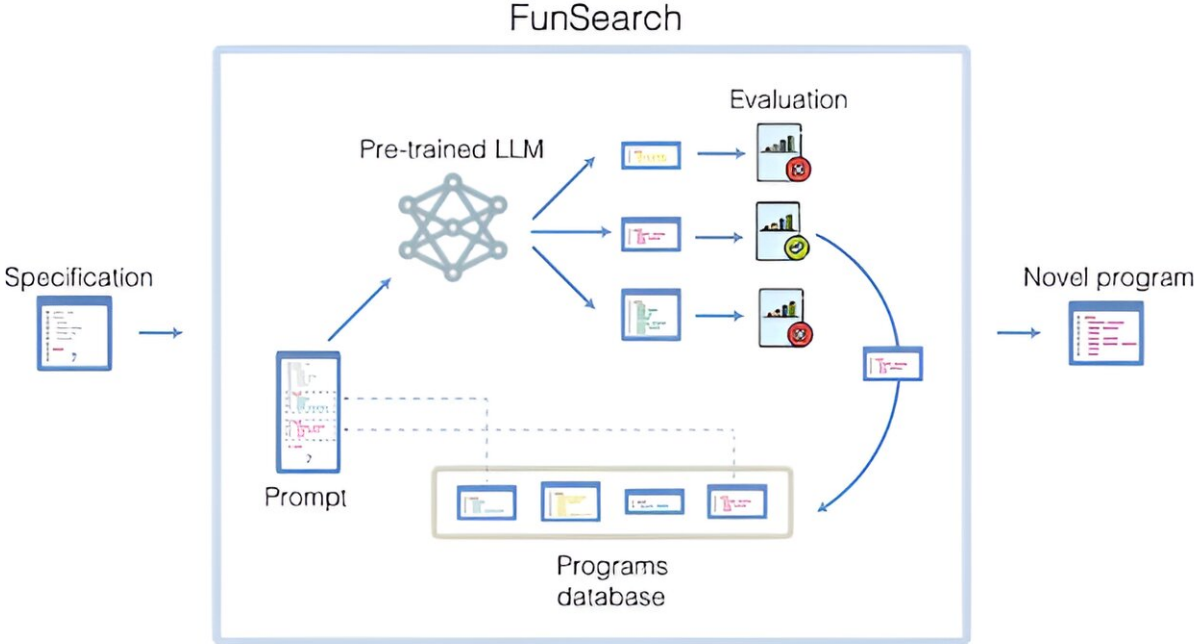


A means for searching for new solutions in mathematics and computer science using an LLM and an evaluator

December 15 2023, by Bob Yirka



The FunSearch process. The LLM is shown a selection of the best programs it has generated so far (retrieved from the programs database), and asked to generate an even better one. The programs proposed by the LLM are automatically executed, and evaluated. The best programs are added to the database, for selection in subsequent cycles. The user can at any point retrieve the highest-scoring programs discovered so far. Credit: Deepmind: <https://deepmind.google/discover/blog/funsearch-making-new-discoveries-in-mathematical-sciences-using-large-language-models/>

A team of computer scientists at Google's DeepMind project in the U.K., working with a colleague from the University of Wisconsin-Madison and another from Université de Lyon, has developed a computer program that combines a pretrained large language model (LLM) with an automated "evaluator" to produce solutions to problems in the form of computer code.

In their paper [published](#) in the journal *Nature*, the group describes their ideas, how they were implemented and the types of output produced by the new system.

Researchers throughout the scientific community have taken note of the things people are doing with LLMs, such as ChatGPT, and it has occurred to many of them that LLMs might be used to help speed up the process of scientific discovery. But they have also noted that for that to happen, a method is required to prevent confabulations, answers that seem reasonable but are wrong—they need output that is verifiable. To address this problem, the team working in the U.K. used what they call an automated evaluator to assess the answers given by an LLM.

After the LLM generates an answer, it is sent to the assessor. The assessor then analyzes the answer and then sends it back to the LLM with suggestions on how to improve its results. This process is repeated multiple times with the answer growing increasingly accurate. The research team calls their system FunSearch (short for functional space search). In testing the system, the researchers found that it was capable of providing verifiable results.

To further test FunSearch, the research teams used it to find new discoveries for what is known as the cap set problem—a math problem that involves discovering the largest set of points in a many-dimensional grid where no three points are on the same line. FunSearch was able to generate solutions that had not been found before—all in the form of

computer programs because of the nature of the LLM that they were using.

The research team acknowledges that FunSearch is not suitable for assisting in all types of research efforts, but suggests that it represents a step toward using LLMs to either find solutions to problems or to stimulate researchers looking for new ways to attack old problems.

More information: Bernardino Romera-Paredes et al, Mathematical discoveries from program search with large language models, *Nature* (2023). [DOI: 10.1038/s41586-023-06924-6](https://doi.org/10.1038/s41586-023-06924-6)

Deepmind: deepmind.google/discover/blog/...rge-language-models/

© 2023 Science X Network

Citation: A means for searching for new solutions in mathematics and computer science using an LLM and an evaluator (2023, December 15) retrieved 29 April 2024 from <https://techxplore.com/news/2023-12-solutions-mathematics-science-llm.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.