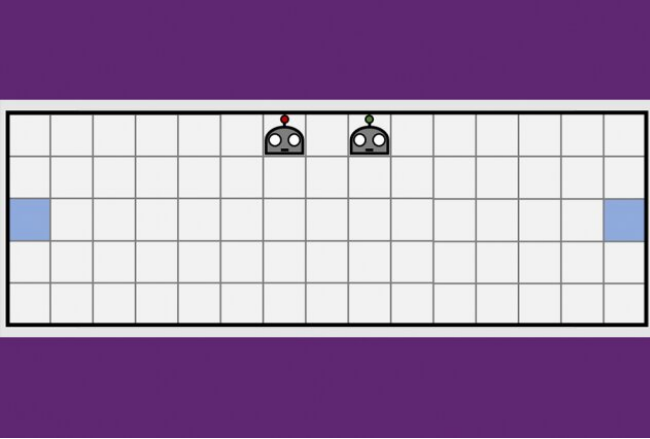


Learning to teach to speed up learning

30 January 2019, by Kim Martineau



Using a new cooperative learning technique, MIT-IBM Watson AI Lab researchers cut by half the time it took a pair of robot agents to learn to maneuver to opposite sides of a virtual room. Credit: Dong-ki Kim

The first artificial intelligence programs to defeat the world's best players at chess and the game Go received at least some instruction by humans, and ultimately, would prove no match for a new generation of AI programs that learn wholly on their own, through trial and error.

A combination of deep learning and reinforcement learning algorithms are responsible for computers achieving dominance at challenging [board games](#) like chess and Go, a growing number of video games, including Ms. Pac-Man, and some card games, including poker. But for all the progress, computers still get stuck the closer a [game](#) resembles real life, with hidden information, multiple players, continuous play, and a mix of short and long-term rewards that make computing the optimal move hopelessly complex.

To get past these hurdles, AI researchers are exploring complementary techniques to help robot agents learn, modeled after the way humans pick up new information not only on our own, but from the people around us, and from newspapers,

books, and other media. A collective-learning strategy developed by the MIT-IBM Watson AI Lab offers a promising new direction. Researchers show that a pair of robot agents can cut the time it takes to learn a simple navigation task by 50 percent or more when the agents learn to leverage each other's growing body of knowledge.

The algorithm teaches the agents when to ask for help, and how to tailor their advice to what has been learned up until that point. The algorithm is unique in that neither agent is an expert; each is free to act as a [student-teacher](#) to request and offer more information. The researchers are presenting their work this week at the AAAI Conference on Artificial Intelligence in Hawaii.

Co-authors on the paper, which received an honorable mention for best student paper at AAAI, are Jonathan How, a professor in MIT's Department of Aeronautics and Astronautics; Shayegan Omidshafiei, a former MIT graduate student now at Alphabet's DeepMind; Dong-ki Kim of MIT; Miao Liu, Gerald Tesauro, Matthew Riemer, and Murray Campbell of IBM; and Christopher Amato of Northeastern University.

"This idea of providing actions to most improve the student's learning, rather than just telling it what to do, is potentially quite powerful," says Matthew E. Taylor, a research director at Borealis AI, the research arm of the Royal Bank of Canada, who was not involved in the research. "While the paper focuses on relatively simple scenarios, I believe the student/teacher framework could be scaled up and useful in multi-player video games like Dota 2, robot soccer, or disaster-recovery scenarios."

For now, the pros still have the edge in Dota2, and other virtual games that favor teamwork and quick, strategic thinking. (Though Alphabet's AI research arm, DeepMind, recently made news after defeating a professional player at the real-time strategy game, Starcraft.) But as machines get better at maneuvering dynamic environments, they may soon be ready for real-world tasks like

managing traffic in a big city or coordinating search-and-rescue teams on the ground and in the air.

"Machines lack the common-sense knowledge we develop as children," says Liu, a former MIT postdoc now at the MIT-IBM lab. "That's why they need to watch millions of video frames, and spend a lot of computation time, learning to play a game well. Even then, they lack efficient ways to transfer their knowledge to the team, or generalize their skills to a new game. If we can train robots to learn from others, and generalize their learning to other tasks, we can start to better coordinate their interactions with each other, and with humans."

The MIT-IBM team's key insight was that a team that divides and conquers to learn a new task—in this case, maneuvering to opposite ends of a room and touching the wall at the same time—will learn faster.

Their teaching algorithm alternates between two phases. In the first, both student and teacher decide with each respective step whether to ask for, or give, advice based on their confidence that the next move, or the advice they are about to give, will bring them closer to their goal. Thus, the student only asks for advice, and the teacher only gives it, when the added information is likely to improve their performance. With each step, the agents update their respective task policies and the process continues until they reach their goal or run out of time.

With each iteration, the algorithm records the student's decisions, the teacher's advice, and their learning progress as measured by the game's final score. In the second phase, a deep reinforcement learning technique uses the previously recorded teaching data to update both advising policies. "With each update the teacher gets better at giving the right advice at the right time," says Kim, a graduate student at MIT.

In a [follow-up paper](#) to be discussed in a workshop at AAIL, the researchers improve on the algorithm's ability to track how well the agents are learning the underlying task—in this case, a box-pushing task—to improve the agents' ability to give and receive advice. It's another step that takes the team closer

to its longer term goal of entering the RoboCup, an annual robotics competition started by academic AI researchers.

"We would need to scale to 11 agents before we can play a game of soccer," says Tesauro, an IBM researcher who developed the first AI program to master the game of backgammon. "It's going to take some more work but we're hopeful."

More information: Shayegan Omidshafiei et al. Learning to Teach in Cooperative Multiagent Reinforcement Learning. arXiv:1805.07830 [cs.MA]. arxiv.org/abs/1805.07830

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology

APA citation: Learning to teach to speed up learning (2019, January 30) retrieved 21 May 2022 from <https://techxplore.com/news/2019-01-learning-to-teach-speed-up.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.