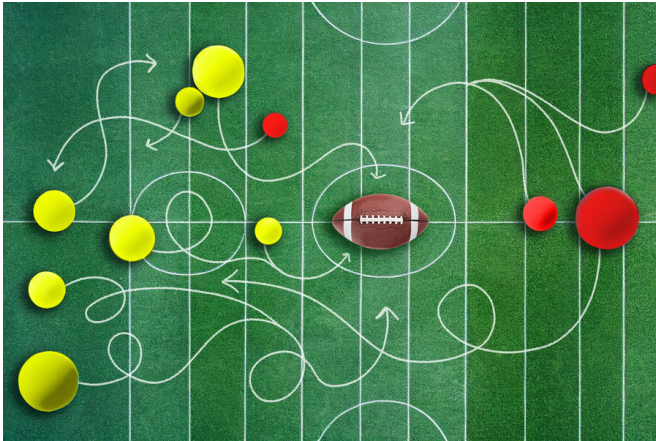


New system can teach a group of cooperative or competitive AI agents to find an optimal long-term solution

23 November 2022, by Adam Zewe



MIT researchers have developed a technique for enabling artificial intelligence agents to think much farther into the future, which can improve the long-term performance of cooperative or competitive AI agents. Credit: Jose-Luis Olivares, MIT, with MidJourney

Picture two teams squaring off on a football field. The players can cooperate to achieve an objective, and compete against other players with conflicting interests. That's how the game works.

Creating artificial intelligence [agents](#) that can learn to compete and cooperate as effectively as humans remains a thorny problem. A key challenge is enabling AI agents to anticipate future behaviors of other agents when they are all learning simultaneously.

Because of the complexity of this problem, current approaches tend to be myopic; the agents can only guess the next few moves of their teammates or competitors, which leads to [poor performance](#) in the long run.

Researchers from MIT, the MIT-IBM Watson AI

Lab, and elsewhere have developed a new approach that gives AI agents a farsighted perspective. Their machine-learning framework enables cooperative or competitive AI agents to consider what other agents will do as time approaches infinity, not just over a few next steps. The agents then adapt their behaviors accordingly to influence other agents' future behaviors and arrive at an optimal, long-term solution.

This framework could be used by a group of autonomous drones working together to find a lost hiker in a thick forest, or by self-driving cars that strive to keep passengers safe by anticipating future moves of other vehicles driving on a busy highway.

"When AI agents are cooperating or competing, what matters most is when their behaviors converge at some point in the future. There are a lot of transient behaviors along the way that don't matter very much in the long run. Reaching this converged [behavior](#) is what we really care about, and we now have a mathematical way to enable that," says Dong-Ki Kim, a graduate student in the MIT Laboratory for Information and Decision Systems (LIDS) and lead author of a paper describing this framework.

More agents, more problems

The researchers focused on a problem known as multiagent reinforcement learning. Reinforcement learning is a form of machine learning in which an AI agent learns by trial and error. Researchers give the agent a reward for "good" behaviors that help it achieve a goal. The agent adapts its behavior to maximize that reward until it eventually becomes an expert at a task.

But when many cooperative or competing agents

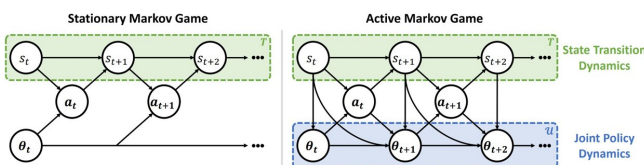
are simultaneously learning, things become increasingly complex. As agents consider more future steps of their fellow agents, and how their own behavior influences others, the problem soon requires far too much computational power to solve efficiently. This is why other approaches only focus on the short term.

"The AIs really want to think about the end of the game, but they don't know when the game will end. They need to think about how to keep adapting their behavior into infinity so they can win at some far time in the future. Our paper essentially proposes a new objective that enables an AI to think about infinity," says Kim.

But since it is impossible to plug infinity into an algorithm, the researchers designed their system so agents focus on a future point where their behavior will converge with that of other agents, known as equilibrium. An equilibrium point determines the long-term performance of agents, and multiple equilibria can exist in a multiagent scenario.

Therefore, an effective agent actively influences the future behaviors of other agents in such a way that they reach a desirable equilibrium from the agent's perspective. If all agents influence each other, they converge to a general concept that the researchers call an "active equilibrium."

The machine-learning framework they developed, known as FURTHER (which stands for FULLY Reinforcing acTive influence with average Reward), enables agents to learn how to adapt their behaviors as they interact with other agents to achieve this active equilibrium.



Within the stationary Markov game setting, agents wrongly assume that other agents will have stationary policies into the future. In contrast, agents in an active Markov game recognize that other agents have non-

stationary policies based on the Markovian update functions Credit: *arXiv* (2022). DOI: 10.48550/arxiv.2203.03535

FURTHER does this using two machine-learning modules. The first, an inference module, enables an agent to guess the [future](#) behaviors of other agents and the learning algorithms they use, based solely on their prior actions.

This information is fed into the reinforcement learning module, which the agent uses to adapt its behavior and influence other agents in a way that maximizes its reward.

"The challenge was thinking about infinity. We had to use a lot of different mathematical tools to enable that, and make some assumptions to get it to work in practice," Kim says.

Winning in the long run

They tested their approach against other multiagent [reinforcement learning](#) frameworks in several different scenarios, including a pair of robots fighting sumo-style and a battle pitting two 25-agent teams against one another. In both instances, the AI agents using FURTHER won the games more often.

Since their approach is decentralized, which means the agents learn to win the games independently, it is also more scalable than other methods that require a central computer to control the agents, Kim explains.

The researchers used games to test their approach, but FURTHER could be used to tackle any kind of multiagent problem. For instance, it could be applied by economists seeking to develop sound policy in situations where many interacting entities have behaviors and interests that change over time.

Economics is one application Kim is particularly excited about studying. He also wants to dig deeper into the concept of an active equilibrium and continue enhancing the FURTHER framework.

The research paper is available on *arXiv*.

More information: Dong-Ki Kim et al, Influencing Long-Term Behavior in Multiagent Reinforcement Learning, *arXiv* (2022). [DOI: 10.48550/arxiv.2203.03535](https://doi.org/10.48550/arxiv.2203.03535)

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: New system can teach a group of cooperative or competitive AI agents to find an optimal long-term solution (2022, November 23) retrieved 1 December 2022 from <https://techxplore.com/news/2022-11-group-cooperative-competitive-ai-agents.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.