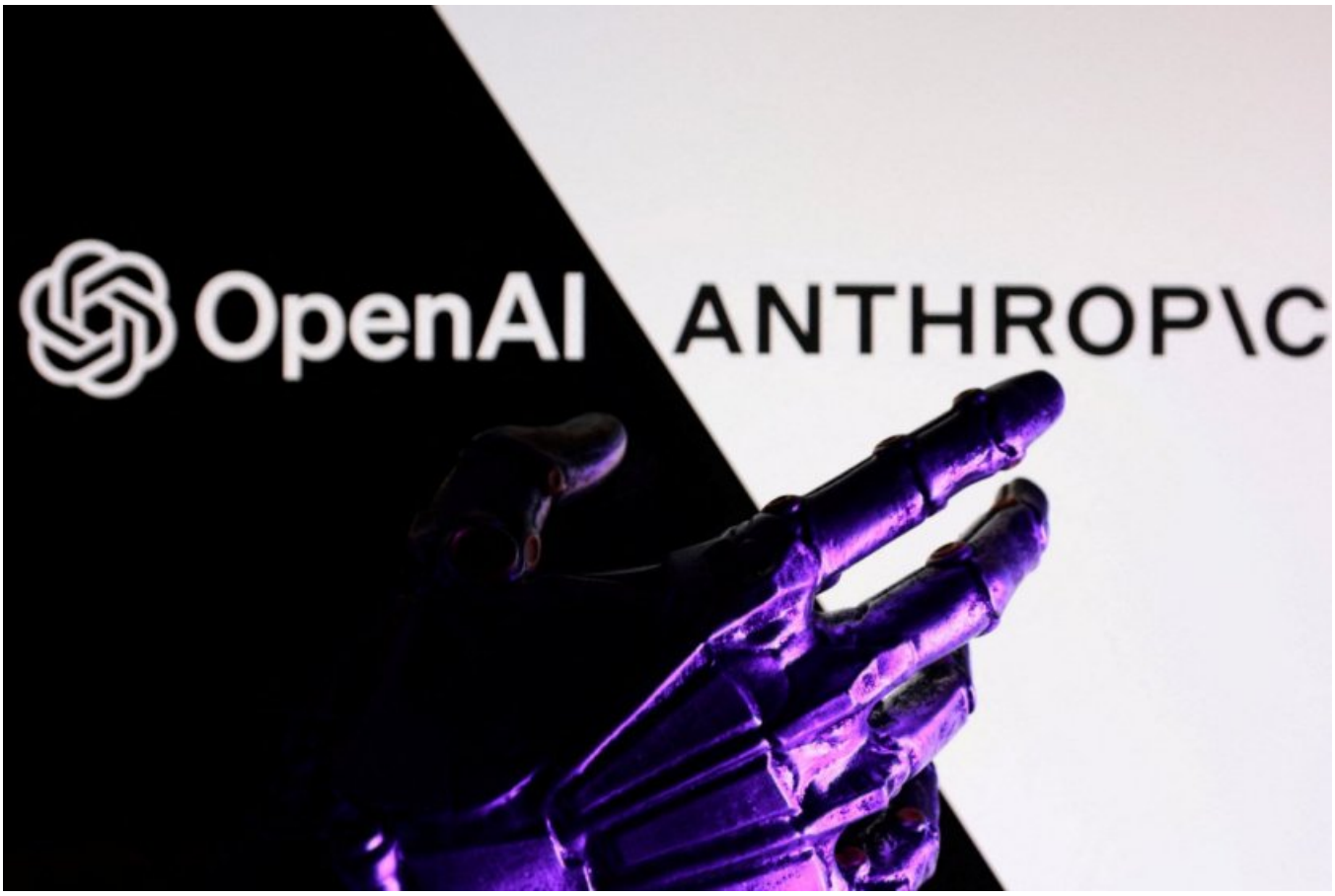
Two AI companies reveal 'Earth-shattering' solutions to major maths problem | New Scientist



OpenAI and Anthropic posted proofs of the VP versus VNP problem

Dado Ruvic/Reuters

Two AI companies have released "Earth-shattering" solutions to a problem related to P versus NP, considered one of the most difficult puzzles in mathematics. Coming a few days after a proof of the Riemann hypothesis, the achievement is a further sign that the pace of progress being made by AI in maths is unlike anything ever seen before.

P versus NP relates to how difficult some problems are to solve using a computer. P problems are relatively simple to solve, but NP problems quickly become drastically more difficult as their size increases. If P could be

proved to be the same as NP, then we would have a radically faster way to solve a host of computationally intensive problems, and there would be a dramatic effect on society: computer encryption would break, logistical and engineering problems could be highly optimised, and new drugs could be discovered more easily.

It is one of the seven Millennium Prize problems published by the Clay Mathematics Institute in 2000, which promise a $1 million award to anyone who can solve them. One of those, the Poincaré conjecture, fell in 2010. Two others have fallen to AI: the Navier-Stokes puzzle in September and the Riemann hypothesis last week.

The new AI results concern VP versus VNP, a problem posed after mathematicians and computer scientists realised the importance of P versus NP, but made little progress towards a solution. The idea was to replace the binary logic of computers with arithmetic, which imposes a more structured problem set. It is supposed to be a stepping stone along the way to P vs NP – not necessarily something that would lead directly to a solution, but something that would widen our understanding and help clear a path to more on P vs NP.

Yesterday, Anthropic quietly published its VP versus VNP finding on its GitHub page with no wider announcement. The work proves that VNP isn't contained within VP and that some classes of problems are simply difficult to solve no matter what. The company said the discovery was made by an internal research version of its Claude model, which also wrote the paper and a formalisation, a conversion of the proof to computer code that allows it to be methodically and logically picked apart and proved correct. "No human mathematician has checked the proof," the company said.

On the same day, Anthropic's rival OpenAI released a paper on GitHub setting out its own solution to the same problem, also including a formalisation. The two solutions are similar in some ways but diverge on key points. Neither company responded to a request for comment, but the

timestamps of the uploads suggest Anthropic beat OpenAI to the prize by about 6 hours.

Tom Gur at the University of Cambridge was surprised by the news and says he thought many other discoveries would be made before VP versus VNP was solved.

"If you'd have asked me in what now feels like ancient history, which would be, I don't know, August, I would have told you that I don't think I will live to see it," says Gur. "We are seeing here decades of progress in days. It's an exciting time to be alive."

Mathematicians argue about the relative difficulty of the two problems, and how far a solution to VP versus VNP gets us to a solution to P versus NP. "But one thing you can say for sure is that this is Earth-shattering progress," says Gur. "It's just so far beyond the range of techniques that we had before."

The difficulty facing mathematicians working on either puzzle – VP versus VNP or P versus NP – is that proving that a problem previously thought to be very hard is actually easy is straightforward: simply show an algorithm that solves it easily. But proving that a problem or set of problems is actually very hard, and will always remain so, requires evidence that there is no algorithm that can be written which can solve it easily.

"The reality is that we have a very, very, very limited arsenal of tools to prove this limitation [that some algorithms are intrinsically hard]," says Gur. The proof of VP versus VNP is an extremely powerful tool, but the question is how generalisable it is, he says. "I think we need, as a community, more time to study these results to understand."

Varun Kanade at the University of Oxford thinks the new proofs will help with – but not directly lead to – a solution for P versus NP. "It does get us closer in the sense that it makes progress on a problem related to the main

one, but my sense is that we will require new techniques for the P v NP question," he says. "At this point, I have to say that I'm not surprised that AI has made progress on this question, having seen the recent announcement of the proof of the Riemann hypothesis."

But Ben Livshits at Imperial College London says not everyone will be buoyed by the discoveries and the way they were achieved.

"AI is now doing research-level mathematics and theoretical computer science, and I don't think we've properly thought through the longer-term effects," says Livshits. "For theoretical computer scientists, many of whom have spent years or whole careers on these problems, it's a very mixed blessing, and several have already said so."