

# A robotic planner that responds to natural language commands

March 10 2020, by Ingrid Fadelli

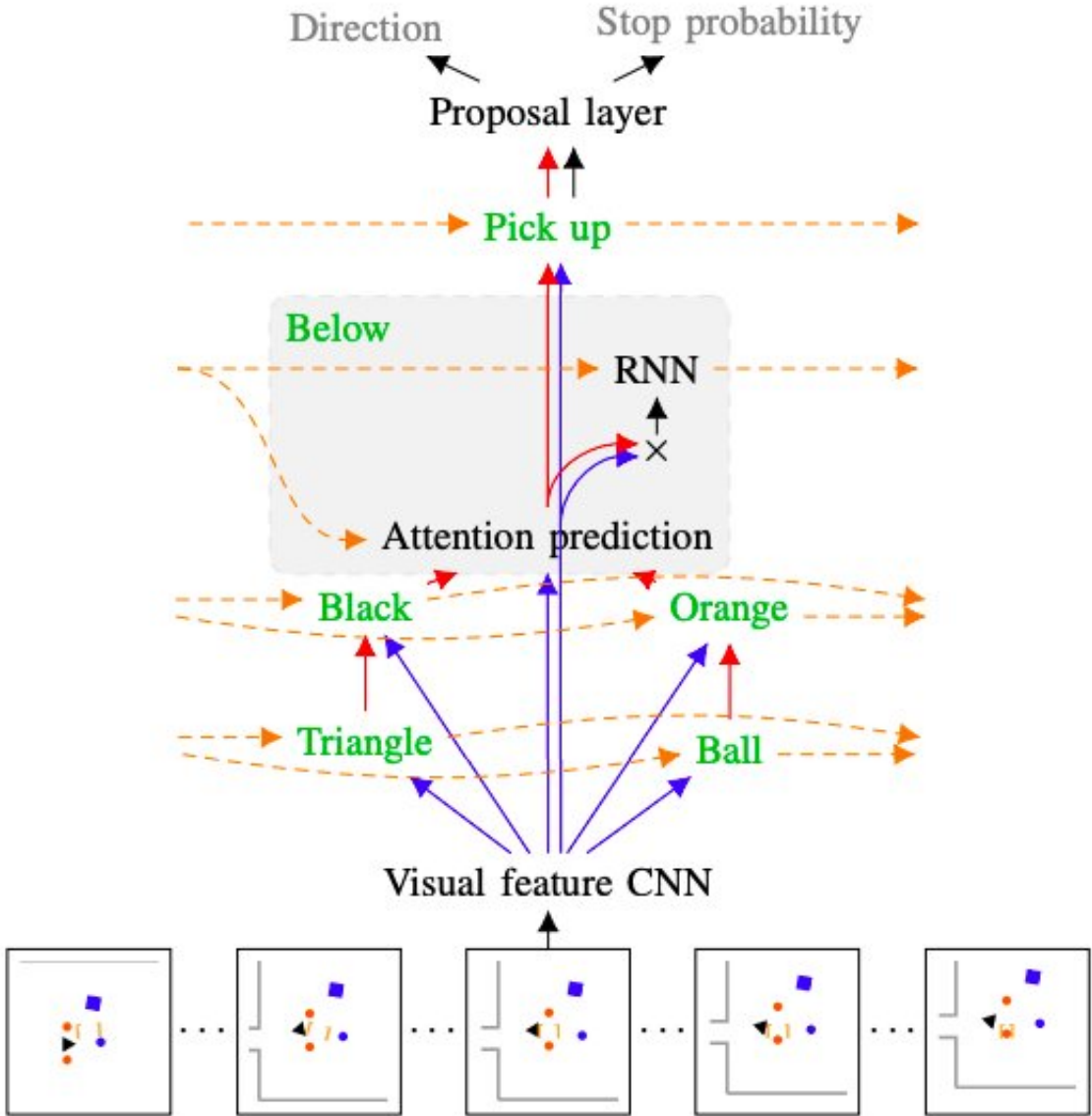


Figure showing how the model developed by the researchers interprets and

follows the command 'pick up the orange ball from below the black triangle'.  
Credit: Kuo, Katz & Barbu.

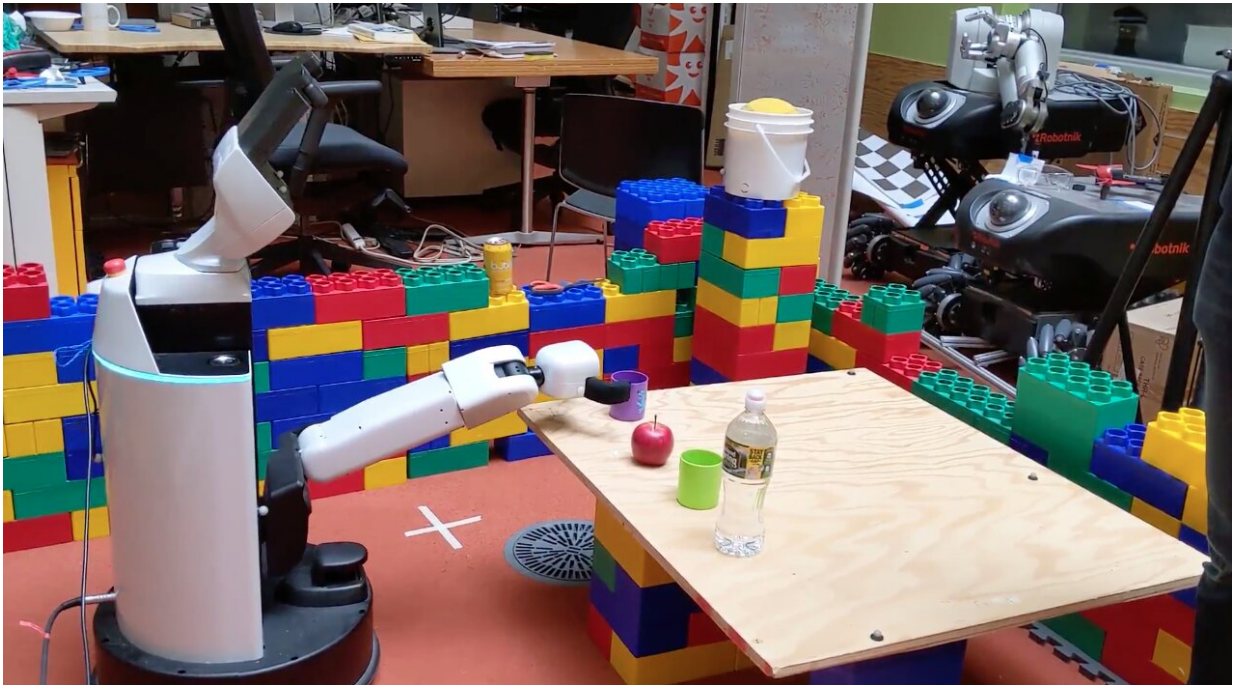
In years to come, robots could assist human users in a variety of ways, both when they are inside their homes and in other settings. To be more intuitive, robots should be able to follow natural language commands and instructions, as this allows users to communicate with them just as they would with other humans.

With this in mind, researchers at MIT's Center for Brains, Minds & Machines have recently developed a sampling-based robotic planner that can be trained to understand sequences of natural language commands. The system they developed, presented in a paper pre-published on arXiv, combines a deep neural [network](#) with a sampling-based planner.

"It's quite important to ensure that future robots in our homes understand us, both for safety reasons and because language is the most convenient interface to ask for what you want," Andrei Barbu, one of the researchers who conducted the study, told TechXplore. "Our work combines three lines of research: robotic planning, deep networks, and our own work on how machines can understand language. The overall objective is to give a [robot](#) only a few examples of what a sentence means and have it follow new commands and new sentences that it never heard before."

The far-reaching goal of the research carried out by Barbu and his colleagues is to better understand body language communication. In fact, while the functions and mechanisms behind spoken communication are now well understood, most communication that takes place among animals and humans is non-verbal.

Gaining a better understanding of body language could lead to the development of more effective strategies for robot-human communication. Among other things, the researchers at MIT have thus been exploring the possibility of translating sentences into robotic motions, and vice versa. Their recent study is a first step in this direction.



Credit: Kuo, Katz & Barbu.

"Robotic planners are amazing at exploring what the robot can do and then having the robot carry out an action," Yen-Ling Kuo, another researcher who carried out the study, told TechXplore. "Our work takes a sentence, breaks it down into pieces, these pieces are translated into small networks, which are recombined back together."

Just like language is made up of words that can be combined into sentences following grammatical rules, the networks developed by Barbu, Kuo and their colleague Boris Katz are comprised of smaller networks trained to understand single concepts. When combined together, these networks can uncover and represent the meaning of whole sentences.

The new robotic planner developed by the researchers has two key components. The first is a recurrent hierarchical [deep neural network](#), which controls how the planner explores the surrounding environment, while also predicting when a planned path is likely to achieve a given goal and estimating the effectiveness of each of the robot's possible moves individually. The second is a sampling-based planner often used in robotics studies, called rapidly exploring random tree (RRT).

"The major advantage of our planner is that it requires little training data," Barbu explained. "If you want to teach a robot, you are not going to give it thousands of examples at home, but a handful are pretty reasonable. Training a robot should entail similar actions to those you might perform if you were training a dog."

While past studies also explored ways of guiding robots via verbal commands, the techniques presented in them often only apply to discrete environments, in which robots can only perform a limited amount of actions. The planner developed by the researchers, on the other hand, can support a variety of interactions with the surrounding environment, even if they involve objects that the robot has never encountered before.

"When our network is confused, the planner portion takes over, figures out what to do and then the network can take over the next time it is confident about what to do," Kuo explained. "The fact that our model is built up out of parts also gives it another desirable property: interpretability."

When they are unable to complete a given task, many existing machine learning models are unable to provide information about what went wrong and the problems they encountered. This makes it harder for developers to identify a model's shortcomings and make targeted changes in its architecture. The deep learning component of the robotic planner created by Barbu, Kuo and Katz, on the other hand, shows its reasoning step by step, clarifying what every word it processes conveys about the world and how it combined the results of its analyses together. This allows the researchers to pinpoint issues that prevented it from successfully completing a given action in the past and make changes in the architecture that could ensure its success in future attempts.

"We are very excited about the notion that robots can quickly learn language and quickly learn new words with very little help from humans," Barbu said. "Normally, deep learning is considered to be very data-hungry; this work reinforces the idea that when you build in the right principles (compositionality) and have agents perform meaningful actions they don't need nearly as much data."

The researchers evaluated the performance of their planner in a series of experiments, while also comparing its performance with that of existing RRT models. In these tests, the planner successfully acquired the meaning of words and used what it learned to represent sequences of sentences that it never encountered before, outperforming all the models it was compared to.

In the future, the model developed by this team of researchers could inform the development of robots that can process and follow natural language commands more effectively. At the moment, their [planner](#) allows robots to process and execute simple instructions such as 'pick up the plate on the table', but is still unable to capture the meaning of more complex ones, such as 'pick up the doll whenever it falls on the floor and clean it'. Barbu, Kuo and Katz are thus currently trying to expand the

range of sentences that the robot can understand.

"Our longer-term future goal is to explore the idea of inverse planning," Kuo said. "That means that if we can turn language into robotic actions, we could then also watch actions and ask the robot 'what was someone thinking when they did this?' We hope that this will serve as a key to unlocking body [language](#) in robots."

**More information:** Deep compositional robotic planners that follow natural language commands. arXiv:2002.05201 [cs.RO].  
[arxiv.org/abs/2002.05201](https://arxiv.org/abs/2002.05201)

© 2020 Science X Network

Citation: A robotic planner that responds to natural language commands (2020, March 10) retrieved 8 September 2024 from <https://techxplore.com/news/2020-03-robotic-planner-natural-language.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.