

Mathematics in the Age of AI

Jose **Brox**

MANOLO**FEST!**

Universidade de Santiago de Compostela



Instituto de Investigación
en Matemáticas



Universidad de Valladolid

25/06/26

A paradigm shift

The future of mathematics?

The everyday mathematician will be more like:

① A **computer programmer** and a **software engineer**.

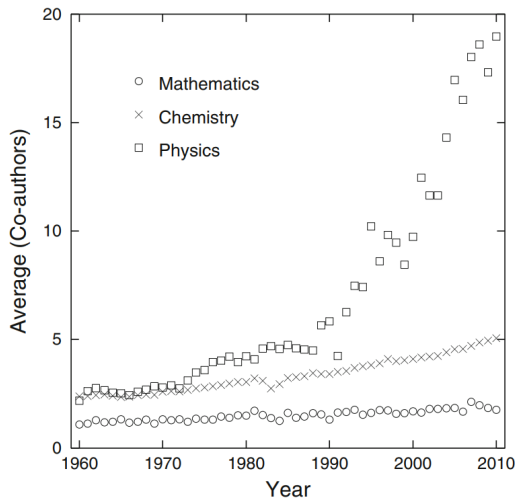
- ▶ For collaboration, exploration, theorem proving, formalization, and reviewing,
- ▶ by using collaborative platforms, interactive math software, automated provers, and proof assistants.

② A **principal investigator** and a **thesis supervisor**.

- ▶ For collaboration with generative AI tools
 - ▶ in all the previously mentioned tasks.
- ▶ Most of you probably already have both feet entering its waters (do you use **Overleaf**? **Mathematica**? Do you ask **ChatGPT**), but haven't entered deep waters yet.

Tao's vision

- How we conduct math research didn't change for centuries



Tao's vision

- **Large-scale and heterogeneous collaboration** are rare in mathematics, due to:
- ① **High entry barriers** (amateurs cannot contribute)
 - ② **High specialization among areas**
(lack of interdisciplinarity)
 - ③ **Rigorous proof standards** (AIs cannot contribute)
 - ④ **Non-scaleable workflows** (1-5 experts in the team)

Having **different contributors** with **different backgrounds and expertise** makes communication and validation difficult, producing a **loss of trust** in both the process and the result

Tao's vision

These handicaps can be surmounted thanks to new tools:

- ① **Collaborative platforms with distributed version control**
- ② **Formal proof assistants with automation**

Hence allowing:

- ① **Citizen mathematics**
- ② **Participation of AI tools** (ATPs, LLMs)
- ③ **Trusted interaction between experts** from separate areas
- ④ **Large-scale** mathematical research projects
- ⑤ **Data-driven** mathematics
- ⑥ **Transparent dissemination** of both process and results

AI×Math

Brox

Intro

The future

Tao's vision

New tools

Collaboration

Formalization

Automatization

Databases

Symbolic AI

ATPs

Generative AI

LLMs

Evidence

Sociology of
the future

Research

Closing

New tools

Collaboration

Informal communication chats

Zulin

The screenshot shows the Zulin web interface. At the top, there's a header with 'Equational' and 'Coordination for the Equational Project'. Below this, there's a search bar and a date '11 DIC 2025'. The main content area shows a chat thread titled 'Submitting the paper'. The text in the chat reads: 'of only 3 readers: the editor and the referees. Everyone else will learn about the work through arXiv, word of mouth, literature search tools like Google Scholar, and citations in other articles. Those who don't have the inclination to read our paper will not change their mind based on which journal it appears in.' Below this, there's a post by 'Jihoon Hyun' with a LaTeX diagram and text: 'I was tinkering around with equations, and figured out that the only magma of size 5 (up to equiv) satisfies both 677 and 47 (dual of 255), but not 2910 (dual of 677). As 47 is somewhat weaker than 2910, I'm searching for magmas satisfying both 677 and 47, but no magmas other than that of size 5 showed up to size 9. Is there some 677 magma which satisfy 47 (or, 2910) in the database? Also, is there something we know about magmas satisfying both 677 and 2910, and in general, magmas satisfying both some equation and its dual? If it shows some sort of symmetry then it may be easier to somewhat construct or analyze the structure.'

On the left sidebar, there's a 'Filter left sidebar' search bar and a list of 'VISTAS' (Views) including 'Conversaciones recientes', 'Bandeja de entrada', 'Feed combinado', 'Menciones', 'Reacciones', 'Mensajes destacados', and 'Borradores'. Below this, there's a 'MENSAJES DIRECTOS' (Direct Messages) section and a 'PINNED CHANNELS' section with 'Equational' and 'FINITE: 677 -> 255'. On the right sidebar, there's a 'Filtrar usuarios' (Filter users) search bar and a 'THIS CHANNEL (310)' list of users including Jose Brox, Andrew Yang, Chris Henson, Kim Morrison, Philip Zucker, Riyaz Ahuja, Ben Eltschig, Bjørn Kjos-Hanssen, David Ledvinka, David Renshaw, Elif Uskulu, Eric Wieser, James Sundstrom, Jared green, Jason Rute, Jireh Loreaux, Joe Watt, Johan Commelin, Junyan Xu, and Kevin Buzzard.

- ▶ The core of the **discussion** happens here
- ▶ Allow **many threads**
- ▶ Allow **writing in LaTeX** and **code snippets**
- ▶ Allow **attaching files** and **links to other platforms**

GIT platforms

with distributed version control:

- ▶ There is a **master copy** of the project, but
- ▶ **each collaborator also owns a full copy** of the project,
- ▶ which she **can modify** at will.
- ▶ **Changes can be merged** to the master copy,
- ▶ **after careful revision**, human or automated.
- ▶ Any **changes can be reverted**.
- ▶ **Conflicts** due to confluent changes **are detected and fixed**

Work in parallel, propose changes for the stable version,
wait for comments and conflicts, revise, merge

GitHub

The screenshot shows the GitHub interface for the repository 'teorth / equational_theories'. The repository is public and was generated from 'leanprover-community/LeanProject'. It has 18 issues, 10 pull requests, 1 project, 1 wiki, and 473 stars. The repository is currently on the 'main' branch, which has 18 branches and 5 tags. The commit history shows a recent commit by 'dependabot[bot]' 3 weeks ago, and a commit by 'teorth' 2 years ago. The repository contains files like '.github', '.vscode', 'blueprint/src', 'commentary', and 'data'. The 'About' section describes the project as a map of relations between different equational theories of Magmas.

teorth / equational_theories

generated from [leanprover-community/LeanProject](#)

Unwatch 6 Fork 89 Starred 473

main 18 Branches 5 Tags

Go to file

Code

dependabot[bot] Bump actions/cache from 4 to 5 5e46624 · 3 weeks ago 1,744 Commits

.github	Bump actions/cache from 4 to 5 (#1425)	3 weeks ago
.vscode	Update settings.json	2 years ago
blueprint/src	Paper dec13 improvements (#1419)	last month
commentary	Paper dec13 improvements (#1419)	last month
data	Higman-Neumann section of the paper (#...	3 months ago

About

A project to map out the relations between different equational theories of Magmas.

[teorth.github.io/equational_theories/](#)

Readme

Apache-2.0 license

Code of conduct

Contributing

Cite this repository

- There is a main **repository**
- Each collaborator has its own **fork** repository
- Changes are proposed in **pull requests**
- Tasks to be done are listed as **issues**

AI×Math

Brox

Intro

The future

Tao's vision

New tools

Colaboration

Formalization

Automatization

Databases

Symbolic AI

ATPs

Generative AI

LLMs

Evidence

Sociology of
the future

Research

Closing

New tools

Formalization

Formal proof assistants

- ▶ Proofs are **written in a formal language**,
- ▶ and then **verified by the assistant**.
- ▶ The **soundness** of the proof **depends on previous definitions and results**.
- ▶ There are **libraries** with previously formalized definitions and results.
- ▶ There are **tactics that can fill routine steps of the proofs**.

Write the proof in the formal language, base it on previous libraries, use tactics, wait for the assistant's verdict, revise

Lean

Powerful automation

Mathematics

```

/-- A prime is a number larger than 1 with no trivial divisors -/
def IsPrime (n : Nat) := 1 < n ∧ ∀ k, 1 < k → k < n → ¬ k | n

/-- Every number larger than 1 has a prime factor -/
theorem exists_prime_factor :
  ∀ n, 1 < n → ∃ k, IsPrime k ∧ k | n := by
  intro n h1
  -- Either `n` is prime...
  by_cases hprime : IsPrime n
  · grind [Nat.dvd_refl]
  -- ... or it has a non-trivial divisor with a prime factor
  · obtain (k, _) : ∃ k, 1 < k ∧ k < n ∧ k | n := by
      simp_all [IsPrime]
      obtain (p, _, _) := exists_prime_factor k (by grind)
      grind [Nat.dvd_trans]

```

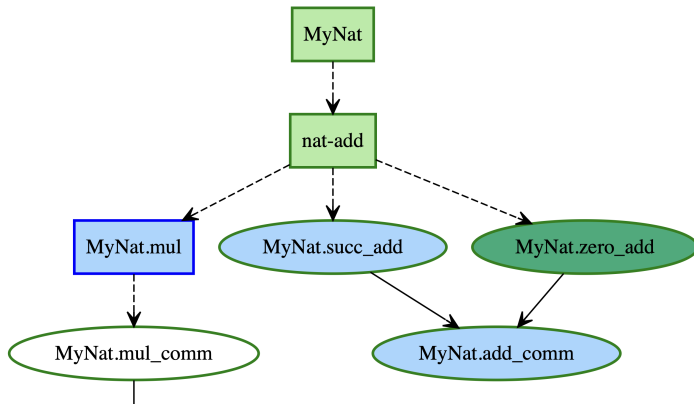
A definition of prime numbers and a proof that there are infinitely many prime numbers.



- ▶ There is big **Mathlib** library (which we don't use)
- ▶ Integrates well with **GitHub**
- ▶ Beware of external tactics, as `native_decide` (we can cheat)

Blueprint

Intermediate step between formal and informal proofs:



- ▶ Allows to follow **what has yet to be formalized**
- ▶ Facilitates **working in parallel**

Benefits of formalization (Jeremy Avigad)

- ① Verify theorems
- ② Correct errors
- ③ Gain intuition
- ④ Build libraries
- ⑤ Search for definitions and theorems
- ⑥ Refactor proofs
- ⑦ Refactor libraries
- ⑧ Design concepts
- ⑨ Communicate
- ⑩ Collaborate
- ⑪ Manage complexity
- ⑫ Manage the literature
- ⑬ Teach
- ⑭ Improve access
- ⑮ Use mathematical computation
- ⑯ Use automated reasoning
- ⑰ Use machine learning
- ⑱ Support a synthesis of machine learning and symbolic AI



J. Avigad (2024). *Mathematics and the formal turn*. **Bull. AMS** 61

AI×Math

Brox

Intro

The future

Tao's vision

New tools

Colaboration

Formalization

Automatization

Databases

Symbolic AI

ATPs

Generative AI

LLMs

Evidence

Sociology of
the future

Research

Closing

New tools

Automatization

Dynamic databases

Number	Description	Best lower bound	Best upper
1a	Sidon set autocorrelation constant	1.2802	1.502862
1b	Erdős minimum overlap constant	0.379005	0.380868
2	Crouzeix constant	2	$1 + \sqrt{2} \approx 2.4142$
3a	Gyamarti-Hennecart-Ruzsa sum-difference constant	1.1740744	1.33333
3b	Takeya sums-differences constant	> 1.77898	1.83333
3c	4-slope Takeya-type sum-difference constant	1.67473389	1.75
4a	Cap set constant	2.2202	2.756
4b	Furstenberg-Sárközy square-difference constant	0.733412	1
5a	Sidon set size constant	0	0.97633
5b	Sidon set density inside (4.5) sets	0.5294	0.5714

- Problems are **ordered by logical dependency**
- A **new value** of one constant **updates all** downstream values

AI×Math

Brox

Intro

The future

Tao's vision

New tools

Colaboration

Formalization

Automatization

Databases

Symbolic AI

ATPs

Generative AI

LLMs

Evidence

Sociology of
the future

Research

Closing

Automatization

Symbolic AI

Automated theorem provers (ATPs)

- ▶ We enter some axioms and a hypothesis or conjecture
- ▶ ATPs **generate proofs**
- ▶ They are also able to **find counterexamples**
- ▶ They **work with first-order logic**, or perhaps higher-order logic
- ▶ They can produce **proofs with many formulas, very long ones**,
- ▶ which can be **difficult for humans to understand**

There are interfaces that translate the output of ATPs into Lean format

Vampire

Intro

 The future
 Tao's vision

New tools

 Colaboration
 Formalization
 Automatization
 Databases
 Symbolic AI
 ATPs
 Generative AI
 LLMs
 Evidence

 Sociology of
 the future

Research

Closing

```

\C(base) josebrox@HISOKA:~/vampire5 -t 10 -sa lrs -t 10 -awr 1:500 -fde none -nwc 0.5 -p tptp
% Running in auto input_syntax mode. Trying TPTP
% Refutation found. Thanks to Tanya!
% SZS status Theorem for TPTP 102744082 implies right mult injective
% SZS output start Proof for TPTP 102744082 implies right mult injective
Fof(f488040,plain,(
  $false),
  inference(subsumption_resolution,[],[f487741,f15])).
Fof(f15,plain,(
  sk1 != sk2),
  inference(cnf_transformation,[],[f11])).
Fof(f11,plain,(
  sk1 != sk2 & m(sk1,sk0) = m(sk2,sk0)),
  inference(skolemisation,[status(esa),new_symbols(skolem,[sk0,sk1,sk2])],[f8,f10])).
Fof(f10,plain,(
  ? [X0,X1,X2] : (X1 != X2 & m(X1,X0) = m(X2,X0)) => (sk1 != sk2 & m(sk1,sk0) = m(sk2,sk0))),
  introduced(definition,[],[choice_axiom])).
Fof(f8,plain,(
  ? [X0,X1,X2] : (X1 != X2 & m(X1,X0) = m(X2,X0))),
  inference(ennf_transformation,[],[f4])).
Fof(f4,negated_conjecture,(
  ~! [X0,X1,X2] : (m(X1,X0) = m(X2,X0) => X1 = X2)),
  inference(negated_conjecture,[status(cth)],[f3])).
Fof(f3,conjecture,(
  ! [X0,X1,X2] : (m(X1,X0) = m(X2,X0) => X1 = X2)),
  file('TPTP 102744082 implies right mult injective.p',unknown)).
Fof(f487741,plain,(
  sk1 = sk2),
  inference(forward_demodulation,[],[f487109,f481825])).
Fof(f481825,plain,(
  ( ! [X3,X0,X4] : (m(m(X3,X4),m(s(X0),X4)) = X3) ),
  inference(forward_demodulation,[],[f481078,f390316])).
  
```

AI×Math

Brox

Prover9/Mace4

Intro

The future

Tao's vision

New tools

Colaboration

Formalization

Automatization

Databases

Symbolic AI

ATPs

Generative AI

LLMs

Evidence

Sociology of

the future

Research

Closing

MACE4 677 as linear product in nonassociative Jordan ring - Prover9/Mace4

File Preferences View Help

Language Options Formulas Prover9 Options Mace4 Options Additional Input

Assumptions:

%with 2,4,no

%size 35, size 16 when nonunital noncommutative

%Jordan ring

x + y = y + x #label(sum_commutative).

(x+y)+z = x+(y+z) #label(sum_associative).

0 + x = x #label(left_neutral).

x + 0 = x #label(right_neutral).

all x exists y (x + y = 0) #label(opposite).

x @ y = y @ x #label(commutativity).

(x @ y) @ (x @ x) = x @ (y @ (x @ x)) #label(Jordan_identity).

~(all x all y all z (x @ y) @ z = x @ (y @ z)) #label(not_associative).

1 @ x = x #label(left_unit).

x @ 1 = x #label(right_unit).

0 @ x = 0 #label(left_annihilation).

x @ 0 = 0 #label(right_annihilation).

x @ (y + z) = (x @ y) + (x @ z) #label(left_distributivity).

(x + y) @ z = (x @ z) + (y @ z) #label(right_distributivity).

%Problem

exists a exists b all x all y (x * y = (a @ x) + (b @ y)) #label(linear_product_from_semirring).

x = y * (x * ((y * x) * y)) #label(677).

x = (y*x)*((y*(y*x))*y) #label(consequence677).

%Finiteness clauses

x*y = x*z -> y = z #label(left_multiplication_injective).

all x all y exists z (x*z = y) #label(left_multiplication_surjective).

%exists x ~(exists y (y*x = x)) #label(weak_right_multiplication_not_surjective).

Goals:

x = ((x * x) * x) * x #label(255).

Show Current Input

Proof Search

Prover9

Time Limit: 180 seconds.

Start

Kill

State: Ready

Info

Show/Save

Model/Counterexample Search

Mace4

Time Limit: 60 seconds.

Start

Kill

State: Ready

Info

Show/Save

AI×Math

Brox

Intro

The future

Tao's vision

New tools

Colaboration

Formalization

Automatization

Databases

Symbolic AI

ATPs

Generative AI

LLMs

Evidence

Sociology of
the future

Research

Closing

Automatization

Generative AI

Large language models (LLMs)

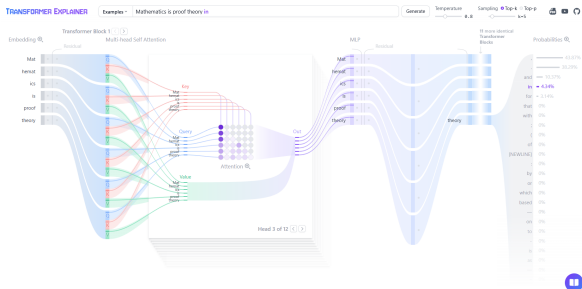


- ▶ Originally, **words were read sequentially**
- ▶ The **processing of a word depended on the immediately preceding one** and on the state of the system

The stochastic-parrots fallacy

Transformer

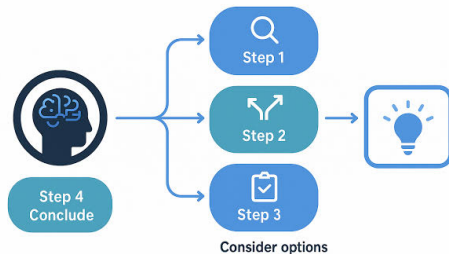
- Each word is **compared with the rest** simultaneously



- Scales with input data, computing capacity, and number of parameters (neuron weights)

Chain of thought

CHAIN OF THOUGHT



- ▶ The LLM **writes out intermediate reasoning steps**
- ▶ It **minimizes the error in each of them**
- ▶ It **self-checks and backtracks** on its own

Coding AI



- ▶ The LLM is **able to write code**
- ▶ It correctly **decides the code given the specification**
- ▶ It decides **its own specification**

Agentic AI



- ▶ We give the AI **access to the environment**
- ▶ It can receive **feedback from its own actions**
- ▶ We can span **multiple agents** and make them **interact**

Primes poem

Prompt: Can you write a proof that there are infinitely many primes, with every line that rhymes?

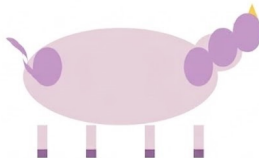
GPT-4:

```
Yes, I think I can,
Though it might take a clever plan.
I'll start by noting Euclid's proof,
Which shows that primes aren't just aloof.
Assume we have a finite list,
Of primes, and that none have been missed.
Multiply them all together,
And add one, just to be clever.
The result will be a number,
That has no prime factors, oh wonder!
But every number has a prime,
That divides it, it's just a matter of time.
So we've found a contradiction,
And our finite list needs eviction.
There must be infinitely many primes,
And that's the end of my rhyming lines.s
```



S. Bubeck et al (2023). Sparks of artificial general intelligence. [arXiv](#)

TikZ Unicorn

GPT-3.5**GPT-4****GPT-4.5****GPT-5****GPT-5.2****GPT-5.5**

Year	Reasoning & Discovery	Formalization & Verification
2020	GPT-3: reliable 2-digit arithmetic (emergent)	Liquid Tensor Experiment launched — Clausen–Scholze condensed math (done 2022)
2021	—	miniF2F: first cross-system formal benchmark
2022	Chain-of-thought + Minerva (~50% MATH); AlphaTensor matmul records	First LLM autoformalization: ~25% of competition problems to Isabelle
2023	FunSearch: first genuinely new construction (cap-set bound)	Polynomial Freiman–Ruzsa conjecture formalized in Lean (Tao et al., ~3 weeks)
2024	AlphaProof + AlphaGeometry 2: IMO silver (28/42); FrontierMath; o1/o3	AlphaProof: ~100M autoformalized theorems

AI×Math

Brox

Intro

The future
Tao's vision

New tools

Colaboration
Formalization
Automatization
Databases
Symbolic AIATPs
Generative AI
LLMs
EvidenceSociology of
the future
Research

Closing

Reasoning & Discovery

Formalization & Verification

2025

May AlphaEvolve: 4×4 matmul in 48
mults (**improves Strassen**)Jul **IMO gold in natural language**
(Gemini, OpenAI): 35/42Sept- **LLMs saturate Putnam**
Dec**Equational Theories Project**
(Tao): 4694 laws**Harmonic's Aristotle: IMO gold
in formal Lean**Math Inc.'s Gauss: **Prime Number
Theorem** formalized

2026

Jan Erdős "long tail" solved
autonomously (GPT-5.2),
Tao-checkedFeb **First Proof** (Hairer et al): 6-8/10
on fresh research lemmasGauss + Viazovska: **sphere
packing** formalized, dims 8 & 24Mar AlphaEvolve: lower bounds for 9
classical **Ramsey numbers**; Claude
solves **Knuth problem**; Aletheia
solves a **Kirby list problem**Apr **Erdős #1196** (GPT-5.4 Pro):
good mathematicians tried 7 yearsMay **Gower's PhD problem solved**;
Unit-distance conjecture
disprovedAleph Prover: unit-distance
disproof formalized in Lean 4

AI×Math

Brox

Intro

The future

Tao's vision

New tools

Colaboration

Formalization

Automatization

Databases

Symbolic AI

ATPs

Generative AI

LLMs

Evidence

Sociology of
the future

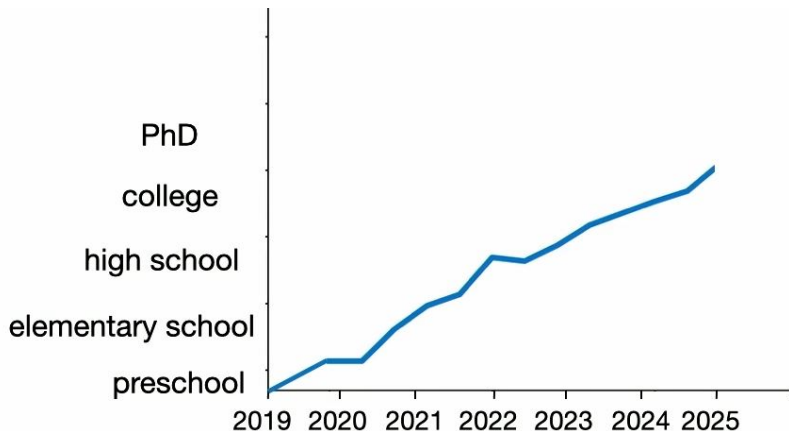
Research

Closing

Sociology of the future

Consequences of AI for research

LLMs math skills

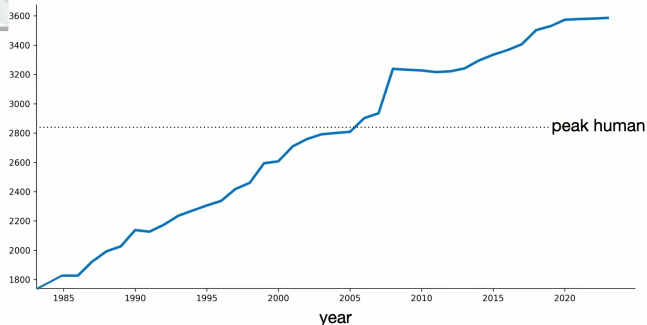


LLMs chess skills



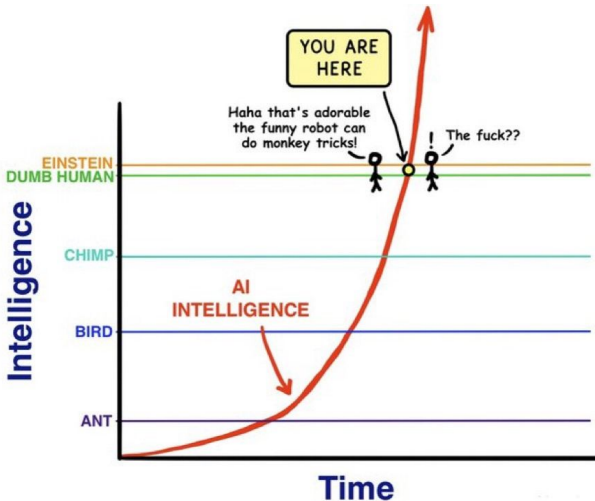
strength
(SSDF Elo)

strength of best chess bot on desktop computer



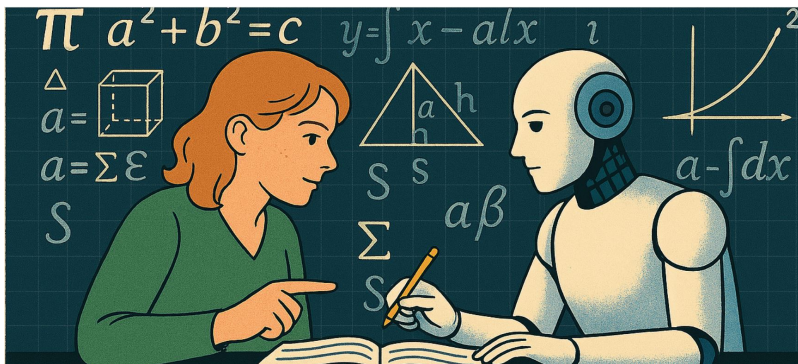
TOY → TOOL → CENTAUR → SUPERHUMAN

Stochastic parrots, monkey tricks?



Consequences for research

Option 1:



Consequences for research. Option 1

- ▶ **Losing funding** sources
- ▶ **Journals requiring formalized proofs**
- ▶ Repositories cannot keep up
- ▶ **AI-generated preprint servers**
- ▶ **AI reviewers**
- ▶ Do **papers and journals make sense?**
- ▶ Does **attribution make sense?**
- ▶ **Learning very quickly** from AI
- ▶ **Mathematicians working on more areas**
- ▶ Will **high-level LLMs be available to everyone**, reserved for a few privileged universities, or only for the company and its clients?
- ▶ Do **PhDs have a future?**

Consequences for research

Option 2:



- Let us not be naive about our chances in this case...

CONGRATULATIONS, MANOLO!

