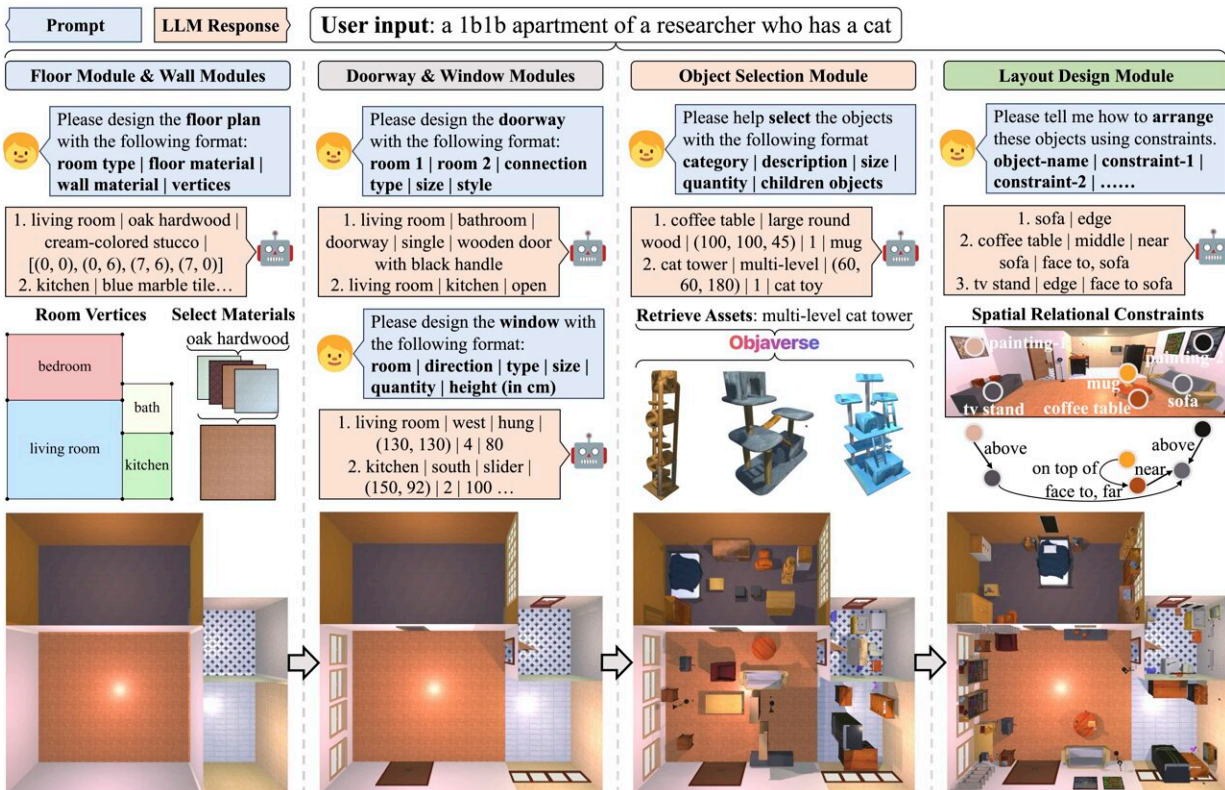# Engineers recreate Star Trek's Holodeck using ChatGPT and video game assets

April 11 2024, by Ian Scheffler

Essentially, Holodeck engages a large language model (LLM) in a conversation, building a virtual environment piece by piece. Credit: Yue Yang

In "Star Trek: The Next Generation," Captain Picard and the crew of the U.S.S. Enterprise leverage the Holodeck, an empty room capable of generating 3D environments, of preparing for missions and entertaining

them, simulating everything from lush jungles to the London of Sherlock Holmes.

Deeply immersive and fully interactive, Holodeck-created environments are infinitely customizable, using nothing but language; the crew has only to ask the computer to generate an environment, and that space appears in the Holodeck.

Today, virtual interactive environments are also used to train robots prior to real-world deployment in a process called "Sim2Real." However, virtual interactive environments have been in surprisingly short supply.

"Artists manually create these environments," says Yue Yang, a doctoral student in the labs of Mark Yatskar and Chris Callison-Burch, Assistant and Associate Professors in Computer and Information Science (CIS), respectively. "Those artists could spend a week building a single environment," Yang adds, noting all the decisions involved, from the layout of the space to the placement of objects to the colors employed in rendering.

That paucity of virtual environments is a problem if you want to train robots to navigate the real world with all its complexities. Neural networks, the systems powering today's AI revolution, require massive amounts of data, which in this case means simulations of the physical world.

"Generative AI systems like ChatGPT are trained on trillions of words, and image generators like Midjourney and DALL-E are trained on billions of images," says Callison-Burch. "We only have a fraction of that amount of 3D environments for training so-called 'embodied AI.' If we want to use generative AI techniques to develop robots that can safely navigate in real-world environments, then we will need to create millions or billions of simulated environments."

Enter [Holodeck](#), a system for generating interactive 3D environments co-created by Callison-Burch, Yatskar, Yang and Lingjie Liu, Aravind K. Joshi Assistant Professor in CIS, along with collaborators at Stanford, the University of Washington, and the Allen Institute for Artificial Intelligence (AI2). Named for its Star Trek forebear, Holodeck generates a virtually limitless range of indoor environments, using AI to interpret users' requests.

The paper is [published](#) on the *arXiv* preprint server.

"We can use language to control it," says Yang. "You can easily describe whatever environments you want and train the embodied AI agents."

Holodeck leverages the knowledge embedded in [large language models](#) (LLMs), the systems underlying ChatGPT, and other chatbots. "Language is a very concise representation of the entire world," says Yang. Indeed, LLMs turn out to have a surprisingly high degree of knowledge about the design of spaces, thanks to the vast amounts of text they ingest during training. In essence, Holodeck works by engaging an LLM in conversation, using a carefully structured series of hidden queries to break down user requests into specific parameters.

Just like Captain Picard might ask Star Trek's Holodeck to simulate a speakeasy, researchers can ask Penn's Holodeck to create "a 1b1b apartment of a researcher who has a cat." The system executes this query by dividing it into multiple steps: First, the floor and walls are created, then the doorway and windows.

Next, Holodeck searches [Objaverse](#), a vast library of premade digital objects, for the sort of furnishings you might expect in such a space: a coffee table, a cat tower, and so on. Finally, Holodeck queries a layout module, which the researchers designed to constrain the placement of objects so that you don't wind up with a toilet extending horizontally

from the wall.

To evaluate Holodeck's abilities, in terms of their realism and accuracy, the researchers generated 120 scenes using both Holodeck and ProcTHOR, an earlier tool created by AI2, and asked several hundred Penn Engineering students to indicate their preferred version, not knowing which scenes were created by which tools. For every criterion—asset selection, layout coherence, and overall preference—the students consistently rated the environments generated by Holodeck more favorably.

The researchers also tested Holodeck's ability to generate scenes that are less typical in robotics research and more difficult to manually create than apartment interiors, like stores, public spaces, and offices. Comparing Holodeck's outputs to those of ProcTHOR, which were generated using human-created rules rather than AI-generated text, the researchers found once again that human evaluators preferred the scenes created by Holodeck. That preference held across a wide range of indoor environments, from science labs to art studios, locker rooms to wine cellars.

Finally, the researchers used scenes generated by Holodeck to "fine-tune" an embodied AI agent. "The ultimate test of Holodeck," says Yatskar, "is using it to help robots interact with their environment more safely by preparing them to inhabit places they've never been before."

Across multiple types of virtual spaces, including offices, daycares, gyms and arcades, Holodeck had a pronounced and positive effect on the agent's ability to navigate new spaces.

For instance, whereas the agent successfully found a piano in a music room only about 6% of the time when pre-trained using ProcTHOR (which involved the agent taking about 400 million virtual steps), the

agent succeeded over 30% of the time when fine-tuned using 100 music rooms generated by Holodeck.

"This field has been stuck doing research in residential spaces for a long time," says Yang. "But there are so many diverse environments out there—efficiently generating a lot of environments to train robots has always been a big challenge, but Holodeck provides this functionality."

In June, the researchers will present Holodeck at the 2024 Institute of Electrical and Electronics Engineers (IEEE) and Computer Vision Foundation (CVF) Computer Vision and Pattern Recognition (CVPR) Conference in Seattle, Washington.

More information: Yue Yang et al, Holodeck: Language Guided Generation of 3D Embodied AI Environments, *arXiv* (2023). DOI: 10.48550/arxiv.2312.09067

GitHub: yueyang1996.github.io/holodeck/

Provided by University of Pennsylvania