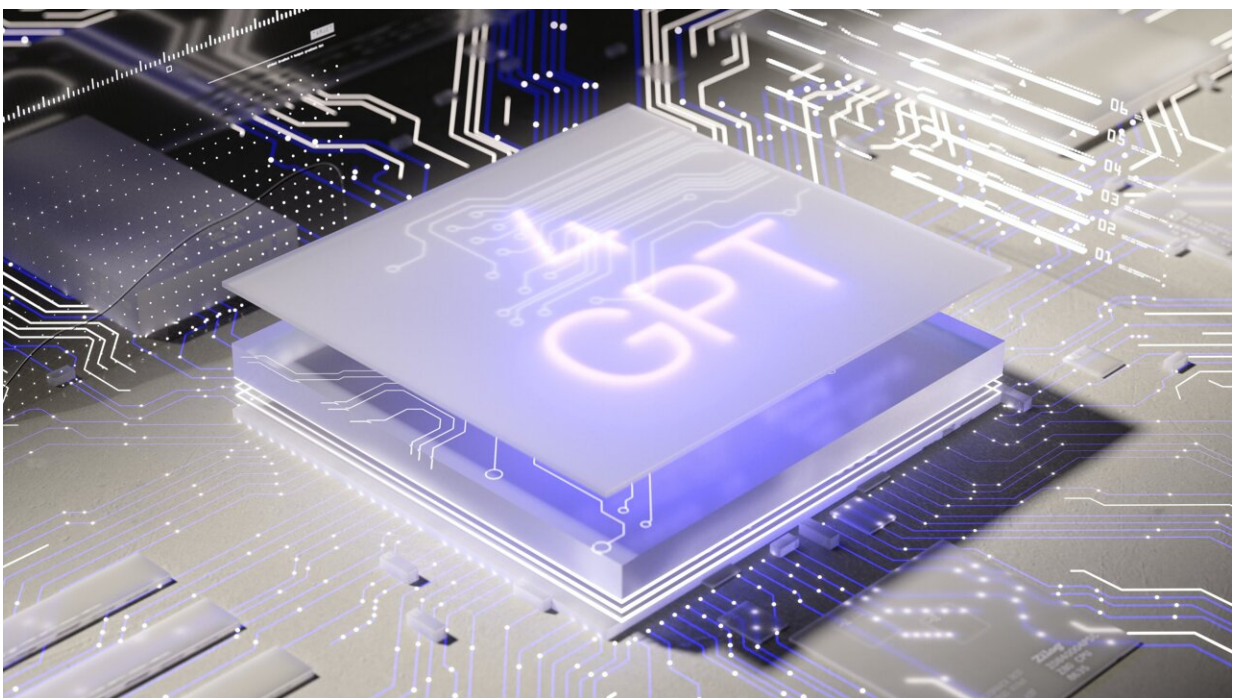


Why OpenAI developing an artificial intelligence that's good at math is such a big deal

December 4 2023, by Tom Oliver



Credit: Unsplash/CC0 Public Domain

With the [recent sacking and swift rehiring](#) of Sam Altman by OpenAI, debates around the development and use of artificial intelligence (AI) are once again in the spotlight. What's more unusual is that a prominent theme in media reporting has been the ability of [AI systems to do math](#).

Apparently, some of the drama at OpenAI was related to the company's development of a new [AI algorithm called Q*](#). The system has been talked about as a significant advance and one of its salient features was a capability to reason mathematically.

But isn't mathematics, the foundation of AI? How could an AI system have trouble with mathematical reasoning, given that computers and calculators can perform mathematical tasks?

AI is not a single entity. It's a patchwork of strategies for performing computation without direct instruction from humans. As we'll see, some AI systems are competent at math.

However, one of the most important current technologies, the [large language models](#) (LLMs) behind AI chatbots such as ChatGPT, has struggled so far to emulate mathematical reasoning. This is because they have been designed to concentrate on language.

If the company's new Q* algorithm can solve unseen mathematical problems, then that might well [be a significant breakthrough](#). Mathematics is an ancient form of human reasoning that [large language models \(LLMs\)](#) have so far struggled to emulate. LLMs are the technology that underlies systems such as [OpenAI's ChatGPT](#).

At the time of writing, the details of the Q* algorithm and its capabilities are limited, but highly intriguing. So there are various subtleties to consider before deeming Q* a success.

For example, math is a subject with which everyone engages to varying extents, and the level of mathematics that Q* is competent at remains unclear. However, there has been published [academic work](#) that uses alternative forms of AI to advance research-level mathematics ([including some written by myself](#)), and one written by a team of mathematicians in

collaboration with researchers at Google DeepMind).

These AI systems could be described as competent at math. However, it's likely that Q^* is not being used to help academics in their work but rather is intended for another purpose.

Nevertheless, even if Q^* is incapable of pushing the boundaries of cutting-edge research, there is very likely some significance to be found in the way it has been built that could raise tantalizing opportunities for future development.

Increasingly comfortable

As a society, we are increasingly comfortable with specialist AI being used to solve predetermined types of problem. For example, [digital assistants](#), [facial recognition](#), and [online recommendation systems](#) will be familiar to most people. What remains elusive is a so-called "[artificial general intelligence](#)" (AGI) that has broad reasoning capabilities comparable to those of a human.

Mathematics is a basic skill that we aspire to teach to every school child, and would surely qualify as a fundamental milestone in the search for AGI. So how else would mathematically competent AI systems be of help to society?

The mathematical mindset is relevant to a multitude of applications, for example coding and engineering, and so mathematical reasoning is a vital transferable skill for both human and artificial intelligence. One irony is that AI is, at a fundamental level, based upon mathematics.

For example, many of the techniques implemented by AI algorithms ultimately boil down to a mathematical area known as [matrix algebra](#). A matrix is simply a grid of numbers, of which a [digital image](#) is a familiar

example. Each pixel is [nothing more than numerical data](#).

Large language models are also inherently mathematical. Based on a huge sample of text, a machine can learn the probabilities for the words that are [most likely to follow a prompt \(or question\) from the user](#) to the chatbot. If you want a pre-trained LLM to specialize in a [particular topic](#), then it can be fine tuned on mathematical literature, or any other domain of learning. A LLM can generate text that reads as if it understands mathematics.

Unfortunately, doing so produces a LLM that is good at bluffing, but bad at detail. The issue is that a mathematical statement is, by definition, one that may be assigned an [unambiguous Boolean value](#) (that is, true or false). Mathematical reasoning amounts to the logical deduction of new mathematical statements from those previously established.

Devil's advocate

Naturally, any approach to [mathematical reasoning](#) that relies on linguistic probabilities is going to be driving outside its lane. One way around this could be to incorporate some system of formal verification into the architecture (exactly how the LLM is built), which continuously checks the logic behind the leaps made by the large language model.

A clue that this has been done could be in the name Q^* , which could plausibly refer to [an algorithm developed all the way back in the 1970s](#) to help with deductive [reasoning](#). Alternatively, Q^* could refer to Q-learning, in which a model can improve over time by testing for and rewarding conclusions that are correct.

But several challenges exist to building mathematically able AIs. For instance, some of the most interesting mathematics consists of highly unlikely events. There are many situations in which one may think that a

pattern exists based on small numbers, but it unexpectedly breaks down when one checks enough cases. This capability is difficult to incorporate into a machine.

Another challenge may come as a surprise: mathematical research can be highly creative. It has to be, because practitioners need to invent new concepts and yet stick within the [formal rules of an ancient subject](#).

Any AI methodology trained only to find patterns in pre-existing mathematics could presumably never create genuinely new mathematics. Given the pipeline between mathematics and technology, this seems to preclude the conception of new technological revolutions.

But let's play devil's advocate for a moment, and imagine whether AI could indeed create new mathematics. The previous argument against this has a flaw, in that it could also be said that the best human mathematicians were also trained exclusively on pre-existing [mathematics](#). Large language models have surprised us before, and will do so again.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: Why OpenAI developing an artificial intelligence that's good at math is such a big deal (2023, December 4) retrieved 28 April 2024 from <https://techxplore.com/news/2023-12-openai-artificial-intelligence-good-math.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.