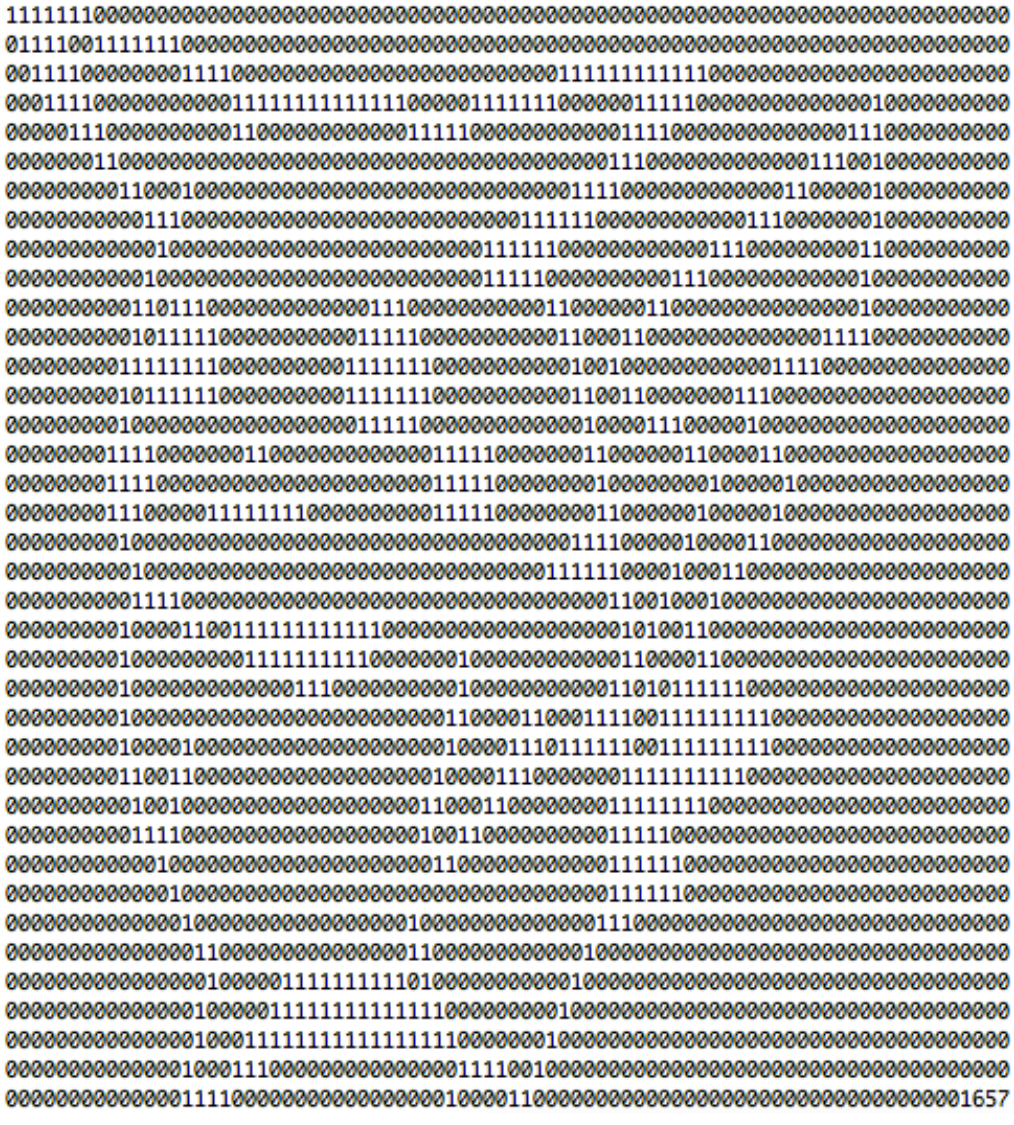
AI and Collatz Conjecture

David Malone

2026-08-01

Neat Trick: Pikaprime

How did I make this?



Neat Trick: Pikaprime

- Make a pattern from digits.
- We know density of primes is about $1/\ln n$.
- So the average gap is about $\ln n$.
- So, never too far to the next prime.
- We also have good tests for primality: Miller-Rabin.
- So, just increment until you find a prime.

Terence Tao: Mathematics in the age of AI

- Talk at the IMC on July 24.
- Could AI cause a crisis in practice or values of mathematics?
- Community question:

  How should the mathematical community respond to the advent of modern AI technologies, and their real and/or claimed capabilities to perform mathematical tasks?

- Question is not really a mathematical question.
- Check full slides at `https://teorth.github.io/tao-web/slides/age-of-ai-icm-2026.pdf`
- Quick summary...

- Template conjecture about capability of AI in mathematics.
- Does an analysis under working hypothesis:

  AI tools will, reasonably soon, become capable of performing a reasonable fraction of research-level mathematical tasks, with reasonable levels of success, quality, supervision, and cost.

- Mathematics has a mix of explicit and implicit goals (e.g. solve problems, make new theories, understand world, build community, . . .)
- AI proofs may result in these goals not being aligned (e.g. solving problems and increasing understanding)
- More to the mathematics than just having a proof.
- Have a look at the *Leiden Declaration*, `https://leidendeclaration.ai`

AI and Mathematics: LLMs

- A Markov chain is a system where your next position depends (randomly) on where you are now. E.g. snakes and ladders.
- They have been used to study language for a long time.
- E.g. Shannon's *A Mathematical Theory of Communication* (1948) proposes generating English text using a Markov chain trained by looking at existing English text.
- *Where you are now* might mean the existing letter, pair of letters, word, tuple of words, ...
- An LLM adds a token to the end of existing text by summarising the existing text as a vector and predicting probabilities for the next word using a neural network.
- Neural network trained on lots of text and then *fine tuned*.

Sometimes called *stochastic parrots*. Can help with search (SAR).

From A Mathematical Theory of Communication

1. Zero-order approximation (symbols independent and equiprobable).

   XFOML RXKHRJFFJUJ ZLPWCFWKCYJ FFJEYVKCQSGHYD QPAAMKBZAACIBZL-HJQD.

2. First-order approximation (symbols independent but with frequencies of English text).

   OCRO HLI RGWR NMIELWIS EU LL NBNESEBYA TH EEI ALHENHTTPA OOBTTVA NAH BRL.

3. Second-order approximation (digram structure as in English).

   ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY ACHIN D ILONASIVE TU-COOWE AT TEASONARE FUSO TIZIN ANDY TOBE SEACE CTISBE.

4. Third-order approximation (trigram structure as in English).

   IN NO IST LAT WHEY CRATICT FROURE BIRS GROCID PONDENOME OF DEMONS-TURES OF THE REPTAGIN IS REGOACTIONA OF CRE.

5. First-order word approximation. Rather than continue with tetragram, . . . , n-gram structure it is easier and better to jump at this point to word units. Here words are chosen independently but with their appropriate frequencies.

   REPRESENTING AND SPEEDILY IS AN GOOD APT OR COME CAN DIFFERENT NAT-URAL HERE HE THE A IN CAME THE TO OF TO EXPERT GRAY COME TO FURNISHES THE LINE MESSAGE HAD BE THESE.

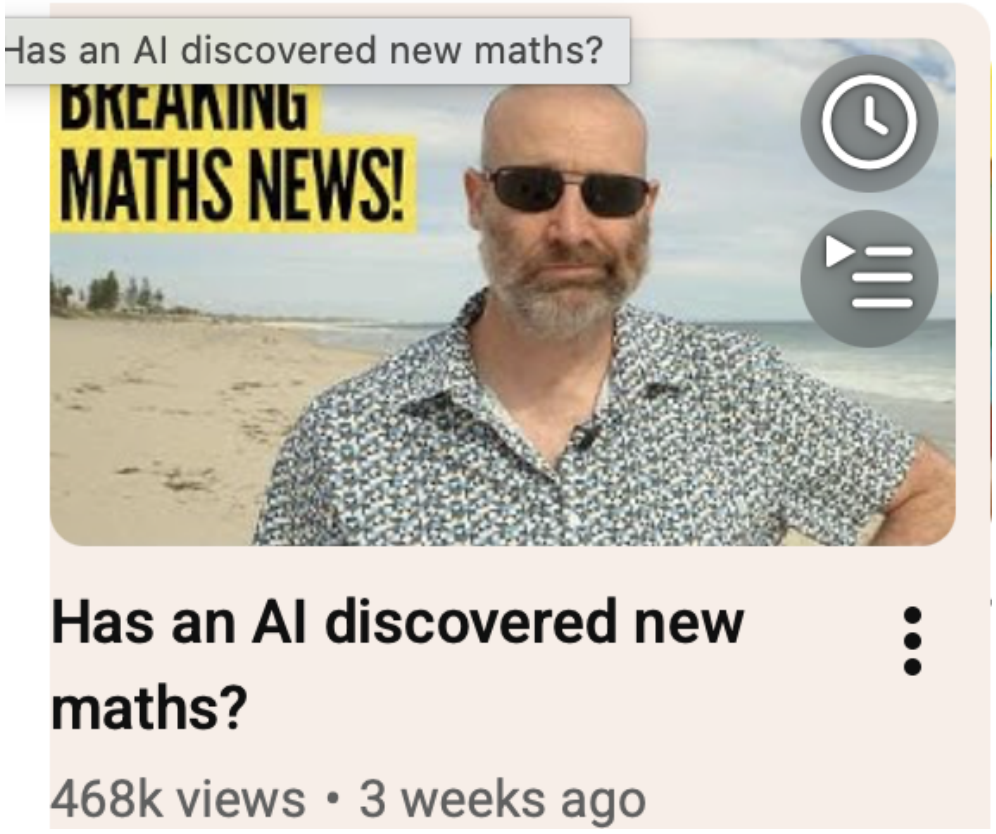
6. Second-order word approximation. The word transition probabilities are correct but no further struc-ture is included.

   THE HEAD AND IN FRONTAL ATTACK ON AN ENGLISH WRITER THAT THE CHAR-ACTER OF THIS POINT IS THEREFORE ANOTHER METHOD FOR THE LETTERS THAT THE TIME OF WHO EVER TOLD THE PROBLEM FOR AN UNEXPECTED.

AI and Mathematics: Checking

- Will a random proof be correct?
- Like generating a random number and hoping it is prime?
- You need to check!
- A human could read it, but...
- Why not output a formal proof and feed it to proof verifier.
- Tools like Lean, Rocq, HOL and others let you do this.

AI and Mathematics: Fast Moving



See Matt Parker's
`https://www.youtube.com/watch?v=ZJ8KThKAfbs`,
or Alex Wilkins *A revolution in maths* (New Scientist, 2026-06-06).

This Week's News: Collatz Conjecture

Pick a natural number a_0 and form a sequence according to the rule: if a_n term is even, $a_{n+1} = a_n/2$ otherwise $a_{n+1} = 3a_n + 1$. Sometimes called *hailstone sequence*.

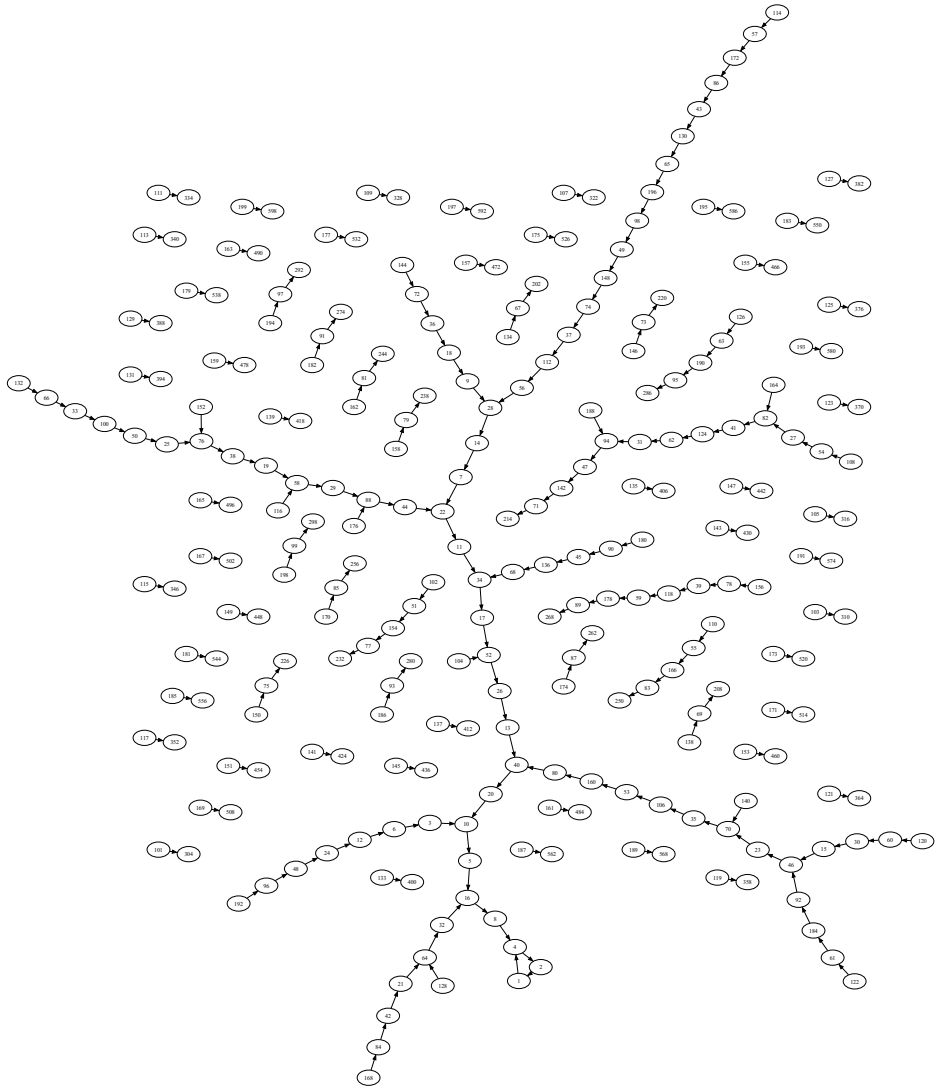
For example, if $a_n = 7$ then the sequence goes:

$$7, 22, 11, 34, 17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, 1, 4, 2, 1, \ldots$$

For every n tested this sequence winds up in this loop. At least all n up to 2.36×10^{21} have been tested.

Collatz conjectured in 1937 that it all numbers wind up in this loop.

Many Visualisations



AI and Collatz

Sat 25 Ramana Kumar uploaded a validating Lean disproof of Collatz conjecture[1].

Sun 26 Soundness bug fixed in Lean[2].

Tue 28 Kiran Gopinathan logs a bug (based on Collatz disproof)[3].

Tue 28 Long discussion kicks off in the Lean chat[4].

Ramana and Kiran get t-shirts for the bug.
Ramana confirms use of LLMs, but is otherwise a bit coy.

---

[1] https://github.com/xrchz/CollatzLean

[2] https://github.com/leanprover/lean-kernel-arena/pull/81

[3] https://github.com/leanprover/lean4/issues/14576

[4] https://leanprover.zulipchat.com/#narrow/channel/270676-lean4/topic/Counterexample.20to.20the.20Lean.20Conjecture.20.28Soundness.20Bug.29

Done!

1. Proof seems to take advantage of bugs in multiple Lean implementations.
2. Ramana is no stranger to verifying theorem provers[5].
3. Things to learn about verifiers and LLM use?
4. Ken Thompson's *Reflections on Trusting Trust*[6].
5. Reflect on human part of mathematics?
6. Coders are asking related questions.

[5] e.g. https://www.repository.cam.ac.uk/items/74f89cc3-4c4a-44d7-bc00-329b641e6bd7 and https://sequent.inc/blog/posts/why-hol-for-deep-verification/

[6] https://www.cs.cmu.edu/~rdriley/487/papers/Thompson_1984_ReflectionsonTrustingTrust.pdf

Miller-Rabin

Test for primality before trying to factor using Miller-Rabin.

Checks Fermat's little theorem and for bad solutions of $z^2 = 1$.

1. We want to check a number n for being prime.
2. Write $n - 1 = 2^k r$, where r is odd.
3. Pick a base a, let $z = a^r \mod n$.
4. If $z = 1 \mod n$ or $z = -1 \mod n$, stop: possibly prime.
5. We know $z \neq \pm 1 \mod n$, so square z up to $k - 1$ times
   - If you get 1, stop: not prime.
   - If you get -1, stop: possibly prime.
6. Now z has been squared $k - 1$ times but we have not got 1 or -1. stop, not prime.

If n is prime, it always says *possibly prime*.

Miller-Rabin Bases

If n is not prime and the algorithm starting at a says *probably prime*, then we call base a a strong liar. At most $(n-1)/4$ bases are strong liars.

If we select a randomly, there is at most a $1/4$ chance the algorithm accidently says *possibly prime*, so just keep retrying until you get a probability of error you are happy with.

Want to find a prime with 100-bit prime? Generate 98 random bits, put a 1 at the start and the end, try (say) 100 Miller-Rabin bases. If it fails, add two and start again.

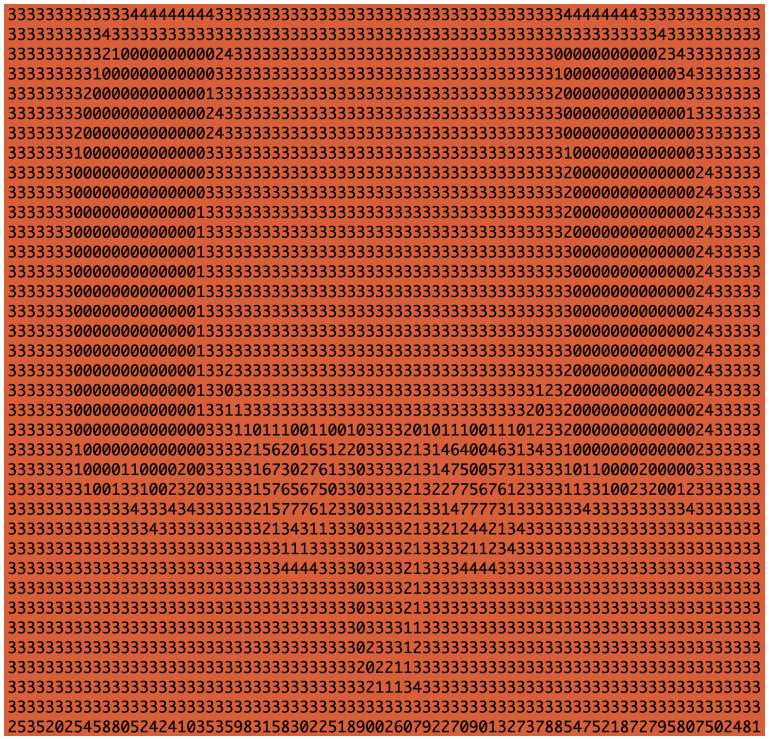
This is efficient and widely used.[7] Apply before using Pollard's Rho.

[7] https://github.com/openssl/openssl/blob/master/crypto/bn/bn_prime.c

Zaltzprime

The Bugle is a satirical podcast hosted by Andy Zaltzman. In his wishlist for scientific breakthroughs for 2024, he hoped for the discovery of a new number that he could use...

Zaltzprime



2960 digit (probably) prime number that had (probably) never been used before.

Zaltzprime

They liked it. . .

It is unquestionably the greatest artistic creation of this or indeed any other millennium.