



Machine-Human Collaborations Accelerate Math Research

Susan D'Agostino

One might think that earning a Fields Medal would quiet any mathematician's nagging doubts about their proof-writing abilities moving forward. But soon after Peter Scholze earned the 2018 award,¹ he questioned one part of the proof of the liquid vector space theorem he co-authored with Dustin Clausen. With humility, Scholze reached out to the math community.

"I spent much of 2019 obsessed with the proof of this theorem, almost going crazy over it," Scholze wrote on the Xena Project blog² in December 2020 about a black box of functional analysis contained within the proof. "I think nobody else has dared to look at the details of this, and so I still have some small lingering doubts. ... this may be my most important theorem to date ... Better be sure it's correct."

Indeed, Adam Topaz, assistant professor of mathematics at the University of Alberta, confessed with nervous

laughter that he and his study group had skipped the proof of the theorem when they read the work in 2020.

"He's a very convincing person," Topaz said during a talk at the February 2023 Machine Assisted Proofs conference³ held at UCLA's Institute for Pure and Applied Mathematics. "If he says something is true, people tend to trust him."

That's why Scholze and a team led by Johan Commelin of the University of Freiburg in Germany launched a large, collaborative project to verify the proof with the free, open-source, Lean Proof Assistant. The community-driven software was developed⁴ principally by Leonardo de Moura when he was at Microsoft Research and is supported by a mathematical library built by the Lean community. Topaz and other mathematicians and computer scientists joined Commelin's effort. By May 2021, they hit their first target, and by July 2022, they completed the project, which required 98,000 lines of code, according to Topaz.

"I find it absolutely insane that interactive proof assistants are now at the level that, within a very reasonable time span, they can formally verify difficult original research," Scholze wrote on his blog.

Susan D'Agostino is a mathematician and Spencer Journalism Fellow at Columbia University. She may be reached through her website: <https://www.susandagostino.com>.

For permission to reprint this article, please contact: reprint-permission@ams.org.

DOI: <https://doi.org/10.1090/noti2780>

¹<https://www.britannica.com/science/Fields-Medal>

²<https://xenaproject.wordpress.com/2020/12/05/liquid-tensor-experiment/>

³<https://www.ipam.ucla.edu/programs/workshops/machine-assisted-proofs/?tab=schedule>

⁴<https://leanprover-community.github.io/>

Carefully checking the proof would have been a “non-starter” without the machine assistance, Topaz said. As a bonus to finding and fixing some inaccuracies and verifying the proof of the main theorem, Topaz reported that the group also proved some auxiliary statements, answered some questions from Scholze’s notes, and filled some gaps in Lean’s mathematics library, specifically in homological algebra, topology, and category theory.

“It was really a human project, not a robot project,” Topaz said.

Since late 2022, academe has been abuzz with talk of ChatGPT, GPT-4, and other AI writing tools that may change research and college writing. But the mathematics community is also grappling with disruption, as evidenced by human-machine collaborations in writing proofs of difficult theorems at the cutting edge of research mathematics. Machines not only verify mathematical results that humans discover but help mathematicians reason, explore new ideas, and even learn new mathematics.

Automated and Interactive Methods

Formal methods, a subfield of computer science, uses techniques based in computational logic to assist with mathematical reasoning, according to Jeremy Avigad, professor of philosophy and mathematical science at Carnegie Mellon University. Within the subfield, there are automated methods and interactive methods. Broadly speaking, in the former, a human asks a question, pushes a button, and gets an answer—or an “I don’t know,” which can leave the mathematician stuck. In the latter, the user works with the computer to explore or check reasoning. That leaves room for possibility.

“It’s early and we’re still figuring out what the technology can do,” Avigad said. “But there’s a lot of promise and a lot of excitement.”

When a mathematician works with an interactive proof assistant, they might “teach” it about a new area of mathematics by entering the basic definitions and theorems. They might later “ask” the machine what it knows about these objects. When formalizing a proof, they will try to get the computer to determine whether something follows from previous statements, providing more information if necessary. In this way, machines and humans collaborate to fill in details and nudge the effort toward rigor.

But distinguishing between automated and interactive methods is something of a false dichotomy, according to Avigad.

“Even if you’re doing math interactively, you want to automate the tedious parts as much as possible,” Avigad said. “And even if you’re using automation, whenever you try something that doesn’t work, you make a change and try again. So, the two sides of the field kind of grow together.”

Michael Kinyon, professor of mathematics at the University of Denver, also blurs this boundary when using the tools. That is, he uses automatic theorem provers, but he uses them as proof assistants. The work has shifted his approach to research questions.

“We may be a little quicker, for example, to try out something using the software first before sitting and staring off into space and trying to figure it out on their own,” Kinyon said. “There’s more of a willingness to be experimental in the early stages of these things.”

That computers “think” differently than humans facilitates the discovery process for humans, according to some.

“That breaks us out of our patterns and deepens our understanding,” Jordan Ellenberg, professor of mathematics at the University of Wisconsin, said. “You can think of formalization as teaching a mathematical idea to a machine, and the machine’s ‘mind’ is very different from ours.” In 2016, Ellenberg and Dion Gijswijt, professor of mathematics at the Delft University of Technology in the Netherlands, solved the cap set conjecture, a result that was later published⁵ in the *Annals of Mathematics*.

Later, in 2019, mathematician Sander Dahmen and computer scientists Johannes Hölzl and Robert Lewis—all from Vrije Universiteit Amsterdam—used the Lean theorem prover to formalize⁶ the proof of Ellenberg and Gijswijt’s theorem. The authors deemed the combinatorial background necessary to formalize the proof as “less intimidating” than proofs from other math subfields. This formalization work provided some reassurance that, under favorable circumstances, one could computer-check a recent paper from the *Annals of Mathematics*. But a vanishingly small proportion of papers in this esteemed journal include short proofs with “less intimidating” prerequisites. Enthusiasts can hope that the software will improve over time.

On Elegance

Some may wonder whether machines reduce proofs to computations. But such reductionism is not at the heart of this work, according to Heather Macbeth, assistant professor of mathematics at Fordham University. That’s because humans write proofs for understanding, and a computation may not offer that.

“You may think that when you move from paper to the computer, we throw [beauty] out the window, that we have decided that we’re going to embrace function over form, and we’re really going to move to a world in which what matters is getting the proof done,” Macbeth said. “That’s not the case.” (Macbeth spoke at the Institute for Pure and Applied Mathematics conference.)

⁵<https://annals.math.princeton.edu/2017/185-1/p08>

⁶<https://drops.dagstuhl.de/opus/vol11texte/2019/11070/>

Macbeth has spent three years writing thousands of lines of Lean code in an effort to formalize proofs. For two years, she has also collaborated with others to maintain the Lean mathematical library. In this role, she has reviewed thousands of lines of code written by others to assess whether the code can be improved and to ensure that it fits with the rest of the library.

"Some of it, I felt, was really beautiful, and some of it I haven't," Macbeth said. "Almost all principles for writing good mathematics on paper extend to the corresponding principles in formal mathematics."

Arguments contained within automated proofs do not always mimic patterns of human thought, according to Macbeth.

"There is such a thing as a good proof that, nonetheless, is not exactly the proof a human would have thought of," Macbeth said. Humans, she observed, in an attempt to be efficient, may prune a search space at various steps in the search for a proof. A machine, however, may check more cases, including those that initially may have seemed less than promising.

"This is really a question about where there are differences and why those differences are interesting," Macbeth said.

Also, some proofs gain elegance during formalization. Thomas Browning and Patrick Lutz, both mathematics graduate students at the University of California, Berkeley, undertook an effort to formalize Galois theory—a subfield many already deem elegant. But at one point in their work, they needed to work with arbitrary finite sequences, which are not among Lean's strengths, according to Browning and Lutz. So, they found a workaround.

"This maneuver is mostly just a way to avoid having to deal with certain types of arguments that don't work very smoothly in Lean," Browning and Lutz wrote in their paper. "But it does have the added benefit that some standard proofs become simpler when rewritten using this new induction principle."

An Accelerating Trend

Decades ago, machine-assisted proofs were viewed as fringe projects by mainstream mathematicians, according to Josef Urban, a distinguished researcher at the Czech Institute of Informatics, Robotics, and Cybernetics in Prague.

"High-profile mathematicians said that people like Tom Hales were wasting their talent by doing the proof of the Kepler conjecture formally," Urban said of the mathematician who, after responding to frustrations with the usual refereeing process, offered a machine-verified proof. "That really changed in the last 15 years, thanks especially to Hales."

To be sure, machine assisted proofs still have naysayers, including Michael Harris, professor of mathematics at

Columbia University. Mathematicians who use theorem provers must learn to code and express problems in terms computers understand, which reduces time spent doing math.

"By the time I've reframed my question into a form that could fit into this technology, I would have solved the problem myself," Harris told⁷ *Quanta Magazine*.

Doron Zeilberger, professor of mathematics at Rutgers University, is also not a fan, though his distaste begins with human-written proofs and extends to machine-assisted proofs.

"We don't need more 'formal versions' of human-generated mathematics," Zeilberger wrote on his blog⁸ last year, while noting his appreciation for the pioneering work of Avigad and others. "Formalizing known proofs is not unlike Pig-Latin. Once you have done it a couple of times, it is no fun anymore. While it is both intellectually and technically challenging, or else such brilliant people would not engage in it, these people are wasting their talents."

"Indeed, our silicon servants, soon to become our masters, can be used much more fruitfully," Zeilberger continues. "Coq [another interactive theorem prover] and Lean continue the pernicious *Greek tradition*, that introduced the axiomatic method and made mathematics a deductive, logic-centric science."

But those views appear to go against the artificial-intelligence headwinds that appear to be driving society in the current moment, including the math community.

Kevin Buzzard, professor of pure mathematics at Imperial College London, gave⁹ a talk, "The Rise of Formalism in Mathematics," at the 2022 International Congress of Mathematicians. There, he dubbed the mood not "the beginning of the end" but the "beginning of the beginning." In his view, mathematical researchers' jobs are not at risk. Machines, however, will increasingly help humans prove mathematical theorems not only by working out examples but also with reasoning. Computers will also help humans find proofs or counterexamples in databases, construct simple proofs, and generally make it easier for humans to learn mathematics, according to Buzzard.

The conference circuit appears to support this claim. In addition to the machine-assisted proofs conference at the Institute for Pure and Applied Mathematics held in February, other respected institutions are devoting time and space this year for mathematicians to gather and discuss this accelerating trend.

In May, the Banff International Research Station for Mathematical Innovation and Discovery hosted a

⁷<https://www.quantamagazine.org/how-close-are-computers-to-automating-mathematical-reasoning-20200827/>

⁸<https://sites.math.rutgers.edu/~zeilberg/Opinion184.html>

⁹<https://www.youtube.com/watch?v=SEID4XYFN7o>

Formalization of Cohomology Theories workshop¹⁰ that brought together formalization experts and subject matter experts to make strides towards cutting-edge research. In June, the Simons Laufer Mathematical Sciences Institute introduced graduate students to the technology and ideas behind it in a Formalization of Mathematics summer graduate school program.¹¹ In July, MIT's Lorentz Center held a Machine-Checked Mathematics workshop¹² targeting mathematicians who have heard about, but not yet tried, the technology.

Outside of the United States, in June, the University of Copenhagen offered a masterclass¹³ called Formalisation of Mathematics to build on the Scholtz's liquid tensor experiment. In July, the Lorentz Center in the Netherlands offered¹⁴ Machine-checked Mathematics, a week-long introductory workshop for interested mathematicians intended to spur collaboration. Also in the summer, the Hausdorff Research Institute for Mathematics offered a program,¹⁵ Prospects of Formal Mathematics, to provide a forum for experts and junior researchers to gather, collaborate, and "interface them better with the mathematical mainstream." The same institution offered¹⁶ a week-long lecture series,¹⁷ Formal Mathematics and Computer-Assisted Proving, in September. Many more may be found on the Lean Community website.¹⁸ Also, Lean is not the only interactive theorem prover. Other systems include Coq, Isabelle/HOL, HOL Light, Agda, cubical Agda, Metamath, and Mizar, according to Buzzard.

"They digitized music—CD, MP3—but at the time nobody foresaw the consequences—Napster, Spotify," Buzzard said during his talk. "We're digitizing mathematics, and I believe this will inevitably change mathematics. You are welcome to join us."



Susan D'Agostino

Credits

Photo of Susan D'Agostino is courtesy of Chris Keeley.

¹⁰<https://www.birs.ca/events/2023/5-day-workshops/23w5124>

¹¹https://www.msri.org/summer_schools/1021

¹²<https://www.lorentzcenter.nl/machine-checked-mathematics.html>

¹³<https://www.math.ku.dk/english/calendar/events/formalisation-of-mathematics/>

¹⁴<https://www.lorentzcenter.nl/machine-checked-mathematics.html>

¹⁵<https://www.him.uni-bonn.de/programs/future-programs/future-trimester-programs/prospects-of-formal-mathematics/description/>

¹⁶<https://www.hsm.uni-bonn.de/events/hsm-schools/formal2023/description/>

¹⁷<https://www.hsm.uni-bonn.de/events/hsm-schools/formal2023/description/>

¹⁸<https://leanprover-community.github.io/events.html>