

Powered by AI, new system makes human-to-robot communication more seamless

November 6 2023, by Juan Siliezar



The new software system, called Lang2LTL, marks an important contribution toward more seamless communications between humans and robots. Photo by Juan Siliezar

The black and yellow robot, meant to resemble a large dog, stood waiting for directions. When they came, the instructions weren't in code but instead in plain English: "Visit the wooden desk exactly two times; in addition, don't go to the wooden desk before the bookshelf."

Four metallic legs whirred into action. The robot went from where it

stood in the room to a nearby bookshelf, and then, after a brief pause, shuffled to the designated wooden desk before leaving and returning for a second visit to satisfy the command.

Until recently, such an exercise would have been nearly impossible for navigation robots like this one to carry out. Most current software for navigation robots can't reliably move from English, or any everyday language, to the [mathematical language](#) that its robots understand and can perform.

And this gets even harder when the software has to make logical leaps based on complex or expressive directions (such as going to the bookshelf before the wooden desk) since that traditionally requires training on thousands of hours of data so that it knows what the robot is supposed to do when it comes across that particular type of command.

Advances in so-called large language models that run on artificial intelligence, however, are changing this. Giving robots newfound powers of understanding and reasoning are not only helping make experiments like this achievable but have [computer scientists](#) excited about transferring this type of success to environments outside of labs, such as people's homes and major cities and towns around the world.

For the past year, researchers at Brown University's Humans to Robots Laboratory have been working on a system with this kind of potential and share it in a [new paper](#) that will be presented at the [Conference](#) on Robot Learning in Atlanta on November 8.

The research marks an important contribution toward more seamless communications between humans and robots, the scientists say, because the sometimes convoluted ways humans naturally communicate with each other usually pose problems when expressed to robots, often resulting in incorrect actions or a long planning lag.

"In the paper, we were particularly thinking about mobile robots moving around an environment," said Stefanie Tellex, a computer science professor at Brown and senior author of the new study. "We wanted a way to connect complex, specific and abstract English instructions that people might say to a robot—like go down Thayer Street in Providence and meet me at the coffee shop, but avoid the CVS and first stop at the bank—to a robot's behavior."

The paper describes how the team's novel system and software makes this possible by using A.I. language models, similar to those that power chatbots like ChatGPT, to devise an innovative method that compartmentalizes and breaks down the instructions to eliminate the need for the training data.

It also explains how the software provides navigation robots with a powerful grounding tool that has the ability to not only take natural language commands and generate behaviors, but is also able to compute the logical leaps a robot may need to make based on both context from the plain-worded instructions and what they say the robot can or can't do and in what order.

"In the future, this has applications for mobile robots moving through our cities, whether a drone, a self-driving car or a ground vehicle delivering packages," Tellex said. "Anytime you need to talk to a robot and tell it to do stuff, you would be able to do that and give it very rich, detailed, precise instructions."

Tellex says the new system, with its ability to understand expressive and rich language, represents one of the most powerful language understanding systems for route directions that has ever been released, since it can essentially start working in robots without the need for [training data](#).

Traditionally, if developers wanted a robot to plot out and complete routes in Boston, for example, they would have to collect different examples of people giving instructions in the city—such as "travel through Boston Common but avoid the Frog Pond"—so the system knows what this means and can compute it to the robot. They have to do that training all over again if they want the robot to then navigate New York City.

The new level of sophistication found in the system the researchers created means it can operate in any new environment without a long training process. Instead, it only needs a detailed map of the environment.

"We basically go from language to actions that are conducted by the robot," said Ankit Shah, a postdoctoral researcher in Tellex's lab at Brown.

To test the system, the researchers put the software through simulations in 21 cities using OpenStreetMap. The simulations showed the system is accurate 80% of the time. The number is far more accurate than other systems similar to it, which the researchers say are only accurate about 20% of the time and can only compute simple waypoint navigation such as going from point A to point B. Such systems also can't account for constraints, like needing to avoid an area or having to go to one additional location before going to point A or point B.

Along with the simulations, the researchers tested their system indoors on Brown's campus using a Boston Dynamics Spot robot. Overall, the project adds to a history of high-impact work coming from Tellex's lab at Brown, which has included research that made robots [better at following spoken instruction](#), an algorithm that improved a robot's ability to [fetch objects](#) and software that helped robots produce [human-like pen strokes](#).

From language to actions

Lead author of the study Jason Xinyu, a computer science Ph.D. student at Brown working with Tellex, says that the success of the new software, called Lang2LTL, is in how it works. To demonstrate, he gives the example of a user telling a drone to go to "the store" on Main Street but only after visiting "the bank."

First, the two locations get pulled out, he explains. The language model then starts to match these abstract locations to specific locations the model knows are in the robot's environment. It also analyzes the metadata that is available on the locations, such as their addresses or what kind of store they are to help the system make its decisions.

In this case, there are a few nearby stores but only one on Main Street, so the system knows to make the leap that "the store" is Walmart and that "the bank" is Chase. The language model then finishes translating the commands to linear temporal logic, which are mathematical codes and symbols that express those commands. The system then takes the now mapped locations and plugs them into the formula it has been creating, telling the [robot](#) to go to point A but only after point B.

"Essentially, our system uses its modular system design and its large language models pre-trained on internet-scaled data to process more complex directional and linear-based natural language commands with different kind of constraints that no robotic system could understand before," Xinyu said. "Previous systems couldn't handle this because they were held back by how they were designed to essentially do this process all at once."

The researchers are already thinking about what comes next in the project.

They plan to release a simulation in November based in OpenStreetMaps on the [project website](#) where users can test out the system for themselves. The demo for web browsers will let users type in natural [language](#) commands that instruct a drone in the simulation to carry out navigation commands, letting the researchers study how their software works for fine-tuning. Soon after, the team hopes to add object manipulation capabilities to the software.

"This work is a foundation for a lot of the work we can do in the future," Xinyu said.

More information: Paper: openreview.net/forum?id=rpWi4SYGXj

GitHub: github.com/h2r/Lang2LTL

Provided by Brown University

Citation: Powered by AI, new system makes human-to-robot communication more seamless (2023, November 6) retrieved 25 May 2024 from <https://techxplore.com/news/2023-11-powered-ai-human-to-robot-communication-seamless.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.