Using large language models to code new tasks for robots
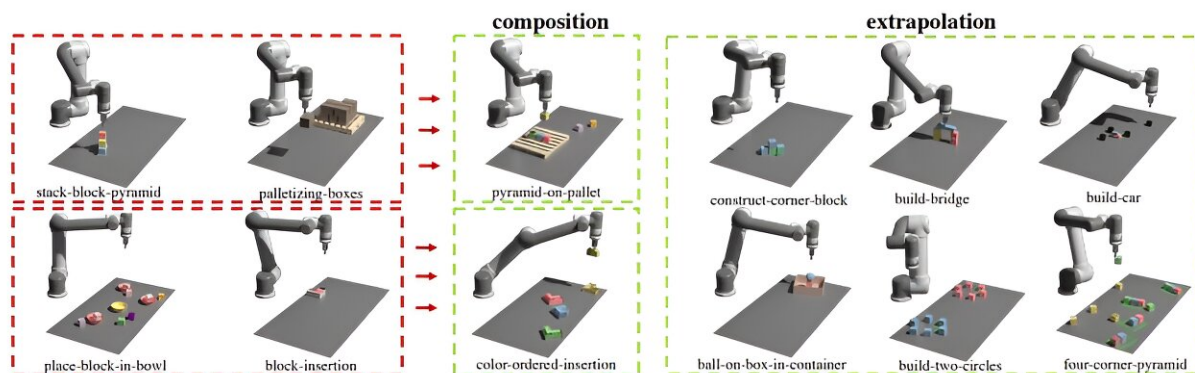
November 28 2023, by Alex Shipps

GenSim demonstrates interesting task-level composition and extrapolation behaviors in code generation for simulation tasks, which are distilled to policy learning through demonstrations. Credit: *arXiv* (2023). DOI: 10.48550/arxiv.2310.01361

You've likely heard that "experience is the best teacher"—but what if learning in the real world is prohibitively expensive? This is the plight of roboticists training their machines on manipulation tasks. Real-world interaction data is costly, so their robots often learn from simulated versions of different activities.

Still, these simulations present a limited range of tasks because each behavior is coded individually by human experts. As a result, many bots cannot complete prompts for chores they haven't seen before. For

example, a robot may not be able to build a toy car because it would need to understand each smaller [task] within that request. Without sufficient, creative simulation data, a robot cannot complete each step within that overarching process (sometimes called long-horizon tasks).

MIT CSAIL's "[GenSim]" attempts to supersize the simulation tasks these machines can be trained on, with a twist. After users prompt [large language models] (LLMs) to automatically generate new tasks or outline each step within a desired behavior, the approach simulates those instructions. By exploiting the code within models like GPT4, GenSim makes headway in helping robots complete each task involved in manufacturing, household chores, and logistics.

The versatile system has goal-directed and exploratory modes. In the goal-directed setting, GenSim takes the chore a user inputs and breaks down each step needed to accomplish that objective. In the exploratory setting, the system comes up with new tasks. For both modes, the process starts with an LLM generating task descriptions and the code needed to simulate the behavior. Then, the model uses a task library to refine the code. The final draft of these instructions can then create simulations that teach robots how to do new chores.

After humans pretrained the system on ten tasks, GenSim automatically generated 100 new behaviors. Meanwhile, comparable benchmarks can only reach that feat by coding each task manually. GenSim also assisted robotic arms in several demonstrations, where its simulations successfully trained the machines to execute tasks like placing colored blocks at a higher rate than comparable approaches.

"In the beginning, we thought it would be amazing to get the type of generalization and extrapolation you find in large language models into robotics," says MIT CSAIL Ph.D. student Lirui Wang, who is a lead author of the paper [posted] to the *arXiv* preprint server.

"So we set out to distill that knowledge through the medium of simulation programs. Then, we bootstrapped the real-world policy based on top of the simulation policies that trained on the generated tasks, and we conducted them through adaptation, showing that GenSim works in both simulation and the real world."

GenSim can potentially aid in kitchen robotics, manufacturing, and logistics, where the approach could generate behaviors for training. In turn, this would enable the machines to adapt to environments with multistep processes, such as stacking and moving boxes to the correct areas.

The system can only assist with pick-and-place activities for now—but the researchers believe GenSim could eventually generate more complex and dexterous tasks, like using a hammer, opening a box, and placing things on a shelf. Additionally, the approach is prone to hallucinations and grounding problems, and further real-world testing is needed to evaluate the usefulness of the tasks it generates. Nonetheless, GenSim presents an encouraging future for LLMs in ideating new robotic activities.

"A fundamental problem in robot learning is where tasks come from and how they may be specified," says Jiajun Wu, Assistant Professor at Stanford University, who is not involved in the work. "The GenSim paper suggests a new possibility: We leverage foundation models to generate and specify tasks based on the common sense knowledge they have learned. This inspiring approach opens up a number of future research directions toward developing a generalist robot."

"The arrival of large language models has broadened the perspectives of what is possible in robot learning and GenSim is an excellent example of a novel application of LLMs that wasn't feasible before," adds Google Deepmind researcher and Stanford adjunct professor Karol Hausman,

who is also not involved in the paper.

"It demonstrates not only that LLMs can be used for asset and environment generation, but also that they can enable the generation of robotic behaviors at scale—a feat previously unachievable. I am very excited to see how scalable simulation behavior generation will impact the traditionally data-starved field of robot learning and I am highly optimistic about its potential to address many of the existing bottlenecks."

"Robotic simulation has been an important tool for providing data and benchmarks to train and assess robot learning models," notes Yuke Zhu, Assistant Professor at The University of Texas at Austin, who is not involved with GenSim. "A practical challenge for using simulation tools is creating a large collection of realistic environments with minimal human effort. I envision generative AI tools, exemplified by large language models, can play a pivotal role in creating rich and diverse simulated environments and tasks.

"Indeed, GenSim shows the promise of large language models in simplifying simulation design through their impressive coding abilities. I foresee great potential for these methods in creating the next generation of robotic simulations at scale."

More information: Lirui Wang et al, GenSim: Generating Robotic Simulation Tasks via Large Language Models, *arXiv* (2023). DOI: 10.48550/arxiv.2310.01361

Github: github.com/liruiw/GenSim

Provided by Massachusetts Institute of Technology

Citation: Using large language models to code new tasks for robots (2023, November 28) retrieved 11 May 2024 from https://techxplore.com/news/2023-11-large-language-code-tasks-robots.html