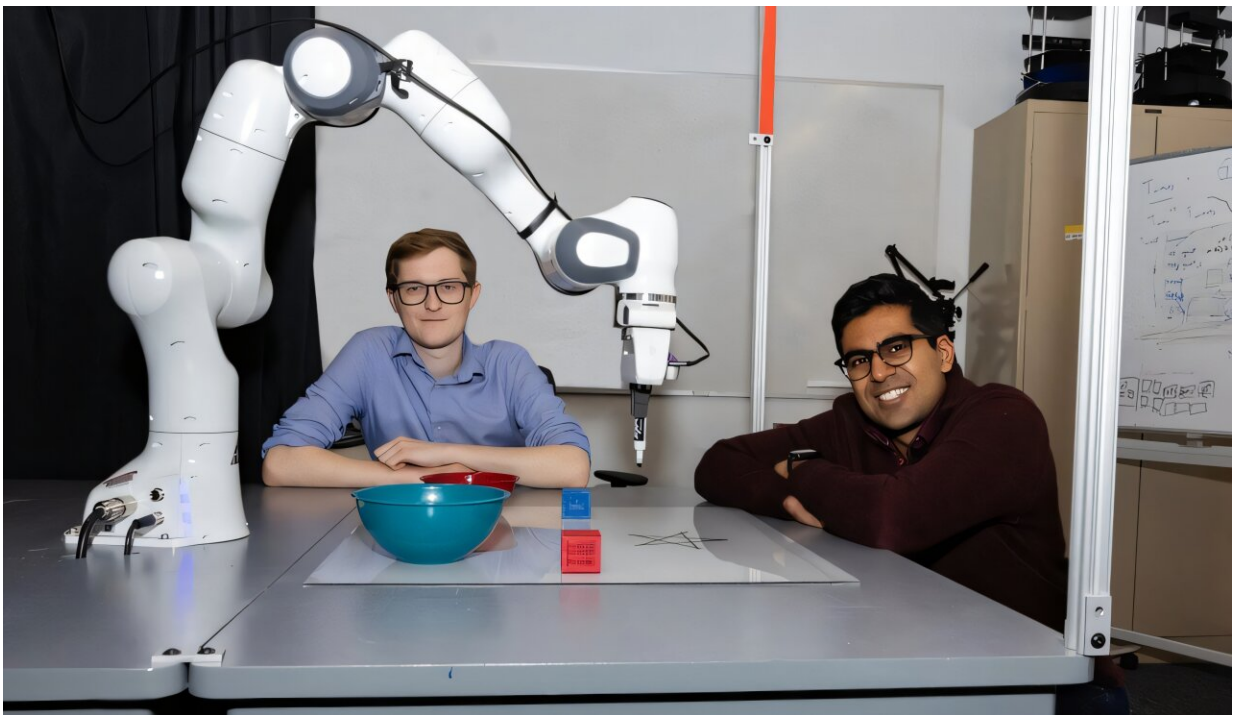# Teaching a robot its limits to complete open-ended tasks safely

December 13 2024, by Alex Shipps

Ph.D. students Aidan Curtis (left) and Nishanth Kumar. To help robots execute open-ended tasks safely, the researchers used vision models to see what's near the machine and model its constraints. Their "PRoC3S" strategy has an LLM sketch up an action plan that's checked in a simulator to ensure it will work in the real world. Credit: Mike Grimmett/MIT CSAIL

If someone advises you to "know your limits," they're likely suggesting you do things like exercise in moderation. To a robot, though, the motto

represents learning constraints, or limitations of a specific task within the machine's environment, to do chores safely and correctly.

For instance, imagine asking a robot to clean your kitchen when it doesn't understand the physics of its surroundings. How can the machine generate a practical multistep plan to ensure the room is spotless? Large language models (LLMs) can get them close, but if the model is only trained on text, it's likely to miss out on key specifics about the robot's physical constraints, like how far it can reach or whether there are nearby obstacles to avoid. Stick to LLMs alone, and you're likely to end up cleaning pasta stains out of your floorboards.

To guide robots in executing these open-ended tasks, researchers at MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL) used vision models to see what's near the machine and model its constraints. The team's strategy involves an LLM sketching up a plan that's checked in a simulator to ensure it's safe and realistic. If that sequence of actions is infeasible, the language model will generate a new plan, until it arrives at one that the robot can execute.

This trial-and-error method, which the researchers call "Planning for Robots via Code for Continuous Constraint Satisfaction" (PRoC3S), tests long-horizon plans to ensure they satisfy all constraints, and enables a robot to perform such diverse tasks as writing individual letters, drawing a star, and sorting and placing blocks in different positions. In the future, PRoC3S could help robots complete more intricate chores in dynamic environments like houses, where they may be prompted to do a general chore composed of many steps (like "make me breakfast").

"LLMs and classical robotics systems like task and motion planners can't execute these kinds of tasks on their own, but together, their synergy makes open-ended problem-solving possible," says Ph.D. student Nishanth Kumar SM '24, co-lead author of a new paper about PRoC3S

posted to the *arXiv* preprint server.

"We're creating a simulation on-the-fly of what's around the robot and trying out many possible action plans. Vision models help us create a very realistic digital world that enables the robot to reason about feasible actions for each step of a long-horizon plan."

The team's work was presented this past month at the Conference on Robot Learning (CoRL) in Munich, Germany.

The researchers' method uses an LLM pre-trained on text from across the internet. Before asking PRoC3S to do a task, the team provided their language model with a sample task (like drawing a square) that's related to the target one (drawing a star). The sample task includes a description of the activity, a long-horizon plan, and relevant details about the robot's environment.

But how did these plans fare in practice? In simulations, PRoC3S successfully drew stars and letters eight out of 10 times each. It also could stack digital blocks in pyramids and lines, and place items with accuracy, like fruits on a plate. Across each of these digital demos, the CSAIL method completed the requested task more consistently than comparable approaches like "LLM3" and "Code as Policies."

The CSAIL engineers next brought their approach to the real world. Their method developed and executed plans on a robotic arm, teaching it to put blocks in straight lines. PRoC3S also enabled the machine to place blue and red blocks into matching bowls and move all objects near the center of a table.

Kumar and co-lead author Aidan Curtis SM '23, who's also a Ph.D. student working in CSAIL, say these findings indicate how an LLM can develop safer plans that humans can trust to work in practice. The

researchers envision a home robot that can be given a more general request (like "bring me some chips") and reliably figure out the specific steps needed to execute it. PRoC3S could help a robot test out plans in an identical digital environment to find a working course of action—and more importantly, bring you a tasty snack.

For future work, the researchers aim to improve results using a more advanced physics simulator and to expand to more elaborate longer-horizon tasks via more scalable data-search techniques. Moreover, they plan to apply PRoC3S to mobile robots such as a quadruped for tasks that include walking and scanning surroundings.

"Using foundation models like ChatGPT to control robot actions can lead to unsafe or incorrect behaviors due to hallucinations," says The AI Institute researcher Eric Rosen, who isn't involved in the research. "PRoC3S tackles this issue by leveraging foundation models for high-level task guidance, while employing AI techniques that explicitly reason about the world to ensure verifiably safe and correct actions.

"This combination of planning-based and data-driven approaches may be key to developing robots capable of understanding and reliably performing a broader range of tasks than currently possible."

  More information: Aidan Curtis et al, Trust the PRoC3S: Solving Long-Horizon Robotics Problems with LLMs and Constraint Satisfaction, *arXiv* (2024). DOI: 10.48550/arxiv.2406.05572

Provided by Massachusetts Institute of Technology

Citation: Teaching a robot its limits to complete open-ended tasks safely (2024, December 13) retrieved 13 December 2024 from https://techxplore.com/news/2024-12-robot-limits-tasks-safely.html