

MIT robot combines vision and touch to learn the game of Jenga

30 January 2019



Jenga! #RoboFail. Credit: Fazeli et al., *Sci. Robot.* 4, eaav3123 (2019)

In the basement of MIT's Building 3, a robot is carefully contemplating its next move. It gently pokes at a tower of blocks, looking for the best block to extract without toppling the tower, in a solitary, slow-moving, yet surprisingly agile game of Jenga.

The robot, developed by MIT engineers, is equipped with a soft-pronged gripper, a force-sensing wrist cuff, and an external camera, all of which it uses to see and feel the [tower](#) and its individual blocks.

As the robot carefully pushes against a block, a computer takes in visual and [tactile feedback](#) from its camera and cuff, and compares these measurements to moves that the robot previously made. It also considers the outcomes of those moves—specifically, whether a block, in a certain configuration and pushed with a certain amount of force, was successfully extracted or not. In real-

time, the robot then "learns" whether to keep pushing or move to a new block, in order to keep the tower from falling.

Details of the Jenga-playing robot are published in the journal *Science Robotics*. Alberto Rodriguez, the Walter Henry Gale Career Development Assistant Professor in the Department of Mechanical Engineering at MIT, says the robot demonstrates something that's been tricky to attain in previous systems: the ability to quickly learn the best way to carry out a task, not just from [visual cues](#), as it is commonly studied today, but also from tactile, physical interactions.

"Unlike in more purely cognitive tasks or games such as chess or Go, playing the game of Jenga also requires mastery of physical skills such as probing, pushing, pulling, placing, and aligning pieces. It requires interactive perception and manipulation, where you have to go and touch the tower to learn how and when to move blocks," Rodriguez says. "This is very difficult to simulate, so the robot has to learn in the real world, by interacting with the real Jenga tower. The key challenge is to learn from a relatively small number of experiments by exploiting common sense about objects and physics."

He says the tactile learning system the researchers have developed can be used in applications beyond Jenga, especially in tasks that need careful physical interaction, including separating recyclable objects from landfill trash and assembling consumer products.

"In a cellphone assembly line, in almost every single step, the feeling of a snap-fit, or a threaded screw, is coming from force and touch rather than vision," Rodriguez says. "Learning models for those actions is prime real-estate for this kind of technology."

The paper's lead author is MIT graduate student

Nima Fazeli. The team also includes Miquel Oller, Jiajun Wu, Zheng Wu, and Joshua Tenenbaum, professor of brain and cognitive sciences at MIT.

Push and pull

In the game of Jenga—Swahili for "build"—54 rectangular blocks are stacked in 18 layers of three blocks each, with the blocks in each layer oriented perpendicular to the blocks below. The aim of the game is to carefully extract a block and place it at the top of the tower, thus building a new level, without toppling the entire structure.

To program a robot to play Jenga, traditional machine-learning schemes might require capturing everything that could possibly happen between a block, the robot, and the tower—an expensive computational task requiring data from thousands if not tens of thousands of block-extraction attempts.

Instead, Rodriguez and his colleagues looked for a more data-efficient way for a robot to learn to play Jenga, inspired by [human cognition](#) and the way we ourselves might approach the game.

The team customized an industry-standard ABB IRB 120 [robotic arm](#), then set up a Jenga tower within the robot's reach, and began a training period in which the robot first chose a random block and a location on the block against which to push. It then exerted a small amount of force in an attempt to push the block out of the tower.

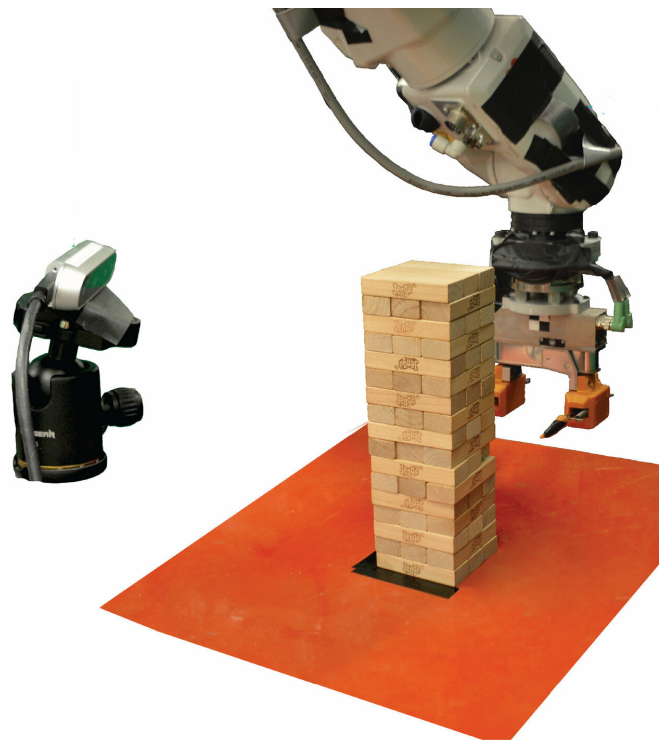
For each block attempt, a computer recorded the associated visual and force measurements, and labeled whether each attempt was a success.

Rather than carry out tens of thousands of such attempts (which would involve reconstructing the tower almost as many times), the robot trained on just about 300, with attempts of similar measurements and outcomes grouped in clusters representing certain block behaviors. For instance, one cluster of might represent attempts on a block that was hard to move, versus one that was easier to move, or that toppled the tower when moved. For each data cluster, the robot developed a simple model to predict a block's behavior given its current visual and tactile measurements.

Fazeli says this clustering technique dramatically increases the efficiency with which the robot can learn to play the game, and is inspired by the natural way in which humans cluster similar behavior: "The robot builds clusters and then learns models for each of these clusters, instead of learning a model that captures absolutely everything that could happen."

Stacking up

The researchers tested their approach against other state-of-the-art machine learning algorithms, in a computer simulation of the game using the simulator MuJoCo. The lessons learned in the simulator informed the researchers of the way the robot would learn in the real world.



Jenga-playing robot. Credit: Fazeli et al., *Sci. Robot.* 4, eaav3123 (2019)

"We provide to these algorithms the same information our system gets, to see how they learn to play Jenga at a similar level," Oller says. "Compared with our approach, these algorithms

need to explore orders of magnitude more towers to learn the game."

Curious as to how their machine-learning approach stacks up against actual human players, the team carried out a few informal trials with several volunteers.

"We saw how many blocks a human was able to extract before the tower fell, and the difference was not that much," Oller says.

But there is still a way to go if the researchers want to competitively pit their robot against a human player. In addition to physical interactions, Jenga requires strategy, such as extracting just the right block that will make it difficult for an opponent to pull out the next block without toppling the tower.

For now, the team is less interested in developing a robotic Jenga champion, and more focused on applying the [robot's](#) new skills to other application domains.

"There are many tasks that we do with our hands where the feeling of doing it 'the right way' comes in the language of forces and tactile cues," Rodriguez says. "For tasks like these, a similar approach to ours could figure it out."

This research was supported, in part, by the National Science Foundation through the National Robotics Initiative.

More information: N. Fazeli et al., "See, feel, act: Hierarchical learning for complex manipulation skills with multisensory fusion," *Science Robotics* (2019). robotics.sciencemag.org/lookup.../scirobotics.aav3123

Provided by Massachusetts Institute of Technology

APA citation: MIT robot combines vision and touch to learn the game of Jenga (2019, January 30) retrieved 1 December 2020 from <https://techxplore.com/news/2019-01-mit-robot-combines-vision-game.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.