Questions Artificial Intelligence Raises for the Mathematics Profession

Version of January 30, 2024 prepared by Heather Macbeth and Emily Riehl

Below we list some questions that AI raises for the mathematics profession. We discuss the role of mathematicians in conversations about AI in the public sphere, the effects AI may have on the economics of the profession, and the role that AI may play in our non-research and research work lives.

What role can mathematicians play in the public conversation around AI?

We argue that mathematicians should play a role in this conversation because:
- We have a domain to offer in which AI success can be measured objectively: Mathematical text and particularly computer formalized mathematical text can be checked for correctness.
- We are more immune than other fields to being intimidated by mathematical or technical language.
- We have a mode of understanding that could be used to create clear standards for messy questions (fairness, explainability, attribution, appropriate use of data).

To illustrate the last point, we point to the role mathematicians have played in quantifying questions of gerrymandering.¹

We think mathematicians should participate in conversations concerning:
- The suitability of AI tools for a particular application in the public sphere.
- The assessment of biases in AI tools under development, and methods by such tools could be regulated.
- Questions on where training data comes from and quantifying attribution for AI generated work.

How might AI affect the economics of the mathematics profession?

The mathematics profession today benefits from the fact that mathematical skills are valued in numerous well-paid jobs from industry to academia. These jobs are the destinations for current and former undergraduate mathematics students, undergraduate mathematics majors, math grad students, postdocs, mathematics educators, and professional mathematicians.²

The decades-long boom in the tech industry has buoyed the mathematics profession as public investment has lessened. If tech work requires less of a mathematical mindset, how does this

---

¹ https://www.quantamagazine.org/how-math-has-changed-the-shape-of-gerrymandering-20230601/
affect the economic opportunities of mathematics students? And if student enrollments fluctuate, does this affect the size of mathematics departments?³

As AI tools for teaching and grading mathematics develop, will this affect the staffing levels required to teach mathematics courses? This could have far-reaching effects on educators at all levels, including the availability of mathematics jobs and graduate teaching stipends.

How can AI assist mathematicians with the non-research aspects of our work and should it?

We could envision AI tools contributing in the following way:

- In the drafting of papers: typesetting handwritten mathematics, drawing diagrams in LaTeX, fixing LaTeX bugs, drafting prose for introduction or background sections, transcribing dictated notes, helping with grammar and spelling (particularly for non-native speakers)⁴
- In improving the accessibility of the mathematics literature: by translating mathematical texts between languages, by transcribing spoken mathematics into text, by narrating written mathematics (particularly in LaTeX or involving diagrams)
- In the review of papers: summarizing the technical contents for editors in a different field, verifying correctness of mathematical arguments (if proofs are accompanied by computer formalizations), identifying related work⁵
- In hiring and admissions: filtered large application pools according to some committee-designed attributes
- In many aspects of teaching: assistance in grading, in lecture note preparation, in creating homework and exams (perhaps modifying for individual students or sections)
- In administrative work: creating documents for internal use, writing emails⁶

To be clear, we do not necessarily endorse all of these potential applications.

It’s quite likely that mathematicians will want to be directly involved in the development of some of these tools, in the way that LaTeX has always been partially developed by hobbyists. We hope for a vibrant ecosystem that is not dependent on a small number of big tech providers or expensive proprietary software.

As a community, we must have a proactive conversation about the ethics of the development of and uses of AI.

⁴ A few AI assistants for Overleaf are available on the chrome web store: see https://chromewebstore.google.com/detail/latex-ai/bcmccneppjkhpcpojebnholbghj or https://chromewebstore.google.com/detail/latex-ai/bcmccneppjkhpcpojebnholbghj
All of the hypothetical software mentioned above will have been trained on the output of countless hours of careful work by mathematics professionals. Does the AMS want to defend the collective output of mathematicians over the decades? What uses of AI software in decisions with human impacts (grading, hiring, admissions, peer review) are fair and appropriate? How do we compensate mathematicians (and tech companies) who develop useful tools for our community while ensuring their broad access? Should we develop norms concerning the disclosure of AI in writing that reflect nuances of different hypothetical use cases? How do we ensure proper attribution of mathematical ideas that are disseminated by AI?

What roles might AI play in mathematics research?

We can envision a few potential uses of AI in mathematics research. Others will undoubtedly emerge.

A computer proof assistant (such as Agda, Coq, HOL Light, Isabelle, Lean, …) is a software program that checks the correctness of mathematical arguments in these logical languages and may provide automation to help construct such arguments. Independent of developments in AI, researchers have been working in the last decades to make such computer proof assistants more powerful and user-friendly and mathematicians are increasingly adopting them to formally verify work in their own areas. Large language models (LLMs) can develop text, written in a conversational language but also in programming languages, and notably in the logical languages of computer proof assistants. One plausible pipeline is that LLMs could iteratively develop such proofs by incorporating the feedback on line-by-line correctness provided by proof assistants.

Some mathematicians are already using AI as a collaborator in mathematical discovery (as opposed to proof). This involves considerable work in developing traditional computational tools to provide the source data in which the AI searches for patterns. There is also art in framing the questions of interest to mathematicians in a way amenable for AI (eg as an error-tolerant optimization problem). Natural uses of AI also suggest themselves in certain needle in a haystack problems that involve searching for interesting mathematical objects with various properties; these can often naturally be framed as optimization problems.

---

7 At least one LLM explicitly acknowledges the arXiv as one source of its training data: [https://arxiv.org/abs/2302.13971](https://arxiv.org/abs/2302.13971)
8 [https://publicationethics.org/cope-position-statements/ai-author](https://publicationethics.org/cope-position-statements/ai-author)
There are essential contributors to these activities that the community may wish to explicitly recognize or support, doing the following mathematical tasks:

- Traditional computational mathematics (development and maintenance of databases, implementation of algorithms). These activities have struggled to secure funding and recognition, but they directly feed into AI-powered mathematics both as training data and for augmentation of reasoning capacity.\(^ {12}\)
- Development of automation tools for computer proof assistants: often this entails making precise patterns of reasoning that mathematicians consider “routine,” and these become used as building blocks in AI output of mathematics in the formal-proof format. There is a two-fold value to this automation: it makes AI tools more powerful but also more human-comprehensible.\(^ {13}\)
- Designing and maintaining libraries of formalized proofs, which both serve as training data and provide the prerequisites to bring (auto)formalization to the frontiers of mathematical research.\(^ {14}\)
- Administering and maintaining curated (and generally high-quality) sources of mathematical knowledge (arXiv, MathSciNet, Zentralblatt, MathOverflow, Stacks Project, nLab, Wikipedia, ….)

There is an important question of proper attribution and credit for all of these activities.

Mathematicians will likely want assistance in developing their understanding of and ability to use these tools, and to contribute to these related activities (as listed above). Does the AMS have a role here?

- The AMS could commission survey lectures, expository articles, mini courses, or blog posts for various venues. This is a fast-moving field with a vast literature, so care must be taken to avoid duplicating effort or producing static resources which date quickly.
- Mathematical reasoning is a prestigious, though small, subfield of AI and there is substantial work being done here, often in or jointly with private industry. How do mathematicians emphasize their own priorities and standards in the development of such systems? Will mathematicians without insider connections have an opportunity to assess and verify the capabilities of a new AI system?
- Right now there seems to be limited opportunities for outside mathematicians to interact with AI developers. Is there a way to incentivize tech companies or computer scientists to provide support for mathematicians who want to use AI in their research?

It seems possible that the availability of AI-powered computer proof assistants could change conventions surrounding the production of and dissemination of mathematical proof. This raises a number of important questions.

- To what extent are computer proof assistants to be trusted? To a large extent, we assume the “kernels” (the small core correctness-checking engines) to be bug-free based on community experience of their plausibility rather than direct inspection.

---

\(^ {12}\) For instance [https://oeis.org/](https://oeis.org/) or [https://simonscollab.iceri.brown.edu/data/](https://simonscollab.iceri.brown.edu/data/)
\(^ {13}\) For instance [https://leanprover-community.github.io/mathlib_docs/tactics.html](https://leanprover-community.github.io/mathlib_docs/tactics.html)
\(^ {14}\) See for example [https://leanprover-community.github.io/mathlib-overview.html](https://leanprover-community.github.io/mathlib-overview.html) or [https://unimath.github.io/agda-unimath/](https://unimath.github.io/agda-unimath/)
• To what extent are human-generated computer formalizations to be trusted? What standards are used to verify that the definitions function as intended?\textsuperscript{15}

• How do we ensure a healthy ecosystem of computer proof assistants and libraries of formalized mathematics?

• Computer proof assistants are a tool that foster large scale collaborations. Mathematicians have not traditionally had to grapple with the problem of credit-allocation in large collaborations the way that other scientific communities have.

• It is hard enough to allocate credit to new formalizations of new mathematical material, but even less visible is the additional work of integrating these contributions into a library and keeping it interoperable with the other contributions, through changes in the library’s idioms and abstractions.

• By what standards do we judge an AI-generated formal proof to be correct? There is a nightmare scenario of such systems operating as oracles only: producing million-line (claimed) proofs of important theorems which give no insight into the mathematical ideas involved. Can we instead develop AI-generating-formal-proof-tools which produce proofs which make transparent the insights involved?

How do we maintain mathematics as an open and collaborative endeavor, in the face of AI tools which can absorb, and then reproduce without credit, any idea mentioned publicly? The current convention is that we cite articles (and conversations, etc) for ideas, as well as for theorems we rely on literally. When our adaptation of an idea is mediated through AI output, we may never learn its original author.

\textsuperscript{15} https://leanprover-community.github.io/blog/posts/lte-examples/