How can artificial intelligence help mathematicians?

Amaury Hayat
CERMICS - Ecole des Ponts - Institut Polytechnique de Paris

AI for Mathematics and Theoretical Computer Science

–

April, 9th

–

Simons Institute and SLMath

Introduction

Social networks

Emails and apps

Virtual assistants

Platforms

Introduction

ChatGPT

MISTRAL AI_

Claude

Gemini

Introduction

From computer scientists awareness...

- Convolutional Neural Networks (LeNet-5, 1998,..., ResNet, 2015)
- Generative adversarial networks (2014)
- Transformer (2017)
- AlphaZero (2018)

... to global awareness

- ChatGPT (followed by Llama, Mistral, etc.)
- Diffusion model (Dall·E, Midjourney, Stable Diffusion)
- 2024 Nobel Prizes in Physics **and** in Chemistry

Introduction

An idea that started taking roots:

How can AI help mathematicians ?

Basic usage:

- asking a math question to an LLM / asking to prove a small lemma

Introduction

An idea that started taking roots:

How can AI help mathematicians ?

Basic usage:

- asking a math question to an LLM / asking to prove a small lemma
  - terrible answers (2022-2023)

Introduction

An idea that started taking roots:

How can AI help mathematicians ?

Basic usage:

- asking a math question to an LLM / asking to prove a small lemma
  - terrible answers (2022-2023)
  - but models are evolving quickly

Introduction

An idea that started taking roots:

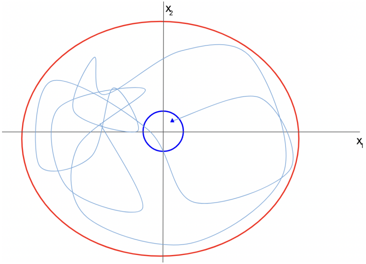
How can AI help mathematicians ?

Basic usages:

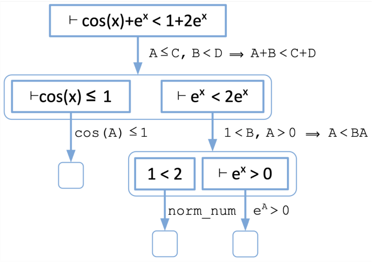
- Finding citations (information retrieval)
- Learning about a new field; explaining a paper (in-context or retrieval)
- Reviewing process (!)

AI for mathematics $\neq$ generic LLM

Outline of the talk

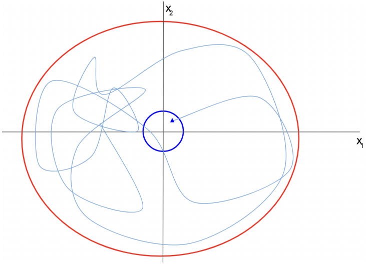


1. AI tools for math discovery

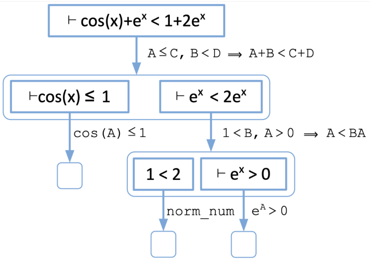


2. When will AI prove theorems?

Outline of the talk



1. AI tools for math discovery



2. When will AI prove theorems?

Solving maths problems with a computer

Solving maths problems with a computer

Conjecture (Euler, 1769)

If there exist integers a_1, a_2,...,a_k, b, and n such that

$$a_1^n + a_2^n + \ldots + a_k^n = b^n,$$

then $k \geq n$.

A problem open for almost 200 years

Solving maths problems with a computer

Lander and Parkin (1966)

Solving maths problems with a computer

Lander and Parkin (1966)

$$27^5 + 84^5 + 110^5 + 133^5 = 144^5$$

Solving maths problems with a computer

COUNTEREXAMPLE TO EULER'S CONJECTURE ON SUMS OF LIKE POWERS

BY L. J. LANDER AND T. R. PARKIN

Communicated by J. D. Swift, June 27, 1966

A direct search on the CDC 6600 yielded

$$27^5 + 84^5 + 110^5 + 133^5 = 144^5$$

as the smallest instance in which four fifth powers sum to a fifth power. This is a counterexample to a conjecture by Euler [1] that at least n nth powers are required to sum to an nth power, $n > 2$.

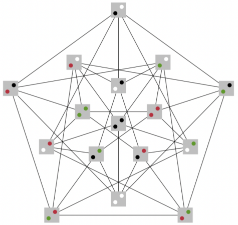
REFERENCE

1. L. E. Dickson, *History of the theory of numbers*, Vol. 2, Chelsea, New York, 1952, p. 648.

Solving maths problems with a computer

Proof of Keller's Conjecture (Brakensiek, Heule, Mackey, Narváez, 2019)

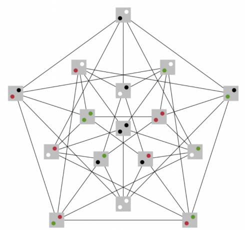
- A proof with many "simple cases" to check
- Many = far too many for a human



Solving maths problems with a computer

Proof of Keller's Conjecture (Brakensiek, Heule, Mackey, Narváez, 2019)

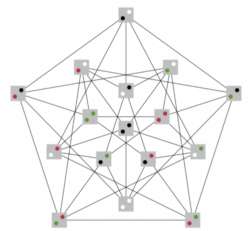
- A proof with many "simple cases" to check
- Many = far too many for a human
- Size of the proof:



Solving maths problems with a computer

Proof of Keller's Conjecture (Brakensiek, Heule, Mackey, Narváez, 2019)

- A proof with many "simple cases" to check
- Many = far too many for a human
- Size of the proof: 200Gb $\sim$ 10 Wikipedias

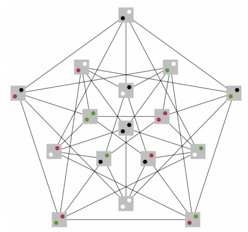


Solving maths problems with a computer

Proof of Keller's Conjecture (Brakensiek, Heule, Mackey, Narváez, 2019)

- A proof with many "simple cases" to check
- Many = far too many for a human
- Size of the proof:

Several impressive results using SAT solvers (see the talk of Giles Gardam and Bernardo Subercaseaux)



Conclusion: Computers have been used to prove theorems for a long time.

Conclusion: Computers have been used to prove theorems for a long time.

Can AI be useful to solve more complicated problems?

Problems where the difficulty is not just a high number of case-checking?

AI in Mathematics Today

Three examples

- Stability of dynamical systems
- Control theory
- Combinatorics

AI in Mathematics Today

Three examples

- Stability of dynamical systems
- Control theory
- Combinatorics

AI in Mathematics Today

Three examples

- Stability of dynamical systems
- Control theory
- Combinatorics

Discovering Lyapunov function
(Alfarano, Charton, H.)

$$\begin{cases} \dot{x}_0 = -7x_0^5 - 4x_0^3x_1^2 - 5x_0^3 \\ \dot{x}_1 = 7x_0^4 - 3x_1 - 2x_2 \\ \dot{x}_2 = -8x_0^2 - 9x_2 \end{cases}$$

$$V(x) = 2x_0^4 + 2x_0^2x_1^2 + 3x_0^2 + 2x_1^2 + x_2^2$$

(see tutorial by Sean Welleck)

Stability of Dynamical Systems

A system of differential equations

$$\dot{x}(t) = f(x(t)),$$

A system of differential equations

$$\dot{x}(t) = f(x(t)),$$

where $x(t) \in \mathbb{R}^n$, $f \in C^1(\mathbb{R}^n)$ and $f(0) = 0$.

A system of differential equations

$$\dot{x}(t) = f(x(t)),$$

$$\text{where } x(t) \in \mathbb{R}^n,\ f \in C^1(\mathbb{R}^n) \text{ and } f(0) = 0.$$

Question (System Stability)

Is it true that for every $\varepsilon > 0$, there exists $\delta > 0$ such that if the initial condition satisfies $\|x(0)\| \leq \delta$ then the solution $x(t)$ exists for all $t \in [0, +\infty)$ and

$$\|x(t)\| \leq \varepsilon,\ \forall\ t \in [0, +\infty).$$

A system of differential equations

$$\dot{x}(t) = f(x(t)),$$

$$\text{where } x(t) \in \mathbb{R}^n,\ f \in C^1(\mathbb{R}^n) \text{ and } f(0) = 0.$$

Question (System Stability)

Is it true that for every $\varepsilon > 0$, there exists $\delta > 0$ such that if the initial condition satisfies $\|x(0)\| \leq \delta$ then the solution $x(t)$ exists for all $t \in [0, +\infty)$ and

$$\|x(t)\| \leq \varepsilon,\ \forall\, t \in [0, +\infty).$$

- Are all solutions arbitrarily bounded if the initial condition is sufficiently small?

A system of differential equations

$$\dot{x}(t) = f(x(t)),$$

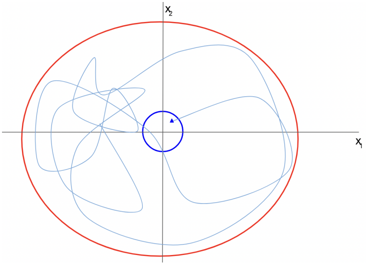
where $x(t) \in \mathbb{R}^n$, $f \in C^1(\mathbb{R}^n)$ and $f(0) = 0$.

Question (System Stability)

Is it true that for every $\varepsilon > 0$, there exists $\delta > 0$ such that if the initial condition satisfies $\|x(0)\| \leq \delta$ then the solution $x(t)$ exists for all $t \in [0, +\infty)$ and

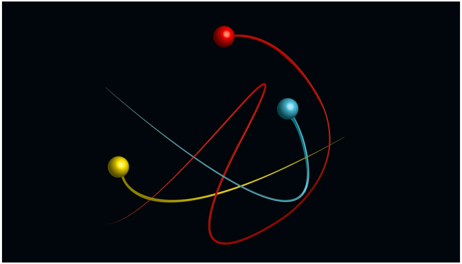
$$\|x(t)\| \leq \varepsilon, \ \forall\ t \in [0, +\infty).$$

- Are all solutions arbitrarily bounded if the initial condition is sufficiently small?

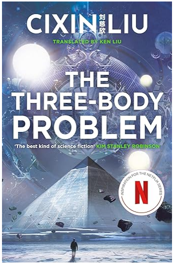


Stability of Dynamical Systems

A problem that has interested mathematicians for over a hundred years.



Stability of Dynamical Systems



Stability of Dynamical Systems

A significant advancement: Lyapunov functions

Theorem

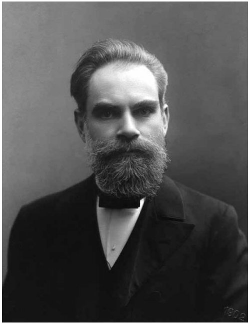
If there exists a function $V \in C^1(\mathbb{R}^n;\mathbb{R})$ *such that for all* $x \in \mathbb{R}^n$

$$V(x) > V(0), \quad \text{and} \quad \nabla V(x) \cdot f(x) \leq 0,$$

and

$$\lim_{\|x\| \to +\infty} V(x) = +\infty,$$

then the system is stable.



A. Lyapunov
(1857-1918)

Stability of Dynamical Systems

A significant advancement: Lyapunov functions

Theorem

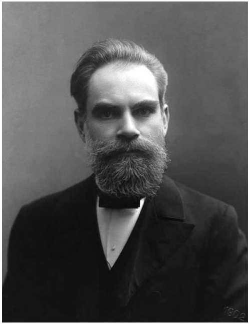
If there exists a function $V \in C^1(\mathbb{R}^n; \mathbb{R})$ *such that for all* $x \in \mathbb{R}^n$

$$V(x) > V(0), \quad \textit{and} \quad \nabla V(x) \cdot f(x) \leq 0,$$

and

$$\lim_{\|x\| \to +\infty} V(x) = +\infty,$$

then the system is stable.



A. Lyapunov
(1857-1918)

Nothing tells us how to find such a function V...

Stability of Dynamical Systems

Nothing tells us how to find such a function V...

And it's not a simple problem.

Stability of Dynamical Systems

Nothing tells us how to find such a function V...

And it's not a simple problem.

$$\dot{x}(t) = \begin{pmatrix} -6x_1^4(t)x_2^5(t) - 3x_1^7(t)x_3^2(t) \\ 3x_1^9(t) - 6x_1^2(t)x_2^5(t)x_3^2(t) \\ -4x_1^2(t)x_3^5(t) \end{pmatrix}$$

Stability of Dynamical Systems

Nothing tells us how to find such a function V...

And it's not a simple problem.

$$\dot{x}(t) = \begin{pmatrix} -6x_1^4(t)x_2^5(t) - 3x_1^7(t)x_3^2(t) \\ 3x_1^9(t) - 6x_1^2(t)x_2^5(t)x_3^2(t) \\ -4x_1^2(t)x_3^5(t) \end{pmatrix}$$

The system is stable, a Lyapunov function is

$$V(x) = x_1^6 + 2(x_2^6 + x_3^4)$$

Stability of Dynamical Systems

A globally asymptotically stable polynomial vector field with no polynomial Lyapunov function

Publisher: IEEE   Cite This   PDF

Amir Ali Ahmadi ; Miroslav Krstic ; Pablo A. Parrilo   All Authors

39 Cites in Papers

399 Full Text Views

Stability of Dynamical Systems

Today, more than a hundred years later, it is still an open question:

there is still no systematic way to construct a Lyapunov function.

Stability of Dynamical Systems

Today, more than a hundred years later, it is still an open question:

there is still no systematic way to construct a Lyapunov function.

$\longrightarrow$ We resort to intuition

Stability of Dynamical Systems

- Intuition, an important concept in mathematics.

Stability of Dynamical Systems

- Intuition, an important concept in mathematics.

- In a number of cases, intuition resembles a kind of pattern recognition. You've seen plenty of examples, and this gives you an idea of how to proceed in a case you've never seen.

Stability of Dynamical Systems

- Intuition, an important concept in mathematics.
- In a number of cases, intuition resembles a kind of pattern recognition. You've seen plenty of examples, and this gives you an idea of how to proceed in a case you've never seen.
- Can we train an AI to have better mathematical intuition than us?

Stability of Dynamical Systems

- Intuition, an important concept in mathematics.
- In a number of cases, intuition resembles a kind of pattern recognition. You've seen plenty of examples, and this gives you an idea of how to proceed in a case you've never seen.
- Can we train an AI to have better mathematical intuition than us?

Task: guessing Lyapunov functions

$$\dot{x}(t) = \begin{pmatrix} -6x_1^4(t)x_2^5(t) - 3x_1^7(t)x_3^2(t) \\ 3x_1^9(t) - 6x_1^2(t)x_2^5(t)x_3^2(t) \\ -4x_1^2(t)x_3^5(t) \end{pmatrix} \rightarrow \quad \text{Yes, } V(x) = x_1^6 + 2(x_2^6 + x_3^4)$$

Stability of Dynamical Systems

Train an AI to have an intuition of Lyapunov functions (Alfarano, Charton, A.H., 2024).

Neural network architecture: Transformer ($\sim$1000 smaller than GPT-3)

Procedure:

1. Generate a set of systems and associated Lyapunov functions.
2. Encode the examples
3. Train the language model (supervised learning)

Stability of Dynamical Systems

Train an AI to have an intuition of Lyapunov functions
(Alfarano, Charton, A.H., 2024).

Neural network architecture: Transformer ($\sim$1000 smaller than GPT-3)

Procedure:

1. Generate a set of systems and associated Lyapunov functions.
2. Encode the examples
3. Train the language model (supervised learning)

Stability of Dynamical Systems

How to generate examples of systems and solutions, for an open problem?

Stability of Dynamical Systems

How to generate examples of systems and solutions, for an open problem?

$\rightarrow$ Use a backward approach: instead of finding a solution from the problem, we find the problem from the solution.

Stability of Dynamical Systems

How to generate examples of systems and solutions, for an open problem?

$\rightarrow$ Use a backward approach: instead of finding a solution from the problem, we find the problem from the solution.

- Find a mathematical way to get:

$$\underbrace{V(x)}_{\text{positive, random}} \quad \rightarrow \text{all systems with Lyapunov function } V$$

  Then sample at random.

- In spirit: finding a Lyapunov function is a hard problem, checking that a function is a Lyapunov function is easier.

Stability of Dynamical Systems

How to generate examples of systems and solutions, for an open problem?

$\rightarrow$ Use a backward approach: instead of finding a solution from the problem, we find the problem from the solution.

- Find a mathematical way to get:

$$\underbrace{V(x)}_{\text{positive, random}} \quad \rightarrow \text{all systems with Lyapunov function } V$$

  Then sample at random.

- In spirit: finding a Lyapunov function is a hard problem, checking that a function is a Lyapunov function is easier.

Limitations: even with a perfect generator, it biases the distribution (and it matters).

Stability of Dynamical Systems

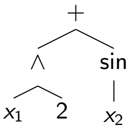
Train an AI to have an intuition of Lyapunov functions
(Alfarano, Charton, A.H., 2024).

Neural network architecture: Transformer ($\sim$1000 smaller than ChatGPT)

Procedure:

1. Generate a set of systems and associated Lyapunov functions.
2. Encode the examples

$(x_1^2+\sin(x_2)) \quad \rightarrow$



$\rightarrow \quad "+","\wedge","x_1","2","\sin","x_2"$

3. Train the language model (supervised learning)

Stability of Dynamical Systems

Results

It works! The AI learns a mathematical intuition of Lyapunov functions.

[1] Existing method for some polynomial systems.

Stability of Dynamical Systems

Results

It works! The AI learns a mathematical intuition of Lyapunov functions.

It works exceptionnally with a priming approach: adding a few high quality examples in the train set (e.g. 50 out of 1M that is 0.005%)

[1] Existing method for some polynomial systems.

Stability of Dynamical Systems

Results

It works! The AI learns a mathematical intuition of Lyapunov functions.

It works exceptionnally with a priming approach: adding a few high quality examples in the train set (e.g. 50 out of 1M that is 0.005%)

Type	n equations	SOS algorithms[1]	AI
polynomial (train distrib.)	2-5	15%	**99%**
polynomial (fwd distrib.)	2-3	47%	**84-93%**
Non-polynomial	2-3	$\sim$ 0%	**87%**

[1] Existing method for some polynomial systems.

Stability of Dynamical Systems

Results

It works! The AI learns a mathematical intuition of Lyapunov functions.

It works exceptionnally with a priming approach: adding a few high quality examples in the train set (e.g. 50 out of 1M that is 0.005%)

Type	n equations	SOS algorithms[1]	AI
polynomial (train distrib.)	2-5	15%	**99%**
polynomial (fwd distrib.)	2-3	47%	**84-93%**
Non-polynomial	2-3	$\sim$ 0%	**87%**

Master students accuracy: $\sim$ 10%

[1] Existing method for some polynomial systems.

Stability of Dynamical Systems

Results

It works! The AI learns a mathematical intuition of Lyapunov functions.

It works exceptionnally with a priming approach: adding a few high quality examples in the train set (e.g. 50 out of 1M that is 0.005%)

Type	n equations	SOS algorithms[2].	AI
polynomial (train distrib.)	2-5	15%	**99%**
polynomial (fwd distrib.)	2-3	47%	**84-93%**
Non-polynomial	2-3	$\sim$ 0%	**87%**

My accuracy (!): $\sim$ 25%

[2] Existing method for some polynomial systems

Meta says its AI could help mathematicians

Tada Images/Shutterstock

An AI system developed by Meta can find solutions to maths problems that have eluded mathematicians for over a century, researchers at the firm claim.

The problems involve mathematical tools called Lyapunov functions, named after mathematician Aleksandr Lyapunov, which analyse whether a system will remain stable over time, meaning its behaviour can be predicted. One famous example of such a system is the motion of three celestial bodies as a result of their mutual gravitational interactions – describing the behaviour of this "three-body problem" is extremely challenging.

Read more Incredible maths proof...

AIがリアプノフ関数の発見に革新！安定性理論に新たなブレイクスルー（2024-10）【論文解説シリーズ】

Summary of the approach

Paradigm: Training a Transformer to have a mathematical intuition on a problem

Key points:

- Generating data in a backward fashion (solution $\rightarrow$ problem)
- Test out-of-distribution on "real" instances of the problem.

Used in many frameworks: explicit solutions to ODE; local controllability; eigenvalues of random matrices; GCD; equilibrium of bio-networks, etc.

Try it yourself:

`https://github.com/ahayat16/Lyapunov/`

and customize on your favorite math problem

AI in Mathematics today

Three examples

- Stability of dynamical systems
- Control theory
- Combinatorics

Control of a Differential System

Evolution of the mosquito population

$$\begin{cases} \dot{E} = \beta_E F \left(1 - \dfrac{E}{K}\right) - (\nu_E + \delta_E) E, \\ \dot{M} = (1-\nu)\nu_E E - \delta_M M, \\ \dot{F} = \nu \nu_E E \dfrac{M}{M + M_s} - \delta_F F, \\ \dot{M}_s = u - \delta_s M_s, \end{cases}$$

$E(t)$ represents mosquito eggs, $F(t)$ fertilized females, $M(t)$ males, $M_s(t)$ sterile males. u is the flow of sterile mosquitoes that we release. This is what is called *control*.

Control of a Differential System

Evolution of the mosquito population

$$\begin{cases} \dot{E} = \beta_E F \left(1 - \frac{E}{K}\right) - (\nu_E + \delta_E) E, \\ \dot{M} = (1-\nu)\nu_E E - \delta_M M, \\ \dot{F} = \nu \nu_E E \frac{M}{M + M_s} - \delta_F F, \\ \dot{M}_s = u - \delta_s M_s, \end{cases}$$

$E(t)$ represents mosquito eggs, $F(t)$ fertilized females, $M(t)$ males, $M_s(t)$ sterile males. u is the flow of sterile mosquitoes that we release. This is what is called *control*.

$$u = f(M + M_s, F + F_s)$$

with $F_s = FM/M_s$.

Control of a Differential System

$E(t)$ represents mosquito eggs, $F(t)$ fertilized females, $M(t)$ males, $M_s(t)$ sterile males. u is the flow of sterile mosquitoes that we release. This is what is called *control*.

$$u = f(M + M_s, F + F_s)$$

Question

Is it possible to find f such that the system is stable and

$$\lim_{t \to +\infty} \|E(t), M(t), F(t)\| = 0 \text{ and } \lim_{t \to +\infty} \|M_s(t)\| = \varepsilon,$$

with ε as small as desired?

Control of a Differential System

$E(t)$ represents mosquito eggs, $F(t)$ fertilized females, $M(t)$ males, $M_s(t)$ sterile males. u is the flow of sterile mosquitoes that we release. This is what is called *control*.

$$u = f(M + M_s, F + F_s)$$

Question

Is it possible to find f such that the system is stable and

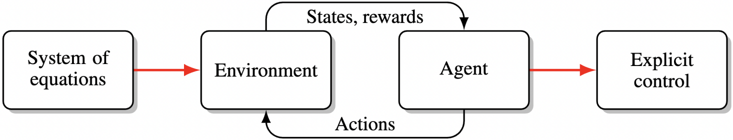
$$\lim_{t\to+\infty} \|E(t), M(t), F(t)\| = 0 \text{ and } \lim_{t\to+\infty} \|M_s(t)\| = \varepsilon,$$

with ε as small as desired?

An open question

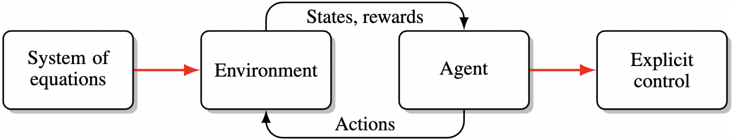
Control of a Differential System

Principle of the approach (Agbo Bidi, Coron, A.H., Lichtlé, 2023)



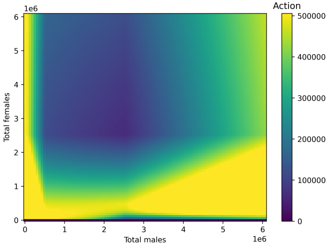
Control of a Differential System

Principle of the approach (Agbo Bidi, Coron, A.H., Lichtlé, 2023)



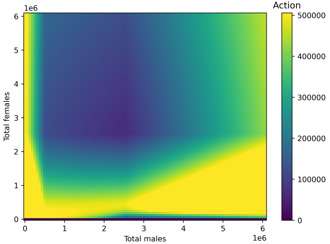
1. **Transform the equations** using a well-chosen numerical scheme.
2. **Train a Reinforcement Learning (RL) model**. The AI trains by trial and error and tries to maximize a well-chosen objective.
3. **Deduce the mathematical control**, from the numerical control.
4. **Verify** that it is a solution to the problem.

Control of a Differential System

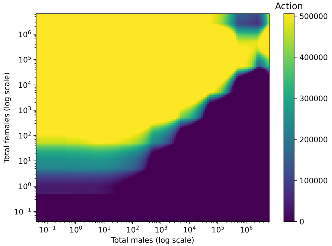


$$u = f(M + M_s, F + F_s)$$

Control of a Differential System



$\xrightarrow{\text{N.Lichtlé}}$



$$u = f(M + M_s, F + F_s)$$

Control of a Differential System

$$u_{\text{reg}}(M+M_s, F+F_s) = \begin{cases} u_{\text{reg}}^{\text{left}}(M+M_s, F+F_s) & \text{if } M+M_s < M^*, \\ u_{\text{reg}}^{\text{right}}(M+M_s, F+F_s) & \text{otherwise,} \end{cases}$$

Control of a Differential System

$$u_{\text{reg}}(M+M_s, F+F_s) = \begin{cases} u_{\text{reg}}^{\text{left}}(M+M_s, F+F_s) & \text{if } M+M_s < M^*, \\ u_{\text{reg}}^{\text{right}}(M+M_s, F+F_s) & \text{otherwise,} \end{cases}$$

$$u_{\text{reg}}^{\text{left}} = \begin{cases} \varepsilon & \text{if } I_1(F+F_s) > \alpha_2, \\ u_{\max}(\alpha_2 - I_1) & \text{if } I_1 \in (\alpha_1, \alpha_2], \\ u_{\max} & \text{otherwise, and} \end{cases}$$

$$u_{\text{reg}}^{\text{right}} = \begin{cases} \varepsilon & \text{if } I_2 > \alpha_2, \\ u_{\max}(\alpha_2 - I_2) & \text{if } I_2 \in (\alpha_1, \alpha_2], \\ u_{\max} & \text{otherwise.} \end{cases}$$

where $I_1(x) = \frac{\log M^*}{\log(F+F_s)}$ and $I_2(x,y) = \frac{\log(M+M_s)}{\log(F+F_s)}$,

Control of a Differential System

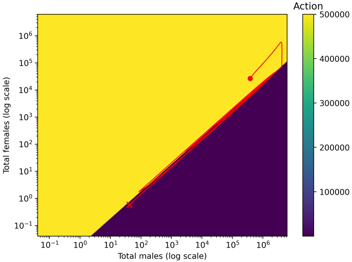
Final control

$$u(t) = \begin{cases} \varepsilon & \text{if } \frac{\log(M+M_s)}{\log(F+F_s)} > \alpha_2, \\ u_{\max} & \text{otherwise,} \end{cases}$$

Control of a Differential System

Final control

$$u(t) = \begin{cases} \varepsilon & \text{if } \frac{\log(M+M_s)}{\log(F+F_s)} > \alpha_2, \\ u_{\max} & \text{otherwise,} \end{cases}$$

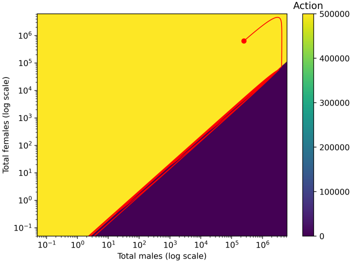


$$\varepsilon > 0$$

Control of a Differential System

Final control

$$u = \begin{cases} \varepsilon & \text{if } \frac{\log(M+M_s)}{\log(F+F_s)} > \alpha_2, \\ u_{\max} & \text{otherwise,} \end{cases}$$

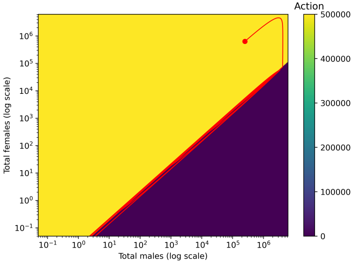


$\varepsilon = 0$

Control of a Differential System

Final control

$$u = \begin{cases} \varepsilon & \text{if } \frac{\log(M+M_s)}{\log(F+F_s)} > \alpha_2, \\ u_{\max} & \text{otherwise,} \end{cases}$$



$\varepsilon = 0$

We can see a mathematical bifurcation with our "AI-augmented intuition".

Control of a Differential System

Robustness of the solution

Can the solution still work if there are uncertainties on the parameters ?

$$\begin{cases} \dot{E} = \beta_E F \left(1 - \dfrac{E}{K}\right) - (\nu_E + \delta_E)E, \\ \dot{M} = (1-\nu)\nu_E E - \delta_M M, \\ \dot{F} = \nu \nu_E E \dfrac{M}{M + M_s} - \delta_F F, \\ \dot{M}_s = u - \delta_s M_s, \end{cases}$$

Control of a Differential System

Robustness of the solution

Can the solution still work if there are uncertainties on the parameters ?

$$\begin{cases} \dot{E} = \beta_E F \left(1 - \dfrac{E}{K}\right) - (\nu_E + \delta_E) E, \\ \dot{M} = (1-\nu)\nu_E E - \delta_M M, \\ \dot{F} = \nu \nu_E E \dfrac{M}{M + M_s} - \delta_F F, \\ \dot{M}_s = u - \delta_s M_s, \end{cases}$$

$$\begin{aligned} \hat{\beta}_E &\sim \mathcal{U}(0.65, 9.03) \\ \hat{\nu}_E &\sim \mathcal{U}(0.01, 0.35) \\ \hat{\delta}_E &\sim \mathcal{U}(0.03, 1.0) \\ \hat{\delta}_F &\sim \mathcal{U}(0.03, 0.1) \\ \hat{\delta}_M &\sim \mathcal{U}(0.09, 0.2) \\ \hat{\nu} &\sim \mathcal{U}(0.39, 1.0) \end{aligned}$$

Yes

AI tools for math discovery

Three examples

- Stability of dynamical systems
- Control theory
- Combinatorics

PatternBoost: finding interesting constructions

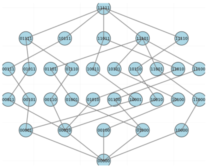
Paradigm: couple a classical algorithm and a transformer. [Charton, Ellenberg, Wagner, Williamson, 2024]

Step 1: a classical algorithm generate "good" constructions from random seeds

Step 2: train a transformer on the best such constructions

Step 3: use the transformer as seeds.

Repeat.



PatternBoost: finding interesting constructions

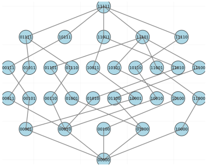
Paradigm: couple a classical algorithm and a transformer. [Charton, Ellenberg, Wagner, Williamson, 2024]

Step 1: a classical algorithm generate "good" constructions from random seeds

Step 2: train a transformer on the best such constructions

Step 3: use the transformer as seeds.

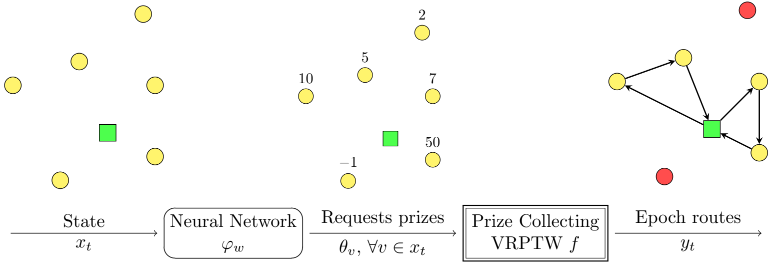
Repeat.



In particular, closed a 30-years conjecture.

Combinatorial Optimization Augmented Machine Learning

Paradigm: couple a combinatorial optimization algorithm and a neural network and train the ensemble [Baty, Jungel, Klein, Parmentier, Schiffer, 2024].



Perform better than existing mathematical algorithm and machine learning algorithm

(winner of EURO NeurIPS challenge 2022)

Many other examples

- in topology feedforward and MPNNs to guess links between different mathematical quantities, in particular hyperbolic and algebraic invariant of knots (a conjecture that was later proved) [Davies et al., 2021]
- in partial differential equations PINNs to find an exact self-similar solution to 3D Euler [Wang, Lai, Gómez-Serrano, Buckmaster, 2023]
- in group theory path-finding for large graphs with ResMLP and RL [Chervov et al. 2025]
- ... and in many others fields

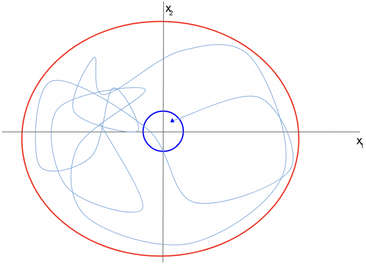
AI tools for math discovery

- AI are already useful in the practice of mathematics and can help solve difficult problems.
- AI are trained to have better intuition than humans on a specific problem.
- This augmented intuition allows us to bypass the difficulty of the problem.

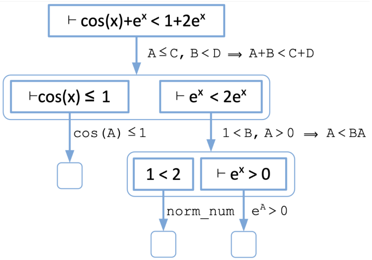
Future of Mathematical AI

Can an AI prove a mathematical result on its own?

Outline of the talk



1. AI tools for math discovery



2. When will AI prove theorems?

AI for Mathematical Proof

Can a trained AI find a proof for a mathematical statement?

- A much harder problem
- Shocking question: calls into question our very vision of mathematics
- A science-fiction future that is probably closer than we imagine

AI for Mathematical Proof

First approach: training a Transformer (GPT-f, Polu, Sutskever, 2020)

Question

Let $a > 0$ and $b > 0$, such that $ab = b - a$, show that

$$\frac{a}{b} + \frac{b}{a} - ab = 2$$

→ GPT → Proof

AI for Mathematical Proof

First approach: training a Transformer (GPT-f, Polu, Sutskever, 2020)

Question

Let $a > 0$ *and* $b > 0$, *such that* $ab = b - a$, *show that*

$$\frac{a}{b} + \frac{b}{a} - ab = 2$$

GPT

Proof

Ilya Sutskever: The OpenAI Genius Who Told Sam Altman He Was Fired

Company's chief scientist led a board coup against one of the most prominent figures in Silicon Valley

Ilya Sutskever

SUIVRE

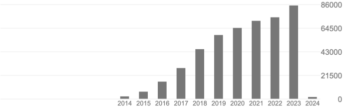
Co-Founder and Chief Scientist of OpenAI

Adresse e-mail validée de openai.com - Page d'accueil

Machine Learning Neural Networks Artificial Intelligence Deep Learning

ARTICLES CITÉE PAR ACCÈS PUBLIC COAUTEURS

	Toutes	Depuis 2019
Citations	462122	356473
indice h	84	79
indice i10	114	110



AI for Mathematical Proof

First approach: training a Transformer (GPT-f, Polu, Sutskever, 2020)

Question

Let $a > 0$ and $b > 0$, such that $ab = b - a$, show that

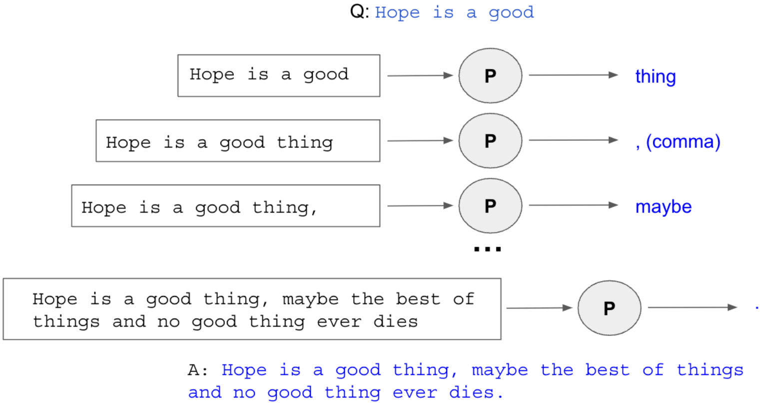
$$\frac{a}{b} + \frac{b}{a} - ab = 2$$

GPT

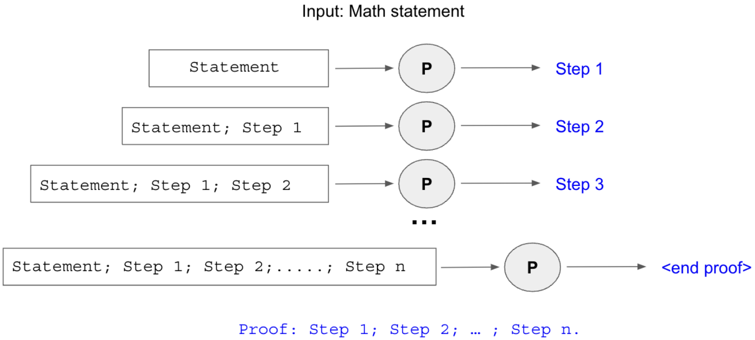
Proof

AI for Mathematical Proof

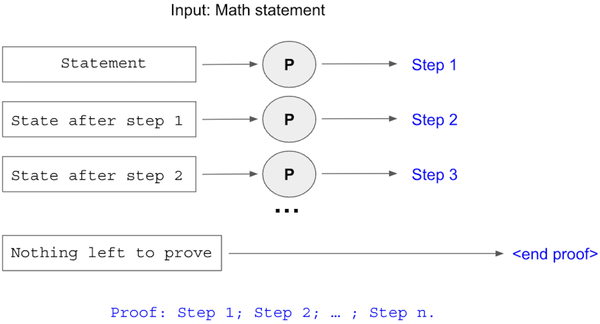
What does an autoregressive transformer do (in a nutshell) ?

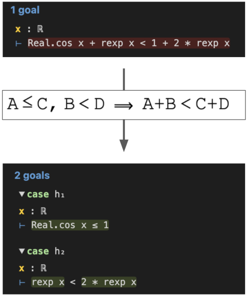


AI for Mathematical Proof



AI for Mathematical Proof





AI for Mathematical Proof

```
theorem Exercice_1
(a b : ℝ)
(h₀: a > 0)
(h₁: b > 0)
(h₂: a*b = b-a) :
a/b+b/a-2*(a*b) = 2 :=
begin
sorry,
end
```

Procedure: train it with examples: (exercises, proofs)

- The hope is that by showing it enough examples, the AI will be capable of learning to reason, just by learning to predict the next step each time.

AI for Mathematical Proof

```
theorem Exercice_1
(a b : ℝ)
(h₀: a > 0)
(h₁: b > 0)
(h₂: a*b = b-a) :
a/b+b/a-2*(a*b) = 2 :=
begin
sorry,
end
```

Proof

Procedure: train it with examples: (exercises, proofs)

- The hope is that by showing it enough examples, the AI will be capable of learning to reason, just by learning to predict the next step each time.

AI for Mathematical Proof

```
theorem Exercice_1
(a b : ℝ)
(h₀: a > 0)
(h₁: b > 0)
(h₂: a*b = b-a) :
a/b+b/a-2*(a*b) = 2 :=
begin
sorry,
end
```

Procedure: train it with examples: (exercises, proofs)

- The hope is that by showing it enough examples, the AI will be capable of learning to reason, just by learning to predict the next step each time.

Enough = sufficiently diverse and sufficiently numerous

→ Limitation: lack of data

AI for Mathematical Proof

We have very few data available (especially formal).

AI for Mathematical Proof

We have very few data available (especially formal).

Question

Let $a > 0$ and $b > 0$, such that $ab = b - a$, show that

$$\frac{a}{b} + \frac{b}{a} - ab = 2$$

informal language

```
theorem Exercice_1
(a b : ℝ)
(h₀: a > 0)
(h₁: b > 0)
(h₂: a*b = b-a) :
a/b+b/a-2*(a*b) = 2 :=
begin
sorry,
end
```

formal language

AI for Mathematical Proof

We have very few data available (especially formal).

Question

Let $a > 0$ and $b > 0$, such that $ab = b - a$, show that

$$\frac{a}{b} + \frac{b}{a} - ab = 2$$

informal language

```
theorem Exercice_1
(a b : ℝ)
(h₀: a > 0)
(h₁: b > 0)
(h₂: a*b = b-a) :
a/b+b/a-2*(a*b) = 2 :=
begin
sorry,
end
```

formal language

Lean: $\sim$150,000 theorems. A large dataset for humans, a small dataset for AI.

AI for Mathematical Proof

We have very few data available (especially formal).

Question

Let $a > 0$ and $b > 0$, such that $ab = b - a$, show that

$$\frac{a}{b} + \frac{b}{a} - ab = 2$$

informal language

```
theorem Exercice_1
(a b : ℝ)
(h₀: a > 0)
(h₁: b > 0)
(h₂: a*b = b-a) :
a/b+b/a-2*(a*b) = 2 :=
begin
sorry,
end
```

formal language

Lean: ~150,000 theorems. A large dataset for humans, a small dataset for AI. → Limit of the approach

AI for Mathematical Proof

We have very few data available (especially formal).

Question

Let $a > 0$ and $b > 0$, such that $ab = b - a$, show that

$$\frac{a}{b} + \frac{b}{a} - ab = 2$$

informal language

```
theorem Exercice_1
(a b : ℝ)
(h₀: a > 0)
(h₁: b > 0)
(h₂: a*b = b-a) :
a/b+b/a-2*(a*b) = 2 :=
begin
sorry,
end
```

formal language

Lean: ~150,000 theorems. A large dataset for humans, a small dataset for AI. → Limit of the approach

LeanLlama F. Glöckle et al. 2023 (Temperature-scaled large language models for Lean proofstep prediction)

AI for Mathematical Proof

Second approach: treat mathematics as a game (Lample, Lachaux, Lavril, Martinet, Hayat, Ebner, Rodriguez, Lacroix, 2022); (Gloeckle, Limperg, Synnaeve, Hayat, 2024)

.

AI for Mathematical Proof

Second approach: treat mathematics as a game (Lample, Lachaux, Lavril, Martinet, Hayat, Ebner, Rodriguez, Lacroix, 2022); (Gloeckle, Limperg, Synnaeve, Hayat, 2024)

.

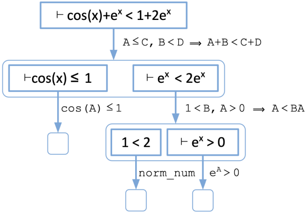
Deepmind (2017)

AI for Mathematical Proof

Second approach: treat mathematics as a game (Lample, Lachaux, Lavril, Martinet, Hayat, Ebner, Rodriguez, Lacroix, 2022); (Gloeckle, Limperg, Synnaeve, Hayat, 2024)

.

Deepmind (2017)

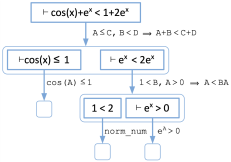


You won !

AI for Mathematical Proof

Main difficulties:

- two-player game vs. solo against a goal.
- In chess, when you play a move you always have a single game. In mathematics: one statement $\to$ multiple statements
- Difficult in mathematics to know automatically in the middle of a proof what the probability of succeeding is.
- The number of possibilities is much, much larger in mathematics

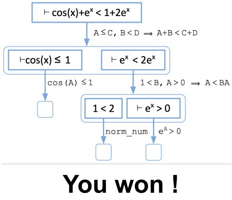


You won !

AI for Mathematical Proof

Main difficulties:

- two-player game vs. solo against a goal.
- In chess, when you play a move you always have a single game. In mathematics: one statement $\rightarrow$ multiple statements
- Difficult in mathematics to know automatically in the middle of a proof what the probability of succeeding is.
- The number of possibilities is much, much larger in mathematics

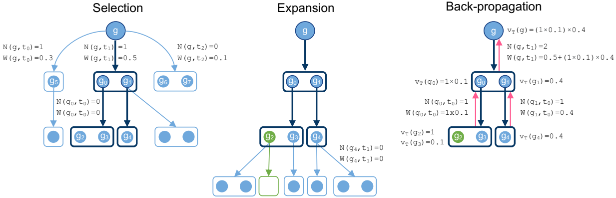


Much more difficult than chess

AI for Mathematical Proof

In practice

- Two transformers: P_θ which predicts a tactic, c_θ which predicts the difficulty of proving a statement (goal, hypothesis, etc.).
- An intelligent proof search that sees the proof as a tree and combines P_θ, c_θ and a tree expansion.



- Continuously training of P_θ and c_θ on successful proofs

AI for Mathematical Proof

Results

Exercises at the undergraduate level...

...30 to 60% of middle school / high school exercises up to Olympiad level...

...and a few exercises from the International Mathematical Olympiads

AI for Mathematical Proof

Results

Exercises at the undergraduate level...

...30 to 60% of middle school / high school exercises up to Olympiad level...

...and a few exercises from the International Mathematical Olympiads

Exercise

Show that for all $n \in \mathbb{N}$, 7 does not divide $2^n + 1$.

AI for Mathematical Proof

Results

Exercises at the undergraduate level...

...30 to 60% of middle school / high school exercises up to Olympiad level...

...and a few exercises from the International Mathematical Olympiads

Exercise

Show that for all $n \in \mathbb{N}$, 7 does not divide $2^n + 1$.

(Lample, Lachaux, Lavril, Martinet, Hayat, Ebner, Rodriguez, Lacroix, 2022)

AI for Mathematical Proof

Results

Exercises at the undergraduate level...

...30 to 60% of middle school / high school exercises up to Olympiad level...

...and a few exercises from the International Mathematical Olympiads

Exercise

Show that for all $n \in \mathbb{N}$, 7 does not divide $2^n + 1$.

(Lample, Lachaux, Lavril, Martinet, Hayat, Ebner, Rodriguez, Lacroix, 2022)

Unexpected outcomes

French AI start-up Mistral reaches unicorn status, marking its place as Europe's rival to OpenAI

> Mistral's value has increased more than sevenfold in six months. It raised nearly €500 million in November and €105 million in its first funding round.

(Euronews 11/12/2023)

MISTRAL AI_

What is next? Autoformalization

What is next? Autoformalization

Solo game $\implies$ still need for a good base supervised model P_θ (trained on data)

What is next? Autoformalization

Solo game $\implies$ still need for a good base supervised model P_θ (trained on data)

How to get more formal data ? One idea: train a model to translate from "natural language proof" to formal and verifiable data.

Exercice 1

Let $a > 0$ and $b > 0$, such that $ab = b - a$, show that

$$\frac{a}{b} + \frac{b}{a} - ab = 2$$

informal language

$\rightarrow$

```
theorem Exercice_1
(a b : ℝ)
(h₀: a > 0)
(h₁: b > 0)
(h₂: a*b = b-a) :
a/b+b/a-2*(a*b) = 2 :=
begin
sorry,
end
```

formal language

What is next? Autoformalization

Solo game $\implies$ still need for a good base supervised model P_θ (trained on data)

How to get more formal data ? One idea: train a model to translate from "natural language proof" to formal and verifiable data.

Exercice 1

Let $a > 0$ and $b > 0$, such that $ab = b - a$, show that

$$\frac{a}{b} + \frac{b}{a} - ab = 2$$

informal language

$\rightarrow$

```
theorem Exercice_1
(a b : ℝ)
(h₀: a > 0)
(h₁: b > 0)
(h₂: a*b = b-a) :
a/b+b/a-2*(a*b) = 2 :=
begin
sorry,
end
```

formal language

Autoformalization: both a means and an end:

- a means: more data to train AI models
- an end: checking the correctness of new mathematical theories

What is next? Autoformalization

Solo game $\implies$ still need for a good base supervised model (trained on data)

How to get more formal data ? One idea: train a model to translate from "natural language proof" to formal and verifiable data.

Autoformalization: both a means and an end:

- a means: more data to train AI models
- an end: checking the correctness of new mathematical theories

Example of contradictory papers:

Schumacher, G., & Tsuji, H. (2004). Quasi-projectivity of moduli spaces of polarized varieties. Annals of mathematics, 597-639.

Kollár, J. (2006). Non-quasi-projective moduli spaces. Annals of mathematics, 1077-1096.

Today, very good statement autoformalization for olympiad style exercise, challenges for proof autoformalization, very far from research level mathematics (e.g. arxiv).

AI and maths

A fast-moving field

2019	2022	2024
$(r_1 - r_2)r_3 = r_1 r_3 - r_2 r_3$	$\forall\ n \in \mathbb{N},\ \neg 7 \mid 2^n + 1$	Silver medal Inter. Math. Olymp.
HOList, LPLG	GPT-f, Thor/DSP, HTPS	Llemma, LeanDojo, InternML, DeepSeek, ABEL, AlphaProof, etc.

AI and maths

A fast-moving field

2019	2022	2024
$(r_1 - r_2)r_3 = r_1 r_3 - r_2 r_3$	$\forall\ n \in \mathbb{N},\ \neg 7 \mid 2^n + 1$	Silver medal Inter. Math. Olymp.
HOList, LPLG	GPT-f, Thor/DSP, HTPS	Llemma, LeanDojo, InternML, DeepSeek, ABEL, AlphaProof, etc.

New approaches post 2024: end-to-end RL and reasoning models (o1, o3, R1, etc.).

Conclusion

- AI methods are already useful in the practice of mathematics
- AI and LLM for proving theorems is only beginning, and there are many ideas... and much to do.
- LLMs will not be the final AI tool which will be used to mathematics.
- The practice of mathematics will probably change... and that's okay.
- AI will not replace mathematicians but will instead enhance them.

Conclusion

Thank you for your attention