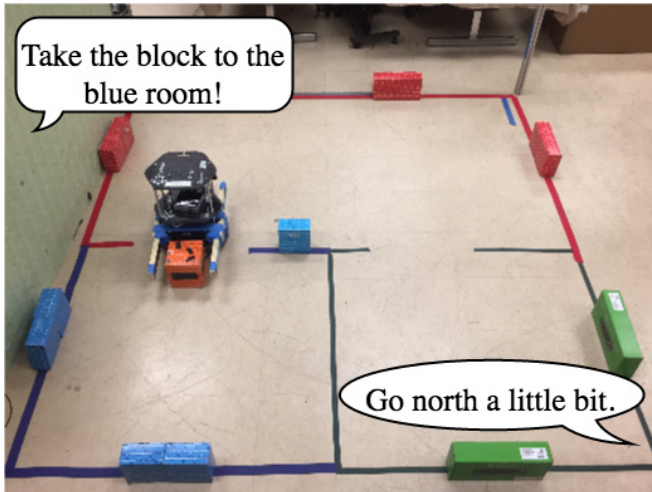


# Research makes robots better at following spoken instructions

13 July 2017, by Kevin Stacey



People give instructions at varying levels of abstraction -- from the simple and straightforward ("Go north a bit.") to more complex commands that imply a myriad of subtasks ("Take the block to the blue room."). A new software system helps robots better deal with instructions whatever their level of abstraction. Credit: Tellex Lab / Brown University

A new system based on research by Brown University computer scientists makes robots better at following spoken instructions, no matter how abstract or specific those instructions may be. The development, which was presented this week at the Robotics: Science and Systems 2017 conference in Boston, is a step toward robots that are able to more seamlessly communicate with human collaborators.

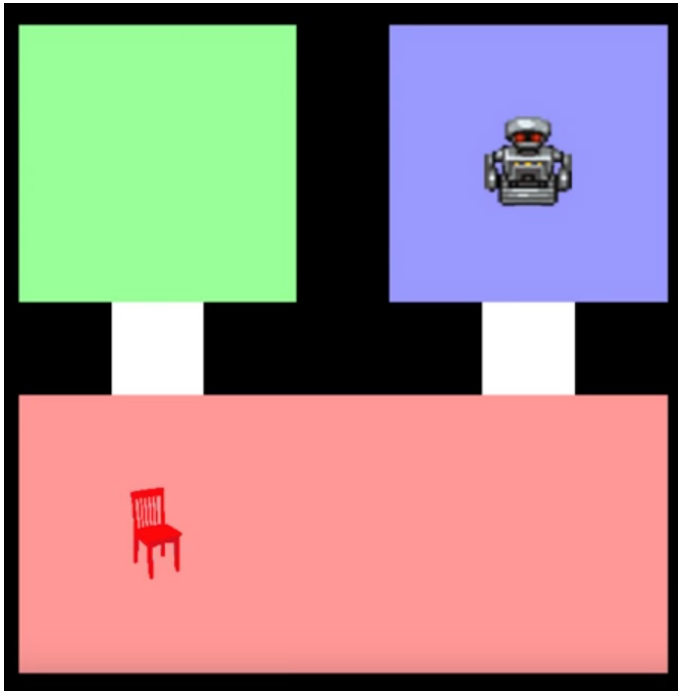
The research was led by Dilip Arumugam and Siddharth Karamcheti, both undergraduates at Brown when the work was performed (Arumugam is now a Brown graduate student). They worked with [graduate student](#) Nakul Gopalan and postdoctoral researcher Lawson L.S. Wong in the lab of Stefanie Tellex, a professor of computer science at Brown.

"The issue we're addressing is language grounding, which means having a robot take natural language commands and generate behaviors that successfully complete a task," Arumugam said. "The problem is that commands can have different levels of [abstraction](#), and that can cause a robot to plan its actions inefficiently or fail to complete the task at all."

For example, imagine someone in a warehouse working side-by-side with a robotic forklift. The person might say to the robotic partner, "Grab that pallet." That's a highly abstract command that implies a number of smaller sub-steps—lining up the lift, putting the forks underneath and hoisting it up. However, other common commands might be more fine-grained, involving only a single action: "Tilt the forks back a little," for example.

Those different levels of abstraction can cause problems for current robot language models, the researchers say. Most models try to identify cues from the words in the command as well as the sentence structure and then infer a desired action from that language. The inference results then trigger a planning algorithm that attempts to solve the task. But without taking into account the specificity of the instructions, the robot might overplan for simple instructions, or underplan for more abstract instructions that involve more sub-steps. That can result in incorrect actions or an overly long planning lag before the robot takes action.

But this new system adds an additional level of sophistication to existing models. In addition to simply inferring a desired task from language, the new system also analyzes the language to infer a distinct level of abstraction.



The virtual Cleanup World provided a training ground for the new language grounding software. Credit: Brown University

"That allows us to couple our task inference as well as our inferred specificity level with a hierarchical planner, so we can plan at any level of abstraction," Arumugam said. "In turn, we can get dramatic speed-ups in performance when executing tasks compared to existing systems."

To develop their new model, the researchers used Mechanical Turk, Amazon's crowdsourcing marketplace, and a virtual task domain called Cleanup World. The online domain consists of a few color-coded rooms, a robotic agent and an object that can be manipulated—in this case, a chair that can be moved from room to room.

Mechanical Turk volunteers watched the robot agent perform a task in the Cleanup World domain—for example, moving the chair from a red room to an adjacent blue room. Then the volunteers were asked to say what instructions they would have given the robot to get it to perform the task they just watched. The volunteers were given guidance as to the level of specificity their

directions should have. The instructions ranged from the high-level: "Take the chair to the blue room" to the stepwise-level: "Take five steps north, turn right, take two more steps, get the chair, turn left, turn left, take five steps south." A third level of abstraction used terminology somewhere in between those two.

The researchers used the volunteers' spoken instructions to train their system to understand what kinds of words are used in each level of abstraction. From there, the system learned to infer not only a desired action, but also the abstraction level of the command. Knowing both of those things, the system could then trigger its hierarchical planning algorithm to solve the task from the appropriate level.

Having trained their system, the researchers tested it in both the virtual Cleanup World and with an actual Roomba-like robot operating in a physical world similar to the Cleanup World space. They showed that when a robot was able to infer both the [task](#) and the specificity of the instructions, it responded to commands in one second 90 percent of the time. In comparison, when no level of specificity was inferred, half of all tasks required 20 or more seconds of planning time.

"We ultimately want to see robots that are helpful partners in our homes and workplaces," said Tellex, who specializes in human-[robot](#) collaboration. "This work is a step toward the goal of enabling people to communicate with robots in much the same way that we communicate with each other."

Provided by Brown University

APA citation: Research makes robots better at following spoken instructions (2017, July 13) retrieved 4 December 2022 from <https://techxplore.com/news/2017-07-robots-spoken.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.*