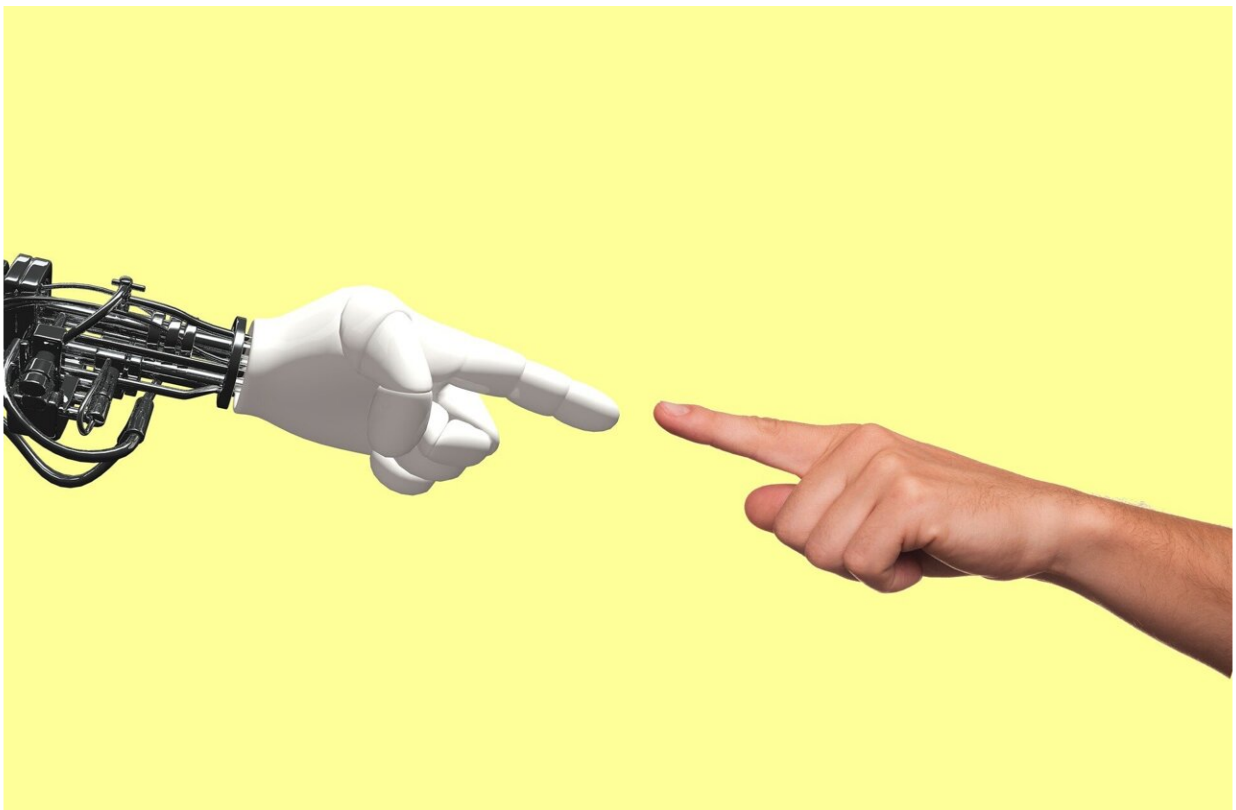
New technique helps user understand why a robot failed, then fine-tune it to perform task

July 18 2023, by Adam Zewe



Credit: Pixabay/CC0 Public Domain

Imagine purchasing a robot to perform household tasks. This robot was built and trained in a factory on a certain set of tasks and has never seen the items in your home. When you ask it to pick up a mug from your kitchen table, it might not recognize your mug (perhaps because this mug is painted with an unusual image, say, of MIT's mascot, Tim the Beaver). So, the robot fails.

" A critical component that is missing from this system is enabling the robot to demonstrate why it is failing so the user can give it feedback," says Andi Peng, an electrical engineering and computer science (EECS) graduate student at MIT.

Peng and her collaborators at MIT, New York University, and the University of California at Berkeley created a framework that enables humans to quickly teach a robot what they want it to do, with a minimal amount of effort.

When a robot fails, the system uses an algorithm to generate counterfactual explanations that describe what needed to change for the robot to succeed. For instance, maybe the robot would have been able to pick up the mug if the mug were a certain color. It shows these counterfactuals to the human and asks for feedback on why the robot failed. Then the system utilizes this feedback and the counterfactual explanations to generate new data it uses to fine-tune the robot.

Fine-tuning involves tweaking a machine-learning model that has already been trained to perform one task, so it can perform a second, similar task.

The researchers tested this technique in simulations and found that it could teach a robot more efficiently than other methods. The robots trained with this framework performed better, while the training process consumed less of a human's time.

This framework could help robots learn faster in new environments without requiring a user to have technical knowledge. In the long run, this could be a step toward enabling general-purpose robots to efficiently perform daily tasks for the elderly or individuals with disabilities in a variety of settings.

Peng, the lead author, is joined by co-authors Aviv Netanyahu, an EECS graduate student; Mark Ho, an assistant professor at the Stevens Institute of Technology; Tianmin Shu, an MIT postdoc; Andreea Bobu, a graduate student at UC Berkeley; and senior authors Julie Shah, an MIT professor of aeronautics and astronautics and the director of the Interactive Robotics Group in the Computer Science and Artificial Intelligence Laboratory (CSAIL), and Pulkit Agrawal, a professor in CSAIL.

The research will be presented at the International Conference on Machine Learning and is available on the pre-print server *arXiv*.

On-the-job training

Robots often fail due to distribution shift—the robot is presented with objects and spaces it did not see during training, and it doesn't understand what to do in this new environment.

One way to retrain a robot for a specific task is imitation learning. The user could demonstrate the correct task to teach the robot what to do. If a user tries to teach a robot to pick up a mug, but demonstrates with a white mug, the robot could learn that all mugs are white. It may then fail to pick up a red, blue, or "Tim-the-Beaver-brown" mug.

Training a robot to recognize that a mug is a mug, regardless of its color, could take thousands of demonstrations.

"I don't want to have to demonstrate with 30,000 mugs. I want to demonstrate with just one mug. But then I need to teach the robot so it recognizes that it can pick up a mug of any color," Peng says.

To accomplish this, the researchers' system determines what specific object the user cares about (a mug) and what elements aren't important for the task (perhaps the color of the mug doesn't matter). It uses this information to generate new, synthetic data by changing these "unimportant" visual concepts. This process is known as data augmentation.

The framework has three steps. First, it shows the task that caused the robot to fail. Then it collects a demonstration from the user of the desired actions and generates counterfactuals by searching over all features in the space that show what needed to change for the robot to succeed.

The system shows these counterfactuals to the user and asks for feedback to determine which visual concepts do not impact the desired action. Then it uses this human feedback to generate many new augmented demonstrations.

In this way, the user could demonstrate picking up one mug, but the system would produce demonstrations showing the desired action with thousands of different mugs by altering the color. It uses these data to fine-tune the robot.

Creating counterfactual explanations and soliciting feedback from the user are critical for the technique to succeed, Peng says.

From human reasoning to robot reasoning

Because their work seeks to put the human in the training loop, the researchers tested their technique with human users. They first conducted a study in which they asked people if counterfactual explanations helped them identify elements that could be changed without affecting the task.

"It was so clear right off the bat. Humans are so good at this type of counterfactual reasoning. And this counterfactual step is what allows human reasoning to be translated into robot reasoning in a way that makes sense," Peng says.

Then they applied their framework to three simulations where robots were tasked with: navigating to a goal object, picking up a key and unlocking a door, and picking up a desired object then placing it on a tabletop. In each instance, their method enabled the robot to learn faster than with other techniques, while requiring fewer demonstrations from users.

Moving forward, the researchers hope to test this framework on real robots. They also want to focus on reducing the time it takes the system to create new data using generative machine-learning models.

"We want robots to do what humans do, and we want them to do it in a semantically meaningful way. Humans tend to operate in this abstract space, where they don't think about every single property in an image. At the end of the day, this is really about enabling a robot to learn a good, human-like representation at an abstract level," Peng says.

More information: Andi Peng et al, Diagnosis, Feedback, Adaptation: A Human-in-the-Loop Framework for Test-Time Policy Adaptation, *arXiv* (2023). DOI: 10.48550/arxiv.2307.06333