AI for Mathematics Research and Education

Jeremy Avigad

Carnegie Mellon University

Institute for Computer-Aided Reasoning in Mathematics

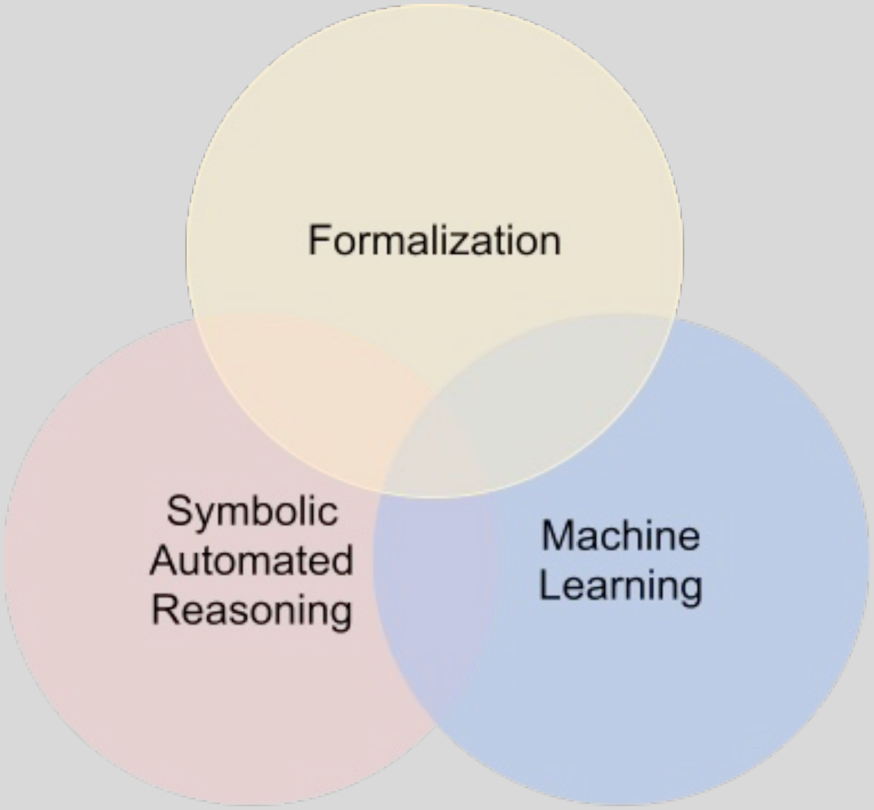
New Technologies for Mathematics

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.



Overview

AI for mathematics research:

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

AI for mathematics education:

- how AI will change mathematics
- how AI will change everything
- what we need to teach our children
- how to teach with AI

Interactive Theorem Proving and Formalization

Interactive Theorem Proving and Formalization

In the early 20th century, logicians developed formal axiomatic systems for mathematics.

It soon became clear that these systems were expressive enough to formalize most mathematics, in principle.

In the early 1970s, the first proof assistants made it possible to formalize and verify proofs in practice.

Today, the practice is known as *interactive theorem proving*. Working with a proof assistant, users construct formal definitions and proofs

LEAN Community

Lean and its Mathematical Library

The Lean theorem prover is a proof assistant developed principally by Leonardo de Moura.

The community recently switched from using Lean 3 to using Lean 4. This website is still being updated, and some pages have outdated information about Lean 3 (these pages are marked with a prominent banner). The old Lean 3 community website has been archived.

The Lean mathematical library, *mathlib*, is a community-driven effort to build a unified library of mathematics formalized in the Lean proof assistant. The library also contains definitions useful for programming. This project is very active, with many regular contributors and daily activity.

You can get a bird's-eye view of what is in the mathlib library by reading the library overview, and read about recent additions on our blog. The design and community organization of mathlib are described in the 2020 article The Lean mathematical library, although the library has grown by an order of magnitude since that article appeared. You can also have a look at our repository statistics to see how the library grows and who contributes to it.

Try it!

You can try Lean in your web browser, install it in an isolated folder, or go for the full install. Lean is free, open source

Learn to Lean!

You can learn by playing a game, following tutorials, or reading books.

Meet the community!

Lean has very diverse and active community. It gathers mostly on a Zulip chat and on GitHub. You

File Edit Selection View Go Run ··· mathlib4

RieszLemma.lean ×

Mathlib > Analysis > NormedSpace > RieszLemma.lean > riesz_lemma

```lean
variable {𝕜 : Type*} [NormedField 𝕜]
variable {E : Type*} [NormedAddCommGroup E] [NormedSpace 𝕜 E]
variable {F : Type*} [SeminormedAddCommGroup F] [NormedSpace ℝ F]

/-- Riesz's lemma, which usually states that it is possible to find a
vector with norm 1 whose distance to a closed proper subspace is
arbitrarily close to 1. The statement here is in terms of multiples of
norms, since in general the existence of an element of norm exactly 1
is not guaranteed. For a variant giving an element with norm in `[1, R]`, see
`riesz_lemma_of_norm_lt`. -/
theorem riesz_lemma {F : Subspace 𝕜 E} (hFc : IsClosed (F : Set E)) (hF : ∃ x : E, x ∉ F) {r : ℝ}
    (hr : r < 1) : ∃ x₀ : E, x₀ ∉ F ∧ ∀ y ∈ F, r * ‖x₀‖ ≤ ‖x₀ - y‖ := by
  classical
    obtain ⟨x, hx⟩ : ∃ x : E, x ∉ F := hF
    let d := Metric.infDist x F
    have hFn : (F : Set E).Nonempty := ⟨_, F.zero_mem⟩
    have hdp : 0 < d :=
      lt_of_le_of_ne Metric.infDist_nonneg fun heq =>
        hx ((hFc.mem_iff_infDist_zero hFn).2 heq.symm)
    let r' := max r 2⁻¹
    have hr' : r' < 1 := by
      simp only [r', max_lt_iff, hr, true_and]
      norm_num
    have hlt : 0 < r' := lt_of_lt_of_le (by simp) (le_max_right r 2⁻¹)
    have hdlt : d < d / r' := (lt_div_iff₀ hlt).mpr ((mul_lt_iff_lt_one_right hdp).2 hr')
    obtain ⟨y₀, hy₀F, hxy₀⟩ : ∃ y ∈ F, dist x y < d / r' := (Metric.infDist_lt_iff hFn).mp hdlt
    have x_ne_y₀ : x - y₀ ∉ F := by
      by_contra h
      have : x - y₀ + y₀ ∈ F := F.add_mem h hy₀F
      simp only [neg_add_cancel_right, sub_eq_add_neg] at this
      exact hx this
    refine ⟨x - y₀, x_ne_y₀, fun y hy => le_of_lt ?_⟩
    have hy₀y : y₀ + y ∈ F := F.add_mem hy₀F hy
    calc
```

Lean InfoView | Lean InfoView ×

▼ RieszLemma.lean:55:4

▼ Tactic state

1 goal

```
𝕜 : Type u_1
inst✝² : NormedField 𝕜
E : Type u_2
inst✝¹ : NormedAddCommGroup E
inst✝ : NormedSpace 𝕜 E
F : Subspace 𝕜 E
hFc : IsClosed ↑F
r : ℝ
hr : r < 1
x : E
hx : x ∉ F
d : ℝ := infDist x ↑F
hFn : (↑F).Nonempty
hdp : 0 < d
r' : ℝ := max r 2⁻¹
hr' : r' < 1
hlt : 0 < r'
hdlt : d < d / r'
y₀ : E
hy₀F : y₀ ∈ F
hxy₀ : dist x y₀ < d / r'
⊢ ∃ x₀ ∉ F, ∀ y ∈ F, r * ‖x₀‖ ≤ ‖x₀
- y‖
```

▶ Messages (1)

▶ All Messages (0)

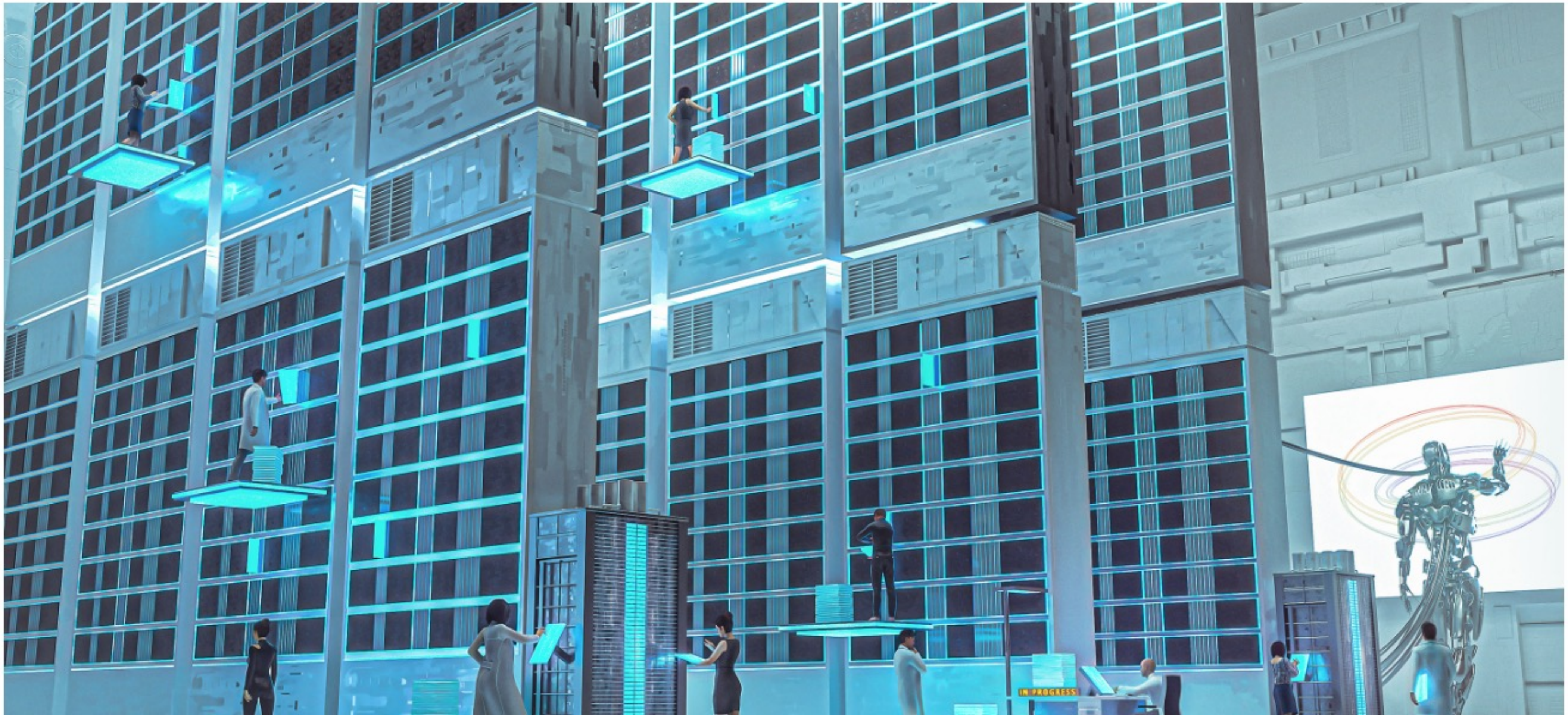
Restart File

FOUNDATIONS OF MATHEMATICS

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*

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

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 a 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

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

What does a homomorphism $f\colon \mathbb{Z}[A^m] \to \mathbb{Z}[A^n]$ 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 $f_A\colon \mathbb{Z}[A^m] \to \mathbb{Z}[A^n]$ such that if $\phi\colon A \to B$ is a group homomorphism inducing $\phi_i\colon \mathbb{Z}[A^i] \to \mathbb{Z}[B^i]$ for each natural number i then the obvious square commutes: $\phi_n \circ f_A = f_B \circ \phi_m$.

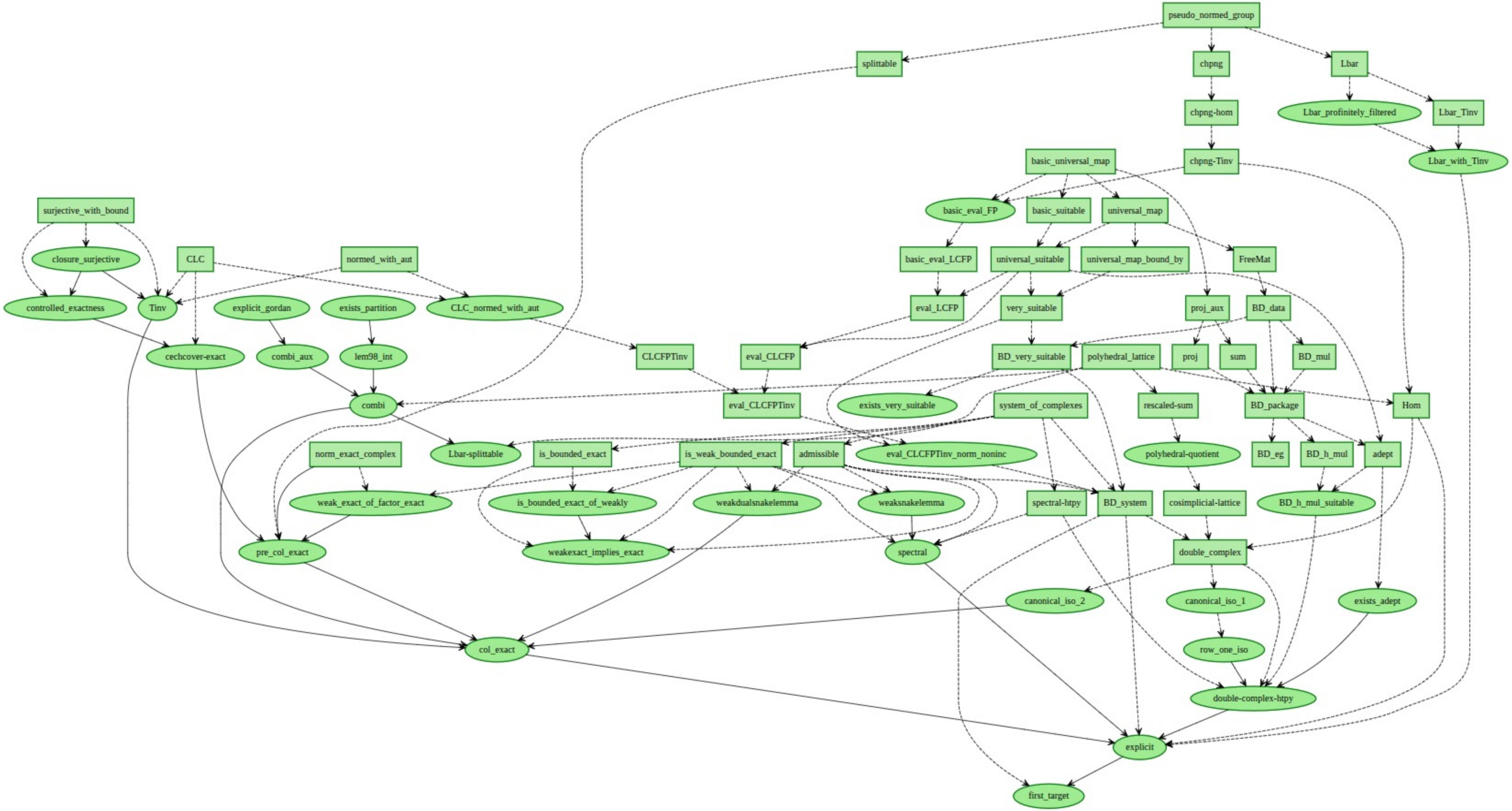
The map f_A is specified by what it does to the generators $(a_1, a_2, a_3, \ldots, a_m) \in A^m$. It can send such an element to an arbitrary element of $\mathbb{Z}[A^n]$, but one can check that universality implies that f_A will be a $\mathbb{Z}$-linear combination of “basic universal maps”, where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

Dependencies

Legend ≡



Automated Reasoning and Symbolic AI

Automated Reasoning and Symbolic AI

Even before computers were invented, logicians were interested in algorithmic procedures to

- decide the truth of mathematical statements, and
- search for proofs.

The first automated provers appeared in the 1950s and 1960s.

Now we have

- first-order provers,
- SAT solvers, and
- SMT solvers.

SAT Solvers

A formula in propositional logic (like $P \vee Q \rightarrow Q \wedge R$) is true or false depending on the truth assignments to the variables.

A satisfiability solver determines whether a formula has a satisfying assignment.

Modern SAT solvers can decide industrial formulas with tens of millions of variables and hundreds of millions of clauses, often in minutes.

Recipe for mathematics:

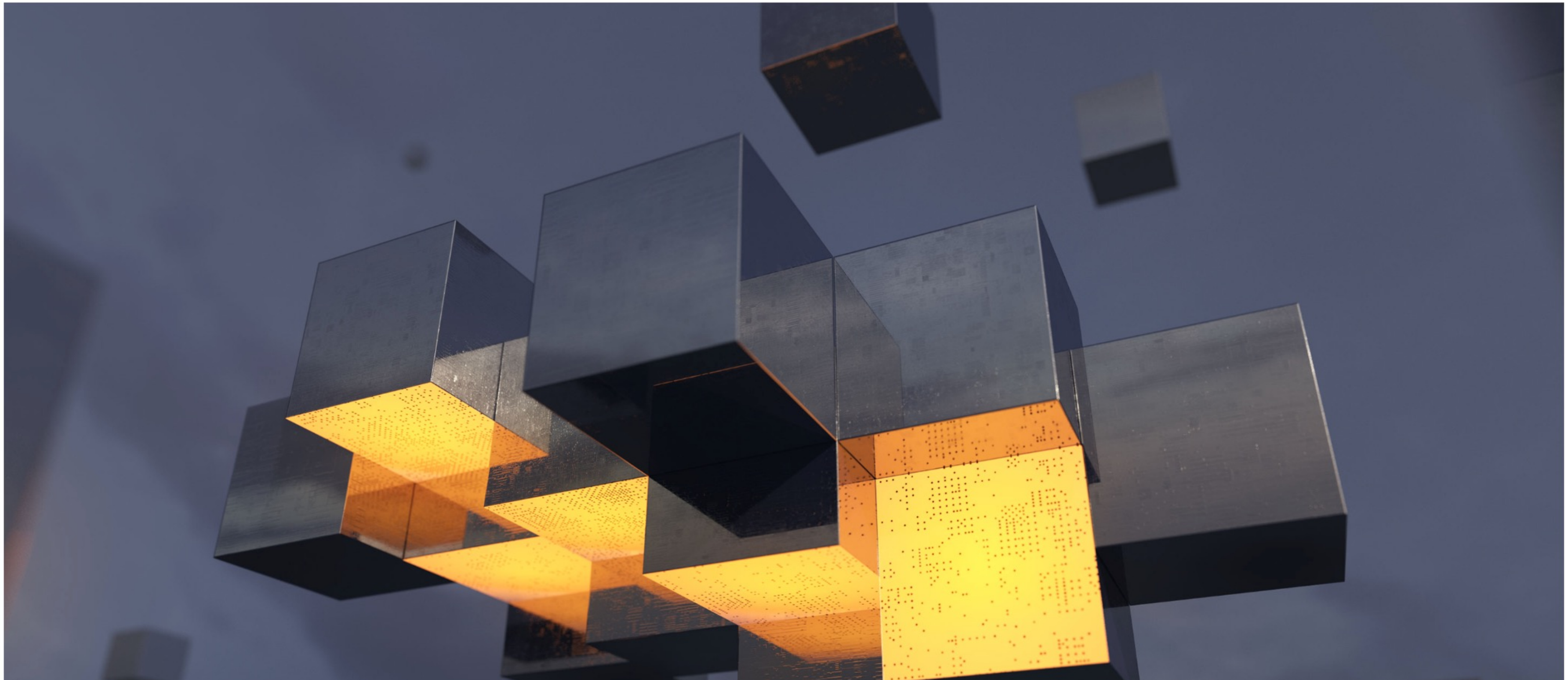
- Encode / reduce a problem to a SAT problem.
- Use a SAT solver.

GEOMETRY

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.

[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.

First- and Higher-Order Theorem Provers

The question as to whether a statement is provable from some hypotheses in first-order logic is equivalent to the halting problem.

The best one can do is to design a complete proof *search.*

There are several such systems available.

Science Times

TUESDAY, DECEMBER 10, 1996

C1

The New York Times

...ors Decide ...on Is Key ...uts' Health

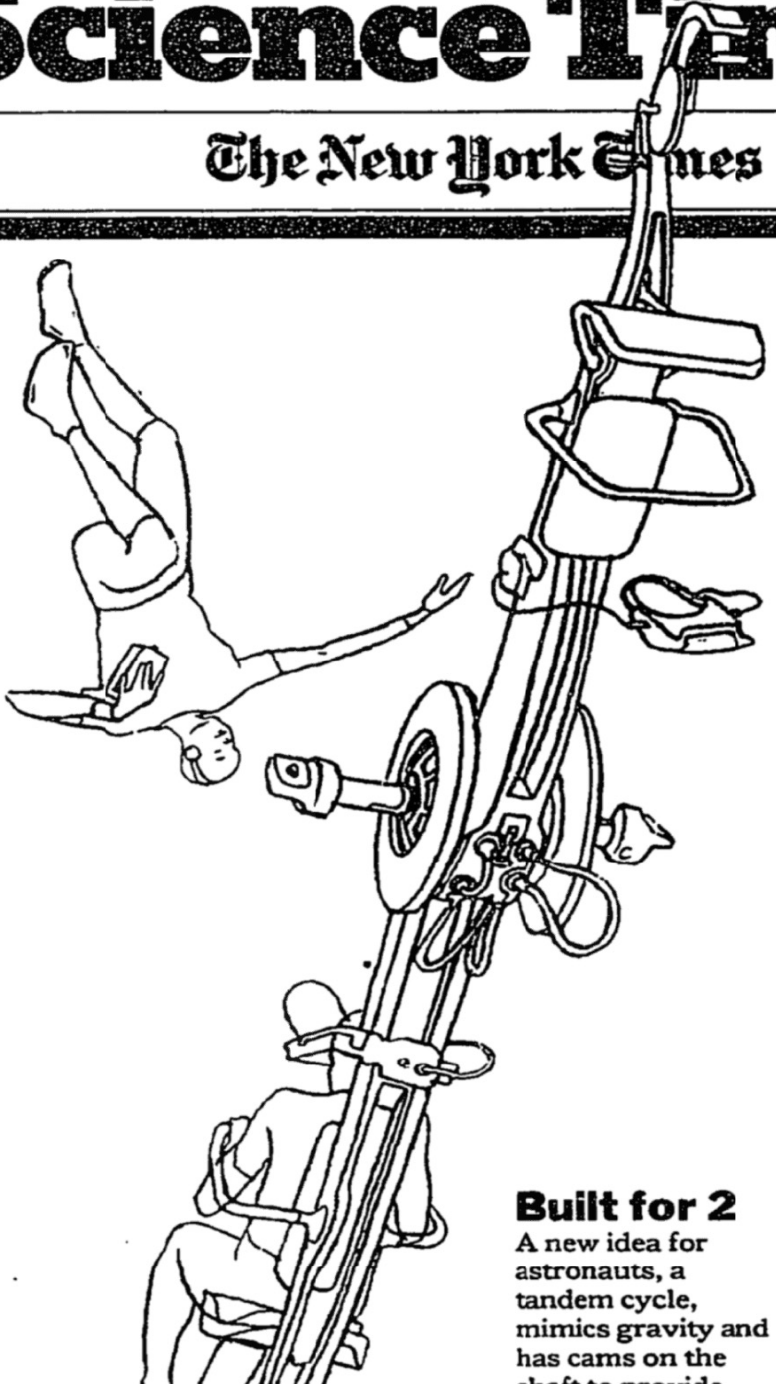
from the hip and lower spine, a trend that if uncorrected over time could prevent long space voyages.

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 toehold 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."

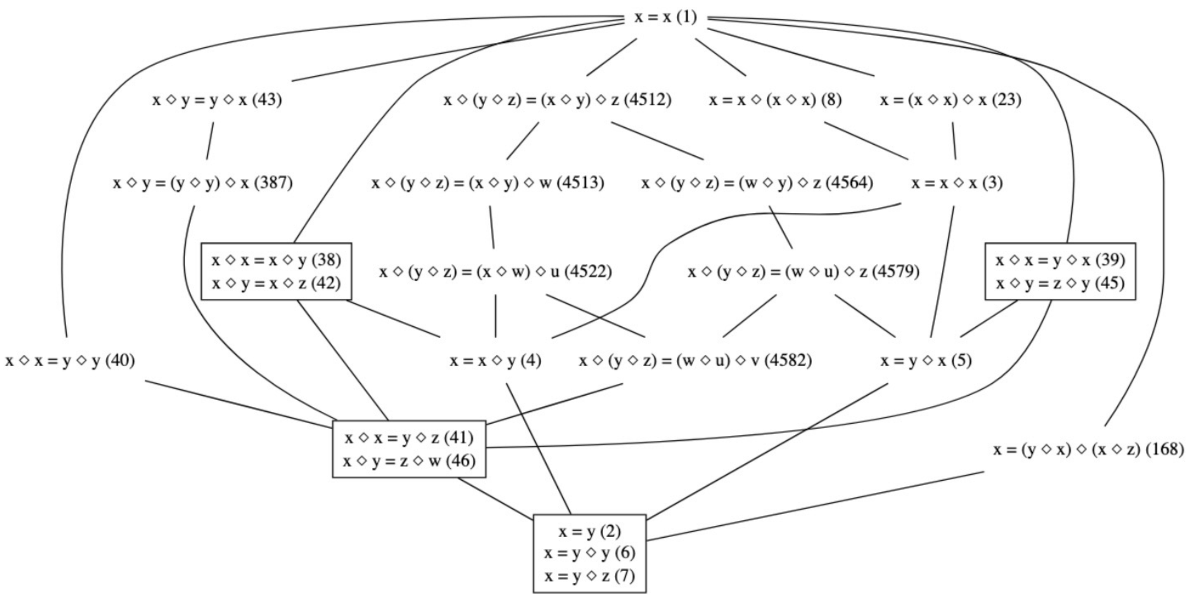
Dr. Wos predicts that the result may mark the beginning of the end for mathematics research as it is now practiced, eventually freeing mathematicians to focus on discovering new conjectures, and leaving the proof to computers.

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

Equational Theories Project

Mapping out the relations between different equational theories of Magmas

Blueprint (web) | Blueprint (pdf) | Paper (pdf) | Documentation | Dashboard | Equation Explorer | Finite Magma Explorer | Graphiti | GitHub



The purpose of this project, launched on Sep 25, 2024, is to explore the space of equational theories of magmas, ordered by implication. To begin with we shall focus only on theories of a single equation, and specifically on the 4694 equational laws involving at most four magma operations, up to symmetry and relabeling (here is the list in Lean and in plain text). This creates 4694*(4694-1) = 22,028,942 implications that need to be proven or disproven, creating both "implications" and "anti-implications".

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 TSYRKLEVICH, 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

Machine Learning

Key approaches:

- **Supervised learning:** the system is presented with (input, output) pairs, and learns a rule connecting them.
- **Unsupervised learning:** the system is presented with data, and learns some sort of structure.
- **Reinforcement learning:** the system acts in a space and is rewarded accordingly; it learns to maximize rewards.

Models can be very simple (linear regression, decision trees) to very complex (neural networks).

Machine Learning and Neural AI

There are several things that machine learning and neural networks can do:

- extract intuitions and detect patterns in data
- compute solutions to PDEs
- construct algebraic expressions
- construct combinatorial objects

Machine Learning and Neural AI

Intuitions from data mining:

- knot invariants
- representation theory
- murmurations

Article

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 Veličković[1], Lars Buesing[1], Sam Blackwell[1], Daniel Zheng[1], Nenad Tomašev[1], Richard Tanburn[1], Peter Battaglia[1], Charles Blundell[1], András Juhász[2], Marc Lackenby[2], Geordie Williamson[3], Demis Hassabis[1] & Pushmeet Kohli[1] ✉

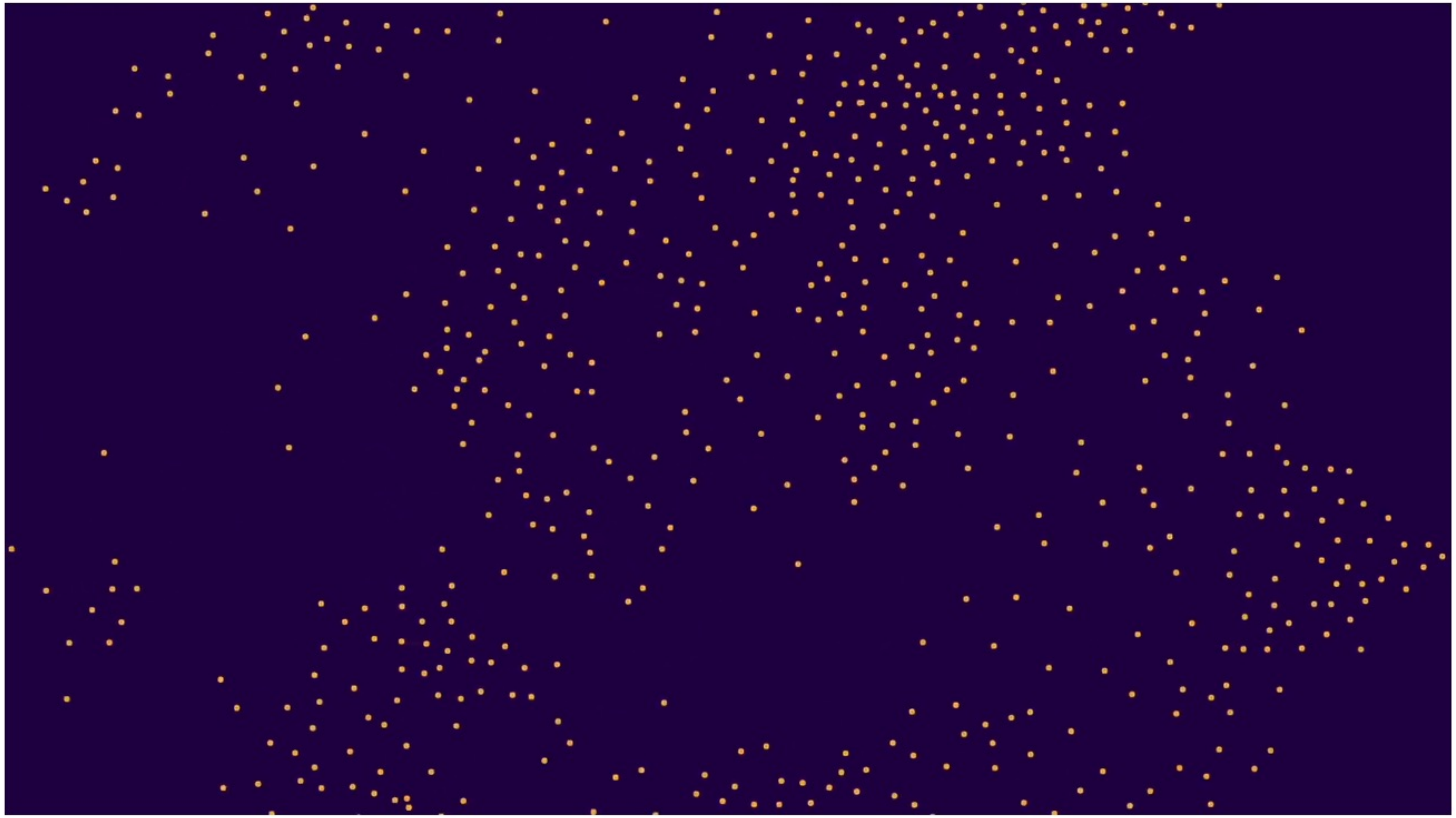
The practice of mathematics involves discovering patterns and using these to formulate and prove conjectures, resulting in theorems. Since the 1960s, mathematicians have used computers to assist in the discovery of patterns and formulation of conjectures[1], most famously in the Birch and Swinnerton-Dyer conjecture[2], a Millennium Prize Problem[3]. 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 groups[4]. 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.

NUMBER THEORY

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*

Machine Learning and Neural AI

Physics-informed neural networks (PINNs) can optimize parameters with respect to an objective function and *constraints* expressed in terms of PDEs.

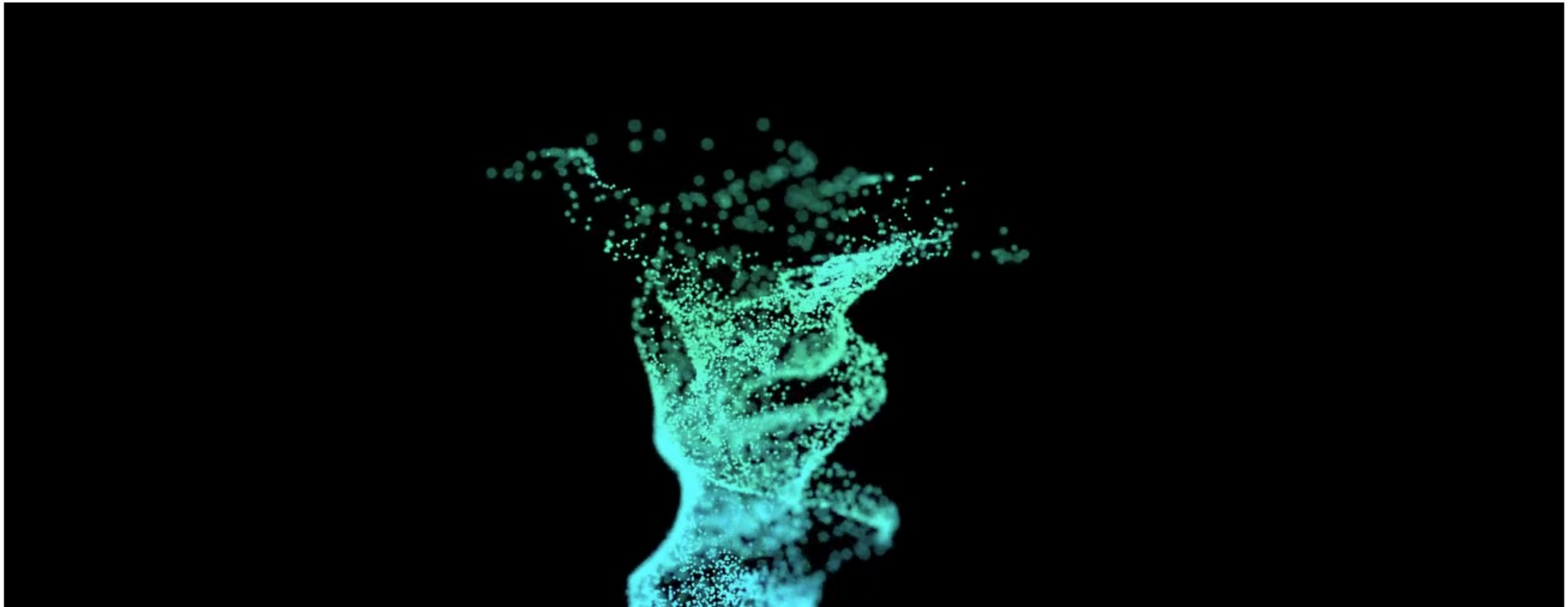
In particular, they can compute approximate solutions to PDEs, and search for regions and parameter settings of interest.

MATHEMATICAL PHYSICS

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.



[Submitted on 17 Sep 2025]

Discovery of Unstable Singularities

Yongji Wang, Mehdi Bennani, James Martens, Sébastien Racanière, Sam Blackwell, Alex Matthews, Stanislav Nikolov, Gonzalo Cao-Labora, Daniel S. Park, Martin Arjovsky, Daniel Worrall, Chongli Qin, Ferran Alet, Borislav Kozlovskii, Nenad Tomašev, Alex Davies, Pushmeet Kohli, Tristan Buckmaster, Bogdan Georgiev, Javier Gómez-Serrano, Ray Jiang, Ching-Yao Lai

Whether singularities can form in fluids remains a foundational unanswered question in mathematics. This phenomenon occurs when solutions to governing equations, such as the 3D Euler equations, develop infinite gradients from smooth initial conditions. Historically, numerical approaches have primarily identified stable singularities. However, these are not expected to exist for key open problems, such as the boundary-free Euler and Navier-Stokes cases, where unstable singularities are hypothesized to play a crucial role. Here, we present the first systematic discovery of new families of unstable singularities. A stable singularity is a robust outcome, forming even if the initial state is slightly perturbed. In contrast, unstable singularities are exceptionally elusive; they require initial conditions tuned with infinite precision, being in a state of instability whereby infinitesimal perturbations immediately divert the solution from its blow-up trajectory. In particular, we present multiple new, unstable self-similar solutions for the incompressible porous media equation and the 3D Euler equation with boundary, revealing a simple empirical asymptotic formula relating the blow-up rate to the order of instability. Our approach combines curated machine learning architectures and training schemes with a high-precision Gauss-Newton optimizer, achieving accuracies that significantly surpass previous work across all discovered solutions. For specific solutions, we reach near double-float machine precision, attaining a level of accuracy constrained only by the round-off errors of the GPU hardware. This level of precision meets the requirements for rigorous mathematical validation via computer-assisted proofs. This work provides a new playbook for exploring the complex landscape of nonlinear partial differential equations (PDEs) and tackling long-standing challenges in mathematical physics.

Comments: 20 pages, 6 figures. Supplementary information will be uploaded in a forthcoming version of the manuscript

Subjects: **Analysis of PDEs (math.AP)**; Fluid Dynamics (physics.flu-dyn)

Cite as: arXiv:2509.14185 **[math.AP]**

9.14185v1 **[math.AP]** for this version)

org/10.48550/arXiv.2509.14185

Machine Learning and Neural AI

Neural networks can construct algebraic and combinatorial objects:

- antiderivatives
- Lyapunov functions
- rewrite sequences in presented groups
- counterexamples in graph theory
- extremal combinatorial objects

[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.

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

ABSTRACTIONS BLOG

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

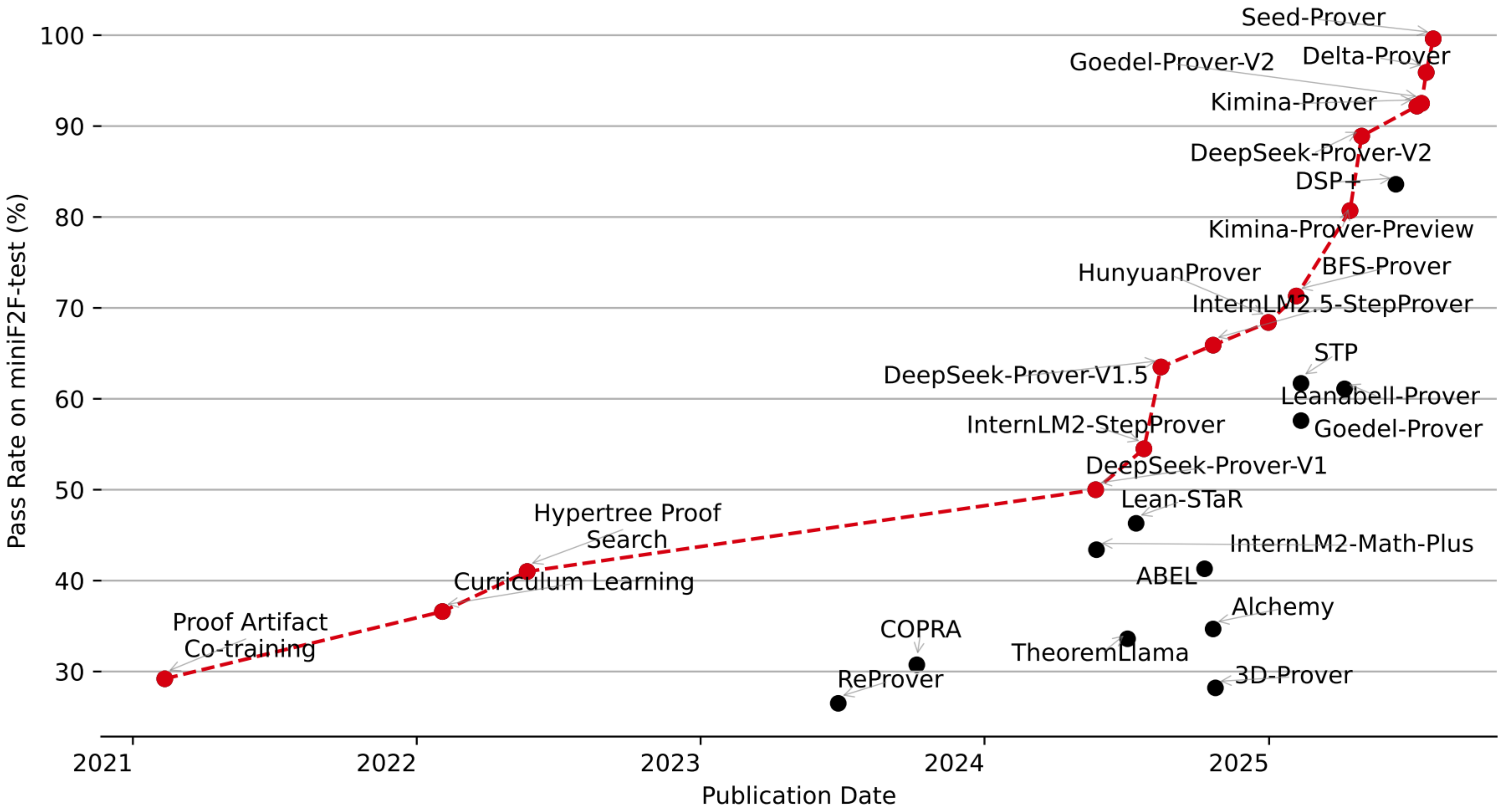
25 JULY 2024

AlphaProof and AlphaGeometry teams

Share

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 Claude 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.).



Pass Rate on miniF2F-test (%)
Publication Date
100
90
80
70
60
50
40
30
2021
2022
2023
2024
2025
Seed-Prover
Delta-Prover
Goedel-Prover-V2
Kimina-Prover
DeepSeek-Prover-V2
DSP+
Kimina-Prover-Preview
BFS-Prover
HunyuanProver
InternLM2.5-StepProver
STP
Leanabell-Prover
Goedel-Prover
DeepSeek-Prover-V1.5
InternLM2-StepProver
DeepSeek-Prover-V1
Lean-STaR
InternLM2-Math-Plus
ABEL
Alchemy
COPRA
TheoremLlama
3D-Prover
ReProver
Hypertree Proof Search
Curriculum Learning
Proof Artifact Co-training

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 someoneone 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 $\mathbf{e} = (e_1, \ldots, e_n)$ 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).

Overview

Overview

AI for mathematics research:

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

AI for mathematics education:

- how AI will change mathematics
- how AI will change everything
- what we need to teach our children
- how to teach with AI

How AI Will Change Mathematics

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 these better than we can?

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?

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.

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?

How AI Will Change Everything

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.

Mathematics can provide us with reasons, justification, and explanation.

Agency

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

This involves the ability to make decisions by *reasoning* and *deliberating*, alone and 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.)

Summary

We need to think about:

- how AI affects our cognition
- how to keep AI reliable and safe
- how to make sure that AI serves our purposes

What We Need to Teach Our Children

The Value of Mathematics

Two reasons to teach children mathematics:

- It's meaningful to them.
- It's useful to them.

Mathematics is:

- an art
- a humanity
- a science

The Aesthetic of Mathematics

There is a strong aesthetic component to mathematics.

It’s also a tradition that stretches back to antiquity.

Precise language and abstraction are fundamental to how to make sense of the world, and how we communicate with one another.

New computing technologies for mathematics challenge us to think about why mathematics is important to us, and what role we have to play going forward.

The Practical Side of Mathematics

Mathematics is fundamental to

- science and technology
- engineering and industry
- finance
- economics and policy

New reasoning technologies can do many of the things we used to do.

We need to think about what roles our children will play in the future.

Teaching with AI

Teaching AI for Mathematics

Two aspects:

- Teaching students how to use AI to do mathematics.
- Using AI to teach mathematics.

The first is easier:

- Several in the Lean community have taught courses on interactive theorem proving.
- Marijn Heule and I have taught a course, *Logic and Mechanized Reasoning*.
- I don't yet know of courses that teach students to use machine learning to solve math problems.

Lean Community

Community

- Zulip chat
- GitHub
- Blog
- Community information
- Community guidelines
- Teams
- Papers about Lean
- Projects using Lean
- Teaching using Lean
- Events

Use Lean

- Online version (no installation)
- Install Lean
- More options

Documentation

- Learning resources (start here)
- Documentation overview
- API documentation
- Declaration search (Loogle)
- Language reference
- Tactic list
- Glossary
- Did you really prove it?
- About MWEs

Library overviews

- Library overview
- Undergraduate maths
- Wiedijk's 100 theorems
- 1000+ theorems

Learning Lean 4

There are many ways to start learning Lean, depending on your background and taste. They are all fun and rewarding, but also difficult and occasionally frustrating. Proof assistants are still difficult to use, and you cannot expect to become proficient after one afternoon of learning.

All the resources listed on this page are about Lean 4. Some have Lean 3 versions, but there is no point learning Lean 3 at this stage.

Hands-on approaches

- Whatever your background, if you want to dive right away, you can play the Natural Number Game. This is an online interactive Lean tutorial focused on proving properties of the elementary operations on natural numbers. The Lean Game Server hosts various learning games including Set Theory, Logic, and Robo (a story about undergrad mathematics).

- For a faster-paced dive, you can get the Glimpse of Lean tutorial. This contains four basic files covering some fundamental aspects of proving using Lean, and then independent topic files about elementary analysis, abstract topology and mathematical logic.

- You can download the tactic cheatsheet (PDF) for a reference of most common tactics.

- If you wish to learn directly from source, the Lean API documentation not only includes `Mathlib`, but also covers `Std`, `Batteries`, `Lake`, and the core compiler. As much of Lean is defined in terms of syntax extensions, this is the closest thing to a comprehensive reference manual that exists.

- If you wish to get your hands dirty and contribute to mathlib, but don't know a good project to start out, then there is a long list of easy issues on the GitHub Issues. If you are working on an issue, please post a reply on the GitHub issue stating that you're working on it, to minimize duplicate effort.

Books

If you prefer reading a book (with exercises), there are a number of freely available Lean books that have proven to be useful to beginners. These are available as HTML or PDFs, but are usually meant to be read interactively in VSCode, doing Lean exercises on the fly:

- The standard mathematics-oriented reference is Mathematics in Lean. You can download it as a PDF, but see also the

Welcome to the Natural Number Game

An introduction to mathematical proof.

In this game, we will build the basic theory of the natural numbers `{0,1,2,3,4,...}` from scratch. Our first goal is to prove that `2 + 2 = 4`. Next we'll prove that `x + y = y + x`. And at the end we'll see if we can prove Fermat's Last Theorem. We'll do this by solving levels of a computer puzzle game called Lean.

Read this.

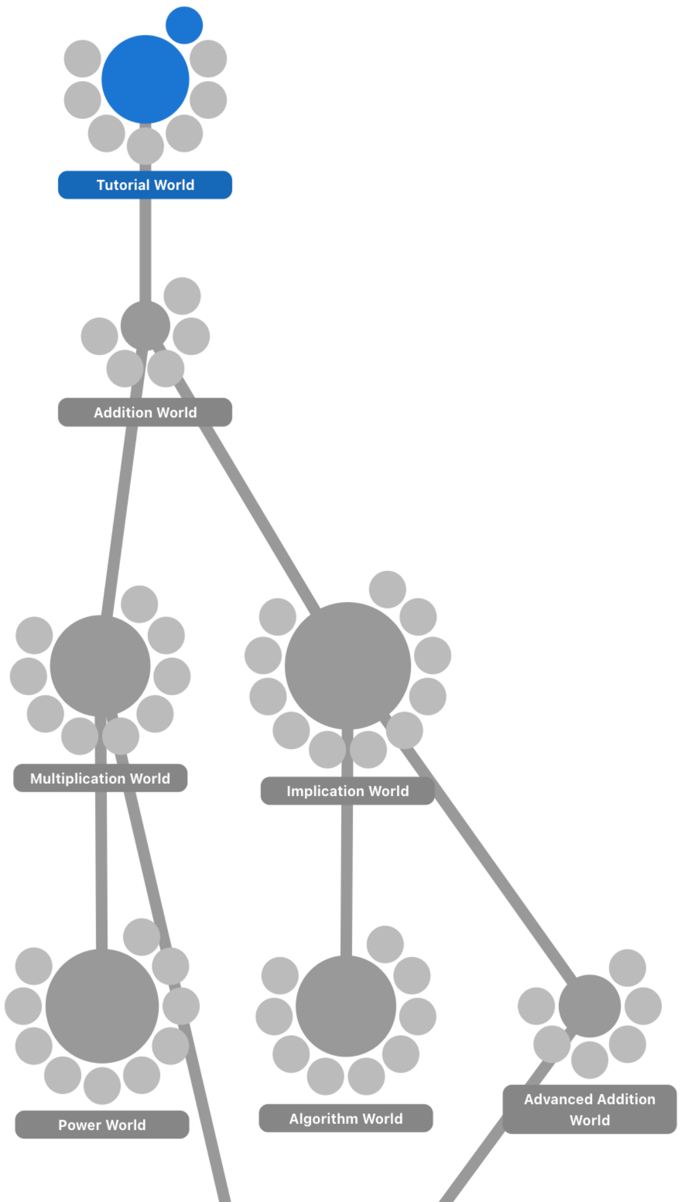
Learning how to use an interactive theorem prover takes time. Tests show that the people who get the most out of this game are those who read the help texts like this one.

To start, click on "Tutorial World".

Note: this is a new Lean 4 version of the game containing several worlds which were not present in the old Lean 3 version. More new worlds such as Strong Induction World, Even/Odd World and Prime Number World are in development; if you want to see their state or even help out, checkout out the issues in the github repo.

More

Click on the three lines in the top right and select



Rules

regular
relaxed
none

Theorems | Tactics | Definitions

apply | cases | contrapose | decide | exact | have | induction | intro | left | rfl | right | rw | simp | simp_add | symm | tauto | trivial | use

1. Introduction
2. Basics
3. Logic
4. Sets and Functions
5. Elementary Number Theory
6. Discrete Mathematics
7. Structures
8. Hierarchies
9. Groups and Rings
10. Linear algebra
11. Topology
12. Differential Calculus
13. Integration and Measure Theory
Index

Mathematics in Lean

- 1. Introduction
  - 1.1. Getting Started
  - 1.2. Overview
- 2. Basics
  - 2.1. Calculating
  - 2.2. Proving Identities in Algebraic Structures
  - 2.3. Using Theorems and Lemmas
  - 2.4. More examples using apply and rw
  - 2.5. Proving Facts about Algebraic Structures
- 3. Logic
  - 3.1. Implication and the Universal Quantifier
  - 3.2. The Existential Quantifier
  - 3.3. Negation
  - 3.4. Conjunction and Iff
  - 3.5. Disjunction
  - 3.6. Sequences and Convergence
- 4. Sets and Functions
  - 4.1. Sets
  - 4.2. Functions
  - 4.3. The Schröder-Bernstein Theorem
- 5. Elementary Number Theory
  - 5.1. Irrational Roots
  - 5.2. Induction and Recursion
  - 5.3. Infinitely Many Primes
  - 5.4. More Induction
- 6. Discrete Mathematics
  - 6.1. Finsets and Fintypes

CONTENTS:

1. Introduction
2. Mathematical Background
3. Lean as a Programming Language
4. Propositional Logic
5. Implementing Propositional Logic
6. Decision Procedures for Propositional Logic
7. Using SAT Solvers
8. Proof Systems for Propositional Logic
9. Using Lean as a Proof Assistant
10. First-Order Logic
11. Implementing First-Order Logic
12. Decision Procedures for Equality
13. Equality and Induction in Lean
14. Decision Procedures for Arithmetic
15. Using SMT solvers
16. Proof Systems for First-Order Logic
17. Using First-Order Theorem Provers
18. Beyond First-Order Logic

Logic and Mechanized Reasoning

Contents:

- 1. Introduction
  - 1.1. Historical background
  - 1.2. An overview of this course
- 2. Mathematical Background
  - 2.1. Induction and recursion on the natural numbers
  - 2.2. Complete induction
  - 2.3. Generalized induction and recursion
  - 2.4. Invariants
  - 2.5. Exercises
- 3. Lean as a Programming Language
  - 3.1. About Lean
  - 3.2. Using Lean as a functional programming language
  - 3.3. Inductive data types in Lean
  - 3.4. Using Lean as an imperative programming language
  - 3.5. Exercises
- 4. Propositional Logic
  - 4.1. Syntax
  - 4.2. Semantics
  - 4.3. Calculating with propositions
  - 4.4. Complete sets of connectives
  - 4.5. Normal forms
  - 4.6. Exercises
- 5. Implementing Propositional Logic
  - 5.1. Syntax
  - 5.2. Semantics
  - 5.3. Normal Forms

Using AI to Teach Mathematics

Using AI to teach mathematics is harder.

- The technology can become the focus, distracting from mathematics.
- Students may rely on AI, rather than learning.
- AI versions of skills may not transfer.

Positives:

- Interaction allows students to try things and see what happens.
- There is constant feedback and correction.
- There are online user communities.
- Students enjoy it and are engaged.

Preface
1. Proofs by calculation
2. Proofs with structure
3. Parity and divisibility
4. Proofs with structure, II
5. Logic
6. Induction
7. Number theory
8. Functions
9. Sets
10. Relations
Index of Lean tactics
Transitioning to mainstream Lean

The Mechanics of Proof

This is a book dealing with how to write careful, rigorous mathematical proofs. The book is paired with code in the computer formalization language Lean. Head over to the associated GitHub repository, https://github.com/hrmacbeth/math2001, to download this code to your own computer or to open it in the cloud on Gitpod.

This book is aimed at the early university level and has been written for the course Math 2001, at Fordham University. Please reach out to the author, Heather Macbeth, with comments and corrections.

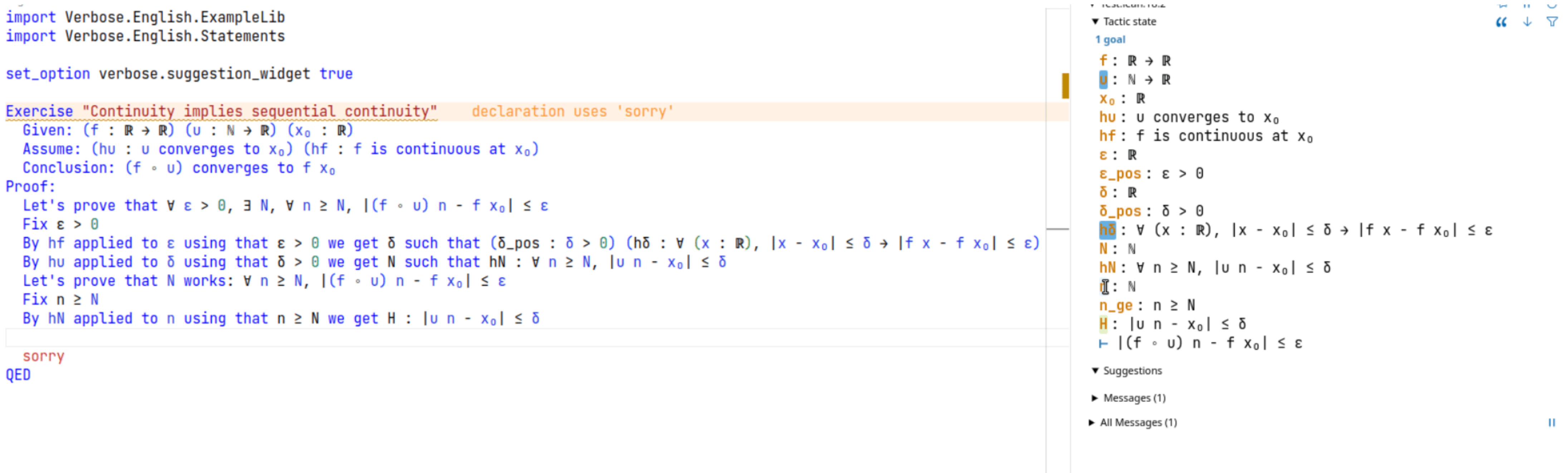
- Preface
  - About this book
  - Why Lean?
  - Contents and prerequisites
  - Note for instructors
  - Acknowledgements
- 1. Proofs by calculation
  - 1.1. Proving equalities
  - 1.2. Proving equalities in Lean
  - 1.3. Tips and tricks
  - 1.4. Proving inequalities
  - 1.5. A shortcut
- 2. Proofs with structure
  - 2.1. Intermediate steps
  - 2.2. Invoking lemmas
  - 2.3. “Or” and proof by cases
  - 2.4. “And”
  - 2.5. Existence proofs
- 3. Parity and divisibility

Verbose Lean 4

This project provides tactics and commands for Lean in a very controlled natural language. The original version of those tactics were written in French for teaching purposes at Université Paris-Saclay in Orsay using Lean 3. The goal is not to make Lean code easier to write, the goal is to make Lean code easier to transfer to a traditional paper proof.

The best way to have a quick look is to read the examples file in English or French, although GitHub obviously misses proper syntax highlighting here.

There is also a point-and-click interface for courses with a low time budget. One can see it in the following animated gif.



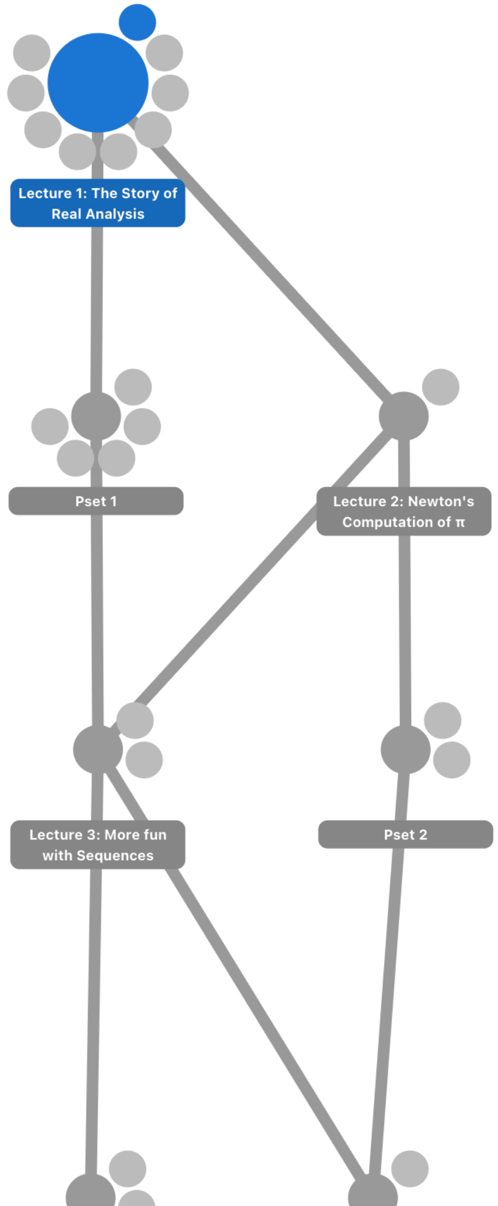
Real Analysis, The Game

Welcome to Real Analysis, The Game! (v0.1)

This course is was developed for Rutgers University Math 311H by Alex Kontorovich.

Follow along with the course lecture notes and videos, available here: https://alexkontorovich.github.io/2025F311H.

This course takes you through an Introduction to the Real Numbers, rigorous ε - δ Calculus, and basic Point-Set Topology. To get started, click on **"Level 1: The Story of Real Analysis"**, and good luck!



Rules ?

- regular
- relaxed
- none

Theorems | **Tactics** | Definitions

apply, bound, by_cases, by_contra, cases', change, choose, contradiction, exact_mod_cast, field_simp, have, induction', intro, left, let, linarith, norm_num, push_cast, push_neg, rewrite, rfl, right, ring_nf, show, specialize, split_ands, use

Teaching with AI

We're in the Wild West.

We need collaborations between:

- mathematicians,
- computer scientists,
- educators,
- education researchers,
- and more.

Institute for Computer-Aided Reasoning in Mathematics

icarm.io

A New Institute

The *Institute for Computer-Aided Reasoning in Mathematics* (ICARM) is a pilot 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.

Conclusions

Conclusions

AI will change:

- the way we do mathematics
- the way we do everything

We need to think about:

- why we do mathematics
- what we need to teach our children
- how to do it

It’s an exciting time for mathematics, but we need to be careful.