Feature

Minecraft's open-ended play environment could be ideal for AI research, some researchers say.

AI VERSUS MINECRAFT

A coding contest aims to spur progress in machine-learning techniques. By Jeremy Hsu

To see the divide between the best artificial intelligence and the mental capabilities of a seven-year-old child, look no further than the popular video game *Minecraft*. A young human can learn how to find a rare diamond in the game after watching a 10-minute demonstration on YouTube. Artificial intelligence (AI) is nowhere close. But in a unique computing competition ending this month, researchers hope to shrink the gap between machine and child — and in doing so, help to reduce the computing power needed to train AIs.

Competitors may take up to four days and use no more than eight million steps to train their AIs to find a diamond. That's still a lot longer than it would take a child to learn, but much faster than typical AI models nowadays.

The contest is designed to spur advances in an approach called imitation learning. This contrasts with a popular technique known as reinforcement learning, in which programs try thousands or millions of random actions in a trial-and-error fashion to home in on the best process. Reinforcement learning has helped generate recommendations for Netflix users, created ways to train robotic arms in factories and even bested humans in gaming. But it can require a lot of time and computing power.

Attempts to use reinforcement learning to create algorithms that can safely drive a car or win sophisticated games such as Go have involved hundreds or thousands of computers working in parallel to collectively run hundreds of years' worth of simulations — something only the most deep-pocketed governments and corporations can afford.

Imitation learning can improve the efficiency of the learning process, by mimicking how humans or even other AI algorithms tackle the task. And the coding event, known as the MineRL (pronounced 'mineral') Competition, encourages contestants to use this technique to teach AI to play the game.

Reinforcement-learning techniques wouldn't stand a chance in this competition on their own, says William Guss, a PhD candidate in deep-learning theory at Carnegie Mellon University in Pittsburgh, Pennsylvania, and head of the MineRL Competition's organizing team. Working at random, an AI might succeed only in chopping down a tree or two in the eight-million-step limit of the competition — and that is just one of the prerequisites for creating an iron pickaxe to mine diamonds in the game. "Exploration is really, really difficult," Guss says. "Imitation learning gives you a good prior about your environment."

Guss and his colleagues hope that the contest, which is sponsored by Carnegie Mellon and Microsoft among others, could have an impact beyond locating *Minecraft* gems, by inspiring coders to push the limits of imitation learning. Such research could ultimately

help to train AI so that it can interact better with humans in a wide range of situations, as well as navigate environments that are filled with uncertainty and complexity. "Imitation learning is at the very core of learning and the development of intelligence," says Oriol Vinyals, a research scientist at Google Deep-Mind in London and a member of the MineRL Competition advisory committee. "It allows us to quickly learn a task without the need to figure out the solution that evolution found 'from scratch'."

Gaming by example

The group behind the competition says that *Minecraft* is particularly good as a virtual training ground. Players of the game showcase many intelligent behaviours. In its popular survival mode, they must defend themselves against monsters, forage or farm food and continually gather materials to build structures and craft tools. New players must learn *Minecraft*'s version of physics, as well as discover recipes to transform materials into resources or tools. The game has become famous for the creativity it unleashes in its players, who construct blocky virtual versions of a wide variety of things: the Eiffel Tower, Disneyland, the Death Star trench run from Star Wars, and even a working computer inside the game.

To create training data for the competition, MineRL organizers set up a public *Minecraft* server and recruited people to complete challenges designed to demonstrate specific tasks, such as crafting various tools. They ultimately captured 60 million examples of actions that could be taken in a given situation and approximately 1,000 hours of recorded behaviour to give to the teams. The recordings represent one of the first and largest data sets devoted specifically to imitation-learning research.

The contest focuses on using imitation to 'bootstrap' learning, so that AIs don't need to spend so much time exploring the environment to find out what is possible from first principles, and instead use the knowledge that humans have built up, says Rohin Shah, a PhD candidate in computer science at the University of California, Berkeley, who runs the AI-focused *Alignment Newsletter*. "To my knowledge, there hasn't been another AI competition focused on this question in particular."

Spurred by cloud computing and an ample supply of data, reinforcement learning has typically generated the lion's share of new AI research papers. But interest in imitation learning is picking up, in part because researchers are grappling with the limits of the trial-and-error approach. Learning in that way requires training data that can showcase all possibilities and consequences of different environmental interactions, says Katja Hofmann, principal researcher at the Game Intelligence group at Microsoft Research in

Cambridge, UK, and a member of the MineRL Competition's organizing committee (Microsoft acquired *Minecraft*'s developer for US$2.5 billion in 2014). Such data can be hard to come by in complex, real-world environments, in which it's not easy or safe to play out all the consequences of bad decisions.

Take self-driving cars, for example. Training them mainly through reinforcement learning would require thousands or millions of trials to work out the differences between safe and reckless driving. But driving simulations cannot include all the possible conditions that could lead to a crash in the real world. And allowing a self-driving car to learn by crashing

> ## "I'm particularly interested in *Minecraft* because it's an example of an environment in which humans actually have diverse goals."

repeatedly on public roads would be downright dangerous. Beyond the safety issues, reinforcement learning can get expensive, demanding computing power worth millions of dollars, Hofmann says.

Unlike pure reinforcement learning's from-scratch approach, imitation learning takes short cuts, getting a head start by learning from example. It has already found a home in uses alongside reinforcement learning. Some of the most celebrated AI demonstrations of the past few years, including the AlphaGo algorithm's 2017 trouncing of human Go masters, combined the two approaches, starting with a foundational model generated using imitation learning.

Imitation learning has limitations, too. One is that it is biased toward solutions that have already been demonstrated in the learning examples. AI trained in this way can therefore be inflexible. "If the AI system makes a mistake, or diverges somewhat from what a human would do, then it ends up in a new setting that's different from what it saw in the demonstrations," Shah says. "Since it hasn't seen this situation, it becomes even more confused, and makes more mistakes, which compound further, leading to pretty bad failures."

Still, a number of researchers see great potential in the technique, especially when it comes to training an AI to pursue specific objectives. "The nice part about imitation learning as opposed to reinforcement learning is you get demonstrations of success," says Debadeepta Dey, principal researcher in the Adaptive Systems and Interaction group at Microsoft Research in Redmond, Washington. "This really helps to speed up learning."

To get to the diamond treasure, the AI-controlled players, or agents, in the MineRL contest have to master a multi-step process.

First, they collect wood and iron to make pickaxes. Then they build torches to light the way. They might also carry a bucket of water to quench underground lava flows. Once all that is prepared, an AI can begin exploring mining shafts and caves, as well as tunnelling its way underground to search for diamond ore.

Competitors must train their AIs with a set of hardware consisting of no more than six central-processing cores and one NVIDIA graphics card — something that most research labs can afford through cloud-computing services. More than 900 teams signed up for the competition's first round and 39 ended up submitting AI agents. The ten groups that made the most progress in training AIs to discover diamonds have advanced to the second and final round. Some of those AIs have managed to obtain iron ore and construct a furnace, two other prerequisites for making an iron pickaxe. But Guss doesn't expect any of the teams' agents to find a diamond — at least in this first competition.

Although the contest targets a specific objective, it could spur wider AI research with *Minecraft*. "I'm particularly interested in *Minecraft* because it's an example of an environment in which humans actually have diverse goals — there's no 'one thing' that humans do in *Minecraft*," Shah says. "This makes it a much more appropriate test bed for techniques that attempt to learn human goals."

And even if the game's graphics and rules do not perfectly reflect physical reality, developing more-efficient ways of training AIs in *Minecraft* could translate to speedier AI learning in areas such as robotics. MineRL "could lead to results which would have an impact in real-world domains, such as robotic assembly of complex objects or any other domain where learning complex behaviour is required", says Joni Pajarinen, a research group leader in the Intelligent Autonomous Systems lab at the Technical University of Darmstadt in Germany.

Once the final round of the competition wraps up on 25 November, Guss and other organizers will review the submissions to determine which AI proves the most advanced diamond hunter. The final results will become public on 6 December, just before NeurIPS (the Conference on Neural Information Processing Systems) in Vancouver, Canada, where all ten finalist teams are invited to present their results.

If the MineRL Competition catches on and becomes a recurring tradition, it could provide a public benchmark for tracking progress in imitation learning. "It seems quite likely that MineRL will encourage more research into imitation learning," Shah says. "Whether imitation learning will have significance for real-world applications remains to be seen, but I am optimistic."

Jeremy Hsu is a freelance journalist based in New York City.