, Constrained differentiable cross-entropy method for safe model-based reinforcement learning, *Proceedings of the 9th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation* (2022). [DOI: 10.1145/3563357.3564055](#)

Provided by Pennsylvania State University

Citation: Engineers help artificial intelligence to learn more safely in the real world (2022, December 19) retrieved 2 June 2023 from <https://techxplore.com/news/2022-12-artificial-intelligence-safely-real-world.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.