Computer-Assisted Mathematics

Automath Seminar

Assia Mahboubi

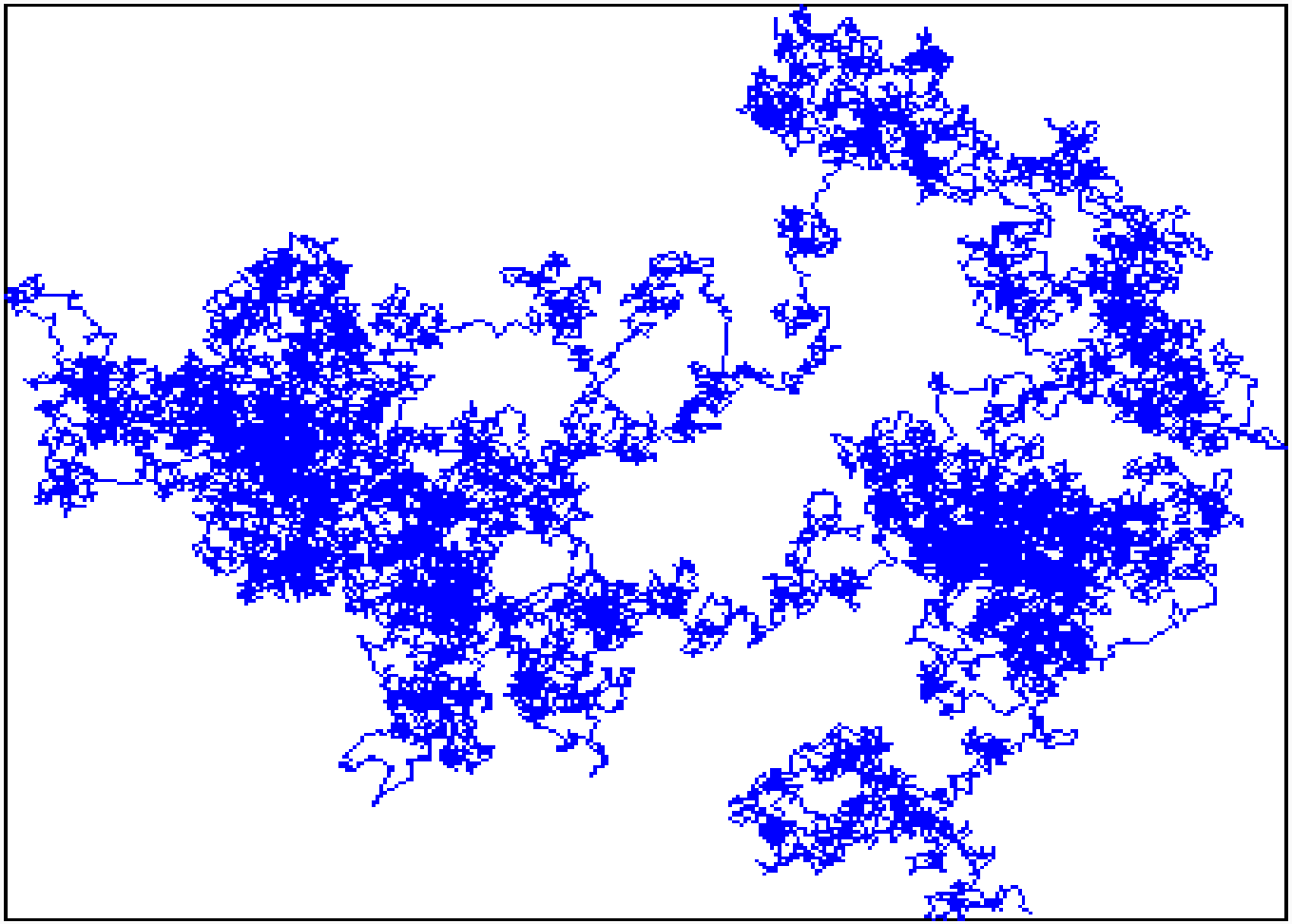
March 19th 2026

Inria, LS2N, Nantes Université, Vrije Universiteit Amsterdam

European Research Council
Established by the European Commission

Explore

Fractals



The notion that these **conjectures** might have been reached by pure thought – with no picture – is simply inconceivable" B. Mandelbrot, 1982

[Mathematics in the Age of the Turing Machine, Thomas Hales, ASL Lecture Notes in Logic. 2013]

A one million dollar problem

CMI

ABOUT PROGRAMS MILLENNIUM PROBLEMS PEOPLE PUBLICATIONS EVEN

Birch and Swinnerton-Dyer Conjecture

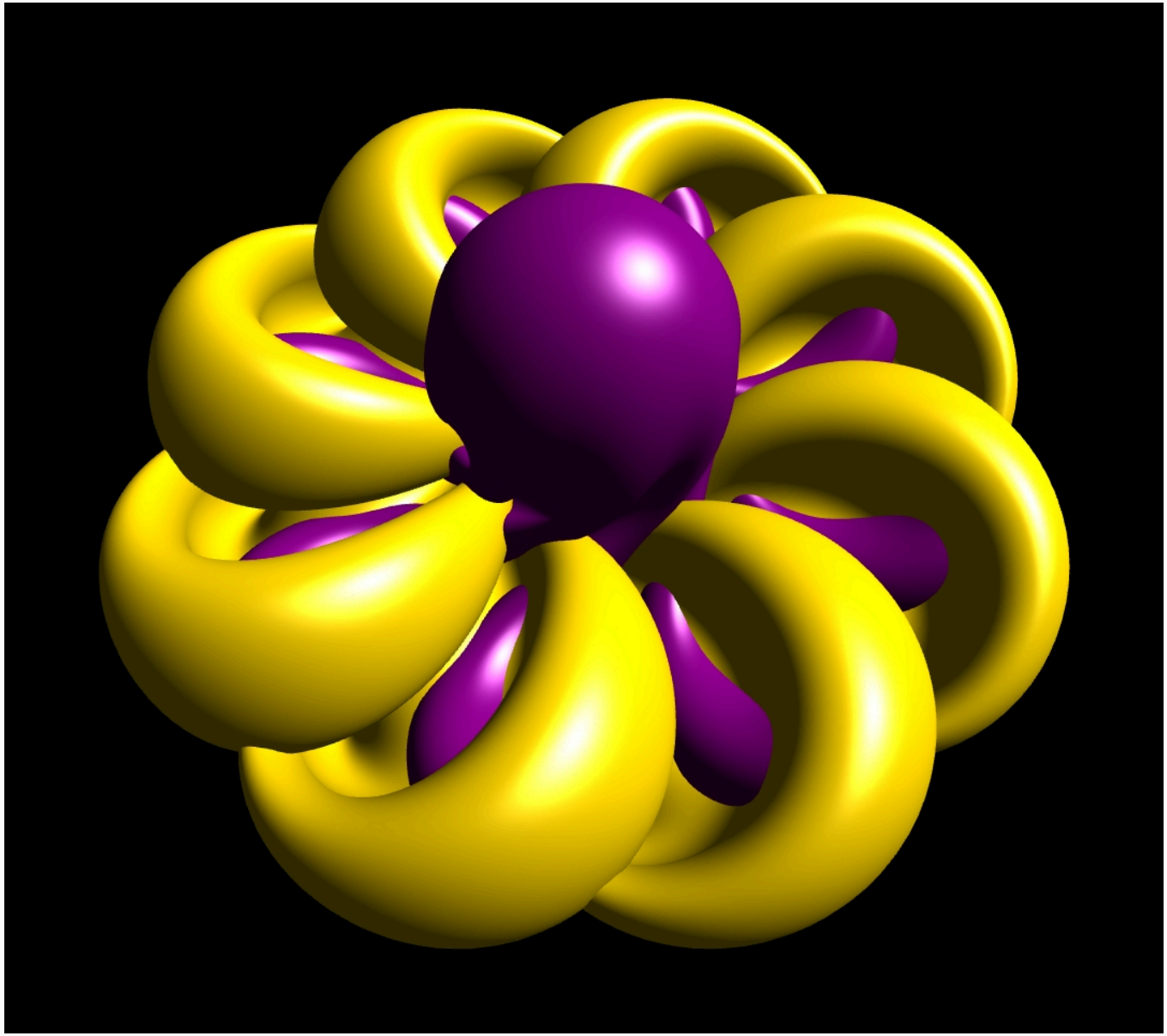
Mathematicians have always been fascinated by the problem of describing all solutions in whole numbers x,y,z to algebraic equations like

$$x^2 + y^2 = z^2$$

Euclid gave the complete solution for that equation, but for more complicated equations this becomes extremely difficult. Indeed, in 1970 Yu. V. Matiyasevich showed that Hilbert's tenth problem is unsolvable, i.e., there is no general method for determining when such equations have a solution in whole numbers. But in special cases one can hope to say something. When the solutions are the points of an abelian variety, the Birch and Swinnerton-Dyer conjecture asserts that the size of the group of rational points is related to the behavior of an associated zeta function $\zeta(s)$ near the point $s=1$. In particular this amazing conjecture asserts that if $\zeta(1)$ is equal to 0, then there are an infinite number of rational points (solutions), and conversely, if $\zeta(1)$ is not equal to 0, then there is only a finite number of such points.

This problem is: Unsolved

Visualizing the invisible



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

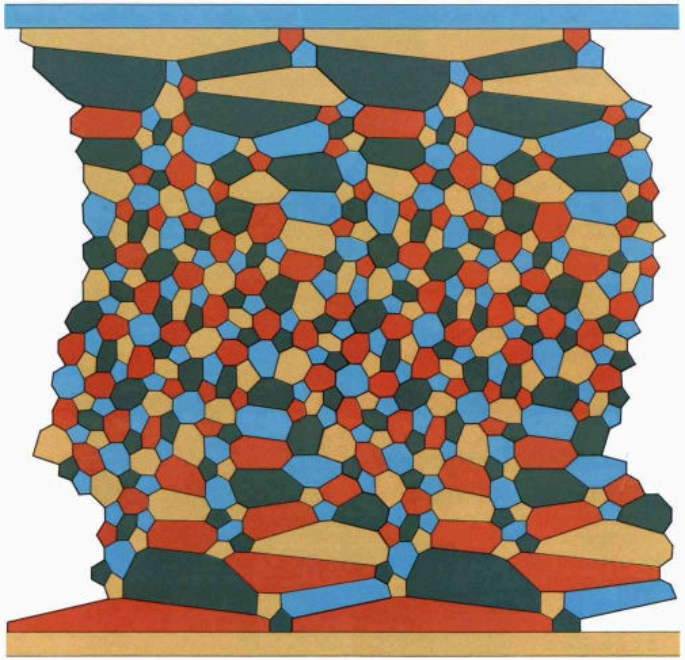
Check for updates

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.

Prove

Four colors suffice



- Conjecture by F. Guthrie (1852)
- Computer-assisted proof : K. Appel, W. Haken (1976)
  ∼1200h computation required by the time of the publication

Kepler conjecture

- Conjectured by J. Kepler, in *trena Seu de Nive Sexangula* (1611)
- Computer-assisted strategy: L. Fejes Tóth (1953)
- Proof: Th. Hales, S. Ferguson (1998)

[A proof of the Kepler conjecture, T. C. Hales, Annals of Mathematics, 2005]

Kepler conjecture

Reviewing process by the Annals of Mathematics:

- 4 years of intensive work by a team of 12 reviewers led by G. Fejes Toth
- Report: 99% certain but unable to completely certify the proof.

Kepler conjecture

Reviewing process by the Annals of Mathematics:

- 4 years of intensive work by a team of 12 reviewers led by G. Fejes Toth
- Report: 99% certain but unable to completely certify the proof.

Report by R. MacPherson, editor of the journal:

- The news from the referees is bad, from my perspective (...)
- They have run out of energy to devote to the problem.
- FT thinks that this situation will occur more and more often in mathematics.
- The mathematical community will have to get used to this state of affairs.

Kepler conjecture

Reviewing process by the Annals of Mathematics:

- 4 years of intensive work by a team of 12 reviewers led by G. Fejes Toth
- Report: 99% certain but unable to completely certify the proof.

Report by R. MacPherson, editor of the journal:

- The news from the referees is bad, from my perspective (...)
- They have run out of energy to devote to the problem.
- FT thinks that this situation will occur more and more often in mathematics.
- The mathematical community will have to get used to this state of affairs.

Epilogue:

[The Annals] will no longer attempt to check the correctness of computer code.

Ternary Goldbach Conjecture

Theorem

Every odd integer greater than 5 can be expressed as the sum of three primes.

$$
\begin{aligned}
7 &= 3 + 2 + 2 \\
9 &= 3 + 3 + 3 \\
11 &= 5 + 3 + 3 \\
13 &= 5 + 5 + 3 \\
15 &= 5 + 5 + 5 \\
17 &= 5 + 5 + 7 \\
19 &= 5 + 7 + 7 \\
21 &= 7 + 7 + 7 \\
23 &= 17 + 3 + 3 \\
\ldots &= \ldots
\end{aligned}
$$

Ternary Goldbach Conjecture

Numerical Verification of the Ternary Goldbach Conjecture up to 8.875e30

H.A. Helfgott, David J. Platt

(Submitted on 14 May 2013 (v1), last revised 1 Apr 2014 (this version, v2))

We describe a computation that confirms the ternary Goldbach Conjecture up to 8,875,694,145,621,773,516,800,000,000,000 (>8.875e30).

Comments: 4 pages

Major arcs for Goldbach's problem

H. A. Helfgott

(Submitted on 13 May 2013 (v1), last revised 14 Apr 2014 (this version, v4))

The ternary Goldbach conjecture states that every odd number $n \geq 7$ is the sum of three primes. The estimation of the Fourier series $\sum_{p \leq x} e(\alpha p)$ and related sums has been central to the study of the problem since Hardy and Littlewood (1923). Here we show how to estimate such Fourier series for α in the so-called major arcs, i.e., for α close to a rational of small denominator. This is part of the author's proof of the ternary Goldbach conjecture. In contrast to most previous work on the subject, we will rely on a finite verification of the Generalized Riemann Hypothesis up to a bounded conductor and bounded height, rather than on zero-free regions. We apply a rigorous verification due to D. Platt; the results we obtain are both rigorous and unconditional. The main point of the paper will be the development of estimates on parabolic cylinder functions that make it possible to use smoothing functions based on the Gaussian. The generality of our explicit formulas will allow us to work with a wide variety of such functions.

Minor arcs for Goldbach's problem

H. A. Helfgott

(Submitted on 23 May 2012 (v1), last revised 30 Dec 2013 (this version, v4))

The ternary Goldbach conjecture states that every odd number n>=7 is the sum of three primes. The estimation of sums of the form \sum_{p\leq x} e(\alpha p), \alpha = a/q + O(1/q^2), has been a central part of the main approach to the conjecture since (Vinogradov, 1937). Previous work required q or x to be too large to make a proof of the conjecture for all n feasible.
The present paper gives new bounds on minor arcs and the tails of major arcs. This is part of the author's proof of the ternary Goldbach conjecture.
The new bounds are due to several qualitative improvements. In particular, this paper presents a general method for reducing the cost of Vaughan's identity, as well as a way to exploit the tails of minor arcs in the context of the large sieve.

Comments: 79 pages; third version. (A couple of explanatory paragraphs have been added.)

[Accepted for publication in Annals of Mathematics, 2015, under revision]

Ternary Goldbach Conjecture

By Cauchy-Schwarz, this is at most

$$\sqrt{\frac{1}{2\pi}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty}\left|\frac{L'(s,\chi)}{L(s,\chi)}\cdot\frac{1}{s}\right|^2|ds|}\cdot\sqrt{\frac{1}{2\pi}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty}|G_\delta(s)s|^2\,|ds|}$$

By (4.12),

$$\begin{aligned}\sqrt{\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty}\left|\frac{L'(s,\chi)}{L(s,\chi)}\cdot\frac{1}{s}\right|^2|ds|} &\le \sqrt{\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty}\left|\frac{\log q}{s}\right|^2|ds|}\\ &\quad+\sqrt{\int_{-\infty}^{\infty}\frac{\left|\frac{1}{2}\log\left(\tau^2+\frac{9}{4}\right)+4.1396+\log\pi\right|^2}{\frac{1}{4}+\tau^2}d\tau}\\ &\le \sqrt{2\pi}\log q+\sqrt{226.844},\end{aligned}$$

where we compute the last integral numerically.[4]

[4] By a rigorous integration from $\tau=-100000$ to $\tau=100000$ using VNODE-LP [Ned06], which runs on the PROFIL/BIAS interval arithmetic package [Knü99].

Computer-authored proofs

2024

■ [j87] Shalosh B. Ekhad, Doron Zeilberger:
Experimenting with Standard Young Tableaux. Math. Comput. Sci. 18(2): 10 (2024)

■ [i12] Tipaluck Krityakierne, Thotsaporn Thanatipanonda, Doron Zeilberger:
von Neumann and Newman Pokers with Finite Decks. CoRR abs/2407.16155 (2024)

2023

■ [j86] Shalosh B. Ekhad, Doron Zeilberger:
Counting Clean Words According to the Number of Their Clean Neighbors. ACM Commun. Comput. Algebra 57(1): 5-9 (2023)

■ [j85] Lucy Martinez, Doron Zeilberger:
How Many Dice Rolls Would It Take to Hit Your Favorite Kind of Number? Maple Trans. 3(3) (2023)

■ [c2] Guy Katriel, Udi Mahanaymi, Christoph Koutschan, Doron Zeilberger, Mike A. Steel, Sagi Snir:
Using Generating Functions to Prove Additivity of Gene-Neighborhood Based Phylogenetics - Extended Abstract. ISBRA 2023: 120-135

2022

■ [j84] Yukun Yao, Doron Zeilberger:
Numerical and Symbolic Studies of the Peaceable Queens Problem. Exp. Math. 31(1): 269-279 (2022)

■ [i11] Robert Dougherty-Bliss, Doron Zeilberger:

▪ number of authors per paper
or build your own?

[–] **Refine list**

showing all 105 records

'20 '10 '00 '90 '80

refine by search term

refine by type
- ☑ Journal Articles (only)
- ☑ Conference and Workshop Papers (only)
- ☑ Parts in Books or Collections (only)
- ☑ Informal and Other Publications (only)

select all | deselect all

refine by coauthor
no coauthors (32)
Shalosh B. Ekhad (8)

(. . .)

1996

■ [j35] Shalosh B. Ekhad, Doron Zeilberger:
The Number of Solutions of X^2=0 in Triangular Matrices Over GF(q). Electron. J. Comb. 3(1) (1996)

■ [j34] Doron Zeilberger:
Proof of the alternating sign matrix conjecture. Electron. J. Comb. 3(2) (1996)

A 200 terabytes proof

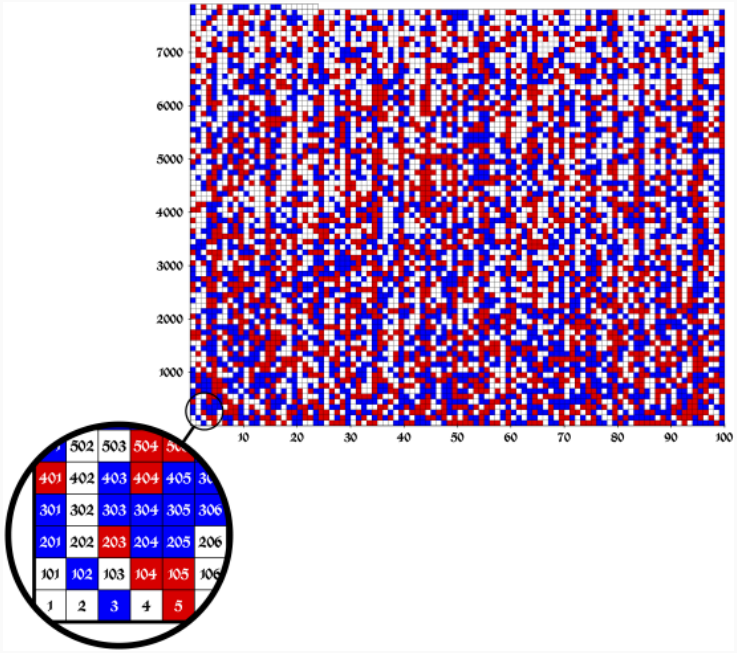
Pythagorean triples problem (Conjecture by R. Graham, 1980)

Is it possible to color each of the positive integers either red or blue, so that no Pythagorean triple of integers a, b, c, satisfying $a^2 + b^2 = c^2$ are all the same color?

A 200 terabytes proof

Pythagorean triples problem (Conjecture by R. Graham, 1980)

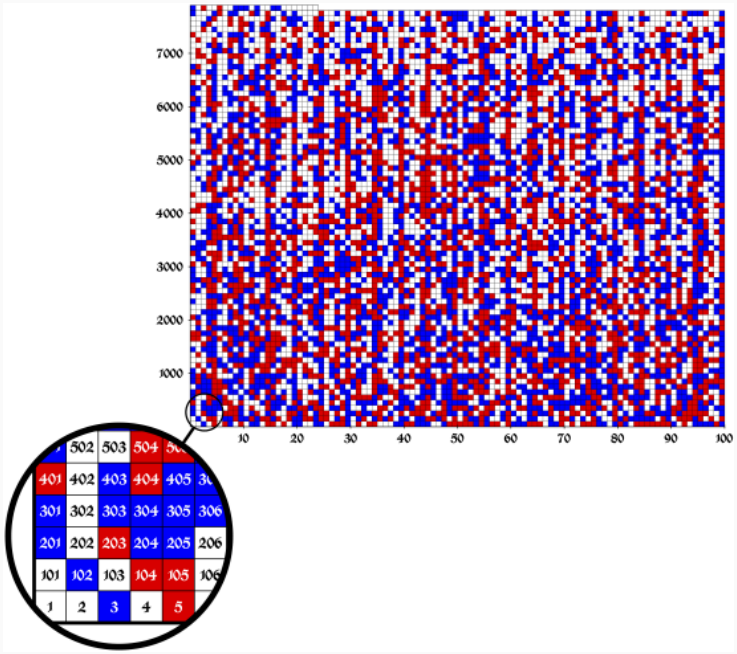
Is it possible to color each of the positive integers either red or blue, so that no Pythagorean triple of integers a, b, c, satisfying $a^2 + b^2 = c^2$ are all the same color?



A 200 terabytes proof

Pythagorean triples problem (Conjecture by R. Graham, 1980)

Is it possible to color each of the positive integers either red or blue, so that no Pythagorean triple of integers a, b, c, satisfying $a^2 + b^2 = c^2$ are all the same color?

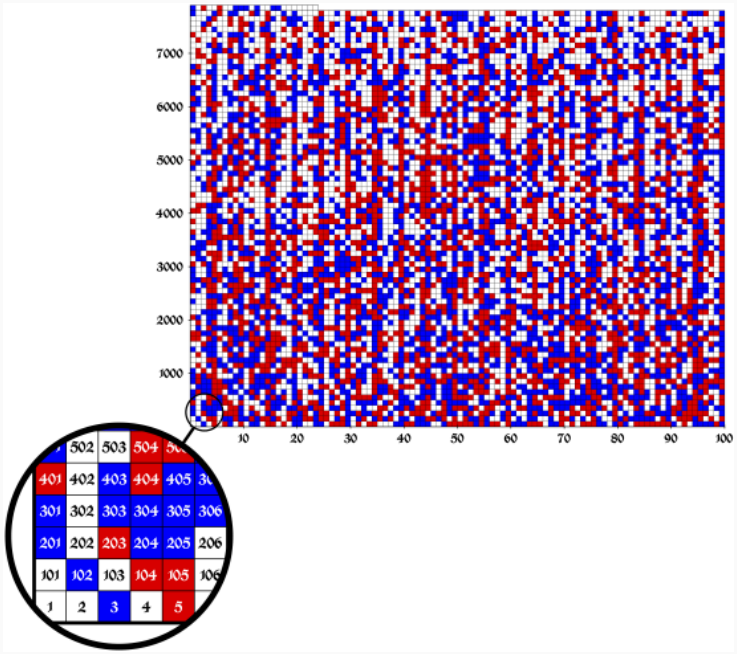


Answer: yes, but only up to 7824.

A 200 terabytes proof

Pythagorean triples problem (Conjecture by R. Graham, 1980)

Is it possible to color each of the positive integers either red or blue, so that no Pythagorean triple of integers a, b, c, satisfying $a^2 + b^2 = c^2$ are all the same color?



Answer: yes, but only up to 7824. Proof: By brute force.

[Solving and Verifying the Boolean Pythagorean Triples problem via Cube-and-Conquer, M. Heule et al. SAT 2016]

Computer-produced handbooks

LMFDB

⌂

Feedback · Hide Menu

LMFDB - The L-functions and Modular Forms Database

Introduction
Overview · Random
Universe · Knowledge

L-functions
Rational · All

Modular forms
Classical · Maass
Hilbert · Bianchi

Varieties
Elliptic curves over $\mathbb{Q}$
Elliptic curves over $\mathbb{Q}(\alpha)$
Genus 2 curves over $\mathbb{Q}$
Higher genus families
Abelian varieties over $\mathbb{F}_q$

Fields
Number fields
p-adic fields

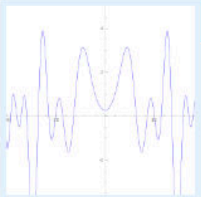
Representations
Dirichlet characters
Artin representations

Groups
Galois groups
Sato-Tate groups

A database

The LMFDB is an extensive database of mathematical objects arising in Number Theory.

Sample lists: L-functions, Elliptic curves, Tables of zeros, Number fields



Hall of fame

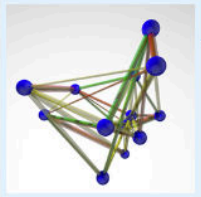
Riemann zeta function
Ramanujan Δ function and its L-function
C277 and its L-function
Gauss elliptic curve and its L-function
Grand Canyon L-function

Search and browse

Search for objects with specific properties, or browse categories.

Browse: L-functions, Modular forms, Elliptic curves, Number fields

See a random object from the database



Visualize data

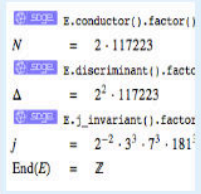
Explore individual plots or view distributions of various objects.

Examples: $\mathrm{GL}(4)$ Level one Maass forms, Isogeny graph of elliptic curve 102.c

Explore and learn

The LMFDB makes visible the connections predicted by the Langlands program. Knowls offer background information when you need it.

LMFDB universe · Knowledge



Code and open software

Download the data, download the code, or see how the data was generated.

GitHub · SageMath · Pari/GP · Magma · Python

This project is supported by grants from the US National Science Foundation, the UK Engineering and Physical Sciences Research Council, and the Simons Foundation.

Contact · Citation · Acknowledgments · Editorial Board · Source · SageMath version 9.2 · LMFDB Release 1.2.1

Computer-produced handbooks

LMFDB

⌂

Feedback · Hide Menu

LMFDB - The L-functions and Modular Forms Database

Introduction
Overview Random
Universe Knowledge

L-functions
Rational All

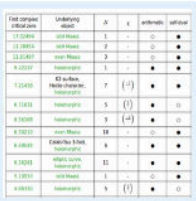
Modular forms
Classical Maass
Hilbert Bianchi

Varieties
Elliptic curves over $\mathbb{Q}$
Elliptic curves over $\mathbb{Q}(\alpha)$
Genus 2 curves over $\mathbb{Q}$
Higher genus families
Abelian varieties over $\mathbb{F}_q$

Fields

Artin representations

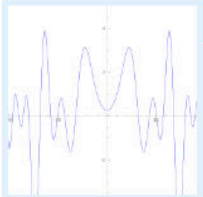
Groups
Galois groups
Sato-Tate groups



A database

The LMFDB is an extensive database of mathematical objects arising in Number Theory.

Sample lists: L-functions, Elliptic curves, Tables of zeros, Number fields



Hall of fame

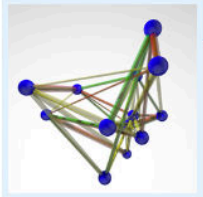
Riemann zeta function
Ramanujan Δ function and its L-function
C277 and its L-function
Gauss elliptic curve and its L-function
Grand Canyon L-function

Search and browse

Search for objects with specific properties, or browse categories.

Browse: L-functions, Modular forms, Elliptic curves, Number fields

See a random object from the database



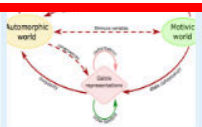
Visualize data

Explore individual plots or view distributions of various objects.

Examples: $\mathrm{GL}(4)$ Level one Maass forms, Isogeny graph of elliptic curve 102.c

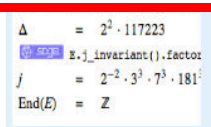
Integral points

These were computed rigorously, using independent implementations in Magma and SageMath which were compared as a consistency check.



by the Langlands program. Knowls offer background information when you need it.

LMFDB universe Knowledge

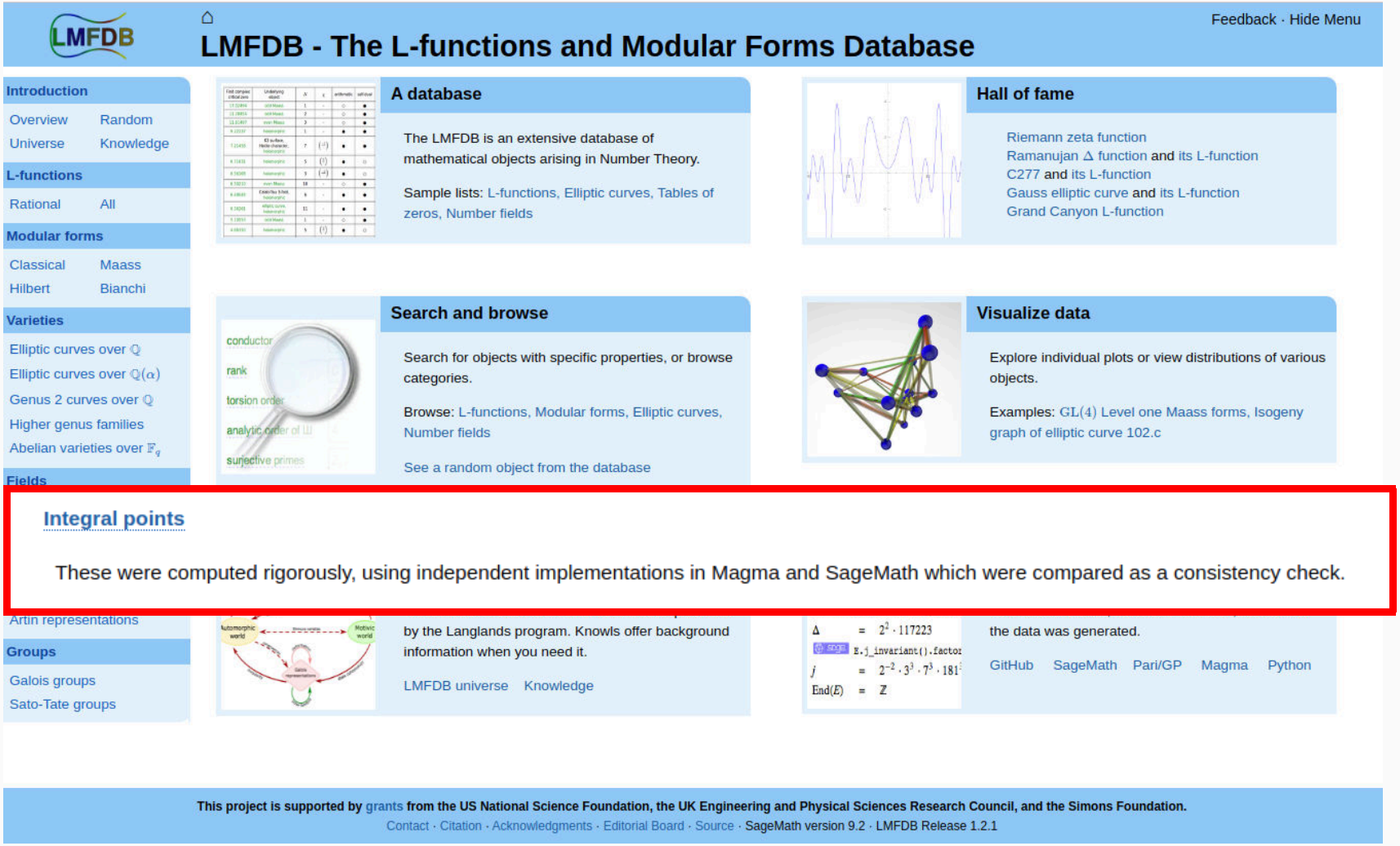


the data was generated.

GitHub SageMath Pari/GP Magma Python

This project is supported by grants from the US National Science Foundation, the UK Engineering and Physical Sciences Research Council, and the Simons Foundation.

Computer-produced handbooks



Problem

Today, no explicit policy for auditing software that produce proof steps.

“Are we just getting the wrong answer faster?”

What is a theorem?

[Alexander Grothendieck at the blackboard during a lesson at IHES, Courtesy of IHES]

What is a theorem?

Bad things happen

This bound is wrong:

MAJOR ARCS FOR GOLDBACH'S PROBLEM 35

By Cauchy-Schwarz, this is at most

$$\sqrt{\frac{1}{2\pi}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty}\left|\frac{L'(s,\chi)}{L(s,\chi)}\cdot\frac{1}{s}\right|^2|ds|}\cdot\sqrt{\frac{1}{2\pi}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty}|G_\delta(s)s|^2\,|ds|}$$

By (4.12),

$$\begin{aligned}\sqrt{\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty}\left|\frac{L'(s,\chi)}{L(s,\chi)}\cdot\frac{1}{s}\right|^2|ds|} &\le \sqrt{\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty}\left|\frac{\log q}{s}\right|^2|ds|}\\ &\quad+\sqrt{\int_{-\infty}^{\infty}\frac{\left|\frac{1}{2}\log\left(\tau^2+\frac{9}{4}\right)+4.1396+\log\pi\right|^2}{\frac{1}{4}+\tau^2}d\tau}\\ &\le \sqrt{2\pi}\log q+\sqrt{226.844},\end{aligned}$$

where we compute the last integral numerically.[4]

[4] By a rigorous integration from $\tau=-100000$ to $\tau=100000$ using VNODE-LP [Ned06], which runs on the PROFIL/BIAS interval arithmetic package [Knü99].

Bad things happen

This bound is wrong:

By Cauchy-Schwarz, this is at most

$$\sqrt{\frac{1}{2\pi}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty}\left|\frac{L'(s,\chi)}{L(s,\chi)}\cdot\frac{1}{s}\right|^2|ds|}\cdot\sqrt{\frac{1}{2\pi}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty}|G_\delta(s)s|^2\,|ds|}$$

By (4.12),

$$\begin{aligned}\sqrt{\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty}\left|\frac{L'(s,\chi)}{L(s,\chi)}\cdot\frac{1}{s}\right|^2|ds|} &\le \sqrt{\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty}\left|\frac{\log q}{s}\right|^2|ds|}\\ &\quad+\sqrt{\int_{-\infty}^{\infty}\frac{\left|\frac{1}{2}\log\left(\tau^2+\frac{9}{4}\right)+4.1396+\log\pi\right|^2}{\frac{1}{4}+\tau^2}d\tau}\\ &\le \sqrt{2\pi}\log q+\sqrt{226.844},\end{aligned}$$

where we compute the last integral numerically.[4]

[4] By a rigorous integration from $\tau=-100000$ to $\tau=100000$ using VNODE-LP [Ned06], which runs on the PROFIL/BIAS interval arithmetic package [Knü99].

Fortunately the proof survives.

Unsolvable bugs

$$\int_0^1 \left|\left(x^4 + 10x^3 + 19x^2 - 6x - 6\right)\ e^x\right| dx \simeq 11.14731055005714$$

Unsolvable bugs

$$\int_0^1 \left|\left(x^4 + 10x^3 + 19x^2 - 6x - 6\right)\, e^x\right| dx \simeq 11.14731055005714$$

May 2016:

- Octave: `quad`/`quadgk`: only 10/9 correct digits;

Unsolvable bugs

$$\int_0^1 \left|\left(x^4 + 10x^3 + 19x^2 - 6x - 6\right)\, e^x\right| dx \simeq 11.14731055005714$$

May 2016:

- Octave: `quad`/`quadgk`: only 10/9 correct digits;
- INTLAB `verifyquad`: false answer, without warning;

Unsolvable bugs

$$\int_0^1 \left|\left(x^4 + 10x^3 + 19x^2 - 6x - 6\right)\, e^x\right| dx \simeq 11.14731055005714$$

May 2016:

- Octave: `quad`/`quadgk`: only 10/9 correct digits;
- INTLAB `verifyquad`: false answer, without warning;
- VNODE-LP: not usable (cf. absolute value).

$\Rightarrow$ INTLAB removed support for the absolute value.

```
53  arb_sqrt(arb_t z, const arb_t x, slong prec)
54  {
55      mag_t rx, zr;
56      int inexact;
57
58      if (mag_is_zero(arb_radref(x)))
59      {
60          arb_sqrt_arf(z, arb_midref(x), prec);
61      }
62      else if (arf_is_special(arb_midref(x)) ||
63               arf_sgn(arb_midref(x)) < 0 || mag_is_inf(arb_radref(x)))
64      {
65          if (arf_is_pos_inf(arb_midref(x)) && mag_is_finite(arb_radref(x)))
66              arb_sqrt_arf(z, arb_midref(x), prec);
67          else
68              arb_indeterminate(z);
69      }
70      else  /* now both mid and rad are non-special values, mid > 0 */
71      {
72          slong acc;
73
74          acc = _fmpz_sub_small(ARF_EXPREF(arb_midref(x)), MAG_EXPREF(arb_radref(x)));
75          acc = FLINT_MIN(acc, prec);
76          prec = FLINT_MIN(prec, acc + MAG_BITS);
77          prec = FLINT_MAX(prec, 2);
78
79          if (acc < 0)
80          {
81              arb_indeterminate(z);
82          }
83          else if (acc <= 20)
84          {
85              mag_t t, u;
86
87              mag_init(t);
88              mag_init(u);
89
90              arb_get_mag_lower(t, x);
91
92              if (mag_is_zero(t) && arb_contains_negative(x))
93              {
```

Code from Arb/FLINT, Fast Library for Number Theory

Symbolic Computation

In SymPy 1.5.1 [1], compare

```
>>> simplify(hyper([n],[m],x).subs({m:-1, n:-1, x:1}))
                                                    2
```

with

```
>>> simplify(hyper([n],[m],x).subs(m, n)).subs({n:-1, x:1})
                                                  E
```

[1] Example suggested by F. Johansson.

Symbolic Computation

In SymPy 1.5.1 [1], compare

```
>>> simplify(hyper([n],[m],x).subs({m:-1, n:-1, x:1}))
                                                    2
```

with

```
>>> simplify(hyper([n],[m],x).subs(m, n)).subs({n:-1, x:1})
                                                  E
```

Wolfram Language (Mathematica) exhibit the exact same phenomenon.

$\Rightarrow$ Cross-verification is not enough.

[1] Example suggested by F. Johansson.

Machine-checked proofs

Program verification

- Recipe:
  - State expected properties on input
  - State desired properties on output
  - Inspect the code to prove implication

- Ingredients:
  - Appropriate specification language, expressive enough
  - (Human insight)
  - Automation

Program verification

- Recipe:
  - State expected properties on input
  - State desired properties on output
  - Inspect the code to prove implication

- Ingredients:
  - Appropriate specification language, expressive enough
  - (Human insight)
  - Automation

$\Rightarrow$ Tony Hoare (1934 - 2026)

Example: in-place inversion of a permutation

- From a permutation array and a few extra slots:

2	4	5	1	3			

- Compute the array of the inverse permutation:

4	1	5	2	3			

Note that this is mathematically trivial:

- Input: $1 \mapsto 2 \quad 2 \mapsto 4 \quad 3 \mapsto 5 \quad 4 \mapsto 1 \quad 5 \mapsto 3$

Example: in-place inversion of a permutation

- From a permutation array and a few extra slots:

- Compute the array of the inverse permutation:

Note that this is mathematically trivial:

- Input: $1 \mapsto 2 \quad 2 \mapsto 4 \quad 3 \mapsto 5 \quad 4 \mapsto 1 \quad 5 \mapsto 3$
- Output: $1 \leftarrow\!\shortmid 2 \quad 2 \leftarrow\!\shortmid 4 \quad 3 \leftarrow\!\shortmid 5 \quad 4 \leftarrow\!\shortmid 1 \quad 5 \leftarrow\!\shortmid 3$

Example: in-place inversion of a permutation

- From a permutation array and a few extra slots:

2	4	5	1	3			

- Compute the array of the inverse permutation:

4	1	5	2	3			

Example: in-place inversion of a permutation

- From a permutation array and a few extra slots:

2	4	5	1	3			

- Compute the array of the inverse permutation:

4	1	5	2	3			

Rules of the game:

- Overwritten data is lost.
- The number of extra slots does not depend on the permutation.

Program verification

An example from the Why3 gallery.

Program verification for mathematics

- Recipe:
  - State expected properties on input
  - State desired properties on output
  - Inspect the code to prove implication
  - Interpret symbolic data

- Ingredients:
  - Appropriate specification language, expressive enough
  - (Human insight)
  - Automation
  - Libraries of mechanized mathematics

Mechanized/formalized mathematics

Representing mathematics in a fixed formal language, i.e.:

- Defining mathematical objects, statements, (calculations,) proofs;
- Verifying the correctness of these proofs by a mechanical process.

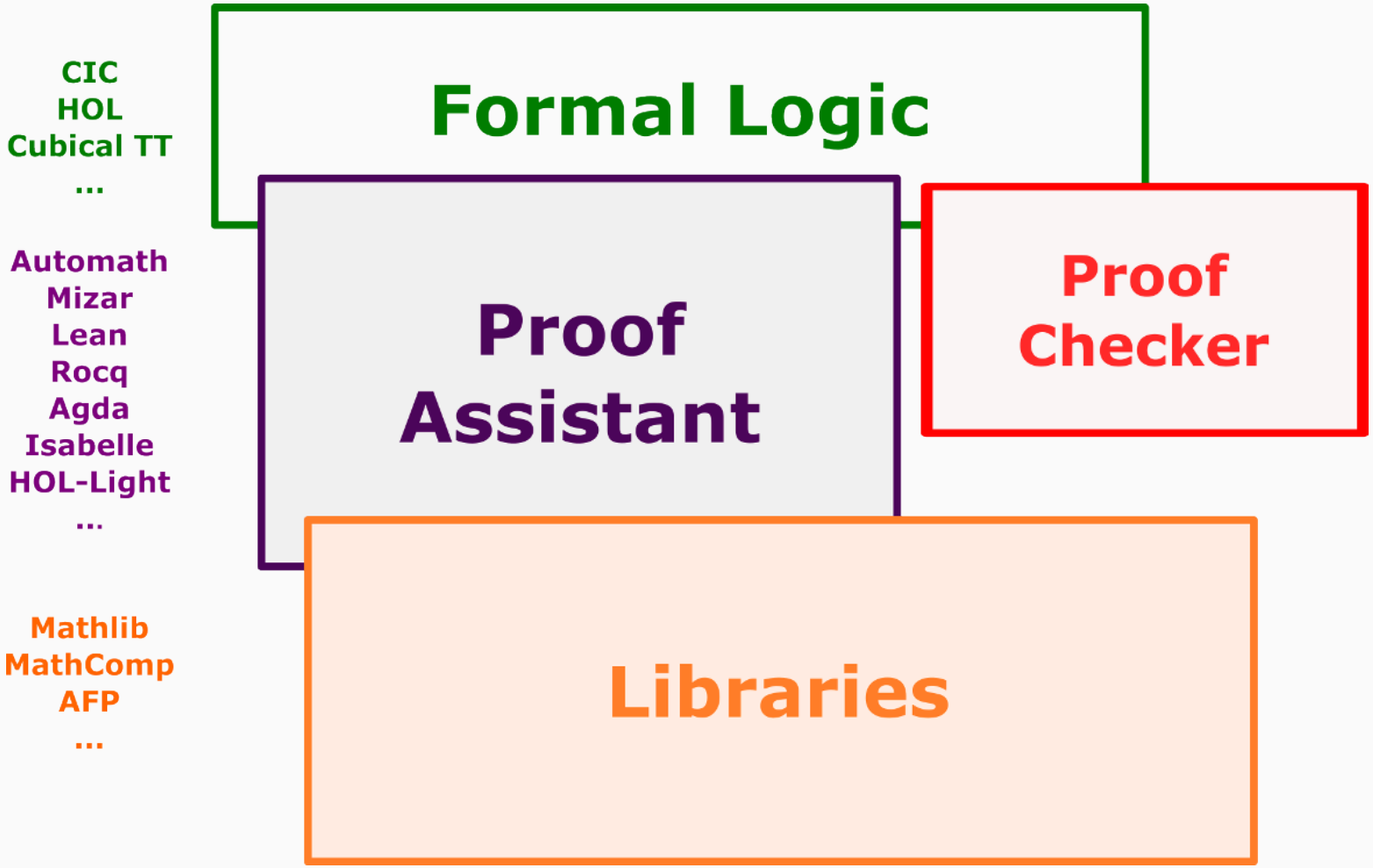
Mechanized/formalized mathematics

Representing mathematics in a fixed formal language, i.e.:

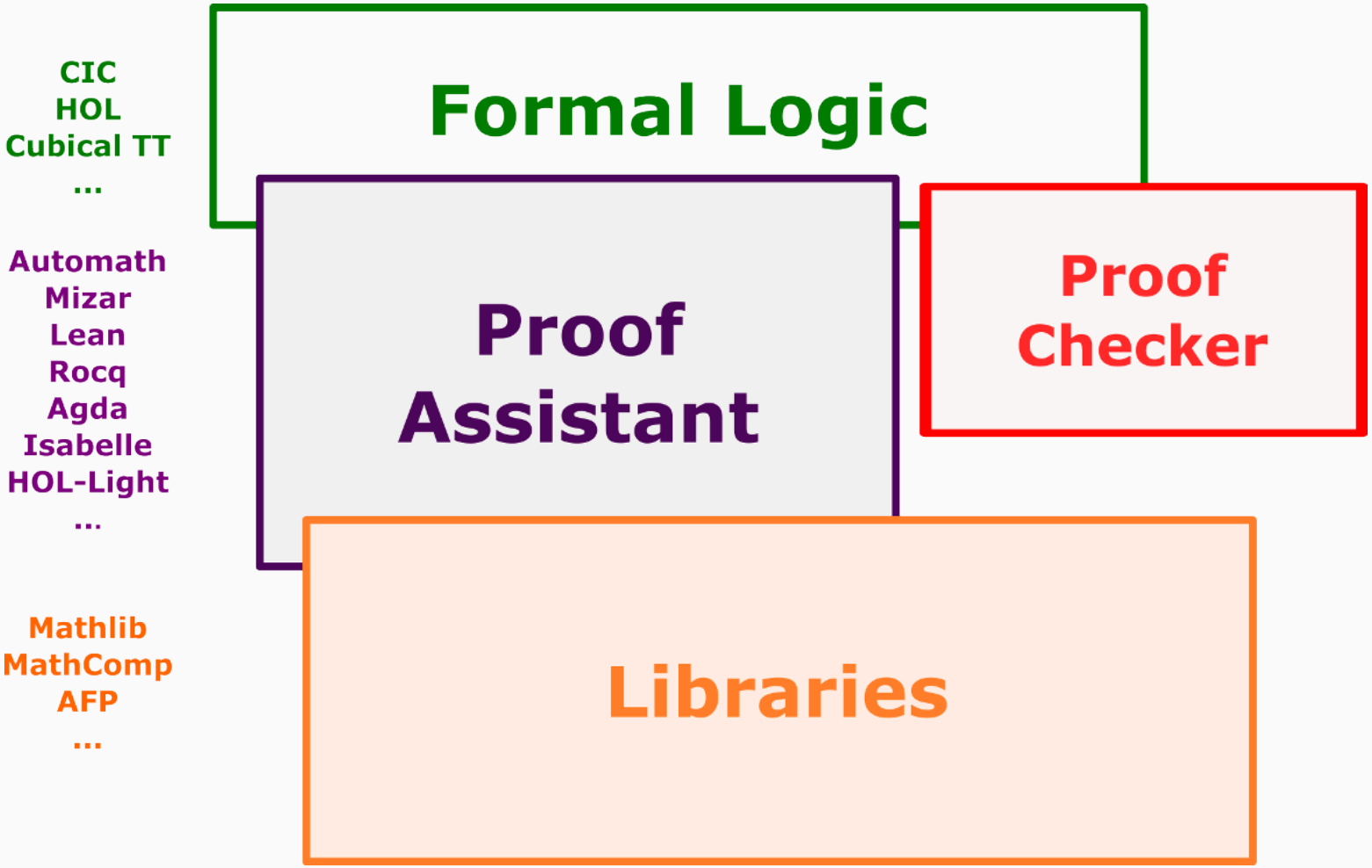
- Defining mathematical objects, statements, (calculations,) proofs;
- Verifying the correctness of these proofs by a mechanical process.

$\Rightarrow$ And use a computer software for doing so.

Skeleton of an interactive theorem prover

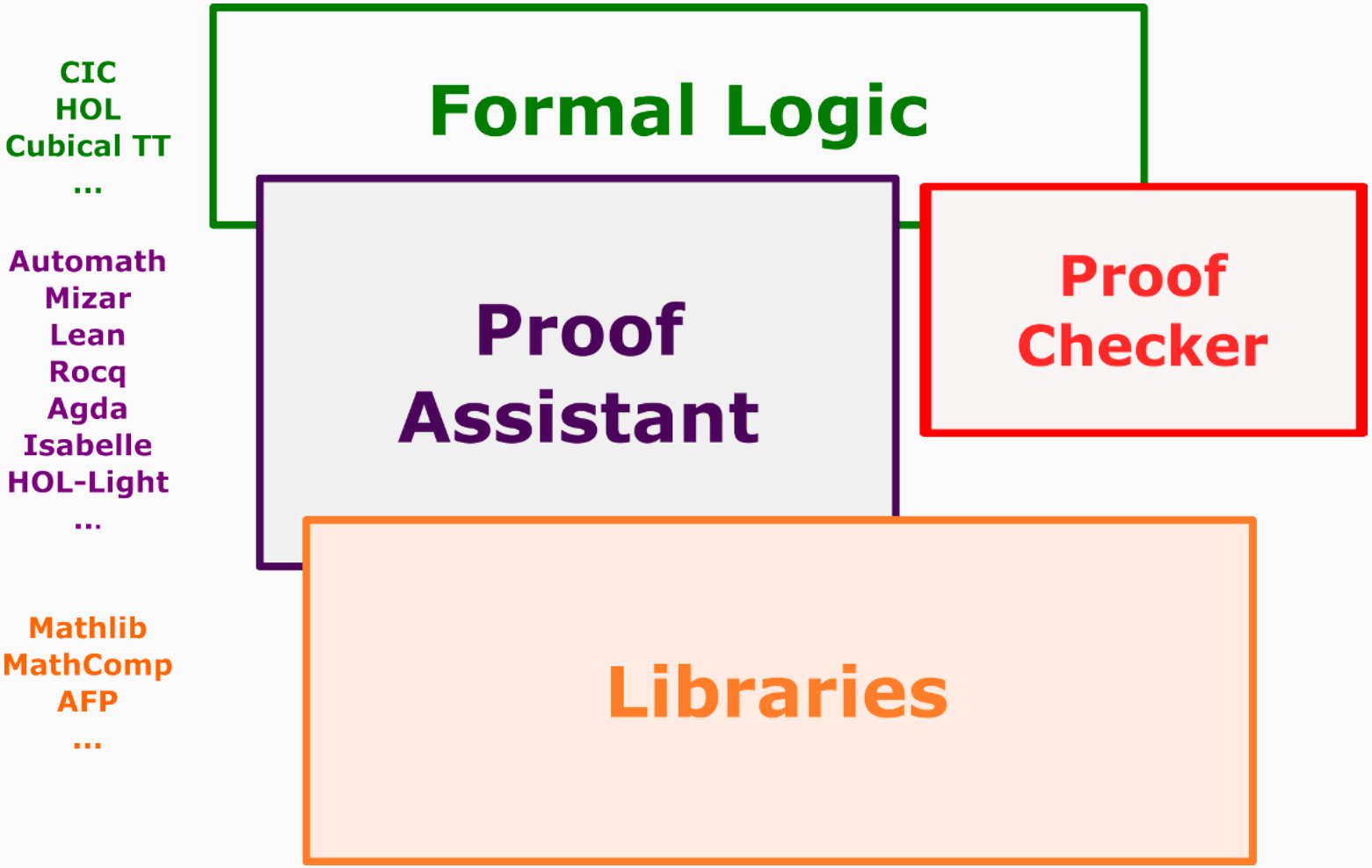


Skeleton of an interactive theorem prover



Acts of faith:

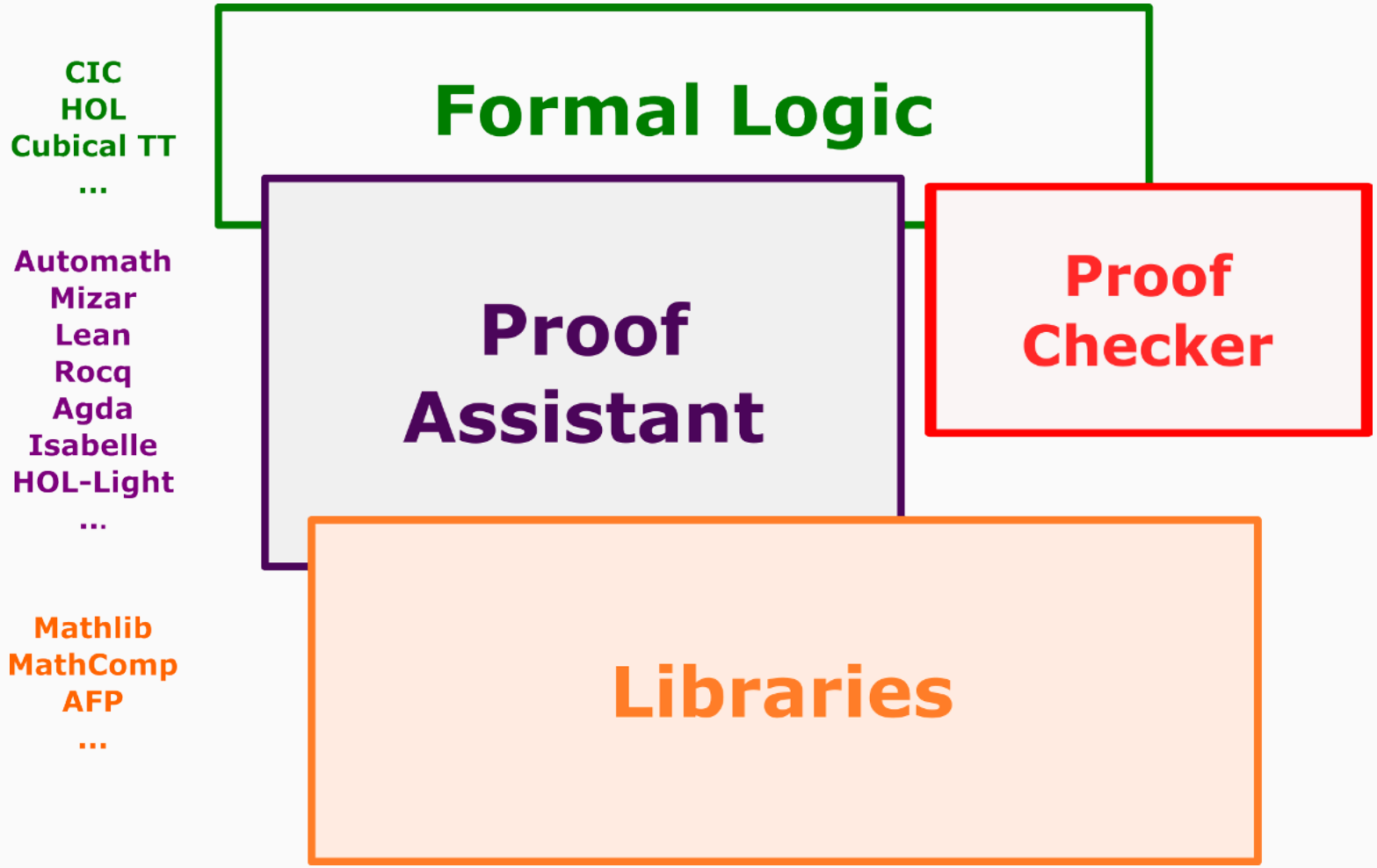
Skeleton of an interactive theorem prover



Acts of faith:

- Correctness of the proof checker

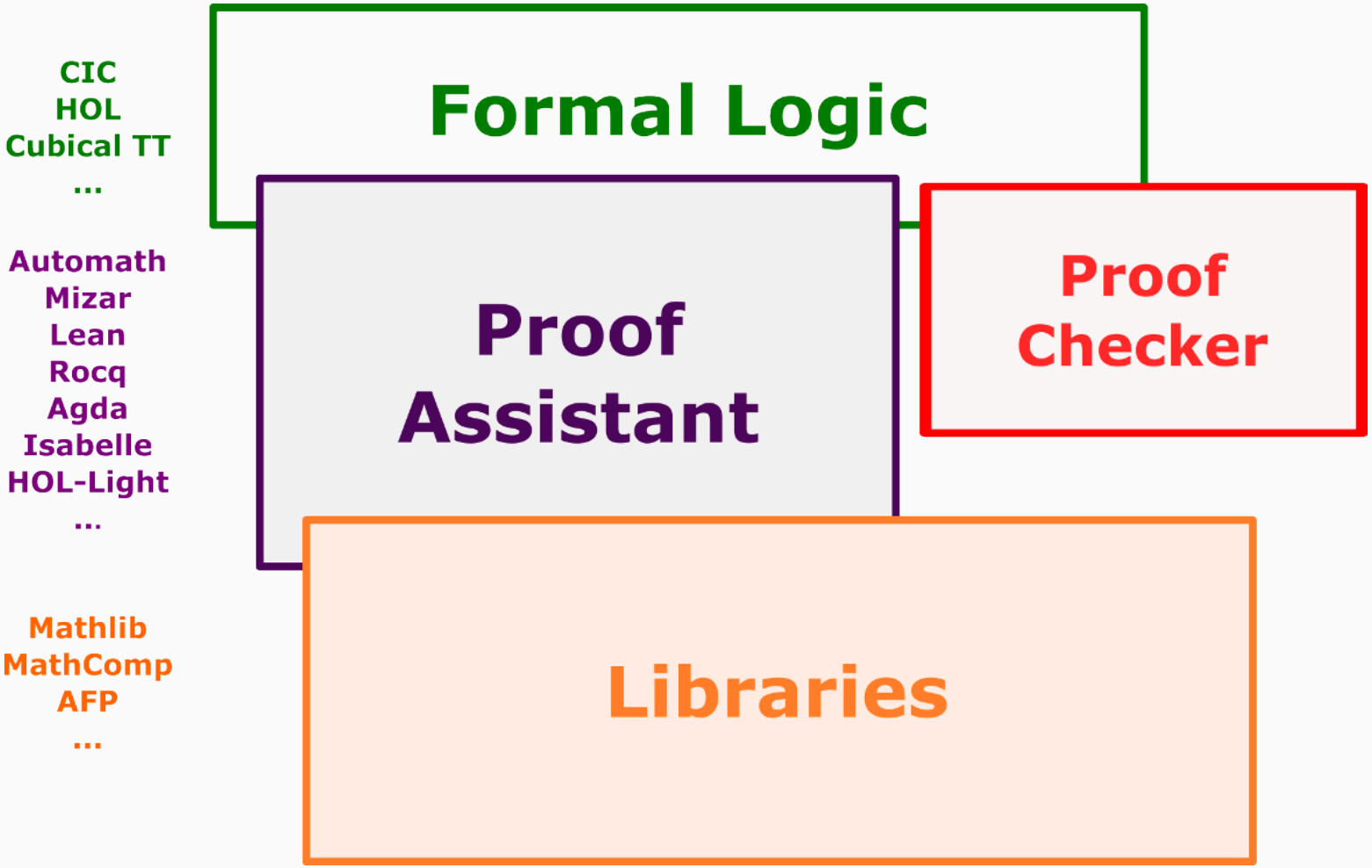
Skeleton of an interactive theorem prover



Acts of faith:

- Correctness of the proof checker
- Soundness of formal definitions

Skeleton of an interactive theorem prover



Acts of faith:

- Correctness of the proof checker
- Soundness of formal definitions X

Proof assistant

Two crucial natures of automation:

- In proofs:

```
have sq (a b : nat) : (a + b) ^2 = a^2 + 2 * a * b + b ^ 2
    by ring.
```

Decision procedures

- In statements:

```
Definition determinant R n (A : 'M[R]_n) : R :=
\sum_(s : 'S_n) (-1) ^+ s * \prod_i A i (s i).
```

Inference of mathematical structures

Foundations have a significant impact on the implementation of these features.

Machine-checked computer proofs

Example: computing rigorous quadratures

Rigorous numerical integration

9

2

I need to evaluate some (one-variable) integrals that neither SAGE nor Mathematica can do symbolically. As far as I can tell, I have two options:

(a) Use GSL (via SAGE), Maxima or Mathematica to do numerical integration. This is really a non-option, since, if I understand correctly, the "error bound" they give is not really a guarantee.

(b) Cobble together my own programs using the trapezoidal rule, Simpson's rule, etc., and get rigorous error bounds using bounds I have for the second (or fourth, or what have you) derivative of the function I am integrating. This is what I have been doing.

Is there a third option? Is there standard software that does (b) for me?

na.numerical-analysis

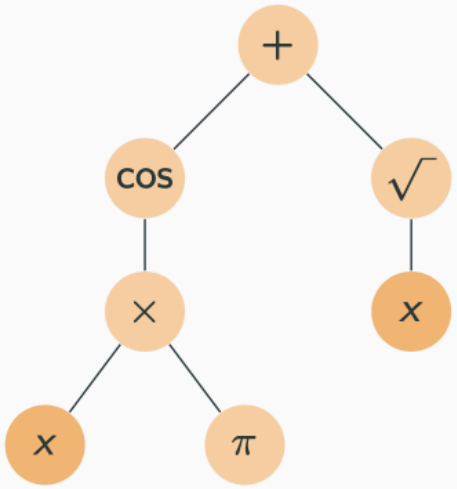
share cite improve this question

asked Mar 5 '13 at 23:03

H A Helfgott

3,620 • 21 • 69

Abstract syntax trees



$[e]_{\mathbb{R}_\perp} : \mathbb{R}_\perp \to \mathbb{R}_\perp$ $\qquad x \mapsto \cos(x \times \pi) + \sqrt{x}$

$[e]_{\mathbb{I}_\perp} : \mathbb{I}_\perp \to \mathbb{I}_\perp$ $\qquad \boldsymbol{x \mapsto \cos(x \times \pi) + \sqrt{x}}$

Correctness theorem of interval extensions:

$$\forall e \in \mathcal{E}, \quad \forall \boldsymbol{i} \in \mathbb{I}_\perp, \quad \forall x \in \boldsymbol{i}, \quad [e]_{\mathbb{R}_\perp}(x) \in [e]_{\mathbb{I}_\perp}(\boldsymbol{i})$$

Formally Verified Approximations

Initial problem:

$$\int_a^b f(x)dx \in [m, M] \quad ?$$

Formally Verified Approximations

Entry in the catalog:

$$\int_{[e_a]_{\mathbb{R}}}^{[e_b]_{\mathbb{R}}} [e_f]_{\mathbb{R}}\, dx \in [m, M] \quad ?$$

Formally Verified Approximations

Verified computation:

$$\int_{[e_a]_{\mathbb{R}}}^{[e_b]_{\mathbb{R}}} [e_f]_{\mathbb{R}}\, dx$$

Formally Verified Approximations

Verified computation:

$$\int_{[e_a]_{\mathbb{R}}}^{[e_b]_{\mathbb{R}}} [e_f]_{\mathbb{R}}\, dx \in \int_{[e_a]_{\mathbb{I}}}^{[e_b]_{\mathbb{I}}} [e_f]_{\mathbb{I}}\, dx$$

Formally Verified Approximations

Verified computation:

$$\int_{[e_a]_{\mathbb{R}}}^{[e_b]_{\mathbb{R}}} [e_f]_{\mathbb{R}}\, dx \in \int_{[e_a]_{\mathbb{I}}}^{[e_b]_{\mathbb{I}}} [e_f]_{\mathbb{I}}\, dx \subseteq [m, M]$$

[Formally Verified Approximations of Definite Integrals, A. Mahboubi, G. Melquiond, Th. Sibut-Pinote, JAR 2018]

Formally Verified Approximations

Verified computation, using rigorous polynomial approximations:

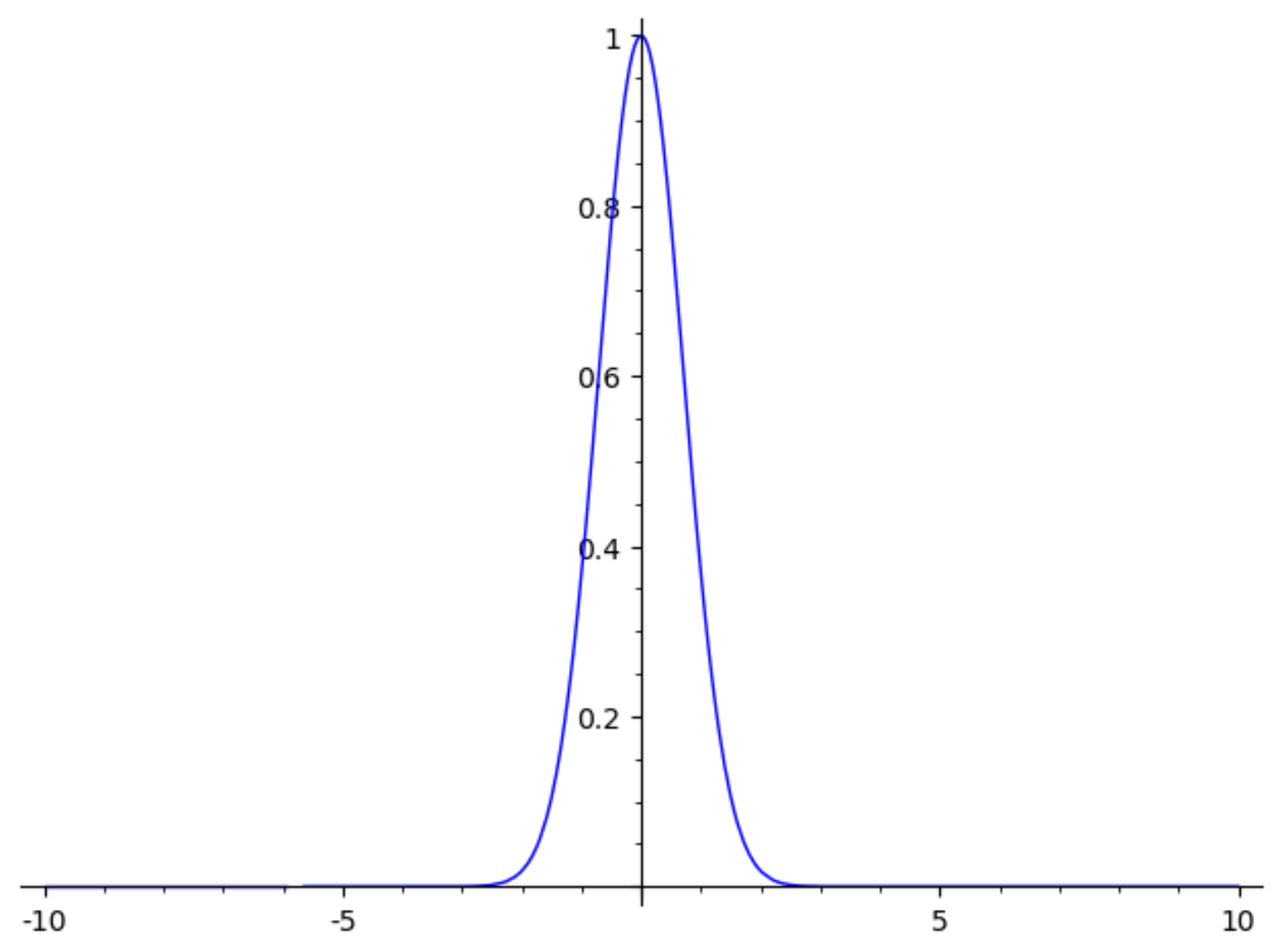
$$\int_{[e_a]_{\mathbb{R}}}^{[e_b]_{\mathbb{R}}} [e_f]_{\mathbb{R}}\, dx \in \int_{[e_a]_{\mathbb{TM}}}^{[e_b]_{\mathbb{TM}}} [e_f]_{\mathbb{TM}}\, dx \subseteq [m, M]$$

[Formally Verified Approximations of Definite Integrals, A. Mahboubi, G. Melquiond, Th. Sibut-Pinote, JAR 2018]

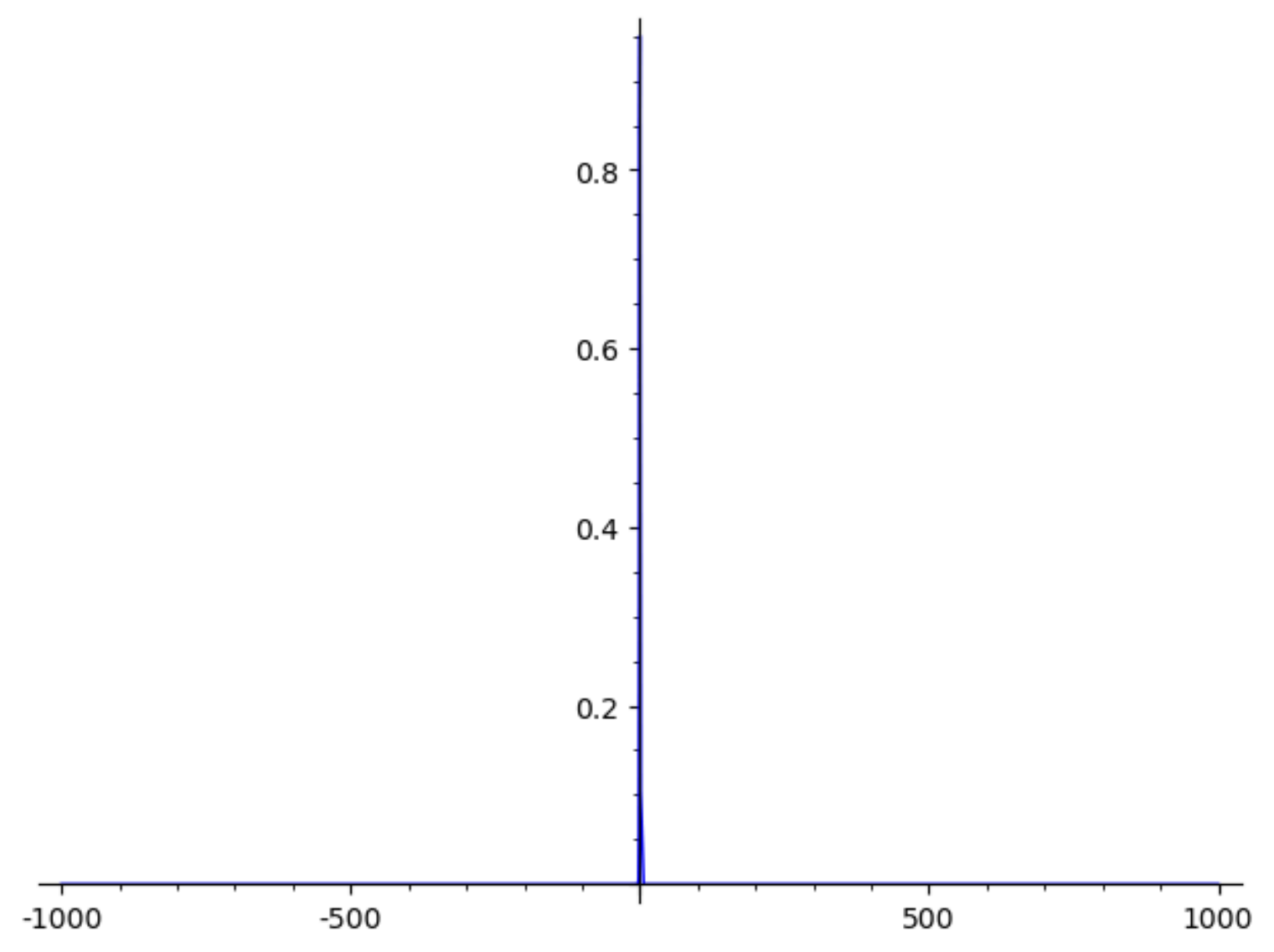
Basic Plots In Two Dimensions

- Sample a given domain;
- Evaluate the function to be plotted at the sample points;
- Color corresponding pixels.

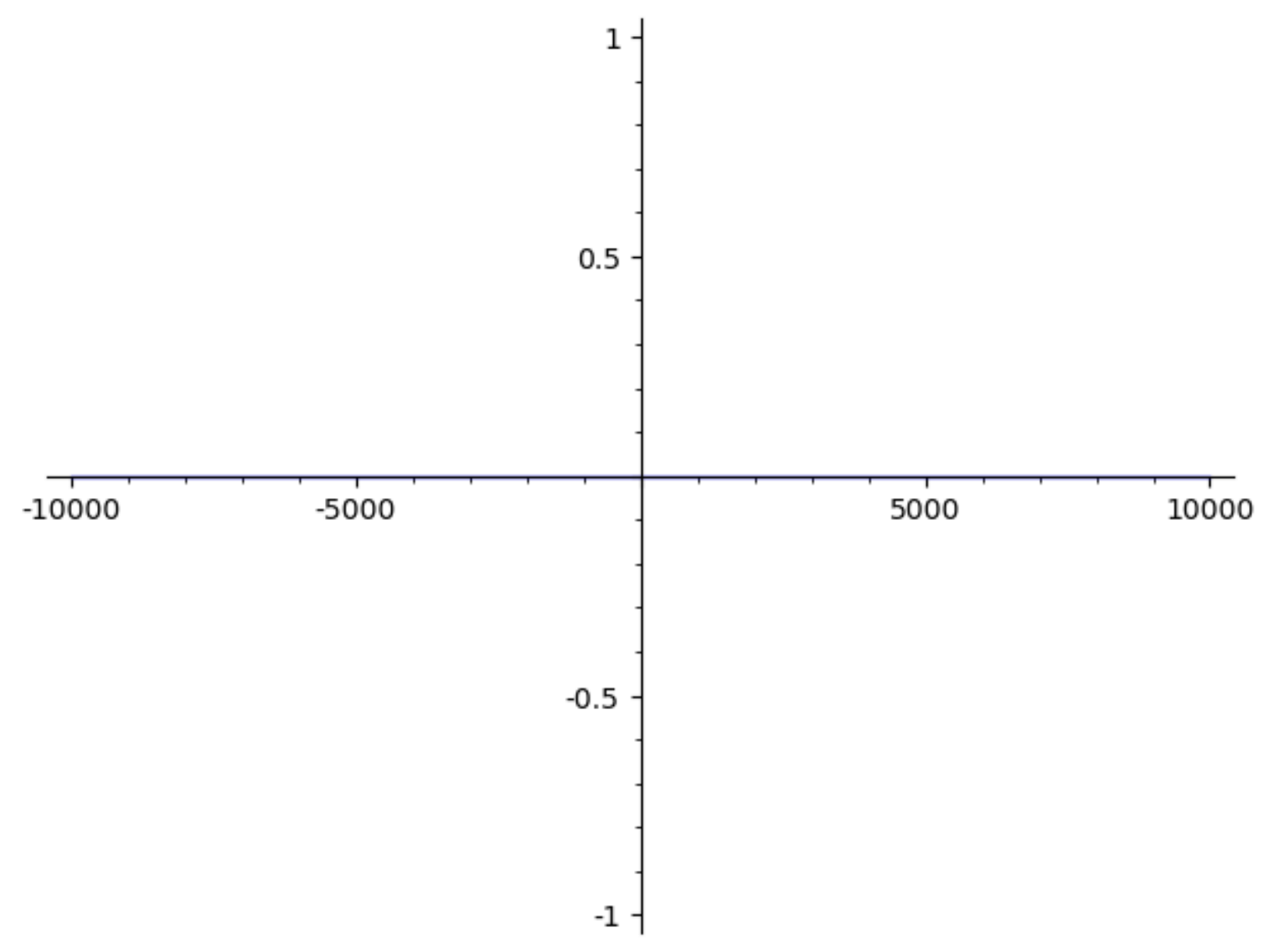
Plotting $exp(-x^2)$ with sagemath



Plotting $exp(-x^2)$ with sagemath

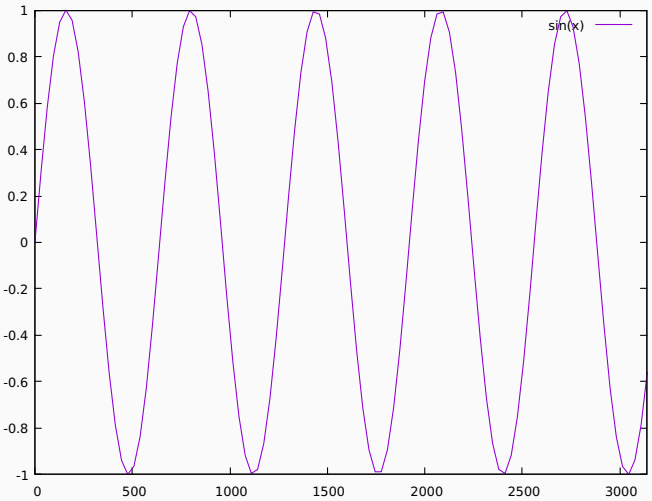


Plotting $exp(-x^2)$ with sagemath



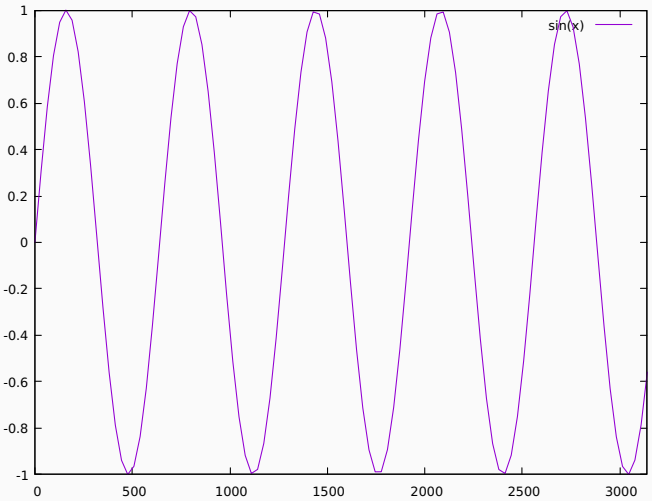
Plotting $sin(x)$ for $x \in [0, 3141]$

Plotting $sin(x)$ for $x \in [0, 3141]$

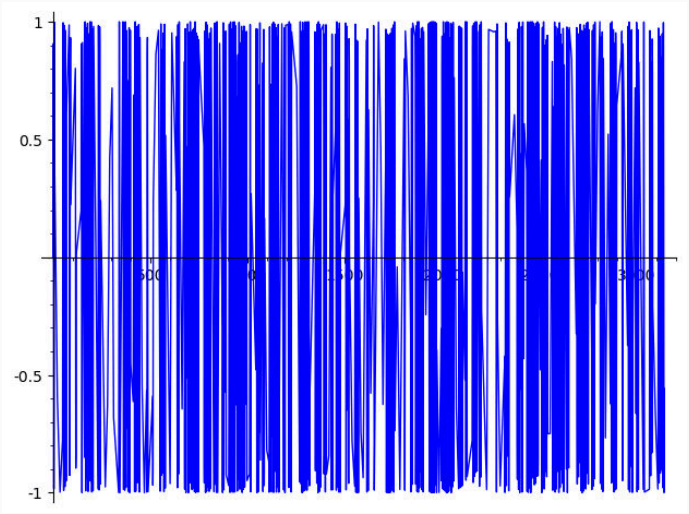


Gnuplot

Plotting $sin(x)$ for $x \in [0, 3141]$



Gnuplot



Sagemath

Faithful plotting is hard

Issues:

- Sampling
- Accuracy
- Bugs

Faithful plotting is hard

Issues:

- Sampling
- Accuracy
- Bugs

Desired properties:

- Correctness: blank pixels are not traversed by the function graph
- Completeness: filled pixels are traversed by the function graph

Faithful plotting is hard

Issues:

- Sampling
- Accuracy
- Bugs

Desired properties:

- Correctness: blank pixels are not traversed by the function graph
- Completeness: filled pixels are traversed by the function graph

$\Rightarrow$ Formally verified plots: guarantee correctness and strive for completeness

Generating formally verified plots

To obtain a verified plot for $f(x)$ for $x \in X$:

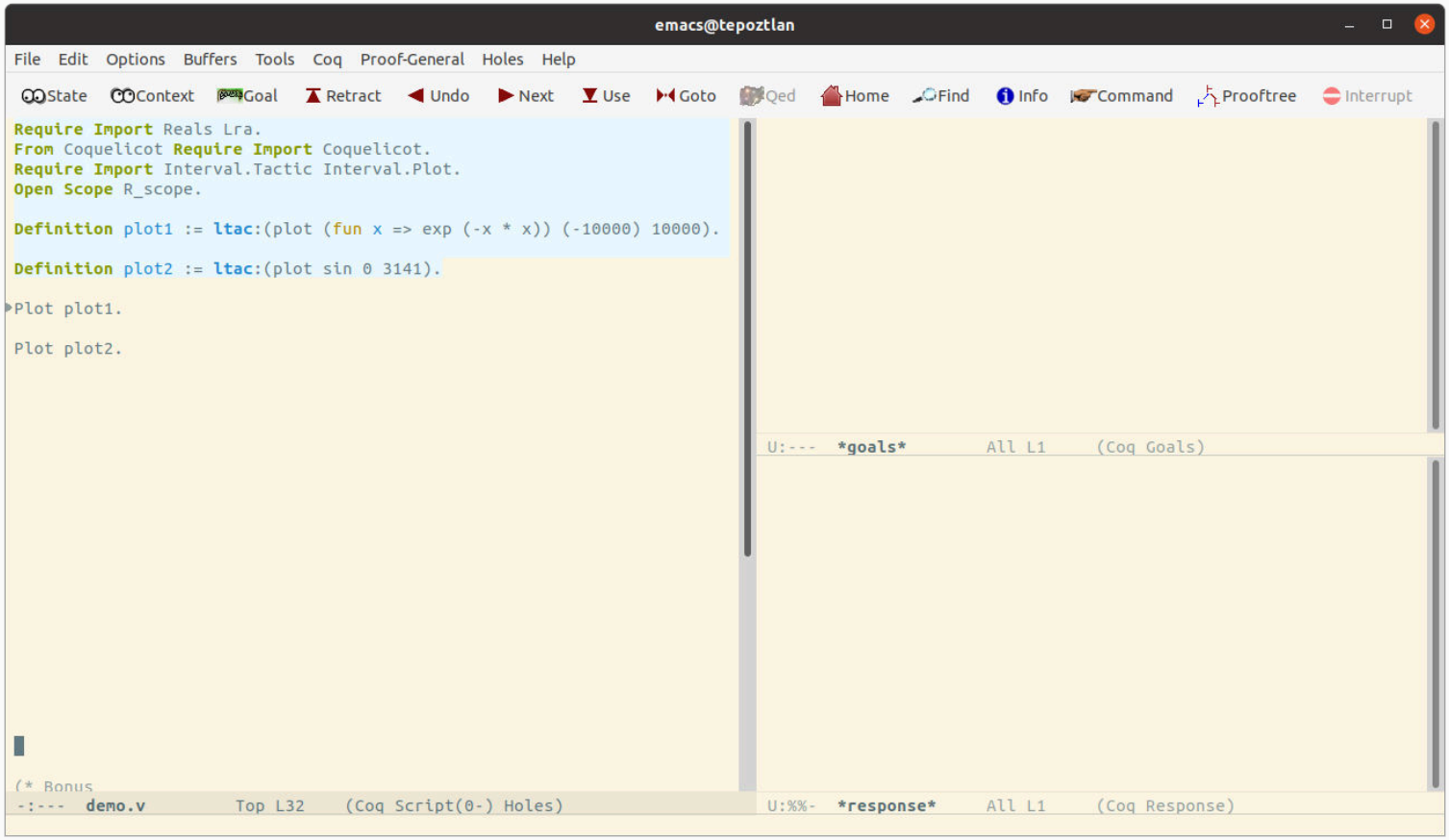
- Partition X in $(X_i)_{i=1\ldots n}$
- Produce a list $(\ell_i)_{i=1\ldots n}$ of intervals
- Ensure (with a formal proof) that for every $i = 1 \ldots n$:

$$\forall x \in X_i, f(x) \in \ell_i$$

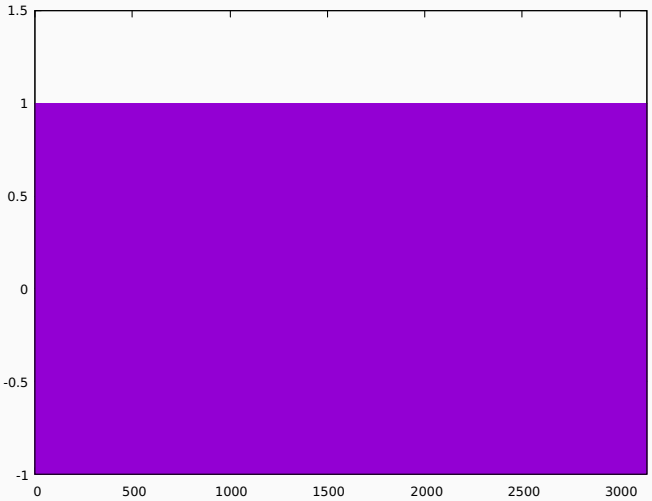
- Fill the corresponding pixels.

Rigorous polynomial approximation make computations efficient enough.

Demo

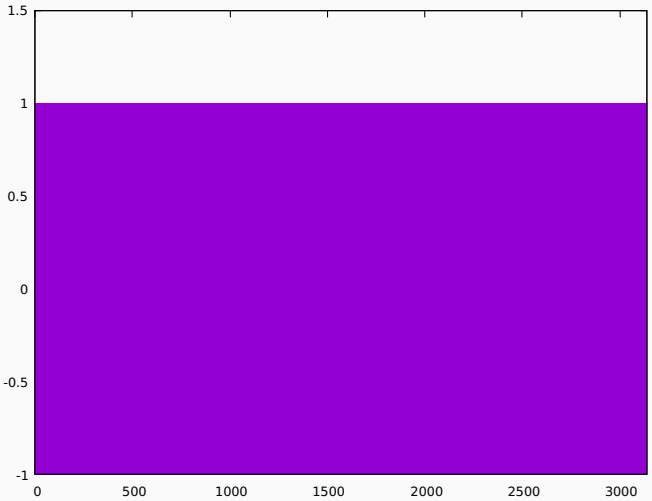

emacs@tepoztlan
File Edit Options Buffers Tools Coq Proof-General Holes Help
State Context Goal Retract Undo Next Use Goto Qed Home Find Info Command Prooftree Interrupt
Require Import Reals Lra.
From Coquelicot Require Import Coquelicot.
Require Import Interval.Tactic Interval.Plot.
Open Scope R_scope.
Definition plot1 := ltac:(plot (fun x => exp (-x * x)) (-10000) 10000).
Definition plot2 := ltac:(plot sin 0 3141).
Plot plot1.
Plot plot2.
(* Bonus
-:--- demo.v Top L32 (Coq Script(0-) Holes)
U:--- *goals* All L1 (Coq Goals)
U:%%- *response* All L1 (Coq Response)

Demo (backup)

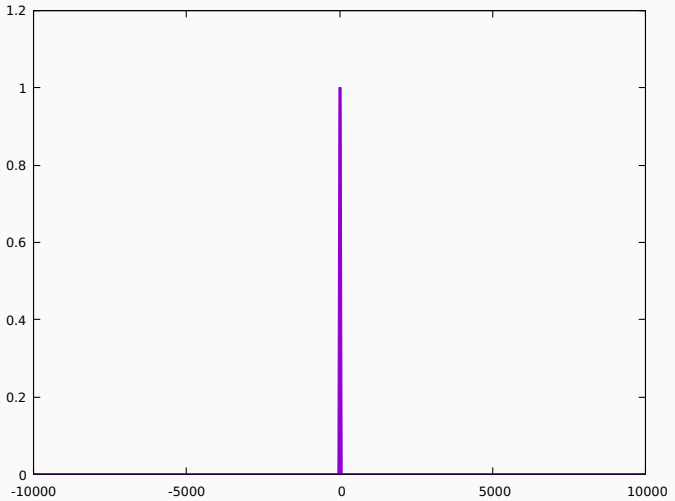


Verified plot of $sin(x)$ for $x \in [0, 3141]$

Demo (backup)



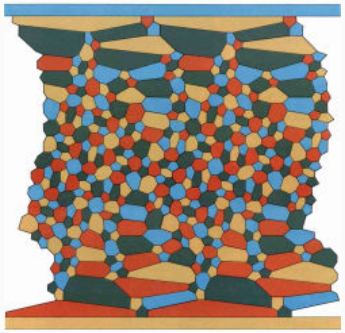
Verified plot of $sin(x)$ for $x \in [0, 3141]$

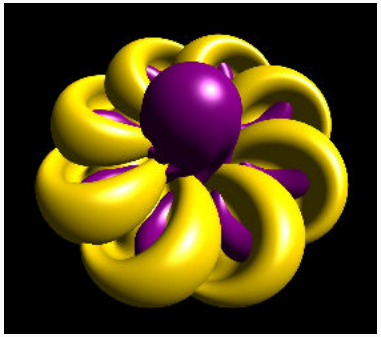


Verified plot of $exp(-x^2)$ for $x \in [-10000, 10000]$

Perspectives

Libraries of machine-checked mathematics





[A computer-checked proof of the four color theorem, G. Gonthier (2003)]

[A Machine-Checked Proof of the Odd Order Theorem, G. Gonthier et al. (2013)]

[A formal proof of the Kepler conjecture Hales et al., Cambridge University Press, (2017)]

[Formalising the h-Principle and Sphere Eversion, F. van Doorn, P. Massot, O. Nash, Procs. of CPP (2023)]

Mathematics in the making

Quoting P. Schölze about the Liquid Tensor experiment:

"(...) This makes the rest of the proof of the Liquid Tensor Experiment considerably more explicit and more elementary, removing any use of stable homotopy theory. I expect that Commelin's complex may become a standard tool in the coming years."

"(...) this made me realize that actually the key thing happening is a reduction from a non-convex problem over the reals to a convex problem over the integers."

The future of mathematics?

The AI and formalization program that is being sold to us as the future of fundamental physics seems to me to come with a big danger. It may turn this field into nothing but an endless investigation of the blind alleys that are what we know about now and have been stuck at for decades. This particular "formalization of QFT" is an expedition headed down such a blind alley.

Peter Woit, on his blog Not Even Wrong

Collaborative mathematics

- 2008 : First commit on the Stacks project (github)
- 2009 : First Polymath (wiki) project
- 2019 : Lean together (Amsterdam)

The Busy Beaver challenge

Tristan Stérin

Computer Scientist

Ventures Research Blog CV

BB(5) paper accepted to STOC 2026

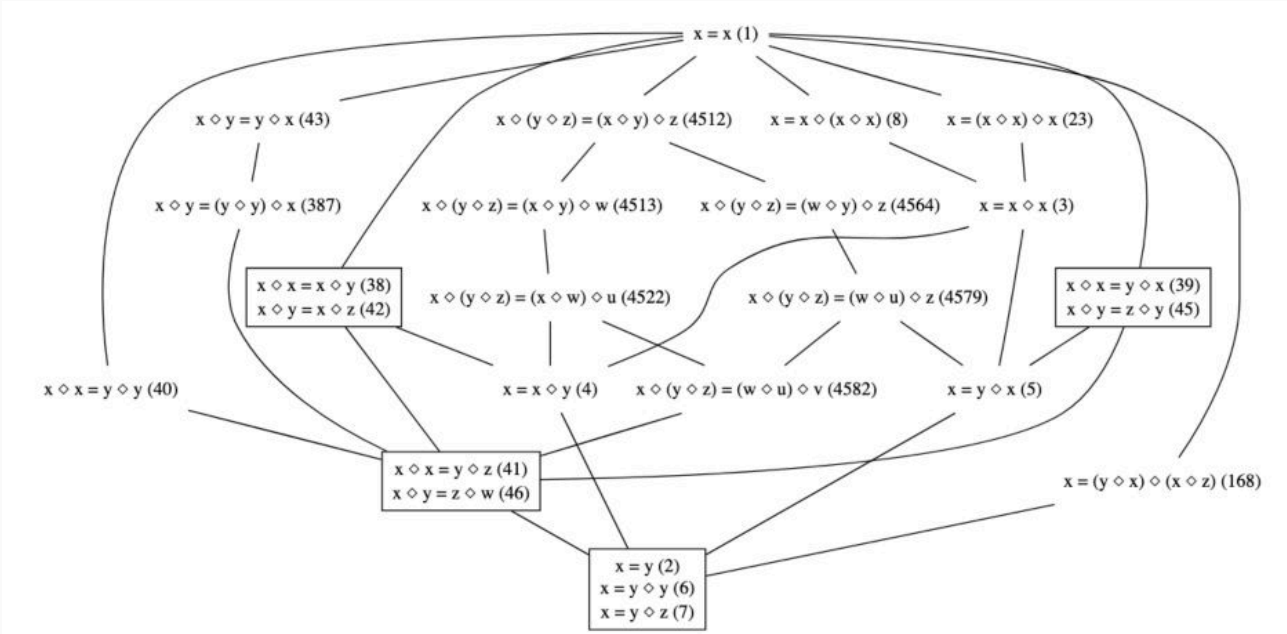
February 2nd 2026

I'm very happy to share that our paper, Determination of the fifth Busy Beaver value, has been accepted to STOC 2026: I will be presenting the paper in Salt Lake City this June!

Over three years in the making: from launching the Busy Beaver Challenge in March 2022, to announcing the Coq proof that BB(5) = 47,176,870 in July 2024, to writing the paper over the past year.

The paper is a "human-readable" version of the Coq proof, describing how we enumerated more than 180M Turing machines and, for each, decided whether it halted or not. If you want to learn more about the project, check out Ben Brubaker's Quanta Magazine article or the associated YouTube video.

The Equational Theories Project



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".

Organized by Pietro Monticone, Shreyas Srinivas, and Terence Tao, with 24 other (human) collaborators.

Large scale collaborations

- ITP provides a bedrock for collaboration to scale.
- Professional and amateur mathematicians can collaborate.
- Computer-produced proofs and human-created proofs have equal status.
- New mathematics have emerged from the projects.

Computer-produced proofs, again

June 8, 2024 | 12 min read Add Us On Google

AI Will Become Mathematicians' 'Co-Pilot'

Fields Medalist Terence Tao explains how proof checkers and AI programs are dramatically changing mathematics

BY CHRISTOPH DRÖSSER EDITED BY GARY STIX

AIs know better how to cheat

In search of falsehood

March 5th 2026

Following up on the previous post where we used Opus 4.6 to prove things in Rocq and Lean4, we turned the tables: what if we asked it to *break* things instead? With the help of expert tips, we pointed Opus 4.6 at the kernels of Rocq and Lean4 (and some of Lean's non-official kernels) and asked it to find soundness bugs – bugs that allow deriving a proof of `False`.

A proof of `False` is arguably the worst possible bug in a proof assistant: from `False` you can prove anything, and the guarantees provided by formal verification vanish entirely.

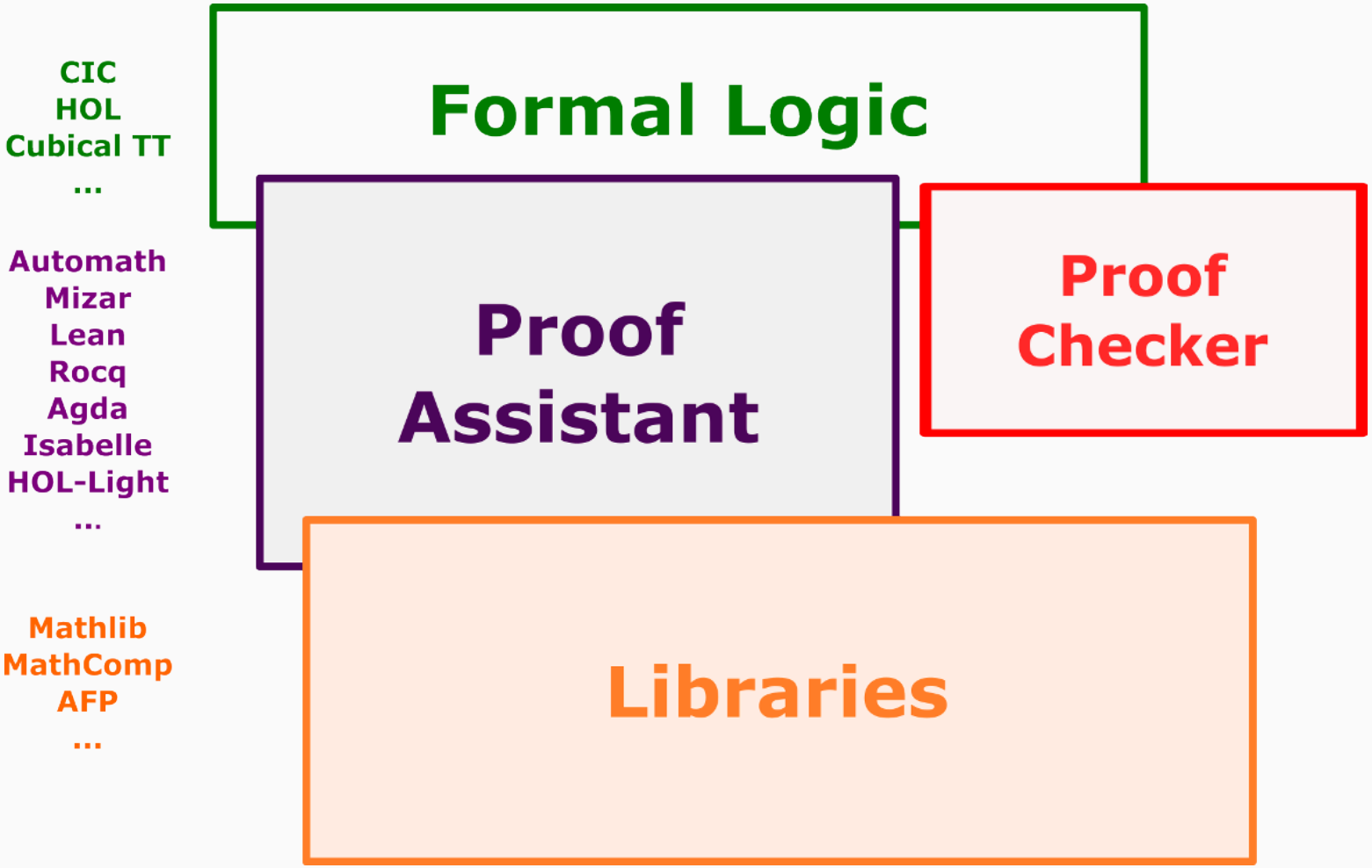
The results:

	Proofs of `False`	Other bugs
Rocq (official)	1, 2, 3, 4, 5, 6, 7	8, 9, 10
Lean4 (official)	0	1, 2, 3, 4
nanoda (Lean, non-official)	1, 2	0
lean4lean (Lean, non-official)	1	0

Important: (a) we arguably got much more precise expert tips along the way on potential bugs in Rocq than Lean, (b) bug 1 in the official Lean4 kernel could, in theory, be exploited into a proof of False.

Tristan Stérin's blog

Skeleton of an interactive theorem prover



Acts of faith:

- Correctness of the proof checker X
- Soundness of formal definitions X

Thank you

The product of mathematics is clarity and understanding. Not theorems, by themselves. (...)

I think of mathematics as having a large component of psychology, because of its strong dependence on human minds. Dehumanized mathematics would be more like computer code, which is very different. Mathematical ideas, even simple ideas, are often hard to transplant from mind to mind. (...)

In short, mathematics only exists in a living community of mathematicians that spreads understanding and breaths life into ideas both old and new.

William Thurston Oct. 30, 2010, on mathoverflow.net