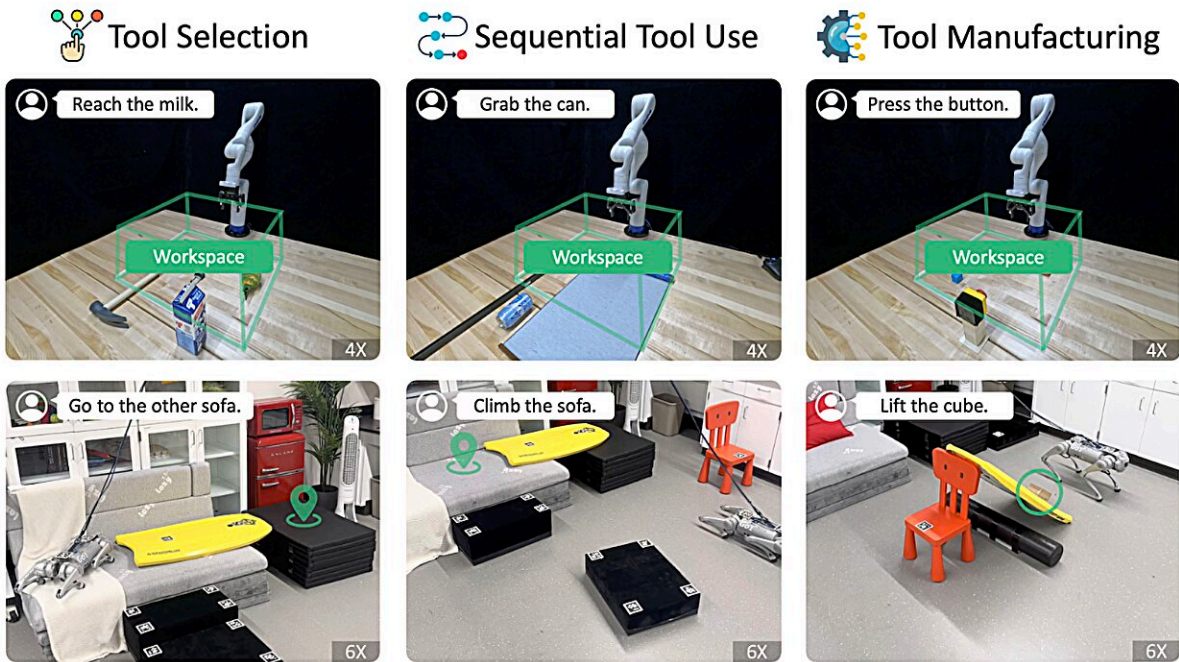


# A system that allows robots to use tools creatively by leveraging large language models

November 14 2023, by Ingrid Fadelli



Robots use tools creatively to solve complex robot tasks that were impossible originally.

Credit: Xu et al.

Researchers at Carnegie Mellon University and Google DeepMind recently developed RoboTool, a system that can broaden the capabilities of robots, allowing them to use tools in more creative ways. This system,

introduced in a paper [published](#) on the *arXiv* preprint server, could soon bring a new wave of innovation and creativity to the field of robotics.

"Tool use is often regarded as the hallmark of advanced intelligence," Mengdi Xu, final-year Ph.D. candidate at Carnegie Mellon University and co-first author of the paper, told Tech Xplore.

"In Wolfgang Koehler's experiments, for instance, apes cleverly stacked crates to access bananas hung out of their reach while crab-eating macaques employed stones as tools to crack open nuts and shells. Beyond using tools for their intended purpose and following established procedures, using tools in creative and unconventional ways provides more flexible solutions but presents far more challenges in cognitive ability."

Robots often complete manual tasks in standard and repetitive ways without exploring alternative approaches. By exploring more creative ways of doing things, however, they could better tackle complex real-world scenarios.

"In robotics, creative [tool use](#) is also a crucial yet very demanding capability because it necessitates the all-around ability to predict the outcome of an action, reason what tools to use, and plan how to use them," Peide Huang, co-first author and Ph.D. candidate, said.

The primary objective of the recent work by Xu, Huang, and their colleagues was to devise a system that allows robots to use tools more creatively. Such a tool could help to tackle numerous real-world problems more effectively, for instance, allowing robots to adapt their strategies when trying to grasp objects that are out of reach or to create stepstones to climb to a target location.

"The rise of large language models (LLMs) has tremendously enhanced

the functionalities of chatbots, coding automation, and visual content creation," Huang explained. "Beyond these digital interfaces, embodied AI could represent the next frontier in intelligence—one that interacts tangibly with the real world. Robots, serving as the physical extensions of LLMs, present an ideal medium for this exploration."

The advent of LLMs and their recent rise in popularity encouraged researchers to explore their use in the field of robotics. Past studies demonstrated the potential of these models for improving various robot capabilities, including their communication with users, as well as their reasoning, planning, and task execution.

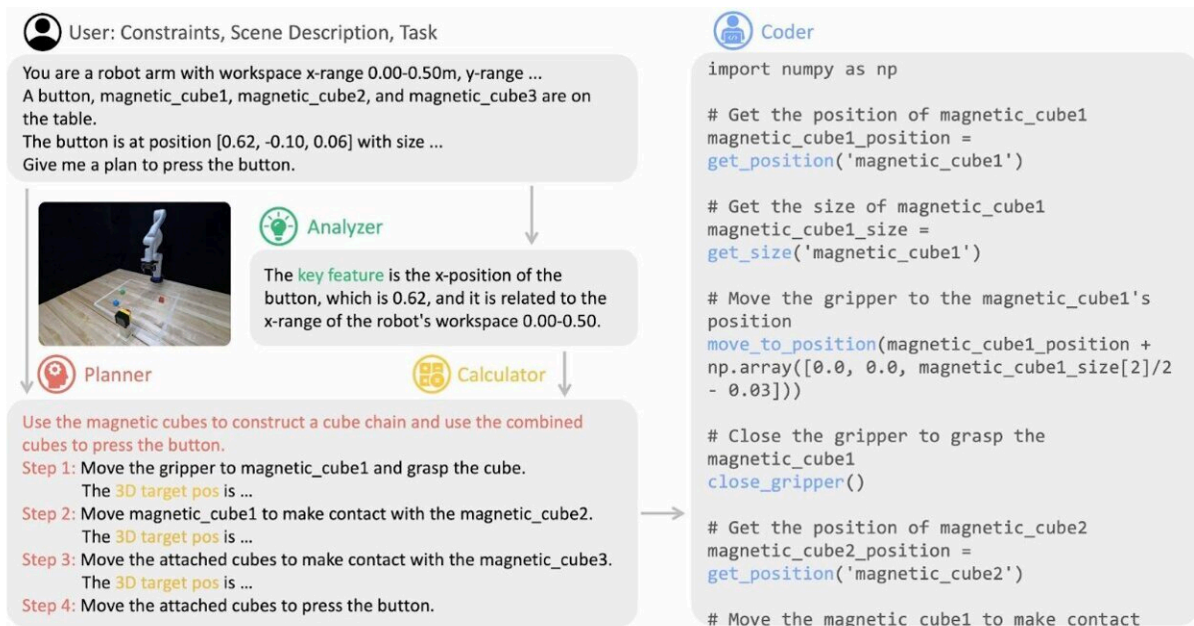
For instance, Google DeepMind's [SayCan](#) tool allows robots to comprehend natural language instructions such as "I spilled my drink, can you help?" and subsequently devise strategies to tackle various domestic chores. Yet, leveraging LLMs to solve problems that require reasoning with implicit constraints set by a robot's body and its surrounding environment remains challenging.

Xu, Huang, and their colleagues set out to explore the use of LLMs to boost the creativity with which robots approach different tasks. In other words, their hope was to create a system that would identify creative ways to make seemingly "impossible" tasks possible.

Their proposed system, dubbed RoboTool, accepts natural language instructions consisting of textual and numerical information about the environment, robot embodiments, and any constraints to follow. It then produces code that applies a robot's parameterized low-level skills to control both simulated and physical robots.

The new tool created by the researchers has four key components: an analyzer, a planner, a calculator, and a coder. The analyzer processes prompts given by users in natural language, identifying key elements that

could affect a requested task's feasibility.



Overview of RoboTool. Credit: Xu et al.

The system's planner component receives both the original language input and the identified key concepts, using them to formulate a comprehensive strategy to complete a task. The calculator component, on the other hand, determines the parameters, such as the target positions required for each parameterized skill.

RoboTool's final component, the coder, converts the comprehensive plan created by the planner and the parameters produced by the calculator into executable code. Notably, all of these components were developed using the GPT-4 model by OpenAI.

RoboTool allows robots to use tools creatively, solving a variety of

complex tasks that they have never encountered before. For example, it could help to create a lever to lift heavy boxes or a stick from magnetic cubes to press an out-of-reach button.

The new tool developed by Xu, Huang and their collaborators could soon be used by roboticists worldwide to broaden the capabilities of their proposed systems. The tool can, for instance, allow robots to perform more complex household tasks, such as unclogging drains or fixing broken furniture using available tools.

"RoboTool could also improve a [robot](#)'s navigation of debris or collapsed structures by improvising with available tools to reach trapped individuals," Xu said. "It could also be applied to construction and maintenance, allowing robots to adaptively fix machinery or structures using whatever tools are on hand, or constructing intricate designs by creatively combining traditional tools."

The researchers have already released demo videos of RoboTool [on the project website](#). In their next studies, they plan to incorporate large vision foundation models into their system, including models that support 3D computer vision, as this could further enhance the sensing and reasoning capabilities of robots in open-world environments.

"We also plan to develop intuitive ways for humans to instruct and collaborate with RoboTool, and to establish [safety measures](#) that for RoboTool that reduce risks when robots are working alongside humans," Ding Zhao, an Associate Professor and the director of CMU Safe AI lab, said.

**More information:** Mengdi Xu et al, Creative Robot Tool Use with Large Language Models, *arXiv* (2023). [DOI: 10.48550/arxiv.2310.13065](#)

© 2023 Science X Network

Citation: A system that allows robots to use tools creatively by leveraging large language models (2023, November 14) retrieved 8 August 2024 from <https://techxplore.com/news/2023-11-robots-tools-creatively-leveraging-large.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.