

Robots let ChatGPT touch the real world thanks to Microsoft



reader comments

64 with

Share this story

Last week, Microsoft researchers announced an experimental framework to control robots and drones using the language abilities of ChatGPT, a popular AI language model created by OpenAI. Using natural language commands, ChatGPT can write special code that controls robot movements. A human then views the results and adjusts as necessary until the task gets completed successfully.

The research arrived in a paper titled “ChatGPT for Robotics: Design Principles and Model Abilities,” authored by Sai

Vemprala, Rogerio Bonatti, Arthur Buckner, and Ashish Kapoor of the Microsoft Autonomous Systems and Robotics Group.

In a demonstration video, Microsoft shows robots—apparently controlled by code written by ChatGPT while following human instructions—using a robot arm to arrange blocks into a Microsoft logo, flying a drone to inspect the contents of a shelf, or finding objects using a robot with vision capabilities.

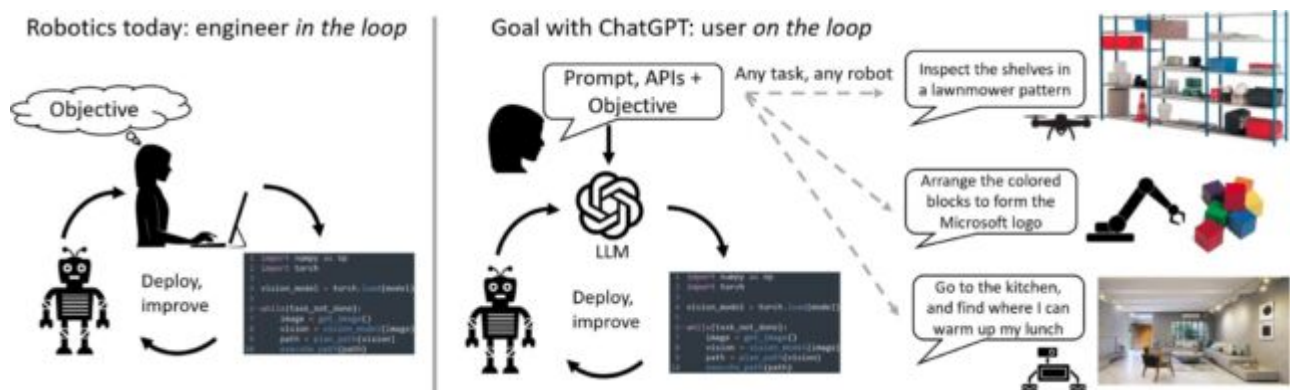
[embedded content]

Microsoft's "ChatGPT for Robotics" demonstration video.

To get ChatGPT to interface with robotics, the researchers taught ChatGPT a custom robotics API. When given instructions like "pick up the ball," ChatGPT can generate robotics control code just as it would write a poem or complete an essay. After a human inspects and edits the code for accuracy and safety, the human operator can execute the task and evaluate its performance.

Advertisement

In this way, ChatGPT accelerates robotic control programming, but it's not an autonomous system. "We emphasize that the use of ChatGPT for robotics is not a fully automated process," reads the paper, "but rather acts as a tool to augment human capacity."



Enlarge / A diagram provided by Microsoft that explains how ChatGPT for Robotics works.

While it appears most of the feedback to ChatGPT (in terms of the success or failure of its actions) comes from humans in

the form of text, the researchers also claim to have had some success with feeding visual data into ChatGPT itself. In one example, researchers tasked ChatGPT with commanding a robot to catch a basketball with feedback from a camera: “ChatGPT can estimate the appearance of the ball and the sky in the camera image using SVG code. This behavior hints at a possibility that the LLM keeps track of an implicit world model going beyond text-based probabilities.”

While the results seem rudimentary for now, they represent early attempts at applying the hottest tech du jour—large language models—to robotic control. According to Microsoft, a ChatGPT interface could open up robotics to a much wider audience in the future.

“Our goal with this research is to see if ChatGPT can think beyond text, and reason about the physical world to help with robotics tasks,” reads a Microsoft Research blog post. “We want to help people interact with robots more easily, without needing to learn complex programming languages or details about robotic systems.”

<https://arstechnica.com/?p=1920803>