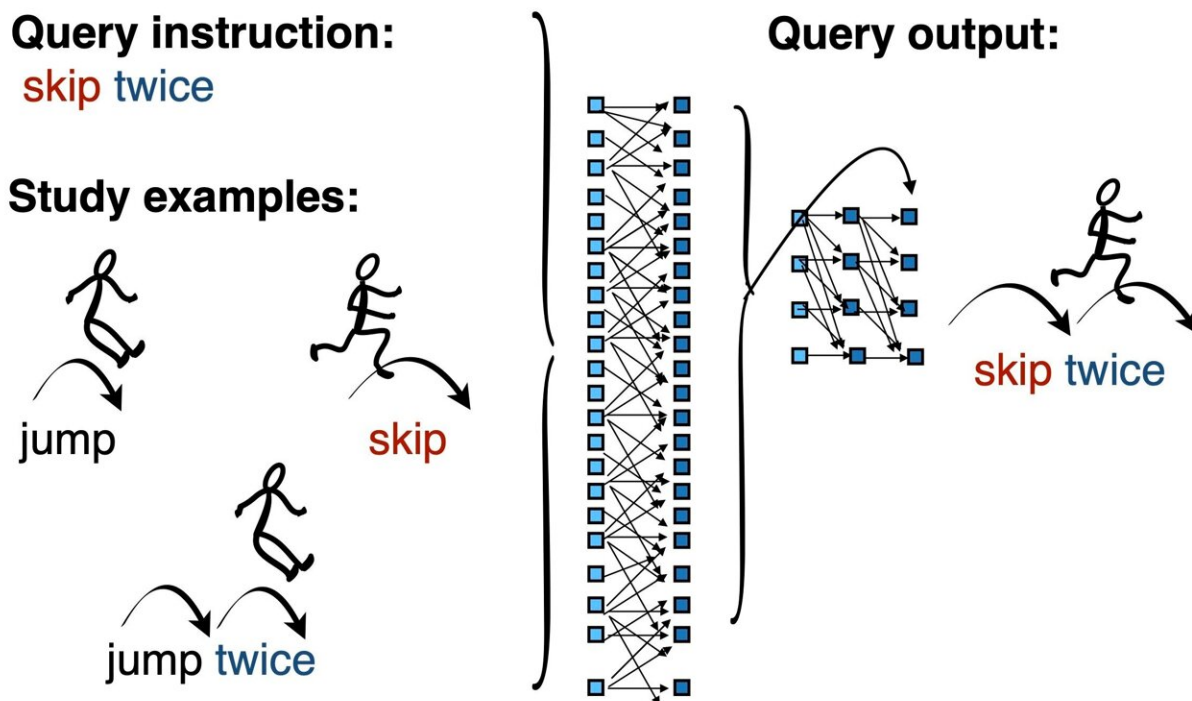


Can AI grasp related concepts after learning only one?

October 25 2023



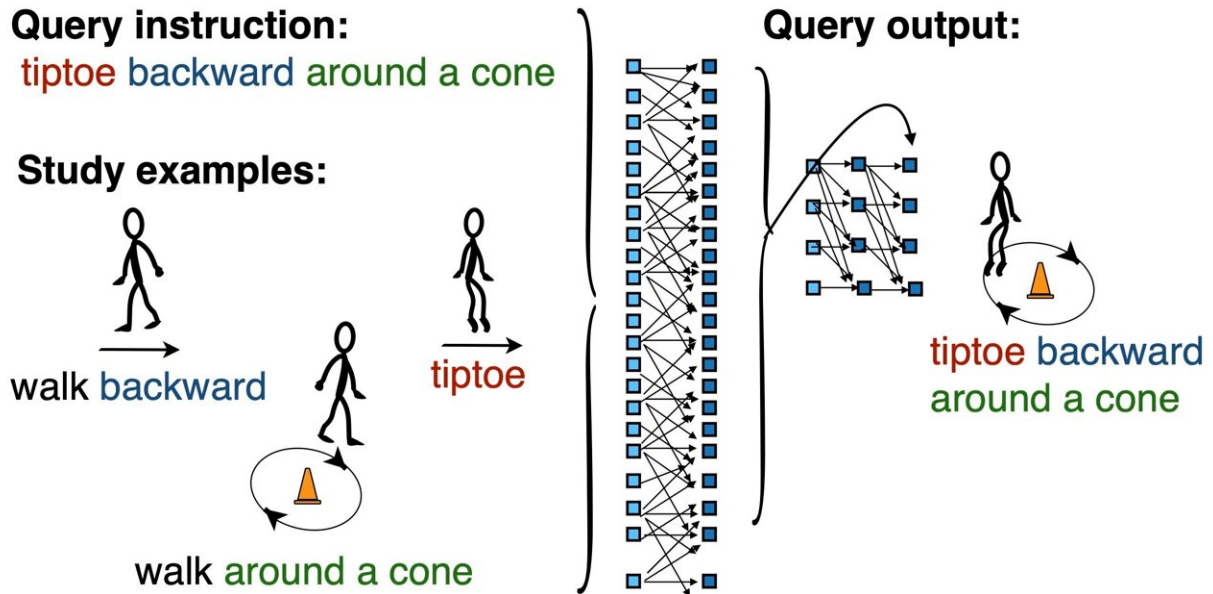
The Meta-Learning for Compositionality (MLC) approach for acquiring compositional skills in neural networks. During a training episode, MLC receives an example of a new word ("skip") and is asked to use it compositionally via the query instruction ("skip twice"). MLC answers the query instruction with an output sequence (symbolic arrows guiding the stick figures), which is compared to a desired target sequence. MLC makes improvements to its parameters. After this episode, a new episode presents another new word, and so on. Previous models, which do not explicitly practice their compositional skills, struggle to learn and use new words compositionally. However, after training, MLC succeeds. Credit: Brenden Lake

Humans have the ability to learn a new concept and then immediately use it to understand related uses of that concept—once children know how to "skip," they understand what it means to "skip twice around the room" or "skip with your hands up."

But are machines capable of this type of thinking? In the late 1980s, Jerry Fodor and Zenon Pylyshyn, philosophers and cognitive scientists, posited that [artificial neural networks](#)—the engines that drive artificial intelligence and [machine learning](#)—are not capable of making these connections, known as "compositional generalizations." However, in the decades since, scientists have been developing ways to instill this capacity in neural networks and related technologies, but with mixed success, thereby keeping alive this decades-old debate.

Researchers at New York University and Spain's Pompeu Fabra University have now developed a technique—[reported](#) in the journal *Nature*—that advances the ability of these tools, such as ChatGPT, to make compositional generalizations.

This technique, Meta-learning for Compositionality (MLC), outperforms existing approaches and is on par with, and in some cases better than, [human performance](#). MLC centers on training neural networks—the engines driving ChatGPT and related technologies for [speech recognition](#) and [natural language](#) processing—to become better at compositional generalization through practice.



The Meta-Learning for Compositionality (MLC) approach for acquiring compositional skills in neural networks. During a training episode, MLC receives an example of a new word ("tiptoe") and is asked to use it compositionally via the query instruction ("tiptoe backward around a cone"). MLC answers the query instruction with an output sequence (symbolic arrows guiding the stick figures), which is compared to a desired target sequence. MLC makes improvements to its parameters. After this episode, a new episode presents another new word, and so on. Previous models, which do not explicitly practice their compositional skills, struggle to learn and use new words compositionally. However, after training, MLC succeeds. Credit: Brenden Lake

Developers of existing systems, including large language models, have hoped that compositional generalization will emerge from standard training methods, or have developed special-purpose architectures in order to achieve these abilities. MLC, in contrast, shows how explicitly practicing these skills allow these systems to unlock new powers, the authors note.

"For 35 years, researchers in [cognitive science](#), [artificial intelligence](#), linguistics, and philosophy have been debating whether neural networks can achieve human-like systematic generalization," says Brenden Lake, an assistant professor in NYU's Center for Data Science and Department of Psychology and one of the authors of the paper. "We have shown, for the first time, that a generic neural network can mimic or exceed human systematic generalization in a head-to-head comparison."

In exploring the possibility of bolstering compositional learning in [neural networks](#), the researchers created MLC, a novel learning procedure in which a neural network is continuously updated to improve its skills over a series of episodes. In an episode, MLC receives a new word and is asked to use it compositionally—for instance, to take the word "jump" and then create new word combinations, such as "jump twice" or "jump around right twice." MLC then receives a new episode that features a different word, and so on, each time improving the network's compositional skills.

Training instructions

input: sequence of artificial words
output: sequence of action symbols (colored circles)

dax ●	lug kiki wif ●●
wif ●	dax kiki lug ●●
lug ●	lug fep kiki wif ●●●●
zup ●	wif kiki dax blicket lug ●●●●
lug fep ●●●	lug kiki wif fep ●●●●
dax fep ●●●	wif blicket dax kiki lug ●●●●
lug blicket wif ●●●	
wif blicket dax ●●●	

Test instructions

input: sequence of artificial words
output: generated by people/neural network (correct answers shown)

zup fep ●●●	zup fep kiki lug ●●●●
zup blicket lug ●●●	wif kiki zup fep ●●●●
dax blicket zup ●●●	lug kiki wif blicket zup ●●●●
zup kiki dax ●●	zup blicket wif kiki dax fep ●●●●●
wif kiki zup ●●	zup blicket zup kiki zup fep ●●●●●

Humans and the MLC model were compared head-to-head on the same task. This few-shot instruction learning task involves responding to instructions (sequence of artificial words) by generating sequences of abstract outputs (colored circles). Participants practiced on the training instructions (left) before

the evaluation on the test instructions (right). Human participants and MLC performed with similar accuracy and made similar mistakes, whereas ChatGPT made many more errors than people. Answer key: - "dax", "wif", "lug", and "zup" are input primitives that map to the output primitives RED, GREEN, BLUE, and YELLOW, respectively. - "fep" takes preceding word as an argument and repeats its output three times ("dax fep" is RED RED RED). - "blicket" takes both the preceding primitive and following primitive as arguments, producing their outputs in a specific alternating sequence ("wif blicket dax" is GREEN RED GREEN). - "kiki" takes both the preceding and following strings as input, processes them, and concatenates their outputs in reverse order ("dax kiki lug" is BLUE RED). Credit: Brenden Lake

To test the effectiveness of MLC, Lake, co-director of NYU's Minds, Brains, and Machines Initiative, and Marco Baroni, a researcher at the Catalan Institute for Research and Advanced Studies and professor at the Department of Translation and Language Sciences of Pompeu Fabra University, conducted a series of experiments with human participants that were identical to the tasks performed by MLC.

In addition, rather than learn the meaning of actual words—terms humans would already know—they also had to learn the meaning of nonsensical terms (e.g., "zup" and "dax") as defined by the researchers and know how to apply them in different ways. MLC performed as well as the human participants—and, in some cases, better than its human counterparts. MLC and people also outperformed ChatGPT and GPT-4, which despite its striking general abilities, showed difficulties with this learning task.

"Large language models such as ChatGPT still struggle with compositional generalization, though they have gotten better in recent years," observes Baroni, a member of Pompeu Fabra University's Computational Linguistics and Linguistic Theory research group. "But

we think that MLC can further improve the compositional skills of large language models."

More information: Brenden Lake, Human-like systematic generalization through a meta-learning neural network, *Nature* (2023).
[DOI: 10.1038/s41586-023-06668-3](https://doi.org/10.1038/s41586-023-06668-3).
www.nature.com/articles/s41586-023-06668-3

Provided by New York University

Citation: Can AI grasp related concepts after learning only one? (2023, October 25) retrieved 9 May 2024 from <https://techxplore.com/news/2023-10-ai-grasp-concepts.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.