

4 Technology

By CADE METZ

FOR decades, elite mathematicians have struggled to solve a collection of thorny problems posed by a 20th-century academic named Paul Erdős.

Last month, an artificial intelligence startup called Harmonic jumped into the mix. Harmonic said its AI technology, Aristotle, had solved an "Erdős problem" with help from a collaborator: OpenAI's latest technology, GPT-5.2 Pro.

For many computer scientists and mathematicians, solving an Erdős problem showed that AI had reached a point where it was capable of doing legitimate academic research. But some experts were quick to point out that the solution generated by AI was not very different from earlier work done by human mathematicians.

"It feels to me like a really clever student who has memorised everything for the test but doesn't have a deep understanding of the concept," said Terence Tao, a professor at UCLA, who is regarded by many as the finest mathematician of his generation. "It has so much background knowledge that it can fake actual understanding."

The debate over what Harmonic's system accomplished was a reminder of two consistent questions about the head-spinning progress of the tech industry's AI development: Did the AI system truly do something brilliant? Or did it merely repeat something that had already been created by brilliant humans?

The answers to those questions could provide a better understanding of the ways AI could transform science and other fields. Whether AI is generating new ideas or not – and whether it may one day do better work than human researchers – it is already becoming a powerful tool when placed in the hands of smart and experienced scientists.

These systems can analyse and store far more information than the human brain, and can deliver information that experts have never seen or have long forgotten.

Dr Derya Unutmaz, a professor at the Jackson Laboratory, a biomedical research institution, said the latest AI systems had reached the point where they would suggest a hypothesis or an experiment that he and his colleagues had not previously considered.

"That is not a discovery. It is a proposal. But it lets you narrow down where you should focus," said Unutmaz, whose research focuses on cancer and chronic diseases. "It allows you to do five experiments rather than 50. That has a profound, accelerating effect."

The excitement around GPT-5's math skills began in October when Kevin Weil, vice president of science at OpenAI, said on social media that the startup's technology had answered several of the mind-bending Erdős problems.

Designed as a way of measuring mathematical ingenuity, the Erdős problems are elaborate conjectures or questions that test the limits of the field. The aim is to prove whether each is right or wrong.

Some problems are enormously difficult to solve, while others are easier. One of the more famous problems asks: If the integer n is greater than or

Reinventing the wheel

equal to 2, can $4/n$ be written as the sum of three positive fractions? In other words, is there a solution to $4/n = 1/x + 1/y + 1/z$?

That problem is still unsolved. But on social media, Weil boasted that GPT-5 had cracked many others. "GPT-5 just found solutions to 10 (!) previously unsolved Erdős problems, and made progress on 11 others," Weil wrote. "These have all been open for decades."

Mathematicians and AI researchers quickly pointed out that the system had identified existing solutions buried in decades of research papers and textbooks. The OpenAI executive deleted his social media post. But even if the initial excitement was overstated, the technology had proved its worth.

"What it was able to do was surprising and useful," said Thomas Bloom, a mathematician

Does AI create new ideas, or just repeat old ones? ●

This visual is human-created, AI-aided

equal to 2, can $4/n$ be written as the sum of three positive fractions? In other words, is there a solution to $4/n = 1/x + 1/y + 1/z$?

That problem is still unsolved. But on social media, Weil boasted that GPT-5 had cracked many others. "GPT-5 just found solutions to 10 (!) previously unsolved Erdős problems, and made progress on 11 others," Weil wrote. "These have all been open for decades."

Mathematicians and AI researchers quickly pointed out that the system had identified existing solutions buried in decades of research papers and textbooks. The OpenAI executive deleted his social media post. But even if the initial excitement was overstated, the technology had proved its worth.

"What it was able to do was surprising and useful," said Thomas Bloom, a mathematician

at the University of Manchester who runs a website dedicated to the Erdős problems. "One of the papers it found was written in German. I never would have found it on my own."

So how has AI made such leaps since ChatGPT was introduced in late 2022? Today's leading AI systems are what scientists called neural networks, which can identify patterns in text, sounds and images and learn to generate this kind of material on their own, including term papers, computer code, voices and diagrams.

About 18 months ago, companies like OpenAI and Google started to hone their systems using a technique called reinforcement learning. Through this process, an AI system can learn behaviour through extensive trial and error.

The result is that these systems can now "reason" through problems in fields like math, science and computer programming. A

system like GPT-5 does not exactly reason like a human, but it can spend additional time working on a problem. Sometimes, this work extends for hours.

(The New York Times has sued OpenAI and Microsoft, claiming copyright infringement of news content related to AI systems. The two companies have denied the suit's claims.)

After Weil's social media post, researchers continued to ask GPT-5 and other AI technologies for solutions to Erdős problems.

Kevin Barreto and Liam Price, two mathematicians in Britain, used GPT-5 this month to crack a problem that was previously unsolved. Then they used Aristotle, the AI system from Harmonic, to verify that the solution was correct. Unlike GPT-5, Aristotle uses a specialised computer programming language to prove whether an answer is right or wrong.

The two mathematicians did play a small role. They nudged OpenAI's system in a new direction when its proof did not quite do what they wanted it to do. But like other experts, they believe that AI has now shown that it can do academic research.

"It's not a very high-level piece of research, but the fact that AI is capable of research of any level is impressive," Bloom said. If a graduate student showed him the same mathematical solution, he added, he would suggest that the student submit it to an academic journal for publication.

But even as some experts applauded what the two AI systems had done, others were less effusive. Tao said the solution was based on methods that were widely known.

"A consensus as to what the problem was even asking was only obtained in the last month so, which might explain why it had not been properly addressed in the literature before," he said.

Though he was impressed, Bloom added he still hadn't seen evidence that AI could generate ideas that humans couldn't. "And I would be surprised if that happened anytime soon," Bloom said.

Nonetheless, scientists say that AI has become a powerful and rapidly improving research tool and that whether it is generating ideas on its own is – for now – a moot point.

When Unutmaz uses AI for his research into chronic diseases, he said, he often feels like he is talking with an experienced colleague. But he acknowledges the machine cannot do its work without a human collaborator. An experienced researcher is still needed to repeatedly prompt the system, explain what it should be looking for and ultimately separate the interesting information from everything else the system produces.

"I am still relevant, maybe even more relevant," he said. "You have to have a very deep expertise to appreciate what it is doing." – ©2026 The New York Times Company

Harmonic's co-founders, Vladimir Tenev and Tudor Achim, in Palo Alto, California. — GABRIELA HASBUN/The New York Times

