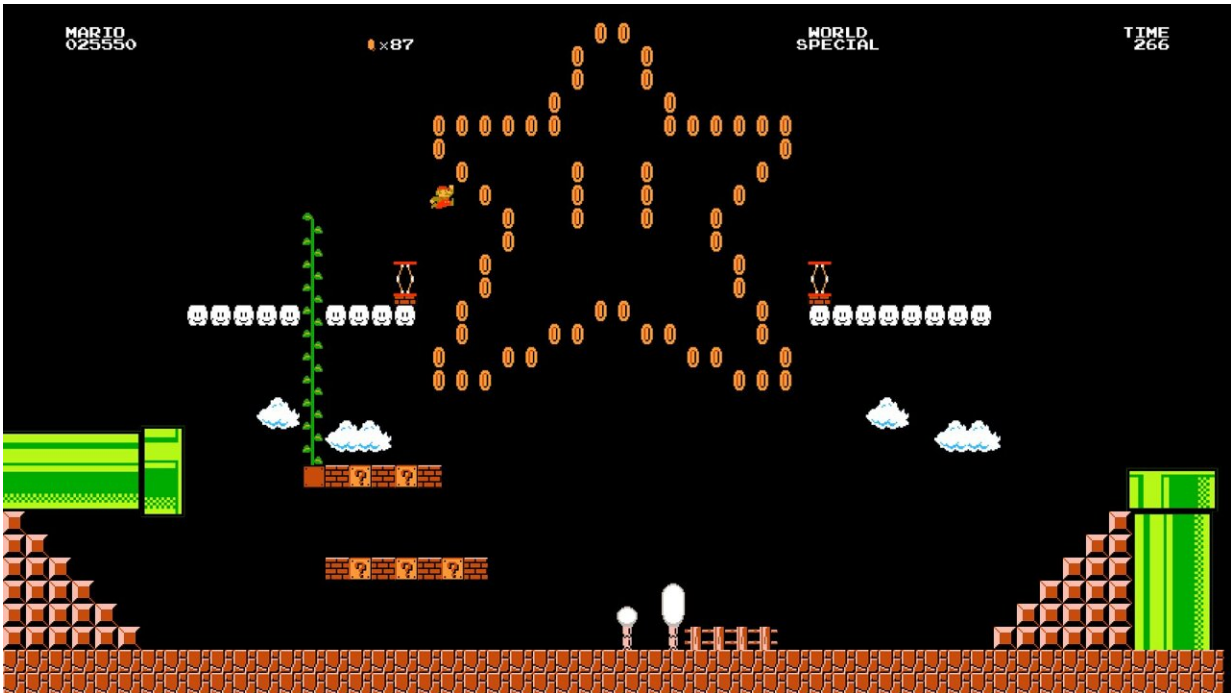


Sometimes it's bad for AI to be too curious

November 10 2022, by Rachel Gordon



Credit: Pixabay/CC0 Public Domain

It's a dilemma as old as time. Friday night has rolled around, and you're trying to pick a restaurant for dinner. (Assuming there's still reservations since you waited until the last minute to book). Anyways, should you go to your most beloved watering hole, or try a new establishment, in the hopes of discovering something superior? Potentially, but that curiosity comes with a risk: you explore, and the food could be worse, or you exploit, and fail to grow out of your narrow pathway.

Curiosity drives AI to explore the world, now in boundless use cases—autonomous navigation, robotic decision making, optimizing [health outcomes](#). Machines, in some cases, use "reinforcement learning" to accomplish a goal, where an AI agent iteratively learns from being rewarded for [good behavior](#) and punished for bad.

Just like the dilemma faced by humans in selecting a restaurant, these agents also struggle with balancing the time spent discovering better actions (exploration) and the time spent taking actions that led to high rewards in the past (exploitation). Too much [curiosity](#) can distract the agent from making good decisions and too little means the agent will never discover good decisions.

In the pursuit of making AI agents with just the right dose of curiosity, researchers from MIT's Improbable AI Laboratory and Computer Science and Artificial Intelligence Laboratory (CSAIL) created an [algorithm](#) that overcomes the problem of AI being too "curious" and getting distracted by the task at hand. Their algorithm automatically increases curiosity when it's needed, and suppresses it if the agent gets enough supervision from the environment to know what to do.

When tested on over sixty video games, the algorithm was able to succeed at both hard and easy exploration tasks, where previous algorithms have only been able to tackle only a hard or easy domain alone. With this method, AI agents use less data for learning decision making rules that maximize incentives.

"If you master the exploration-exploitation trade off well, you can learn the right decision-making rules faster—and anything less will require lots of data, which could mean suboptimal medical treatments, lesser profits for websites, and robots that don't learn to do the right thing," says Pulkit Agrawal, MIT Professor and Director of the Improbable AI Lab, who supervised the research.

"Imagine a website trying to figure out the design or layout of its content that will maximize sales. If one doesn't perform exploration-exploitation well, converging to the right website design or the right website layout will take a long time, which means profit loss. Or in a health care setting, like with COVID-19, there may be a sequence of decisions that need to be made to treat a patient, and if you want to use decision-making algorithms, they need to learn quickly and efficiently—you don't want a suboptimal solution when treating a large number of patients. We hope that this work will apply to real-world problems of that nature."

Curiosity killed the cat

It's hard to encompass the nuances of curiosity's psychological underpinnings—the underlying neural correlates of challenge seeking behavior are a poorly understood phenomena. Attempts to categorize the behavior have spanned studies that have dove deeply into studying our impulses, deprivation sensitivities, and social and stress tolerances.

With reinforcement learning, this process is sort of "pruned" emotionally and stripped down to the bare bones, but it's quite complicated (surprise surprise) on the technical side. Essentially, the agent should only be curious when there's not enough supervision available to try out different things, and if there is supervision, it must adjust curiosity and lower it.

Since a large subset of gaming is little agents running around fantastical environments looking for rewards and performing a long sequence of actions to achieve some goal, it seemed like the logical testbed for the researchers' algorithm. In experiments, with games like Mario Kart and Montezuma's revenge, they divided said games into two different buckets: one where supervision was sparse, meaning the agent had less guidance, which were considered "hard" exploration games, and a second where supervision was more dense, or the "easy" exploration games.

Suppose in Mario Kart, for example, you only remove all rewards so you don't know when an enemy kills you. You're not given any reward when you collect a coin or jump over pipes. The agent is only told in the end how well it did. This would be bucket one with sparse supervision. Algorithms that incentivize curiosity do really well in this scenario.

But now, suppose the agent is provided dense supervision—a reward for jumping over pipes, collecting coins and killing enemies. Here an algorithm without curiosity performs really well because it gets rewarded very often. But instead, if you take the algorithm that also uses curiosity, it learns slowly. It is because the curious agent might attempt to run fast in different ways, dance around, go to every part of the game screen—things which are interesting—but do not help the agent succeed at the game. The team's algorithm, however, consistently performed well, irrespective of what environment it was in.

Future work might involve circling back to the exploration that's delighted and plagued psychologists for years: an appropriate metric for curiosity –no one really knows the right way to mathematically define curiosity.

"Getting consistent good performance on a novel problem is extremely challenging—so by improving exploration algorithms, we can save your effort on tuning an algorithm for your problems of interest. We need curiosity to solve extremely challenging problems, but on some problems it can hurt performance. We propose an algorithm that removes the burden of tuning the balance of exploration and exploitation. Previously what took, for instance, a week to successfully solve the problem. With this new algorithm, we can get satisfactory results in a few hours." says MIT CSAIL Ph.D. student Zhang-Wei Hong, co-lead author along with Eric Chen, MIT CSAIL MEng '22, on a new paper about the work.

"Intrinsic rewards like curiosity are fundamental to guiding agents to

discover useful diverse behaviors, but this shouldn't come at the cost of doing well at the given task. This is an important problem in AI and the paper provides a way to balance that tradeoff. It would be interesting to see how such methods scale beyond games to real world robotic agents," says Deepak Pathak, Faculty at Carnegie Mellon University.

"One of the greatest challenges for current AI and cognitive science is how to balance exploration and exploitation—the search for information versus the search for reward. Children do this seamlessly, but it is challenging computationally," notes Alison Gopnik, Distinguished Professor of Psychology and Affiliate Professor of Philosophy at UC Berkeley, who was not involved with the project.

"This paper uses impressive new techniques to accomplish this automatically, designing an agent that can systematically balance curiosity about the world and the desire for reward, [thus taking] another step towards making AI agents (almost) as smart as children."

More information: Eric R Chen, Zhang-Wei Hong, Joni Pajarinen, Pulkit Agrawal, Redeeming intrinsic rewards via constrained policy optimization. openreview.net/forum?id=36Yz37cEN_Q

Provided by MIT Computer Science & Artificial Intelligence Lab

Citation: Sometimes it's bad for AI to be too curious (2022, November 10) retrieved 4 May 2024 from <https://techxplore.com/news/2022-11-bad-ai-curious.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.
