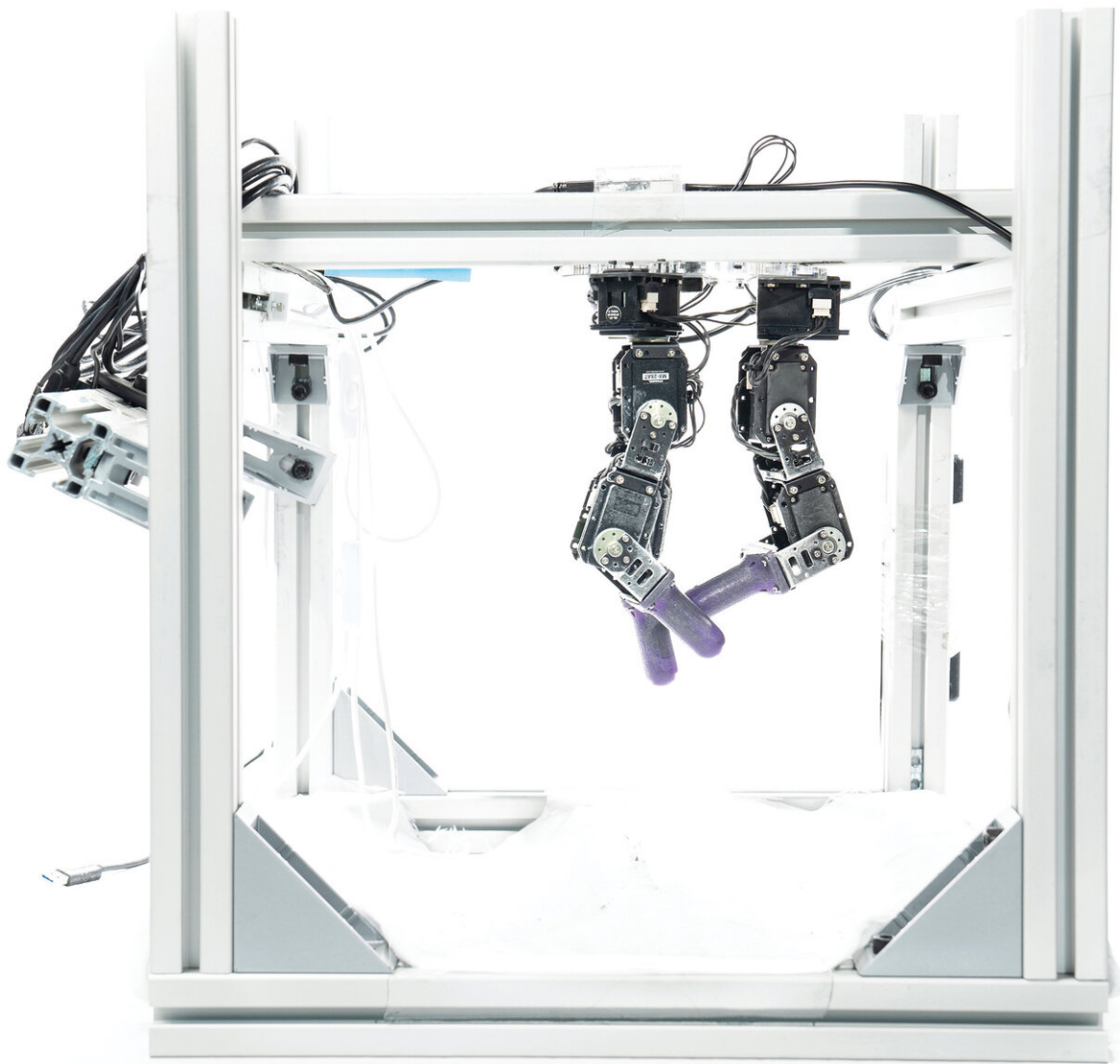


Deep reinforcement learning: Teaching robots like children

April 23 2020, by Thomas Lee



Robot in a frame in EECS professor Sergey Levine's lab. Credit: Adam Lau

When children play with toys, they learn about the world around them—and today's robots aren't all that different. At UC Berkeley's Robot Learning Lab, groups of robots are working to master the same kinds of tasks that kids do: placing wood blocks in the correct slot of a shape-sorting cube, connecting one plastic Lego brick to another, attaching stray parts to a toy airplane.

Yet the real innovation here is not what these robots are accomplishing, but rather how they are doing it, says Pieter Abbeel, professor of electrical engineering and computer sciences and director of the Robot Learning Lab.

Taking inspiration from the way that children instinctively learn and adapt to a wide range of unpredictable environments, Abbeel and assistant professor Sergey Levine are developing algorithms that enable robots to learn from [past experiences](#)—and even from other robots. Based on a principle called [deep reinforcement learning](#), their work is bringing robots past a crucial threshold in demonstrating human-like intelligence, with the ability to independently solve problems and master new tasks in a quicker, more efficient manner.

"If you see a [robot](#) do something through reinforcement learning, it means it actually knows how to acquire a new skill from its own trial and error," Abbeel says. "That's a much more important accomplishment than the specific task it completed."

And while today's most advanced robots still can't match the brain power of a toddler, these researchers are poised to equip robots with cutting-edge artificial intelligence (AI) capabilities, allowing them to generalize between tasks, improvise with objects and manage unexpected challenges in the world around them.

Making "good" decisions

Over the past 80 years, seemingly unrelated innovations in mathematics, [economic theory](#) and AI have converged to push robots tantalizingly close to something approaching human intelligence.

In 1947, mathematician John von Neumann and economist Oskar Morgenstern developed a theorem that formed the basis of something called expected utility theory. In a nutshell, the theory holds that when given a set of choices left to chance, a person will choose the option that produces an outcome with the maximum level of individual satisfaction. Moreover, we can represent that desired outcome, the "reward," with a numeric value.

"That number represents what they want," Abbeel says. "So the theorem shows that having a reward is fully universal. The only thing you need is a number."

Researchers then applied this theory to computers by giving them numerical incentives to learn how to play board games.

Take chess. If the computer's goal is to checkmate its opponent as quickly as it can, that outcome is assigned the highest number in the game. The computer explores which moves to make to achieve checkmate: a "good" move earns the computer a high number while a "bad" move produces a low number.



Abbeel working with BRETT, a deep-learning robot. Credit: Adam Lau

Since choices that represent higher numbers mean the computer will reach its goal more quickly, the computer becomes proficient at chess by systematically learning, through trial and error, to make "good" decisions while avoiding "bad" ones.

Using this reinforcement learning technique, researchers created computers that could defeat human champions in checkers, chess and even Atari video games. In 2017, AlphaGo, an AI program invented by Google, beat the world's best player at Go, an abstract strategy game much more complicated than chess and checkers—cracking a new threshold in AI.

Neural networks

Teaching a computer to win a video game is one thing. Teaching a robot to perform a physical action is much harder.

For one thing, software code exists in the virtual world, which means AI programs enjoy unlimited space to explore and learn. Robots, however, are physical objects operating in physical space. Training a robot to grasp and manipulate objects or navigate spaces without crashing into a filing cabinet requires painstaking and tedious programming work.

Researchers must feed the robot with a vast database of images and train it to recognize patterns so that it can distinguish pictures of chairs from pictures of cats. That way, when a robot rolls into a room, its sensors, or "eyes," can detect an object blocking its path. The robot compares the visual data to similar images in its database before it can successfully conclude that object is indeed a chair.

"Such trial and error takes a long time," Levine says.

But the use of artificial [neural networks](#) has allowed robots to process and analyze information at much faster rates. These networks consist of connected units or nodes that resemble the neurons in human brains. Each node can signal other nodes to connect to it, allowing the robots to establish relationships between different types of data.

Using this approach, Berkeley researchers have been able to do things like teach robots how to run, both in computer simulations and in real life. The robot learns the optimal neural connections it must make to apply the right amount of force to the motors in its arms, hips and legs.

"Through different runs, the robot tries different strengths of connections between the neurons," Abbeel says. "And if one connection

pattern is better than the others, the robot might retain it and do a variation on that connection, and then repeat, repeat, repeat."

The robots understand that certain neural connections earn them rewards, so they continue along that path until achieving the objective, which is to run across a room without falling down or veering off in the wrong direction. The algorithms the Berkeley researchers ultimately produced allowed the robots to not only remember what they learned from trial and error but also to build upon their experiences.



Blue, a low-cost, human-friendly robot designed by EECS professor Pieter Abbeel. Credit: Adam Lau

"The strength of connections between the neurons, and which neurons are connected, is essentially how we internalize experience," Abbeel says. "You need algorithms that look at those experiences and rewires those connections in the network to make the robot perform better."

Eventually, Levine says, researchers might be able to create what he calls "lifelong robotics systems," in which robots improve themselves by continuously analyzing their previous individual triumphs and mistakes and those from other robots.

"When faced with complex tasks, robots will turn their observations into action," he says. "We provide them with the necessary ingredient for them to make those connections."

Multi-tasking robots

Most robots today still require humans to set a reward. But what if robots could set their own goals, unsupervised, similar to the way children explore their environment?

Advances in unsupervised deep reinforcement learning could lead to gains not yet realized in supervised settings. Unlike in other areas of deep learning, robotics researchers lack the large data sets needed to train robots on a broad set of skills. But autonomous exploration could help robots learn a variety of tasks much more quickly.

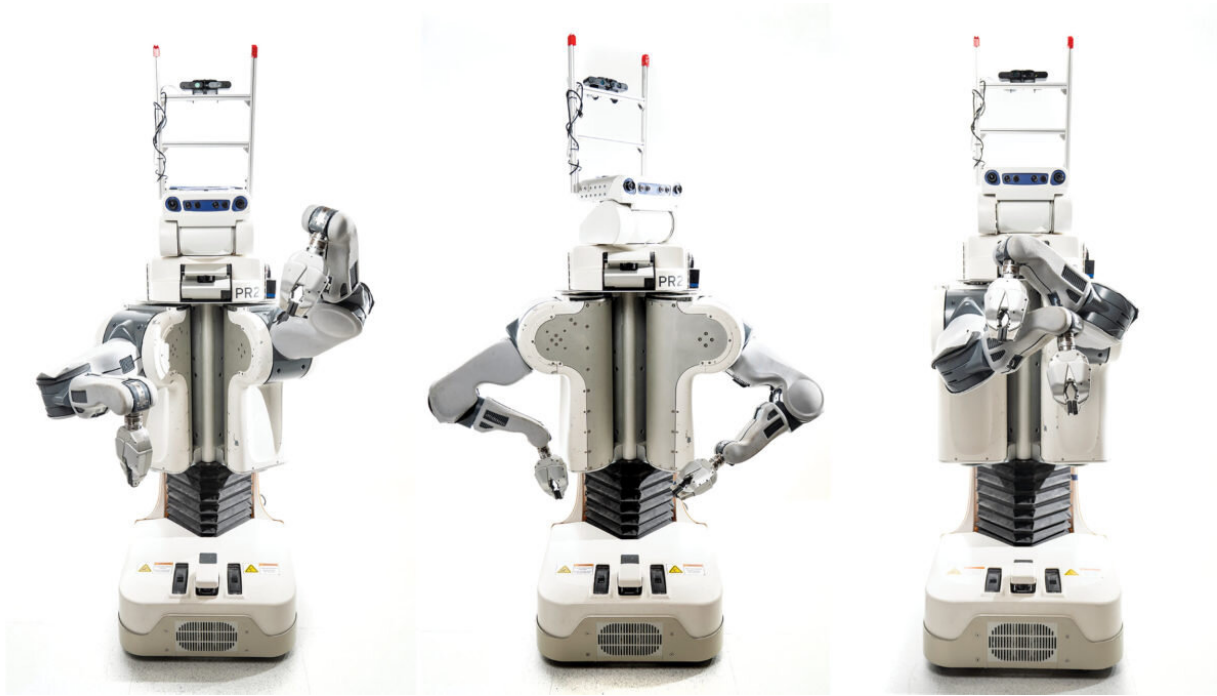
Work coming out of Berkeley has shown what this might look like in robotic systems developed by Abbeel, Levine and Chelsea Finn (Ph.D. EECS), now an assistant professor at Stanford University, as well as student researchers. Robots, drawing on their own data and human demos, can experiment independently with objects. Some of the skills mastered include pouring items from one cup to another, screwing a cap onto a bottle and using a spatula to lift an object into a bowl.

Robots even taught themselves to use an everyday object, such as a water bottle, as a tool to move other items across a surface, demonstrating that they can improvise. Further research by Finn and Levine, collaborating with researchers from the University of Pennsylvania, showed that robots could learn how to use tools by watching videos of humans using tools with their hands.

"What's significant is not the raw skills these robots can do," Finn says, "but the generality of these skills and how they can be applied to many different tasks."

One of the main challenges that researchers are contending with is how to fully automate self-supervised deep reinforcement learning. Robots might be learning like a toddler, but they don't have comparable motor skills.

"In practice, it's very difficult to set up a robotic learning system that can learn continually, in real-world settings, without extensive manual effort," Levine says. "This is not just because the underlying algorithms need to be improved, but because much of the scaffolding and machinery around robotic learning is manual."



BRETT, the Berkeley Robot for the Elimination of Tedious Tasks. Credit: Adam Lau

For example, he says, if a robot is learning to adjust an object in its hand and drops that object, or if a robot is learning to walk and then falls down, a [human needs](#) to step in and fix that. But in the real world, humans are constantly learning on their own, and every mistake becomes a learning opportunity.

"Potentially, a multi-task view of learning could address this issue, where we might imagine the robot utilizing every mistake as an opportunity to instantiate and learn a new skill. If the coffee-delivery robot drops the coffee, it should use that chance to practice cleaning up spilled coffee," Levine says.

"If this is successful, then what we will see over the next few years is

increasingly more and more autonomous learning, such that robots that are actually situated in real-world environments learn continually on the job."

Entering the real world

Some of these advances in deep reinforcement learning for robotics are already making their way out of the lab and into the workplace.

Obeta, a German electronics parts manufacturer, is using the technology developed by Covariant.AI, a company co-founded by Abbeel, for robots to sort through bins of thousands of random gadgets and components that pass through the conveyor belt at its warehouse. The robot can pick and sort more than 10,000 different items it has never seen before with more than 99% accuracy, according to Covariant. There's no need to pre-sort items, making this technology a game-changer for manufacturing.

"I think we will likely see robots gradually permeating more and more real-world settings, but starting on the 'back end' of the commercial sector, and gradually radiating from there toward less and less structured environments," Levine says.

We might see robots transitioning from industrial settings like factories and warehouses to outdoor environments or retail shops. Imagine robots weeding, thinning and spraying crops on farms; stocking grocery store shelves; and making deliveries in hotels and hospitals. Eventually, robots could be deployed in more outward-facing roles, such as janitorial work in large commercial enterprises.

Levine says the fully consumer-facing home robot is some ways away, as robots must first master more complex domains with variability. This will also require a cadre of human experts, out in the field, before it can

be fully realized.

Whatever the domain, these researchers aim to use robots to work collaboratively with people and enhance productivity, as opposed to displacing people from jobs. The attendant scientific, political and economic factors merit serious consideration and are integral to the work they do.

"Every technology has potential for both positive and negative outcomes, and as researchers it's critical for us to be cognizant of this," Levine says. "Ultimately, I believe that ever more capable robots have tremendous potential to make peoples' lives better, and that possibility makes the work worthwhile."

Provided by University of California - Berkeley

Citation: Deep reinforcement learning: Teaching robots like children (2020, April 23) retrieved 23 April 2024 from <https://techxplore.com/news/2020-04-deep-robots-children.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.