“The Longest Ever Proof”: Computer-Assisted Provers and Organising Pythagorean Triples

One of the many merits of mathematics is the ability to declare a statement true with certainty.

For millennia, mathemeticians have sought to prove these statements – often called theorems – using logical, step-by-step arguements, with the hopes of expanding the tree of knowledge that is at their disposal.

However, as maths advances into increasingly thorny areas, these proofs become incredibly elaborate; new breakthroughs require time, effort, and unique ways of thinking, often from many researchers working together. All this needs to be done while ensuring proofs are thorough and unblemished, yet clear and concise. So, it is easy to see how the concept of new methods of proving is enticing!

The future of proofs now looks towards computers; these machines have been shown to execute vast quantities of calculations incredibly quickly, and this can be utilized to try and create proofs in ways that humans couldn’t possibly do.

Automated theorem provers, or ATPs, are completely independent computer programs that often rely on their sheer calculation speed to try and deal with theorems and derive a proof. Other provers simply act as ‘proof assistants’, which have varying levels of independence and are used in conjunction with real mathmeticians. Between these two, there is a lot of room for computers to help out in discovering new proofs – with differing amounts of success so far.

We will look at one of the most prominent proofs to come from a computer-assisted proof, which, at the time, was dubbed the ‘largest ever’ by the journal *Nature*, due to its mammoth size of 200 terabytes. That’s over 250,000 times the size of the human genome!

“The theorem can be likened to a pearl, and the method of proof to an oyster. The pearl is prized for its luster and simplicity; the oyster is a complex living beast whose innards give rise to this mysteriously simple gem.”

- Douglas Hostadter, Godel Escher Bach: An Eternal Golden Braid

---

The Boolean Pythagorean Triples Problem

This problem considers Pythagorean triples – that is, positive integers a, b, c, that satisfy the equation $a^2 + b^2 = c^2$.
There are an infinite amount of Pythagorean triples (you can easily find multiples of any existing triple), which is where the Boolean Pythagorean triples problem arises;

Can you group all positive integers into two groups, so that no group contains a complete Pythagorean triple?

This can also be thought of as colouring all positive integers either red or blue, and assuring that no Pythagorean triple can be found that is all one colour. For example, in the Pythagorean triple $20^2 + 21^2 = 29^2$, all of these are valid colourings:

20 21 29
20 21 29

20 21 29
20 21 29

20 21 29
20 21 29

In the 1980s, mathmetician Ronald Graham (of "Graham's number" fame) pressed for a solution to this problem, offering a $100 prize. What many researchers now set out to do was use computing power to try and find the most integers you can colour in this system, or if it seems like there is a way to colour integers correctly forever. Simple enough, right?

A Numbers Game

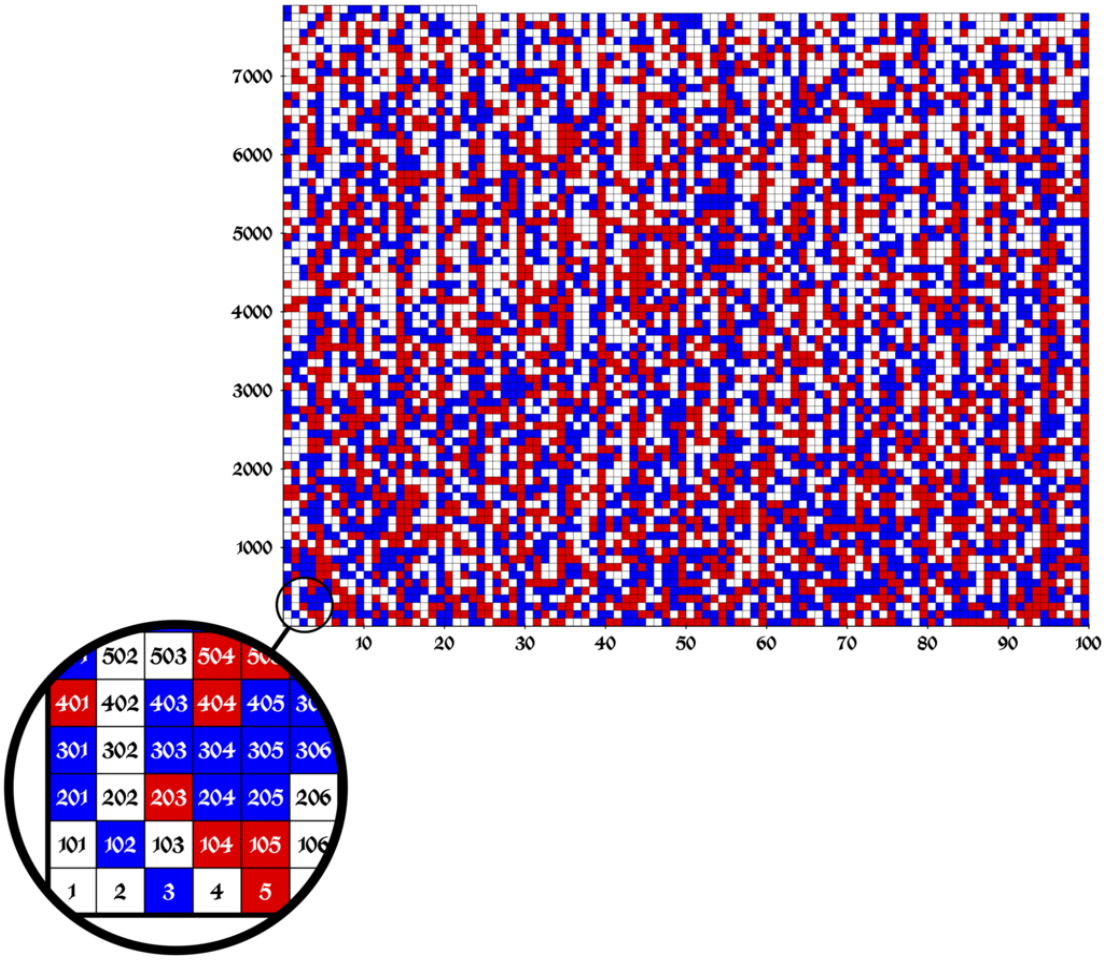
As there are two possible ways to colour each integer (red or blue), for the first n numbers, there would be 2^n possible colourings. This means that checking every single combination to try and solve this problem would grow incredibly difficult for large numbers. In 2015, researchers Cooper and Overstreet found a way to colour the first 7664 numbers correctly. This was the furthest they could go with the computational power available to them, showing just how challenging this problem is to solve in a brute-force style way.

In 2016, computer scientists Heule, Kullmann and Marek tried to find the greatest number you could go up to, before it was impossible to colour every number correctly. The process of finding the greatest number is much more difficult: instead of just needing to find one valid colouring for a given number, you need to test *all possibilities* to say with certainty that it is impossible. However, they could make this process significantly more managable, by using ideas from Ramsey theory to eliminate some of these possibilities beforehand.

Ramsey theory considers how large a structure has to be before it can be grouped by a certain property. In our case, we are looking at the greatest number before a two colour grouping is impossible. Making this connection is an important one – once the problem had been translated into Ramsey theory, they used well-established theorems to eliminate certain possibilities. Other ideas can also be used; for example, if an integer isn't involved in any Pythagorean triples, the colour of it is negligible.

These ideas allowed the range of possibilities to be optimized, producing a ‘brute-reasoning’ method, that meant the researchers were able to examine a reasonable quantity of possibilities.

The researchers then translated the problem into Boolean, so that they could use a ‘Boolean satisfiability solver’ - a proof assistant capable of calculating the solutions to these sorts of problems very efficiently. So, they ran all the remaining combinations on the University of Texas’ *Stampede* supercomputer, and in a surprisingly anti-climactic two days, they had 200 terabytes of computer calculations, and within that, their solution.



The above diagram shows the greatest possible valid solution of 7824, with the red and blue numbers coloured appropriately, and white numbers being those that can be coloured with either (many being the numbers that are not found in any Pythagorean triples). However, if you try to expand this solution to 7825, you encounter this:

$$625^2 + 7800^2 = 7825^2$$

$$5180^2 + 5865^2 = 7825^2$$

This means the number 7825 cannot be coloured red or blue, and hence, 7824 is our limit.

Bigger quesitons

This 2016 paper successfully solved the problem, and Ronald Graham was able to award Heule with his $100 prize (I'm not sure how much of the electricity bill that covered).

However, many mathmeticians were unsatisfied with this solution, with some reluctant to label it as a proof. As mentioned earlier, the clarity and concise nature of proofs is highly valued, and this method runs heavily against those ideals. It also doesn't provide us with any insight into future questions; What is the significance of the number 7825? What would happen if you had three colours to work with, instead of just two? These are the sorts of intriguing follow-up questions you are used to asking when faced with a proof, but this paper has not given any insight into their answers.

This sort of contention is not unfamiliar. In 1998, Thomas Hales presented a computer-assisted proof to the Kepler conjecture, where mathmetician and science-fiction writer Ian Stewart compared it to Andrew Wiles' proof of Fermat's last theorem, stating: "Wiles' proof of Fermat resembles War and Peace, but Hales' proof of Kepler resembles a telephone directory". It is important for mathematical proofs to be checked over and verified, but computer-assisted proofs, especially those that fall into the category of ATPs, can take several years to be verified by humans.

There is another potential risk if computers are the primary contributors to proofs and papers. Who would a mathmetician consult to ask questions about the paper, or to explain esoteric or complex ideas in that area?

A potential way to circumvent this issue uses neural networks, that could analyse existing huma-written proofs, to produce more understandable works themselves. This could allow us to harness the computational power that computers have to create proofs, while still having the clarity of those written by experts.

There is still a long way to go, and these advancements may take decades before mathmeticians are able to professionally use them. Nevertheless, there are important ideas about what constitutes a proof that must be considered now, as we are already seeing many problems be solved with computational approaches. It is vital that the transparency of mathematics is maintained, and we must ensure that computer-assisted proofs do not obfuscate new research, but instead enrich it.

Sources

https://www.quantamagazine.org/how-close-are-computers-to-automating-mathematical-reasoning-20200827/https://plus.maths.org/content/brief-introduction-proofs

Copper and Overstreet, https://arxiv.org/abs/1505.02222

Heule, Kullmann and Marek https://arxiv.org/abs/1605.00723

Nature, https://doi.org/10.1038/nature.2016.19990