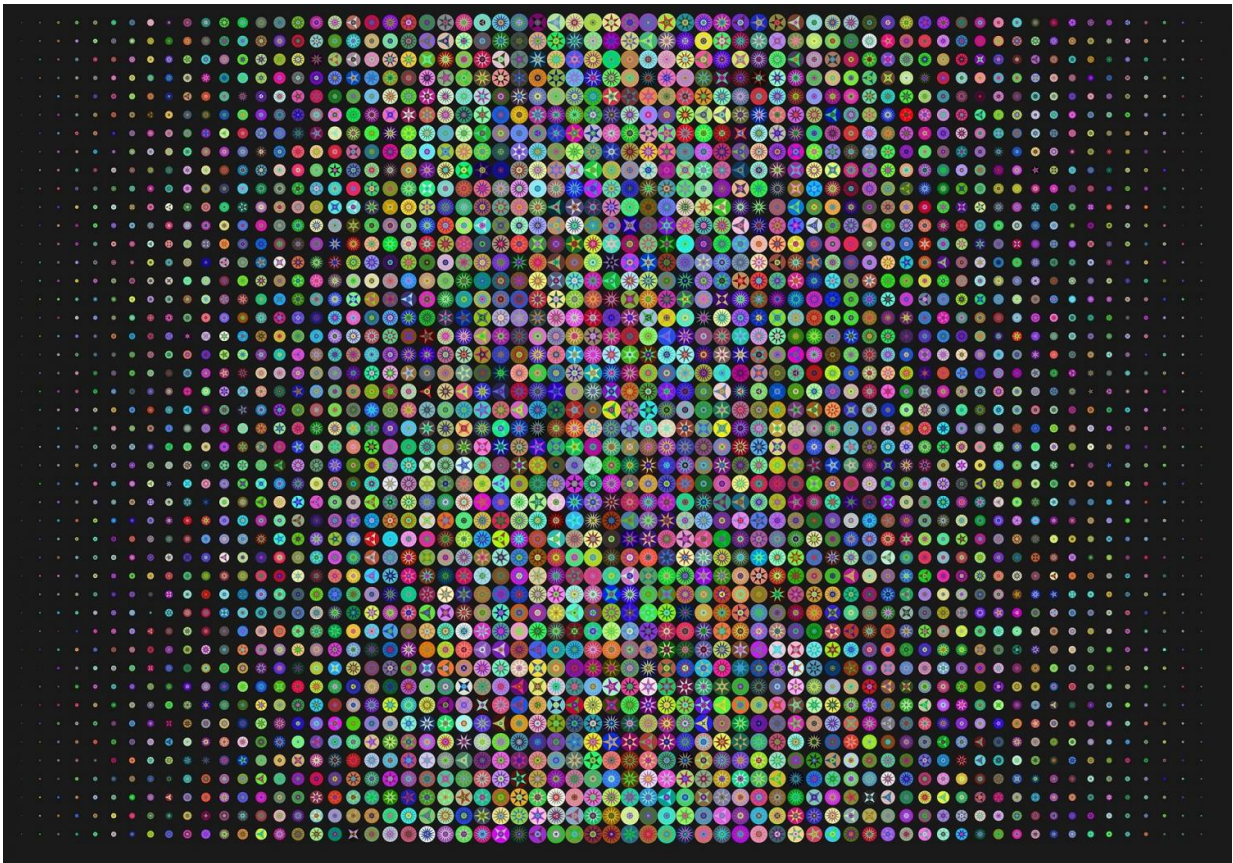# Researchers use video games to unlock new levels of AI

November 5 2018

Credit: CC0 Public Domain

Expectations for artificial intelligences are very real and very high. [An analysis in](#) Forbes projects revenues from A.I. will skyrocket from $1.62 billion in 2018 to $31.2 billion in 2025. The report also included a

survey revealing 84 percent of enterprises believe investing in A.I. will lead to competitive advantages.

"It is exciting to see the tremendous successes and progress made in recent years," says Daniel Jiang, assistant professor of industrial engineering at the University of Pittsburgh Swanson School of Engineering. "To continue this trend, we are looking to develop more sophisticated methods for algorithms to learn strategies for optimal decision making."

Dr. Jiang designs algorithms that learn decision strategies in complex and uncertain environments. By testing algorithms in simulated environments, they can learn from their mistakes while discovering and reinforcing strategies for success. To perfect this process, Dr. Jiang and many researchers in his field require simulations that mirror the real world.

"As industrial engineers, we typically work on problems with an operational focus. For example, transportation, logistics and supply chains, energy systems and health care are several important areas," he says. "All of those problems are high-stakes operations with real-world consequences. They don't make the best environments for trying out experimental technologies, especially when many of our algorithms can be thought of as clever ways of repeated 'trial and error' over all possible actions."

One strategy for preparing advanced A.I. to take on real-world scenarios and complications is to use historical data. For instance, algorithms could run through decades' worth of data to find which decisions were effective and which led to less than optimal results. However, researchers have found it difficult to test algorithms that are designed to learn adaptive behaviors using only data from the past.

Dr. Jiang explains, "Historical data can be a problem because people's actions fix the consequences and don't present alternative possibilities. In other words, it is difficult for an algorithm to ask the question 'how would things be different if I chose door B instead of door A?' In historical data, all we can see are the consequences of door A."

Video games, as an alternative, offer rich testing environments full of complex decision making without the dangers of putting an immature A.I. fully in charge. Unlike the real world, they provide a safe way for an algorithm to learn from its mistakes.

"Video game designers aren't building games with the goal to test models or simulations," Dr. Jiang says. "They're often designing games with a two-fold mission: to create environments that mimic the real world and to challenge players to make difficult decisions. These goals happen to align with what we are looking for as well. Also, games are much faster. In a few hours of real time, we can evaluate the results of hundreds of thousands of gameplay decisions."

To test his algorithm, Dr. Jiang used a genre of video games called Multiplayer Online Battle Arena or MOBA. Games such as League of Legends or Heroes of the Storm are popular MOBAs in which players control one of several "hero" characters and try to destroy opponents' bases while protecting their own.

A successful algorithm for training a gameplay A.I. must overcome several challenges, such as real-time decision making and long decision horizons—a mathematical term for when the consequences of some decisions are not known until much later.

"We designed the algorithm to evaluate 41 pieces of information and then output one of 22 different actions, including movement, attacks and special moves," says Dr. Jiang. "We compared different training

methods against one another. The most successful player used a method called Monte Carlo tree search to generate data, which is then fed into a neural network."

Monte Carlo tree search is a strategy for decision making in which the player moves randomly through a simulation or a video game. The algorithm then analyzes the game results to give more weight to more successful actions. Over time and multiple iterations of the game, the more successful actions persist, and the player becomes better at winning the game.

"Our research also gave some theoretical results to show that Monte Carlo tree search is an effective strategy for training an agent to succeed at making difficult decisions in real-time, even when operating in an uncertain world," Dr. Jiang explains.

Dr. Jiang published his research in a paper co-authored with Emmanuel Ekwedike and Han Liu and presented the results at the 2018 International Conference on Machine Learning in Stockholm, Sweden this past summer.

At the University of Pittsburgh, he continues to work in the area of sequential decision making with Ph.D. students Yijia Wang and Ibrahim El-Shar. The team focuses on problems related to ride-sharing, energy markets, and public health. As industries prepare to put A.I. in charge of critical responsibilities, Dr. Jiang ensures the underlying algorithms stay at the top of their game.

  More information: Feedback-Based Tree Search for Reinforcement Learning, arXiv:1805.05935 [cs.AI] arxiv.org/abs/1805.05935