

AI AGENTS PLAY HIDE- AND-SEEK

An OpenAI project demonstrated “emergent behavior” by AI players

▶ **After 25 million games, the AI agents playing hide-and-seek with each other had mastered four basic game strategies.** The researchers expected that part.

But after a total of 380 million games, the AI players developed strategies that the researchers didn’t know were possible in the game environment. That was the part that surprised the team at OpenAI, a research company based in San Francisco.

The AI players learned everything via a machine-learning technique known as reinforcement learning. In this learning method, AI agents start out by taking random actions. Sometimes those random actions produce desired results, which earn the AI players rewards. Via trial and error on a massive scale, they can learn sophisticated strategies.

In the context of games, this process can be abetted by having the AI play

against another version of itself, ensuring that the opponents will be evenly matched. It also locks the AI into a process of one-upmanship, where any new strategy that emerges forces the opponent to search for a countermeasure. Over time, this “self-play” amounts to what the researchers call an “auto-curriculum.”

According to OpenAI researcher Igor Mordatch, this experiment shows that self-play “is enough for the agents to learn surprising behaviors on their own—it’s like children playing with each other.”

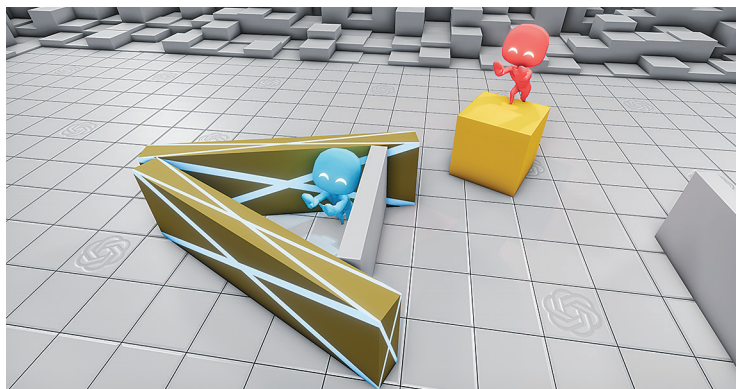
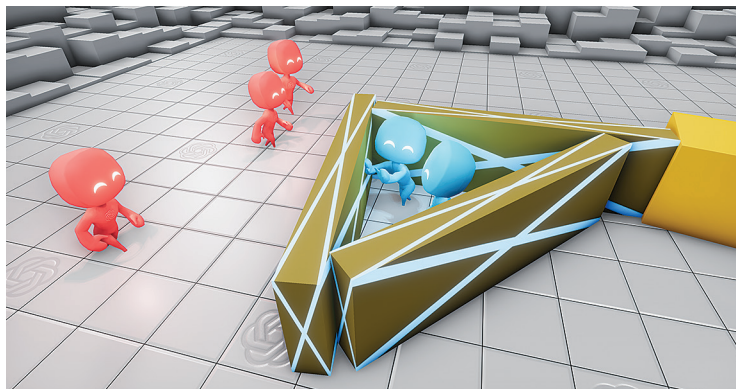
Reinforcement learning is a hot field of AI research right now. The OpenAI team used the technique when they trained bots to play the video game *Dota 2*. That team of bots squashed a world-champion human team last April. The Alphabet subsidiary DeepMind used it to triumph in the ancient board game Go.

Aniruddha Kembhavi, a researcher at the Allen Institute for Artificial Intelligence (AI2) in Seattle, says games such as hide-and-seek offer a good way for AI agents to learn “foundational skills.” He worked on a team that taught an AI to play Pictionary with humans, viewing the gameplay as a way for the AI to work on commonsense reasoning and communication. “We are, however, quite far away from being able to translate these preliminary findings in highly simplified environments into the real world,” says Kembhavi.

In OpenAI’s game of hide-and-seek, both the hidiers and the seekers received a reward only if they won the game, leaving the AI players to develop their own strategies. Within a simple 3D environment containing walls, blocks, and ramps, the players first learned to run around and chase each other. Next, the hidiers learned to move the blocks around to build forts, and then the seekers learned to move the ramps, enabling them to jump inside the forts. Then the hidiers learned to move all the ramps into their forts before the seekers could use them.

The two strategies that surprised the researchers came next. First the seekers learned that they could jump onto a box and “surf” it over to a fort, allowing them to jump in—a maneuver that the researchers hadn’t realized was physically possible in the game environment. So as a final countermeasure, the hidiers learned to lock all the boxes into place so they weren’t available for use as surfboards.

In this circumstance, having AI agents behave in an unexpected way wasn’t a problem—they found different paths to their rewards, but didn’t cause any trouble. However, you can imagine situations in which the outcome would be rather serious. Consider Nick Bostrom’s famous example of a paper clip factory run by an AI, whose goal is to make as many paper clips as possible. As Bostrom told *IEEE Spectrum* back in 2014, the AI might realize that “human bodies consist of atoms, and those atoms could be used to make some very nice paper clips.”



NEW TRICKS: Blue AI agents build a fort [top] to keep opponents out. A red agent uses a box to “surf” into the fort [bottom] during a game developed by OpenAI.

Bowen Baker, another member of the OpenAI research team, notes that it’s hard to predict all the ways an AI agent may act inside an environment—even a simple one. “Building these environments is hard,” he says. “The agents will come up with these unexpected behaviors, which will be a safety problem down the road when you put them in more complex environments.”

AI researcher Katja Hofmann at Microsoft Research Cambridge, in England, has seen a lot of gameplay by AI agents: She started a competition that uses *Minecraft* as the playing field. She says the emergent behavior seen in this game, and in prior experiments by other researchers, shows that games can be a useful tool for studies of safe and responsible AI.

“I find demonstrations like this, in games and gamelike settings, a great

way to explore the capabilities and limitations of existing approaches in a safe environment,” says Hofmann. “Results like these will help us develop a better understanding on how to validate and debug reinforcement learning systems—a crucial step on the path towards real-world applications.”

Baker says there’s also a hopeful takeaway from the surprises in the hide-and-seek experiment. “If you put these agents into a rich enough environment, they will find strategies that we never knew were possible,” he says. “Maybe they can solve problems that we can’t imagine solutions to.”

—ELIZA STRICKLAND

A version of this article appears in our Tech Talk blog.

POST YOUR COMMENTS at <https://spectrum.ieee.org/hideandseek1119>