

How to Worry about Mathematics in the Age of AI

Jeremy Avigad

Carnegie Mellon University



Institute for Computer-Aided
Reasoning in Mathematics

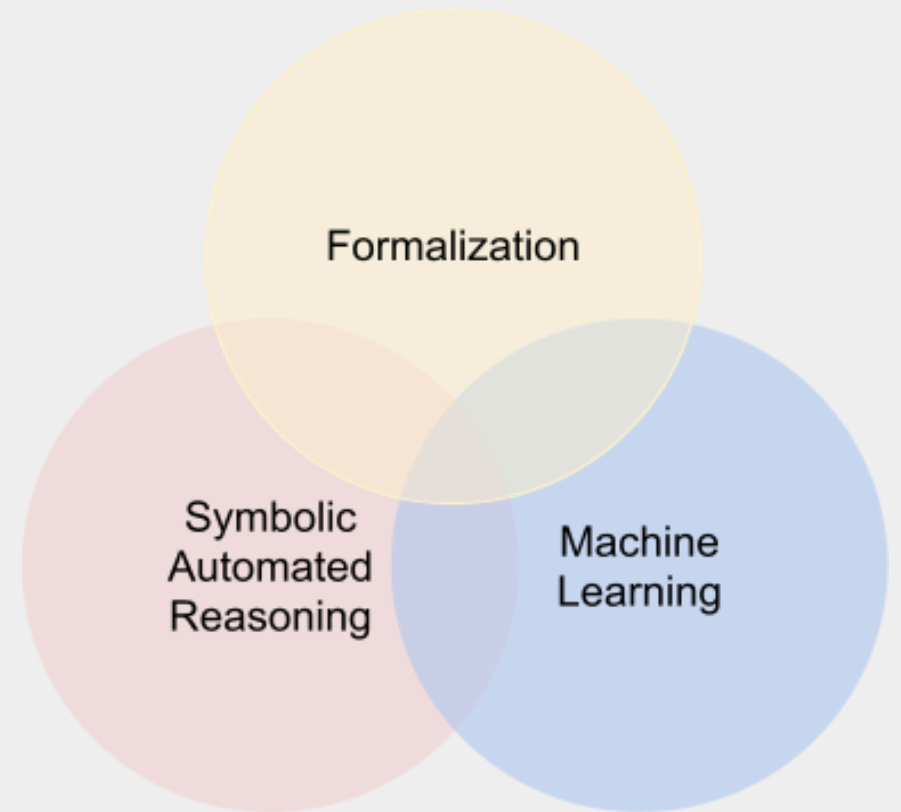
New technologies for mathematics

New reasoning technologies:

- interactive theorem proving and formalization
- automated reasoning and symbolic AI
- machine learning and neural AI

Call these, collectively, “AI for Mathematics.”

All three come together in neurosymbolic theorem proving.



New technologies for mathematics

Goals of this talk:

- survey the technologies (quickly!)
- raise some concerns

I am generally optimistic, but we should also be cautious and thoughtful.

Interactive theorem proving and formalization

- There are several proof assistants in use today (Coq, Isabelle, HOL Light, Lean, ...).
- Lean's Mathlib has more than two million lines of code.
- The Lean Zulip chat has 13K+ subscribers, 1,200+ active in any two-week period.
- There have been several notable formalizations and collaborations:
 - the liquid tensor experiment
 - the sphere eversion theorem
 - the polynomial Freiman-Ruzsa conjecture
 - improved upper bounds on Ramsey's theorem
 - Carleson's theorem
 - the FLT project

Building the Mathematical Library of the Future

24 |

A small community of mathematicians is using a software program called Lean to build a new digital repository. They hope it represents the future of their field.



Liquid tensor experiment

Posted on [December 5, 2020](#) by [xenaproject](#)

This is a guest post, written by Peter Scholze, explaining a liquid real vector space mathematical formalisation challenge. For a pdf version of the challenge, see [here](#). For comments about formalisation, see section 6. Now over to Peter.

1. The challenge

I want to propose a challenge: Formalize the proof of the following theorem.

Theorem 1.1 (Clausen-S.) *Let $0 < p' < p \leq 1$ be real numbers, let S be a profinite set, and let V be a p -Banach space. Let $\mathcal{M}_{p'}(S)$ be the space of p' -measures on S . Then*

$$\mathrm{Ext}_{\mathrm{Cond}(\mathrm{Ab})}^i(\mathcal{M}_{p'}(S), V) = 0$$

for $i \geq 1$.

 Comment

- Introduction
- 1 First part ▼
 - 1.1 Breen–Deligne data**
 - 1.2 Variants of normed groups
 - 1.3 Spaces of convergent power series
 - 1.4 Some normed homological algebra
 - 1.5 Completions of locally constant functions
 - 1.6 Polyhedral lattices
 - 1.7 Key technical result
- 2 Second part ►
- 3 Bibliography
- Section 1 graph
- Section 2 graph

1.1 Breen–Deligne data

The goal of this subsection is to give a precise statement of a variant of the Breen–Deligne resolution. This variant is not actually a resolution, but it is sufficient for our purposes, and is much easier to state and prove.

We first recall the original statement of the Breen–Deligne resolution.

Theorem(Breen–Deligne)

For an abelian group A , there is a resolution, functorial in A , of the form

$$\dots \longrightarrow \bigoplus_{i=1}^{n_i} \mathbb{Z}[A^{r_{ij}}] \longrightarrow \dots \longrightarrow \mathbb{Z}[A^3] \oplus \mathbb{Z}[A^2] \longrightarrow \mathbb{Z}[A^2] \longrightarrow \mathbb{Z}[A] \longrightarrow A \longrightarrow 0.$$

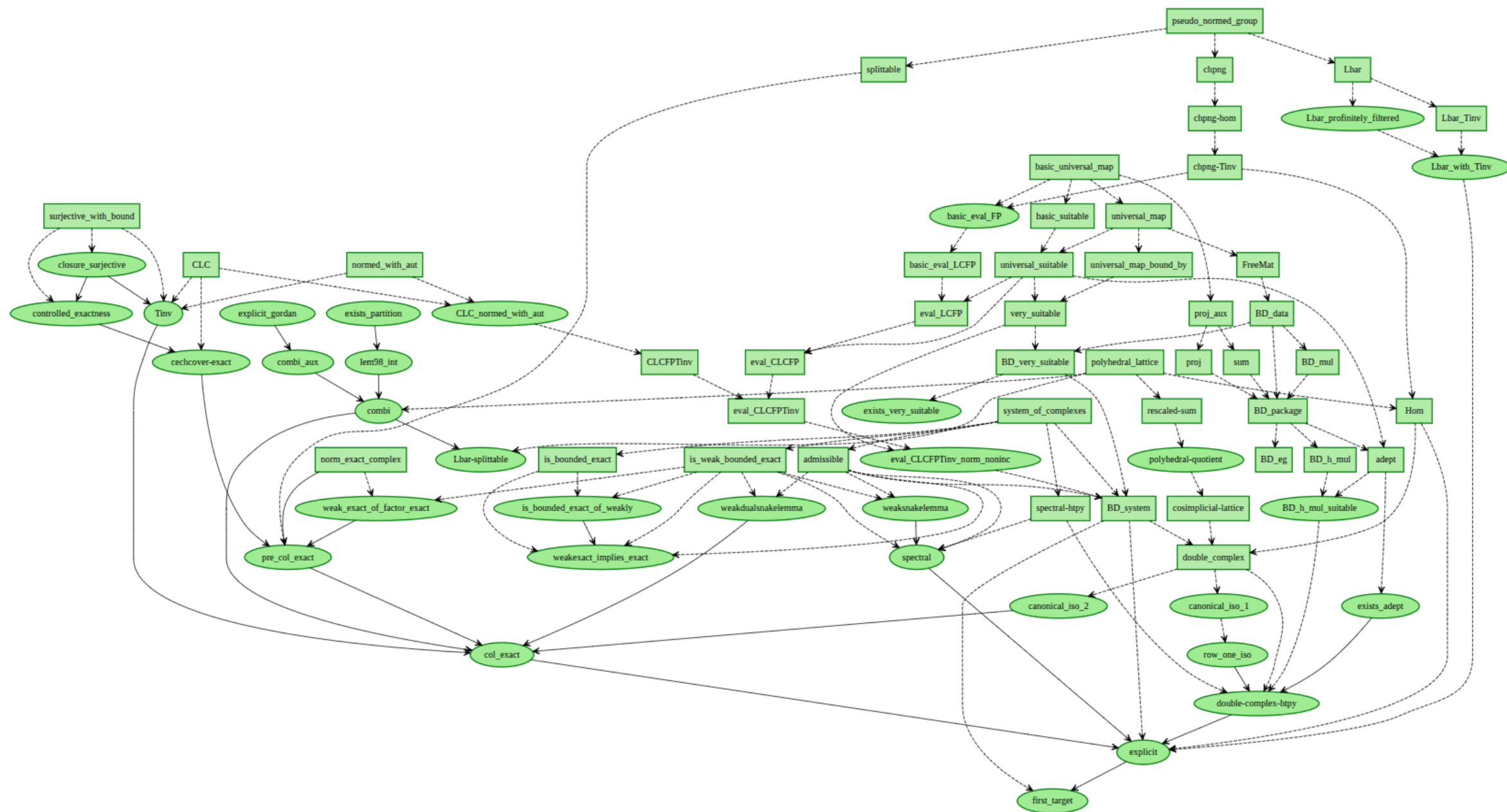
What does a homomorphism $f: \mathbb{Z}[A^m] \rightarrow \mathbb{Z}[A^n]$ that is functorial in A look like? We should perhaps say more precisely what we mean by this. The idea is that m and n are fixed, and for each abelian group A we have a group homomorphism $f_A: \mathbb{Z}[A^m] \rightarrow \mathbb{Z}[A^n]$ such that if $\phi: A \rightarrow B$ is a group homomorphism inducing $\phi_i: \mathbb{Z}[A^i] \rightarrow \mathbb{Z}[B^i]$ for each natural number i then the obvious square commutes: $\phi_n \circ f_A = f_B \circ \phi_m$.

The map f_A is specified by what it does to the generators $(a_1, a_2, a_3, \dots, a_m) \in A^m$. It can send such an element to an arbitrary element of $\mathbb{Z}[A^n]$, but one can check that universality implies that f_A will be a \mathbb{Z} -linear combination of “basic universal maps”, where a “basic universal map” is one that sends (a_1, a_2, \dots, a_m) to (t_1, \dots, t_n) , where t_i is a \mathbb{Z} -linear combination $c_{i,1} \cdot a_1 + \dots + c_{i,m} \cdot a_m$. So a “basic universal map” is specified by the $n \times m$ -matrix c .

Definition 1.1.1 ✓

A basic universal map from exponent m to n , is an $n \times m$ -matrix with coefficients in \mathbb{Z} .

Legend ≡



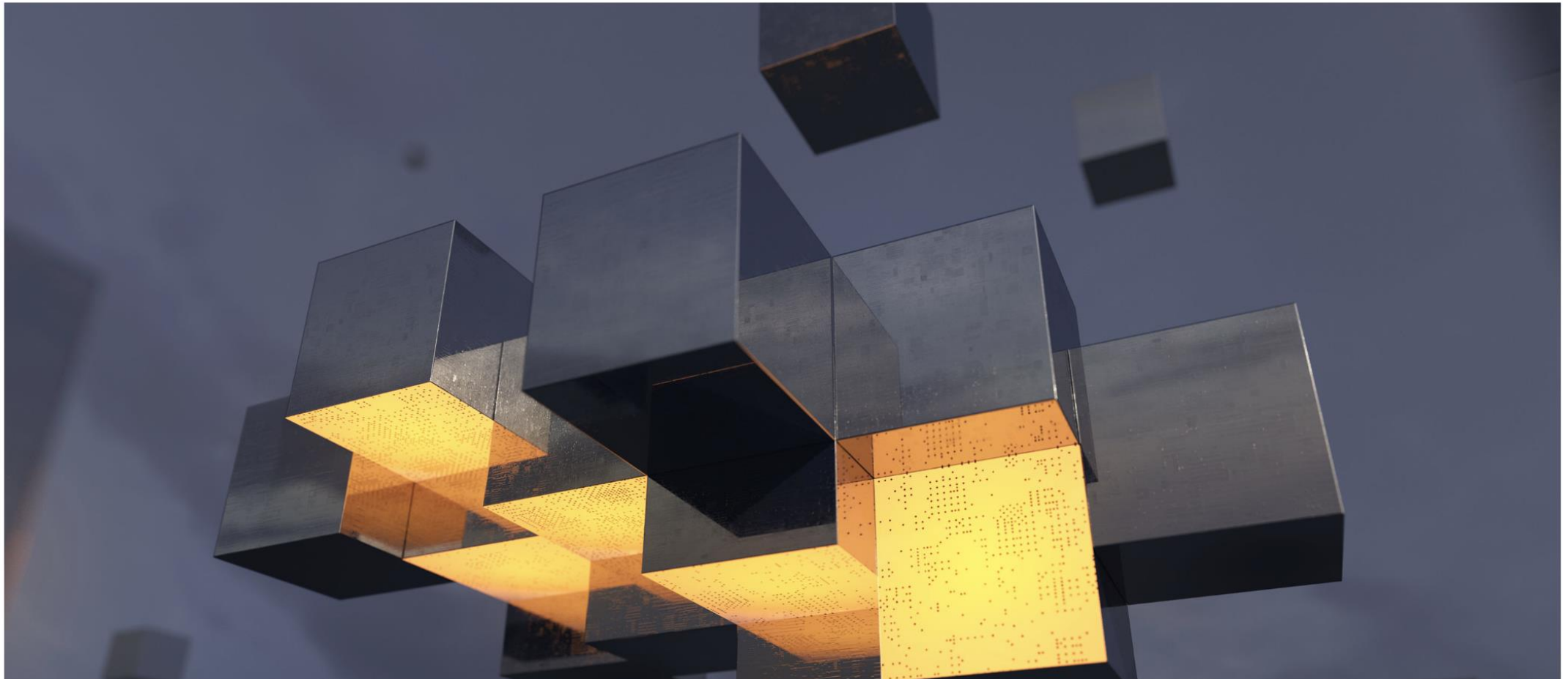
Automated reasoning and symbolic AI

- Several theorems proved with the help of SAT solvers
 - Arithmetic / combinatorial problems (Schur numbers, Pythagorean triples)
 - Geometric theorems (Keller's conjecture, the happy ending theorem)
 - Gardam's refutation of the Kaplansky unit conjecture
- Decision procedures and automated reasoning used for formalization
- First-order provers and model finders used in the Equational Theories Project
- Verified computer algebra and symbolic computation
- Verified numeric computation

Computer Search Settles 90-Year-Old Math Problem

10 |

By translating Keller's conjecture into a computer-friendly search for a type of graph, researchers have finally resolved a problem about covering spaces with tiles.



A counterexample to the unit conjecture for group rings

By GILES GARDAM

To the memory of Willem Henskens

Abstract

The unit conjecture, commonly attributed to Kaplansky, predicts that if K is a field and G is a torsion-free group, then the only units of the group ring $K[G]$ are the trivial units, that is, the non-zero scalar multiples of group elements. We give a concrete counterexample to this conjecture; the group is virtually abelian and the field is order two.

Computer Science > Computational Geometry

[Submitted on 1 Mar 2024]

Happy Ending: An Empty Hexagon in Every Set of 30 Points

Marijn J.H. Heule, Manfred Scheucher

Satisfiability solving has been used to tackle a range of long-standing open math problems in recent years. We add another success by solving a geometry problem that originated a century ago. In the 1930s, Esther Klein's exploration of unavoidable shapes in planar point sets in general position showed that every set of five points includes four points in convex position. For a long time, it was open if an empty hexagon, i.e., six points in convex position without a point inside, can be avoided. In 2006, Gerken and Nicolás independently proved that the answer is no. We establish the exact bound: Every 30-point set in the plane in general position contains an empty hexagon. Our key contributions include an effective, compact encoding and a search-space partitioning strategy enabling linear-time speedups even when using thousands of cores.

ors Decide on Is Key ts' Health

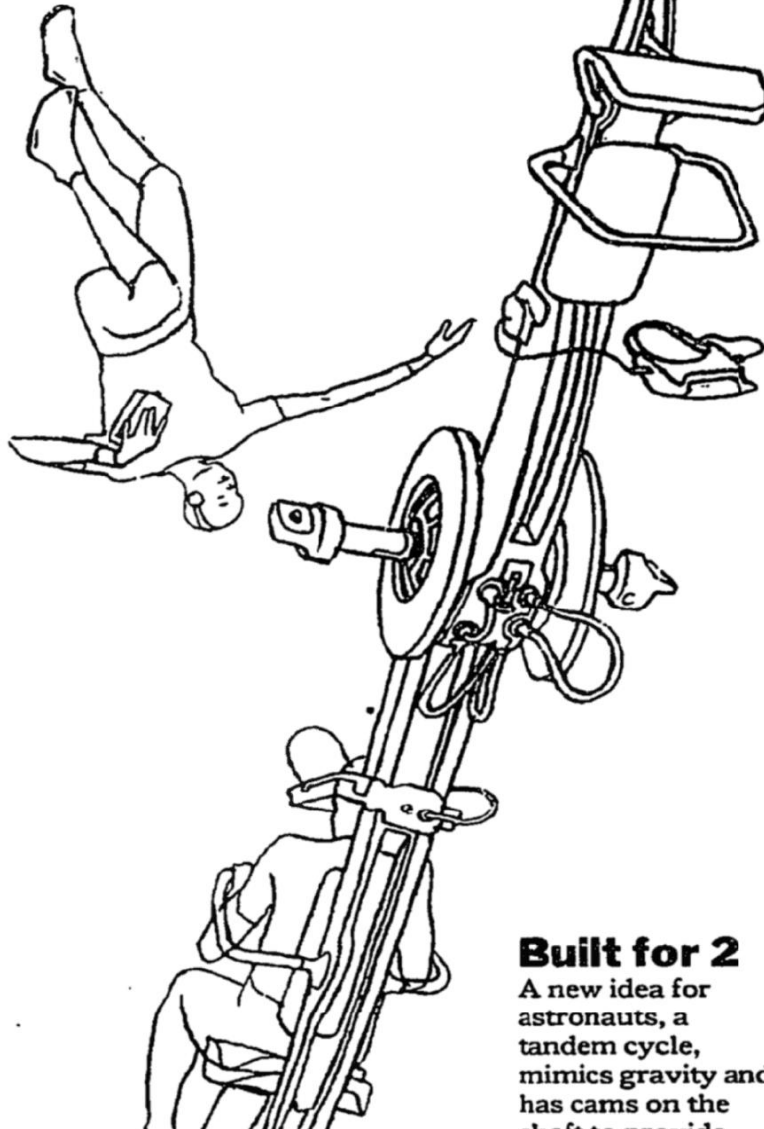
from the hip and lower spine, a trend that if uncorrected over time could prevent long space voyages.

Experts say a trip to Mars, a year or two each way, carries the risk of leaving an astronaut crippled upon return.

"We've learned that bone loss from selected sites on the skeleton is a problem that we still don't have a solution to," Dr. Frank M. Sulzman, director of life science research at the National Aeronautics and Space Administration, said in an interview.

But NASA and its advisers say they are on the verge of finding what may be a simple way to prevent a wide range of space illnesses: nothing fancy or high-tech, it boils down to hard exercise, the orbital equivalent of pumping iron.

Astronauts now tend to do endurance types of exercise, including cycling, rowing and walking on a treadmill, that stress aerobics and stamina. But a wide consensus is developing among space physiologists and NASA officials that this approach is wrong and needs to be supplemented by strenuous workouts that increase



Built for 2

A new idea for astronauts, a tandem cycle, mimics gravity and has cams on the

With Major Math Proof, Brute Computers Show Flash of Reasoning Power

The achievement would have been called creative if a human had done it.

By GINA KOLATA

COMPUTERS are whizzes when it comes to the grunt work of mathematics. But for creative and elegant solutions to hard mathematical problems, nothing has been able to beat the human mind. That is, perhaps, until now.

A computer program written by researchers at Argonne National Laboratory in Illinois has come up with a major mathematical proof that would have been called creative if a human had thought of it. In doing so, the computer has, for the first time, got a foothold into pure mathematics, a field described by its practitioners as more of an art form than a science. And the implications, some say, are profound, showing just how powerful computers can be at reasoning itself, at mimicking the flashes of logical insight or even

those conjectures were easy to prove. The difference this time is that the computer has solved a conjecture that stumped some of the best mathematicians for 60 years. And it did so with a program that was designed to reason, not to solve a specific problem. In that sense, the program is very different from chess-playing computer programs, for example, which are intended to solve just one problem: the moves of a chess game.

"It's a sign of power, of reasoning power," said Dr. Larry Wos, the supervisor of the computer reasoning project at Argonne. And with this result, obtained by a colleague, Dr. William McCune, he said, "We've taken a quantum leap forward."

Dr. Wos predicts that the result may mark the beginning of the end for mathematics research as it is now practiced, eventually freeing mathematicians to focus on discovering new conjectures, and leaving the proof to computers.

But the result also may challenge the very notion of creative thinking, raising the possibility that computers could take a parallel path to reach the same conclusions as great human thinkers. Or it may be that since no one has any idea how humans think, the magnificent bursts of

THE EQUATIONAL THEORIES PROJECT: ADVANCING COLLABORATIVE MATHEMATICAL RESEARCH AT SCALE

MATTHEW BOLAN, JOACHIM BREITNER, JOSE BROX, MARIO CARNEIRO, MARTIN DVORAK,
ANDRÉS GOENS, AARON HILL, HARALD HUSUM, ZOLTAN KOCSIS, BRUNO LE FLOCH,
LORENZO LUCCIOLI, DOUGLAS MCNEIL, ALEX MEIBURG, PIETRO MONTICONE, PACE
NIELSEN, GIOVANNI PAOLINI, MARCO PETRACCI, BERNHARD REINKE, DAVID RENSHAW,
MARCUS ROSSEL, CODY ROUX, JÉRÉMY SCANVIC, SHREYAS SRINIVAS, ANAND RAO
TADIPATRI, TERENCE TAO, VLAD TSARKLEVICH, DANIEL WEBER, FAN ZHENG

ABSTRACT. We report on the *Equational Theories Project* (ETP), an online collaborative pilot project to explore new ways to collaborate in mathematics with machine assistance. The project successfully determined all 22 028 942 edges of the implication graph between the 4694 simplest equational laws on magmas, by a combination of human-generated and automated proofs, all validated by the formal proof assistant language *Lean*. As a result of this project, several new constructions of magmas obeying specific laws were discovered, and several auxiliary questions were also addressed, such as the effect of restricting attention to finite magmas.

Machine learning and neural AI

- Intuitions from data mining:
 - knot invariants
 - representation theory
 - murmurations
- PDEs and physics-informed neural networks
- Neural networks constructing algebraic expressions
 - antiderivatives
 - Lyapunov functions
- Combinatorial constructions
 - counterexamples in graph theory
 - AlphaEvolve, OpenEvolve, ShinkaEvolve

Advancing mathematics by guiding human intuition with AI

<https://doi.org/10.1038/s41586-021-04086-x>

Received: 10 July 2021

Accepted: 30 September 2021

Published online: 1 December 2021

Open access



Check for updates

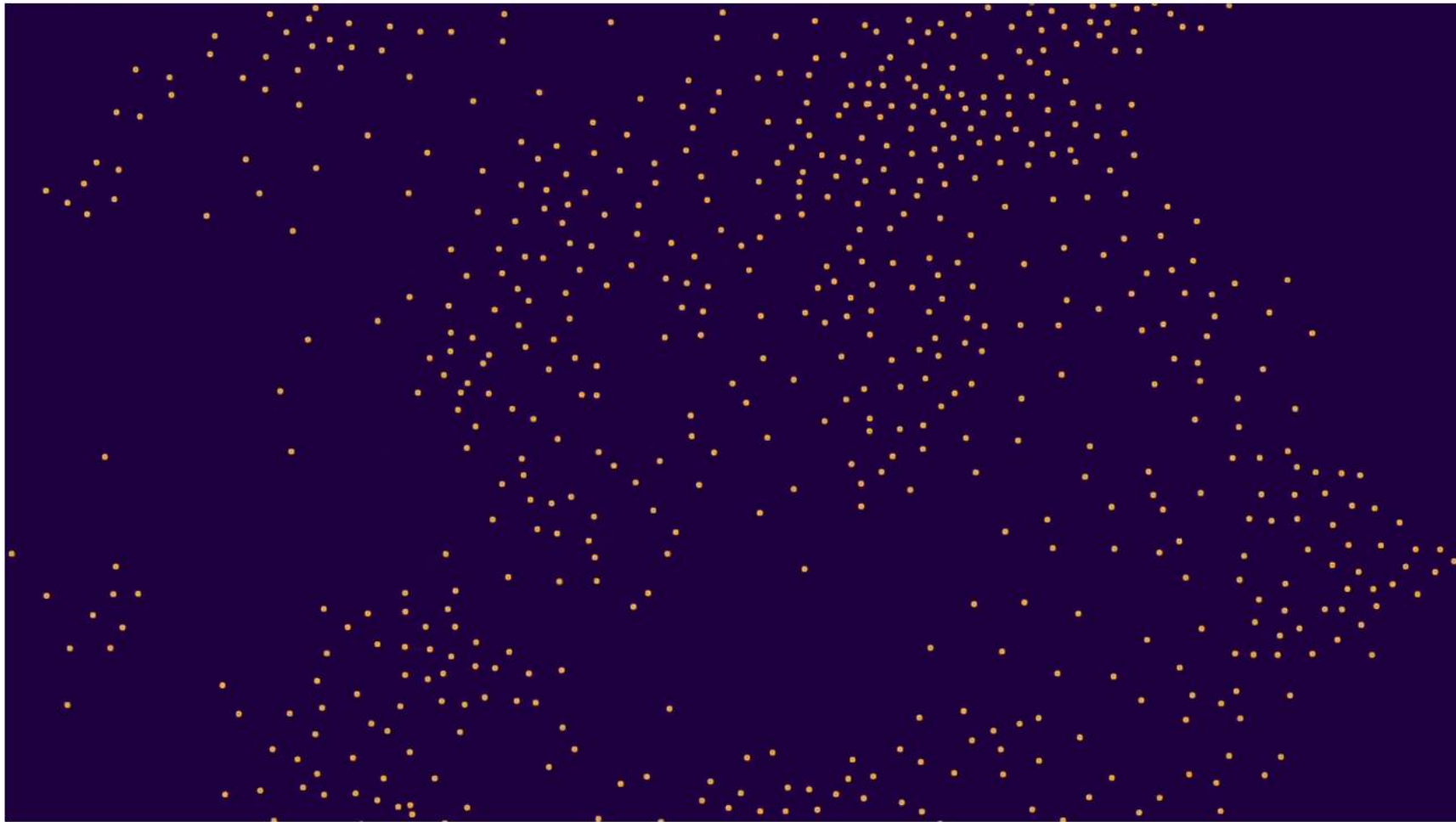
Alex Davies^{1✉}, Petar Veličković¹, Lars Buesing¹, Sam Blackwell¹, Daniel Zheng¹, Nenad Tomašev¹, Richard Tanburn¹, Peter Battaglia¹, Charles Blundell¹, András Juhász², Marc Lackenby², Geordie Williamson³, Demis Hassabis¹ & Pushmeet Kohli^{1✉}

The practice of mathematics involves discovering patterns and using these to formulate and prove conjectures, resulting in theorems. Since the 1960s, mathematicians have used computers to assist in the discovery of patterns and formulation of conjectures¹, most famously in the Birch and Swinnerton-Dyer conjecture², a Millennium Prize Problem³. Here we provide examples of new fundamental results in pure mathematics that have been discovered with the assistance of machine learning—demonstrating a method by which machine learning can aid mathematicians in discovering new conjectures and theorems. We propose a process of using machine learning to discover potential patterns and relations between mathematical objects, understanding them with attribution techniques and using these observations to guide intuition and propose conjectures. We outline this machine-learning-guided framework and demonstrate its successful application to current research questions in distinct areas of pure mathematics, in each case showing how it led to meaningful mathematical contributions on important open problems: a new connection between the algebraic and geometric structure of knots, and a candidate algorithm predicted by the combinatorial invariance conjecture for symmetric groups⁴. Our work may serve as a model for collaboration between the fields of mathematics and artificial intelligence (AI) that can achieve surprising results by leveraging the respective strengths of mathematicians and machine learning.

Elliptic Curve ‘Murmurations’ Found With AI Take Flight

6 |

Mathematicians are working to fully explain unusual behaviors uncovered using artificial intelligence.



When viewed the right way, elliptic curves can flock like birds.

Paul Chaikin for *Quanta Magazine*

Deep Learning Poised to 'Blow Up' Famed Fluid Equations

6 |

For centuries, mathematicians have tried to prove that Euler's fluid equations can produce nonsensical answers. A new approach to machine learning has researchers betting that "blowup" is near.



Computer Science > Machine Learning

[Submitted on 10 Oct 2024]

Global Lyapunov functions: a long-standing open problem in mathematics, with symbolic transformers

Alberto Alfarano, François Charton, Amaury Hayat

Despite their spectacular progress, language models still struggle on complex reasoning tasks, such as advanced mathematics. We consider a long-standing open problem in mathematics: discovering a Lyapunov function that ensures the global stability of a dynamical system. This problem has no known general solution, and algorithmic solvers only exist for some small polynomial systems. We propose a new method for generating synthetic training samples from random solutions, and show that sequence-to-sequence transformers trained on such datasets perform better than algorithmic solvers and humans on polynomial systems, and can discover new Lyapunov functions for non-polynomial systems.

MATHEMATICAL EXPLORATION AND DISCOVERY AT SCALE

BOGDAN GEORGIEV, JAVIER GÓMEZ-SERRANO, TERENCE TAO, AND ADAM ZSOLT WAGNER

ABSTRACT. `AlphaEvolve` [223] is a generic evolutionary coding agent that combines the generative capabilities of LLMs with automated evaluation in an iterative evolutionary framework that proposes, tests, and refines algorithmic solutions to challenging scientific and practical problems. In this paper we showcase `AlphaEvolve` as a tool for autonomously discovering novel mathematical constructions and advancing our understanding of long-standing open problems.







To demonstrate its breadth, we considered a list of 67 problems spanning mathematical analysis, combinatorics, geometry, and number theory. The system rediscovered the best known solutions in most of the cases and discovered improved solutions in several. In some instances, `AlphaEvolve` is also able to *generalize* results for a finite number of input values into a formula valid for all input values. Furthermore, we are able to combine this methodology with `Deep Think` [148] and `AlphaProof` [147] in a broader framework where the additional proof-assistants and reasoning systems provide automated proof generation and further mathematical insights.

These results demonstrate that large language model-guided evolutionary search can autonomously discover mathematical constructions that complement human intuition, at times matching or even improving the best known results, highlighting the potential for significant new ways of interaction between mathematicians and AI systems. We present `AlphaEvolve` as a powerful new tool for mathematical discovery, capable of exploring vast search spaces to solve complex optimization problems at scale, often with significantly reduced requirements on preparation and computation time.

What's new

Updates on my research and expository papers, discussion of open problems, and other maths-related topics. By Terence Tao

RECENT COMMENTS

-  Anonymous on 275A, Notes 0: Foundations of...
-  Anonymous on What are the odds, II: the Ven...
-  Anonymous on The maximal length of the Erdő...
-  Sam on 245B, notes 1: Signed measures...
-  Terence Tao on The maximal length of the Erdő...
-  Terence Tao on What are the odds, II: the Ven...

The maximal length of the Erdős–Herzog–Piranian lemniscate in high degree

15 December, 2025 in math.CV, paper | Tags: Erdos, lemniscate, polynomials | by Terence Tao | 19 comments

I’ve just uploaded to the arXiv my preprint [The maximal length of the Erdős–Herzog–Piranian lemniscate in high degree](#). This paper resolves (in the asymptotic regime of sufficiently high degree) an old question about the [polynomial lemniscates](#)

$$\partial E_1(p) := \{z : |p(z)| = 1\}$$

attached to monic polynomials p of a given degree n , and specifically the question of bounding the arclength $\ell(\partial E_1(p))$ of such lemniscates. For

I recently explored this problem with the optimization tool *AlphaEvolve*, where I found that when I assigned this tool the task of optimizing $\ell(\partial E_1(p))$ for a given degree n , that the tool rapidly converged to choosing p to be equal to p_0 (up to the rotation and translation symmetries of the problem). This suggested to me that the conjecture was true for all n , though of course this was far from a rigorous proof. AlphaEvolve also provided some useful visualization code for these lemniscates which I have incorporated into the paper (and this blog post), and which helped build my intuition for this problem; I view this sort of “vibe-coded visualization” as another practical use-case of present-day AI tools.

Neurosymbolic theorem proving

- AlphaProof / AlphaGeometry earn silver medal score on 2024 IMO.
- Four systems achieve gold medal score on 2025 IMO (two formal, two informal).
- Several open-source provers are available for Lean: DeepSeek, Kimina, Goedel Prover, ...
- Code pilots like Clause Sonnet are helpful with formalization.
- Corporate models (OpenAI, Google) are becoming good at informal mathematics.
- Autoformalizers and provers made available to mathematicians: Aristotle (Harmonic), AlphaProof (Google DeepMind), Gauss (Math Inc.).

At the Math Olympiad, Computers Prepare to Go for the Gold

15 |

Computer scientists are trying to build an AI system that can win a gold medal at the world's premier math competition.



SCIENCE

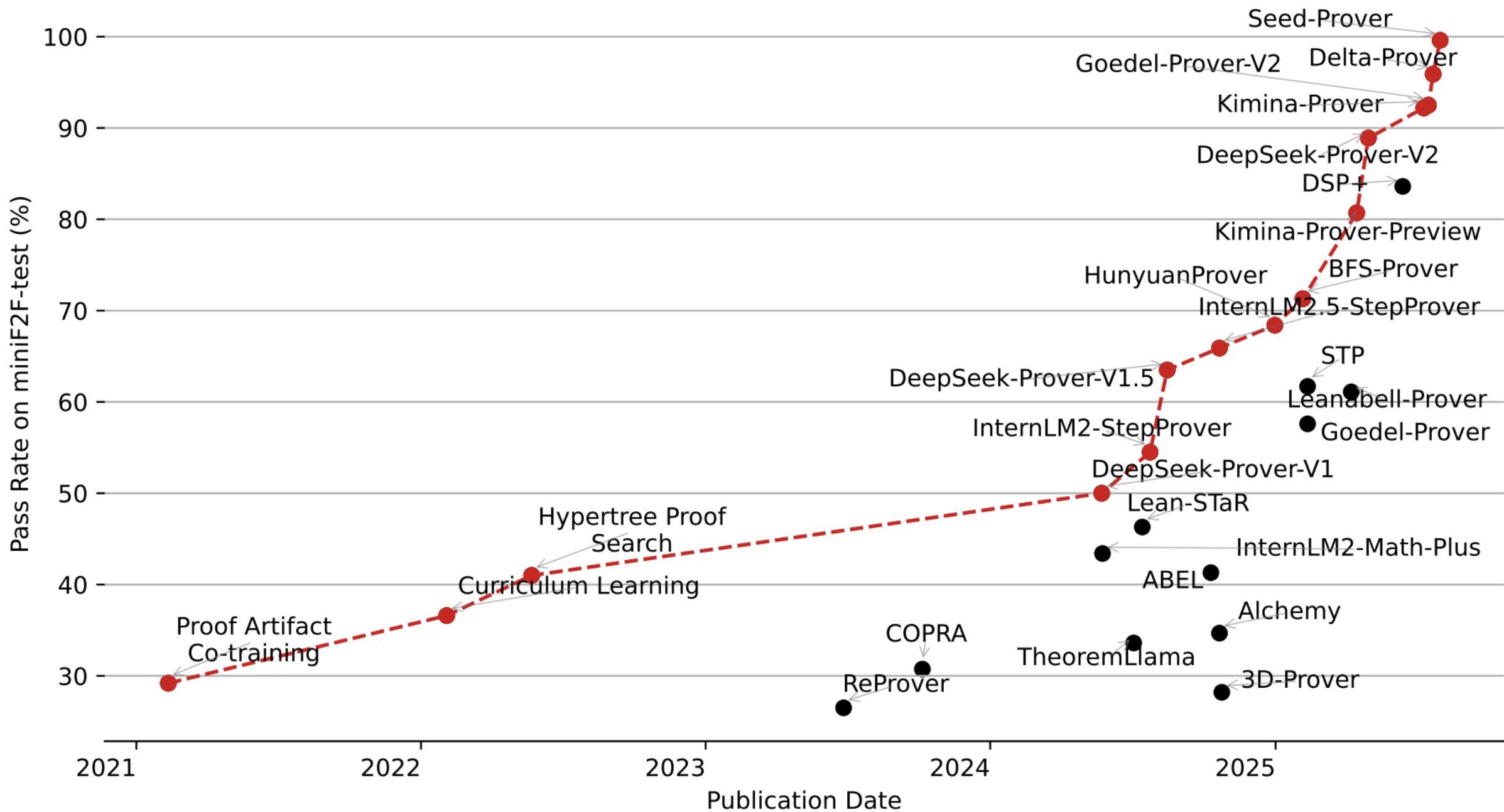
AI achieves silver-medal standard solving International Mathematical Olympiad problems

25 JULY 2024

AlphaProof and AlphaGeometry teams

[Share](#)







Terence Tao
@tao@mathstodon.xyz

Over at the Erdos problem website, AI assistance is now becoming routine. Here is what happened recently regarding Erdos problem #367 erdosproblems.com/367 :

1. On Nov 20, Wouter van Doorn produced a (human-generated) disproof of the second part of this problem, contingent on a congruence identity that he thought was true, and was "sure someone here is able to verify... does indeed hold".
2. A few hours later, I posed this problem to Gemini Deepthink, which (after about ten minutes) produced a complete proof of the identity (and confirmed the entire argument): gemini.google.com/share/81a65a... . The argument used some p-adic algebraic number theory which was overkill for this problem. I then spent about half an hour converting the proof by hand into a more elementary proof, which I presented on the site. I then remarked that the resulting proof should be within range of "vibe formalizing" in Lean.
3. Two days later, Boris Alexeev used the Aristotle tool from Harmonic to complete the Lean formalization, making sure to formalize the final statement by hand to guard against AI exploits. This process took two to three hours, and the output can be found at borisalexeev.com/t/Erdos367.le...

EDIT: after making this post, I decided to round things out by making AI literature searches on this problem, which (after about fifteen minutes) turned up some related literature on consecutive powerful numbers, but nothing directly relating to #367.
chatgpt.com/share/6921427d-9dc...

EXTREMAL DESCENDANT INTEGRALS ON MODULI SPACES OF CURVES: AN INEQUALITY DISCOVERED AND PROVED IN COLLABORATION WITH AI

JOHANNES SCHMITT

ABSTRACT.

HUMAN

For the pure ψ -class intersection numbers $D(\mathbf{e}) = \langle \tau_{e_1} \cdots \tau_{e_n} \rangle_g$ on the moduli space $\overline{\mathcal{M}}_{g,n}$ of stable curves, we determine for which choices of $\mathbf{e} = (e_1, \dots, e_n)$ the value of $D(\mathbf{e})$ becomes extremal. The intersection number is minimal for powers of a single ψ -class (i.e. all e_i but one vanish), whereas maximal values are obtained for balanced vectors ($|e_i - e_j| \leq 1$ for all i, j). The proof uses the nefness of the ψ -classes combined with Khovanskii–Teissier log-concavity.

AUTHOR’S NOTE.

HUMAN

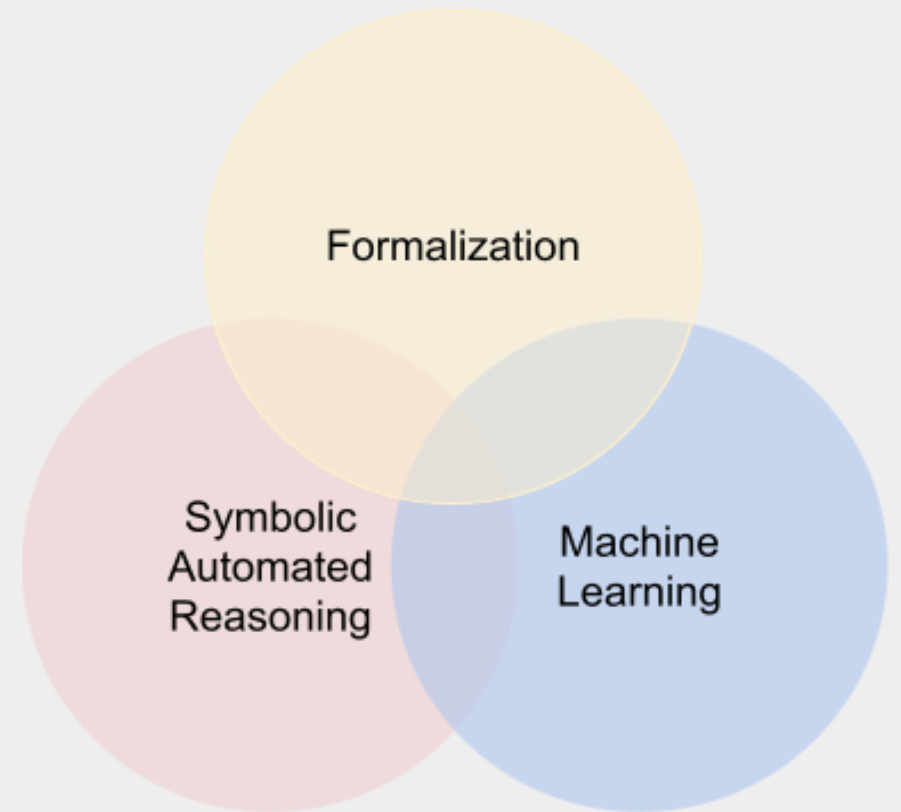
The question of finding extremal values of the ψ -intersection numbers first occurred to the author when looking for a toy problem to explore using the software OpenEvolve [Sha25]. The conjecture that balanced exponents lead to the maximal values is a natural guess, and was indeed discovered quickly by the tested model. To the author’s knowledge, this optimization-style problem was novel and not covered by existing literature: it is a simple and natural question, but somewhat orthogonal to the questions usually studied in enumerative geometry. After some experimental verification and presenting the conjecture to several colleagues (who confirmed its open status), it was submitted as a problem to the IMProofBench project [SBD⁺25]. This project collects research level mathematics questions and tests them against a range of AI models. As part of this evaluation, the conjecture was independently proven by several such models, without human intervention (see Appendix A for further details).

New technologies for mathematics

We have discussed:

- interactive theorem proving and formalization
- automated reasoning and symbolic AI
- machine learning and neural AI
- neurosymbolic theorem proving.

The technologies are still niche, but they are promising, and likely to have an impact.



New technologies for mathematics

Goals of this talk:

- survey the technologies
- raise some concerns

Things to worry about

AI in mathematics

- changes to mathematical practice
- access to research mathematics
- access to mathematics in general
- the role of industry

AI in society

- the effects of AI on cognition
- reliability and transparency
- agency

Changes to mathematical practice

We have always been proud of the fact that mathematics relies on pure thought.

How will the experience of doing mathematics change?

We are proud of our ability to:

- construct complex, rigorous arguments
- detect subtle patterns and connections

What will happen when AI can do those better than we can?

Changes to mathematical practice

These have to do with:

- our ability to do mathematics
- the enjoyment of mathematics

I am optimistic that we will find ways to use AI while preserving the essence of the subject, but this is not a foregone conclusion.

Access to research mathematics

We don't need

- expensive hardware
- large budgets
- project managers

What happens if/when mathematics requires acquiring and managing resources?

Will this limit access to mathematical research?

Access to mathematics

New technologies offer new opportunities for learning:

- interactive systems with correction and feedback
- online communities and social media

Experience shows that taking advantage of them requires:

- money: computing resources, schools, after-school and summer activities
- connections: parents, teachers, mentors who know how to take advantage of the technologies

Will technology lead to greater democratization or greater disparities?

The role of industry

Several big-tech companies and startups are working on AI for math:

- applications to coding
- applications to finance
- applications to science, engineering, and modeling
- applications to other things
- advertisement and PR

They are very good at what they do.

The goals are not necessarily aligned with research mathematics.

The role of industry

What will happen if the best mathematics is coming out of corporations rather than universities and research centers?

What will happen when companies lose interest in research mathematics?

- Will we still rely on their computational resources?
- Will we still rely on their tools?
- Will we be able to modify them and continue to develop them?
- Will we still have jobs?

Things to worry about

AI in mathematics

- changes to mathematical practice
- access to research mathematics
- access to mathematics in general
- the role of industry

AI in society

- the effects of AI on cognition
- reliability and transparency
- agency

The effects of AI on cognition

Students are using corporate models to do their homework.

It's generally easier to ask ChatGPT or Gemini to solve a problem than to do it ourselves.

How will this impact their lives?

The effects of AI on cognition

Compare to concerns about the effects of:

- iPhones and social media
- video games
- computers
- calculators
- television

Is the reliance on AI qualitatively different?

Reliability and transparency

Generally, when we ask AI a question, we want the answer to be

- reliable,
- aligned with our interests,
- likely to help us achieve our goals.

We worry about:

- safety and security
- values and morals.

Reliability and transparency

We want transparency:

- reasons
- explanations
- justification

This provides one role for mathematics: it provides us with artifacts we can query and audit.

Agency

Being *rational* is an important part of our identity.

This involves the ability to make decisions by *reasoning* and *deliberating* with others.

If we are not careful, AI will take us out of the deliberative process: we ask a question, and we get an answer.

Agency

The solution is to make AI part of *our* deliberative process.

We should ask for explanations and reasons, process these ourselves, and ask more questions.

Mathematics provides a language for precise reasoning and deliberation; that's what it was made for.

See my essay, "Is Mathematics Obsolete?"

(The answer is no. Mathematics is as important as ever.)

Things to worry about

AI in mathematics

- changes to mathematical practice
- access to research mathematics
- access to mathematics in general
- the role of industry

AI in society

- the effects of AI on cognition
- reliability and transparency
- agency



Institute for Computer-Aided Reasoning in Mathematics



icarm.io

A new institute

The *Institute for Computer-Aided Reasoning in Mathematics* (ICARM) is a new NSF MSRI on the campus of Carnegie Mellon University.

Its mission is to

- empower mathematicians to take advantage of new technologies for mathematical reasoning and keep mathematics central to everything we do;
- unite mathematicians of all kinds, computer scientists, students, and researchers to achieve that goal; and
- ensure that mathematics and the new technologies are accessible to everyone.