EXPERIMENTAL MATHEMATICS SEMINAR · OCTOBER 2026

Five Zeilberger Prize Problems, Solved with Opus 5/Fable 5

An account of the human–AI methodology

Jaideep Sai Padhi · Purdue University

arXiv:2608.11290

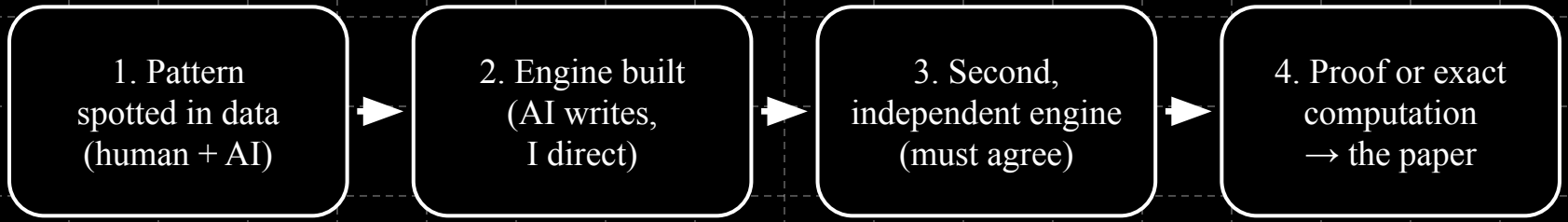
The Five Challenges

- Chomp: a bar with three winning opening moves (Ekhad–Zeilberger, Second Computational Challenge)
- Restricted permutations: holonomicity of $a_{r,\square}$ and $b_{r,\square}$ for all r, s (Spahn–Zeilberger Challenge 3); $a_{2,2}$ operator partial
- Solid SYT: the order-2 recurrence for shape [[n,n],[n,1]] (First Rigorous Challenge); cone exponent partial
- Kauers–Zeilberger Conjectures 2a and 2b: $G(n) \sim C_1 8^n/n^4$ and $H(n) \sim C_2(7+5\sqrt{2})^n/n^4$, proved
- The Geode: the 8,367-digit coefficient $G[1000^{(5)}]$ (Amdeberhan–Kauers–Zeilberger), plus D = 6, …, 10

The Division Of Labor

- Me: chose the problems and when to abandon them, ran every computation on a 256-core rented machine, designed and enforced the verification protocol, adjudicated between proposals, and caught the errors.
- The model: proposed the reductions, constructed the proofs, derived the closed forms, wrote every line of code, and drafted the manuscripts. (I took MVC at the time)
- The rule: nothing is built on until it is checked against an independently known quantity. Internal consistency is not validation.
- §3 of the paper states this division in full; §11, the error ledger, records what the checks caught.
- THIS IS A PATIENCE GAME! (at least for me…)

How A Result Is Produced



Any disagreement → back to step 1.

This loop is where both Chomp track laws died, and where C_1 was caught.

The Failure Mode

- The machine's typical error is to extrapolate a pattern past its evidence and state it fluently. It looks exactly like a result.
- Our example: a law for which thin-tail Chomp shapes (b^p, 1^q) are losing held for b = 2, …, 6. Then we computed b = 7:

$$q(7,3) = 12 \neq 2 \cdot 7 - 3, \qquad q(7,4) = 14 \neq 3 \cdot 7 - 5$$

Two independent laws, five confirmations each, both false. Defense: two engines that must agree, and nothing in the paper without a proof or an exact computation.

The Challenge: Chomp

- Chomp is played on an a × b bar of chocolate. A move picks a square and eats it with everything below and to its right.
- The top-left square is poisoned: whoever eats it loses.
- Gale (1974): the first player always wins on a full bar, but the proof never says which move wins, or how many do.
- Ken Thompson found the 8 × 10 bar has two winning first moves. Ekhad and Dr. Zeilberger then searched every bar up to 14 × 14.
- Their 2018 Second Computational Challenge ($100 pledged to the OEIS): find any bar with three or more winning first moves.

Asked for: one bar with three winning moves. Delivered: 10 × 42, plus a theorem on where winning moves can lie.

Case 1 — Chomp: Three Winning Moves

- Ekhad–Zeilberger (2018): find any a×b bar with three or more winning first moves. Gale asked in 1974 whether the winning move is always unique; Thompson's 8×10 showed it is not, with two.
- Found in the 10×45 box (about 2.9×10^{10} positions, under half an hour on 256 cores), re-derived in the 10×42 box, and countersigned by an independently written engine.

$$\mathcal{W}(10 \times 42) = \{(5, 36),\ (7, 30),\ (8, 26)\}$$

Chomp, In Formal Terms

Positions are order ideals of the a × b poset. The search space for the 10 × 42 bar:

$$\#\{\text{positions inside } a \times b\} = \binom{a+b}{a}, \qquad \binom{52}{10} = 15,820,024,220$$

Retrograde analysis labels every position, sweeping ideals in increasing rank:

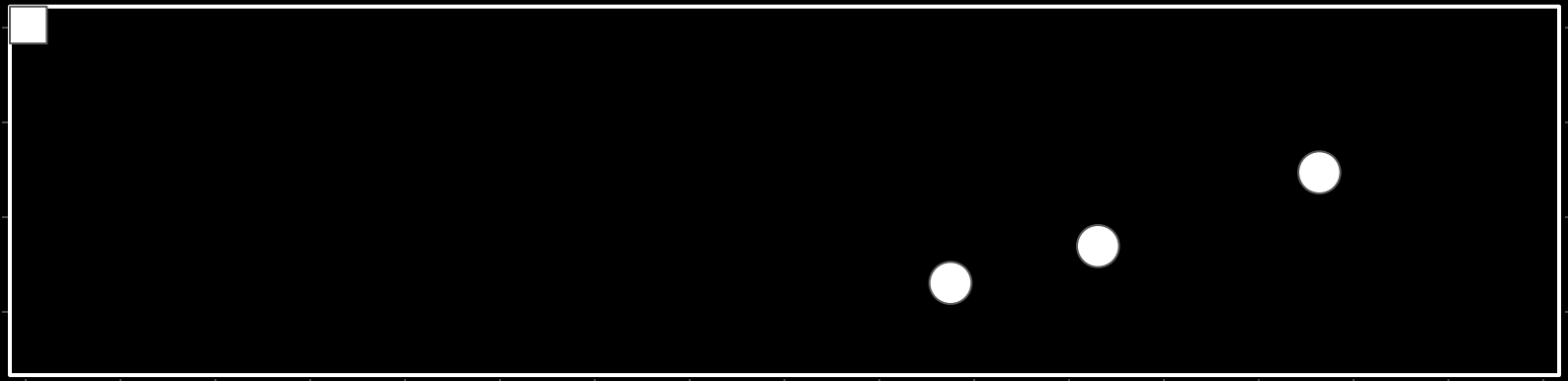
$$\lambda \in \mathcal{P} \iff \text{every move from } \lambda \text{ lands in } \mathcal{N}, \qquad \lambda \in \mathcal{N} \iff \text{some move lands in } \mathcal{P}$$

The full bar is always a first-player win; the question is how many winning first moves there are:

$$\text{Gale: } [a \times b] \in \mathcal{N} \text{ (strategy stealing)}, \qquad \mathcal{W}(a \times b) = \{(i,j) : [a \times b] \setminus \text{bite}(i,j) \in \mathcal{P}\}$$

One bit per state, each move evaluated in O(1) by a hockey-stick identity on the ranking, area-layered in parallel: |W(10 × 42)| = 3.

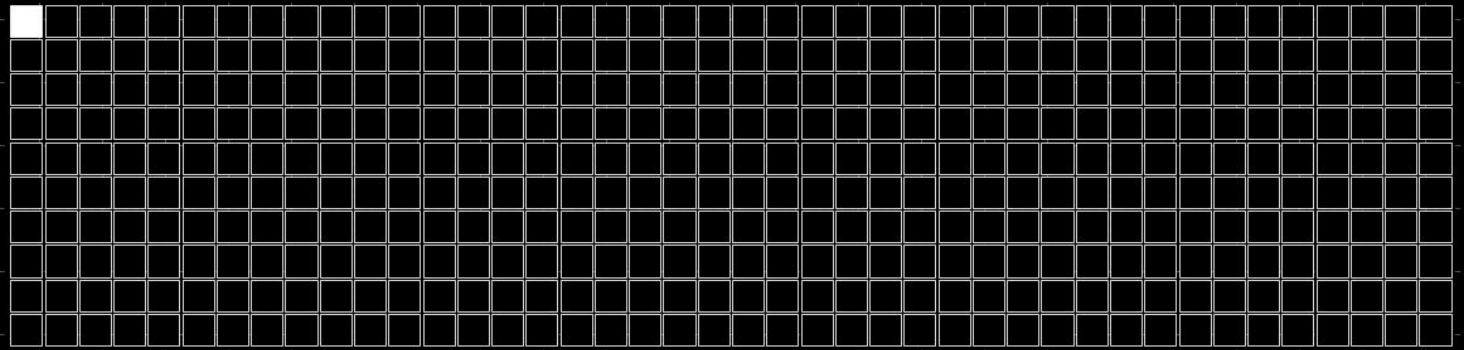
The 10 × 42 Bar: Three Ways To Win



- White square: the poisoned cell. White dots: the three winning bites (5,36), (7,30), (8,26).
- Thompson's 8 × 10 had two. Nobody had found three in eight years.
- Smallest in the searched region; whether it's smallest overall is open.

Watch The Three Winning Bites

Chomp on the 10 × 42 bar. The poisoned square is top-left.



Find a first move that leaves a losing position for the opponent.

- Each bite removes every square below and to its right and leaves the opponent a losing position. In 52 years of the game, this is the first bar found with three such moves.

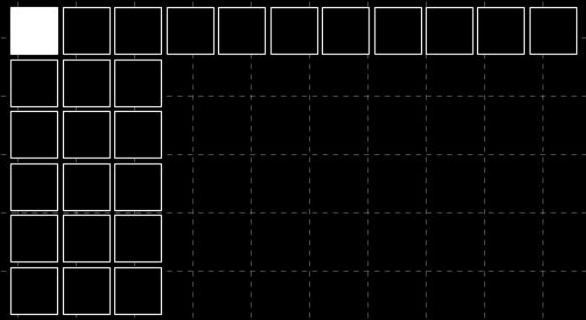
Chomp: What Is And Isn't Claimed

- 10×42 is the smallest triple in the searched region. Smaller-area bars such as 11×20 are unchecked, so minimality is open.
- The conjectured track laws q = 2b−3 and q = 3b−5 each held for five cases, then failed at b = 7.
- What survived is a theorem: the Staircase Theorem, a two-line proof that all 19 bars we found obey (below).

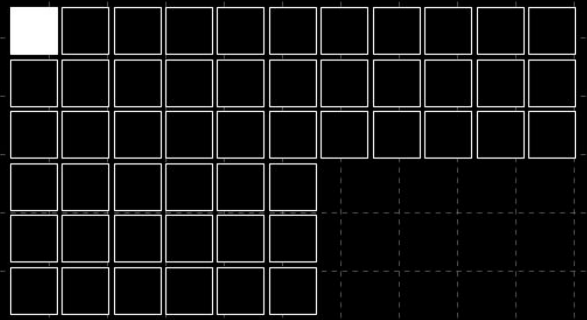
$$(i, j), (i', j') \in \mathcal{W},\ i < i' \implies j > j'$$

Why Winning Bites Must Form A Staircase

Suppose bites (2,4) and (4,7) both win, with $2 < 4$ and $4 \leq 7$



A: after bite (2,4)



B: after bite (4,7)

Both A and B would have to be P-positions (losing for the player to move).

- If one winning bite were weakly above-left of another, one move would join two P-positions, which is impossible. This is the whole proof.

The Challenge: Restricted Permutations

- $a_{r,□}(n)$ counts permutations of 1, …, n with $\pi(i+r) - \pi(i) \neq s$ for every i; $b_{r,□}(n)$ forbids $|\pi(i+r) - \pi(i)| = s$.
- The central question is holonomicity: does each counting sequence satisfy a linear recurrence with polynomial coefficients?
- Challenge 3 ($300): are $a_{r,□}$ and $b_{r,□}$ holonomic for all r, s > 1? Only the case (1, 1) was classical.
- Challenge 1: prove the order-8, degree-11 recurrence for $a_{2,2}$ that Kauers and Koutschan guessed from 35 terms.
- The catch: a decomposition whose number of auxiliary variables grows with n proves nothing.

Asked for: holonomicity in general, and one specific recurrence. Delivered: a proof for every r, s, and a verified reduction of $a_{2,2}$.

Case 2 — Restricted Permutations

- Theorem: for every fixed r, s ≥ 1, both $a_{r,□}$ and $b_{r,□}$ are holonomic.
- Interleave positions mod r and values mod s: a step of r becomes a step of 1, except at (r − 1) + (s − 1) seams, a number independent of n.
- Two-layer inclusion–exclusion turns the count into at most 2^{r-1} rational functions and an umbral integral; closure theorems of Lipshitz and Almkvist–Zeilberger finish it.
- $a_{2,2}$: reduction machine-verified end to end and two order-2 recurrences proved, but the full operator needs about 70,000 unknowns. Partial.

Holonomicity, Formally

The two families, for fixed r and s:

$$a_{r,\,s}(n) = \#\{\pi \in \mathfrak{S}_n : \ \pi_{i+r} - \pi_i \neq s\}, \qquad b_{r,\,s}(n) = \#\{\pi \in \mathfrak{S}_n : \ |\pi_{i+r} - \pi_i| \neq s\}$$

Interleaving bijection σ□ = v⁻¹(π_p(k)): the constraint becomes “no succession”, with a bounded set of excused positions and values:

$$\pi_{i+r} - \pi_i \neq s \iff \sigma_{k+1} - \sigma_k \neq 1 \ \ (k \notin \mathcal{P},\ \sigma_k \notin \mathcal{V}^+), \qquad |\mathcal{P}| = r - 1, \ \ |\mathcal{V}^+| = s - 1$$

Blocks of forced bonds carry a rational factor; arrangements of blocks are counted by an umbral integral:

$$\mathcal{B}(x) = \sum_{\ell \geq 1} (-1)^{\ell-1} x^{\ell} = \frac{x}{1+x}, \qquad \langle z_0^{j_0} \cdots z_q^{j_q} \rangle = \int_{\mathbb{R}_+^{q+1}} z_0^{j_0} \cdots z_q^{j_q}\, e^{-z_0 - \cdots - z_q}\, dz = j_0! \cdots j_q!$$

At most 2^{r-1} terms, at most 1 + r umbral and s − 1 catalytic variables, all independent of n. Coefficient extraction, the umbral integral, specialisation mod lcm(r, s) and interlacing each preserve holonomicity.

The Challenge: Solid Standard Young Tableaux

- A solid tableau fills a 3-D shape with 1, 2, …, N so that entries increase along all three directions.
- Unlike ordinary tableaux, there is no hook-length formula; the counts are usually wild.
- Zeilberger studied the two-layer shape [[n, n], [n, 1]] with 3n + 1 cells; its counts are g(n) = 2, 48, 1038, 22566, …
- From 40 terms he found, empirically, a recurrence of order 2 with polynomial coefficients of degree 12.
- First Rigorous Challenge: prove that recurrence.

Asked for: a proof. Delivered: an explicit closed form for g(n), from which the recurrence follows.

Case 3 — Solid Standard Young Tableaux

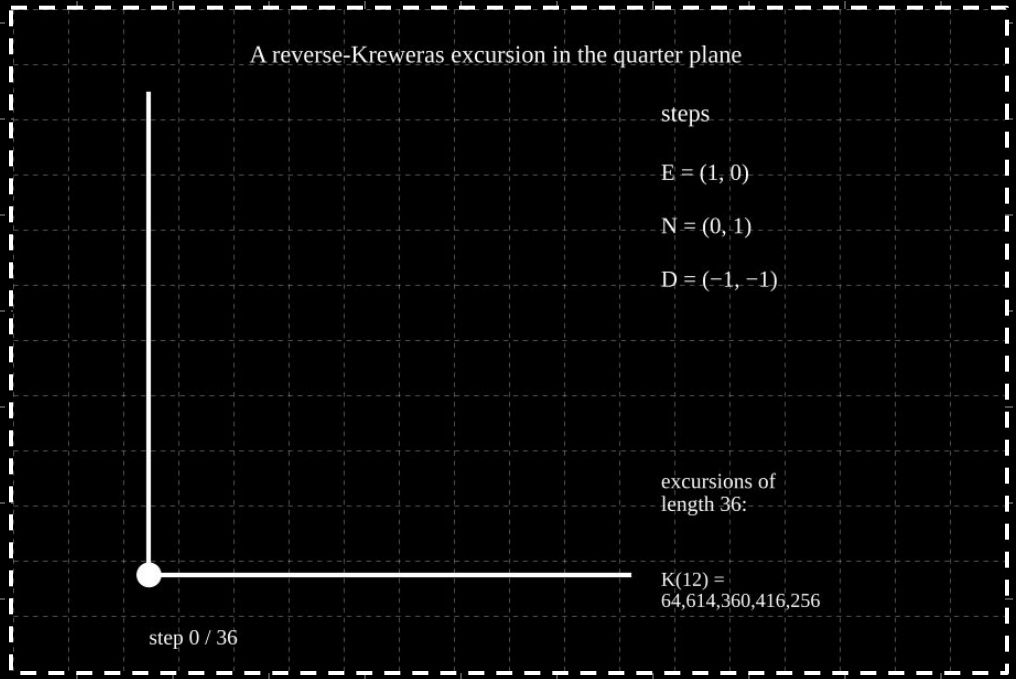
- First Rigorous Challenge: prove Dr. Zeilberger's empirical order-2, degree-12 recurrence for g(n), the number of solid SYT of shape [[n,n],[n,1]].
- Route: a deletion–insertion bijection to reverse-Kreweras walks, then the algebraic kernel method (Bousquet-Mélou, Mishna). The result is a closed form:

$$g(n) = c(n)\,(3n+1)\,K(n) + T(n), \qquad K(n) = \frac{4^n (3n)!}{(n+1)!\,(2n+1)!}$$

$$c(n) = \frac{8n^2 + n - 24}{4(n+2)(2n+3)}, \qquad T(n) = \frac{7n+5}{5} \prod_{k=1}^{n} \frac{6(2k-1)(6k-1)(6k+1)}{(k+1)(4k+3)(4k+5)}$$

Two hypergeometric terms ⇒ a rank-2 Q(n)-module ⇒ the order-2 recurrence follows by Cramer's rule.

Tableaux Become Walks



- Fill a solid tableau one cell at a time and record the differences $u = x_{11} - x_{12}$ and $v = x_{11} - x_{21}$.
- The three row types become steps $(1,1)$, $(-1,0)$, $(0,-1)$, and the shape constraints become the quarter plane. These are exactly Kreweras walks, counted by $K(n)$.
- The one extra cell becomes a marked time τ, and that is where the kernel method takes over.

From One Extra Cell To Two Walk Sums

Complementary counting over the 3n + 1 insertion times of the single cell (Lemma 2.2):

$$g(n) = (3n+1)K(n) - 2S(n) + T(n)$$

Prefix factorisation: the prefix before τ is forced (or is a ballot path), the rest is a reverse-Kreweras walk:

$$T(n) = \sum_{a=0}^{n} k(3n-a;\, a, a), \qquad S(n) = \sum_{0 \le c \le a} \mathrm{Bal}(a,c)\, k(3n-a-c;\, a,\, a-c)$$

The ballot-weighted double sum collapses to a single near-axis count (Theorem 1.3):

$$S(n) = k(3n+1;\, 1, 0) = \frac{3(9n+16)(3n+1)}{8(n+2)(2n+3)}\, K(n)$$

Open question from the paper: a bijective proof that the double sum S(n) counts walks of length 3n + 1 ending at (1, 0).

Inside The Kernel Method

Substituting the Catalan series kills the kernel, and only R(x) survives the constant-term extraction:

$$y^* = \frac{1 - \sqrt{1 - 4t^2 x}}{2t} \quad \Longrightarrow \quad K(x, y^*) = -t\, x^2 y^*$$

Diagonal endpoints admit a closed form, which appears to be new (Theorem 1.2):

$$k(3n-a;\, a, a) = K(n)\, \frac{(a+1)!\,(2a+1)!\,n!\,(3n-a)!}{4^a\,(a!)^3\,(n-a)!\,(3n)!}$$

Summing over a gives T(n), a hypergeometric term with ratio

$$\frac{T(m)}{T(m-1)} = \frac{6(2m-1)(6m-1)(6m+1)(7m+5)}{(m+1)(4m+3)(4m+5)(7m-2)}$$

(3n+1)K(n) and T(n) are two hypergeometric terms with different ratios, so g lies in a rank-2 module and Cramer's rule produces the order-2, degree-12 recurrence.

The Challenge: Tableaux With Restricted Runs

- G(n): standard Young tableaux of shape [n, n, n] in which every run in every row has length ≥ 2. H(n): the same with every run of odd length.
- 17 terms of G and 12 of H were published: G = 0, 1, 1, 5, 15, 69, 304, 1518, …
- Conjectures 2a, 2b (Kauers–Zeilberger 2020): $G(n) \sim C_1 \, 8^n / n^4$ and $H(n) \sim C_2 \, (7 + 5\sqrt{2})^n / n^4$.
- Their own estimates were deliberately rough: $C_1 \approx 0.521286$ and $C_2 \approx 0.63892$.
- The difficulty: these are walks in a cone whose steps are runs of random length, outside the reach of existing local limit theorems.

Asked for: proofs of both asymptotic laws. Delivered: both proved, with C_1 to 11 digits and C_2 to 10.

Case 4 — Kauers–Zeilberger Restricted Runs

- Proof: a new local limit theorem for excursions of run-modulated walks in cones, transporting Denisov–Wachtel–Zhang to Markov-additive processes with a corrector ladder.
- After isotropization the cone is the wedge of angle π/3, giving exponent −4; each constant is C = ϱK/(2ρ) · V(0)², the square of a harmonic function's apex value.

$$G(n) \sim C_1 \, \frac{8^n}{n^4}, \qquad -4 = -1 - \frac{\pi}{\pi/3}$$

$$C_1 = 0.52128605909(2), \qquad C_2 = 0.6389278129(4)$$

An Error Worth Showing

- No closed form is known, so the constants are numerical: an exact DP to n = 500 for G and n = 200 for H, then acceleration of A_N = a(N) N^4 μ^{-N} by Richardson's step:

$$A^{(k+1)}{}_N = \frac{(N+1)\,A^{(k)}{}_{N+1} - (N-k)\,A^{(k)}{}_N}{k+1}$$

An earlier run, to n = 110 at order 6, gave 0.5212860373: wrong from the eighth digit. The column was still rising, and kept rising for twenty more orders. Only agreement with Salzer's method, which converges under different hypotheses, makes the final value trustworthy:

Richardson 0.52128605909**04**, Salzer 0.52128605908**36**, tail fit 0.52128605909**14**

"Looked Converged" Is Not "Converged"

Order-6 Richardson
0.5212860373
stable for several orders, so we believed it



Order 28 + Salzer + tail fit
0.52128605909(2)
three methods agree to ~10^{-11}

The two values split at the 8th decimal. A monotone sequence that is under-extrapolated looks exactly like one that has converged.

The Challenge: The 5D Geode

- Hyper-Catalan series (Wildberger–Rubine): $S = 1 + \Sigma_{k \geq 2} t_k S^k$, in D variables t_2, …, t_{D+1}.
- The Geode G is what is left after dividing out one linear factor: $S - 1 = (t_2 + t_3 + \ldots + t_6) G$.
- Amdeberhan, Kauers and Zeilberger (2025) asked for one coefficient: G(1000, 1000, 1000, 1000, 1000).
- Rubine computed the four-variable value in 35.9 hours, estimated five variables at about 1,000 times the cost, and wrote that new innovations were required.
- The obstacle: written out directly, the coefficient is an alternating sum of $(M + 1)^{D-1} \approx 10^{12}$ terms.

Asked for: one exact integer. Delivered: all 8,367 digits, plus the 6- to 10-variable values.

Case 5 — The 5D Geode

- The direct sum has $1001^4 \approx 10^{12}$ terms. That is the wall Dr. Rubine hit at 4D (35.9 hours) and said 5D needed new ideas to pass.
- Three verified lemmas (constant row-sum, a binomial product, a Beta integral) collapse it to one integral plus one coefficient extraction. That costs O(M²) time and O(M) memory per prime.

$$G[M^D] = \frac{(DM)!}{(M!)^D}\,[z^{DM}]\,(1+z)^E \int_0^1 h(x,z)^M\,dx$$

$$h(x,z) = (1-x)\prod_{k=1}^{D-1}\left((1+z)^k - x\right)$$

$G(1000^5)$ has 8,367 digits. The 4D check reproduced Rubine's value in 168 s, and D = 6…10 were computed as a bonus.

The Three Lemmas Behind The Geode Collapse

Start from the hyper-Catalan numbers, the coefficients of S = 1 + Σ t□ S^k:

$$C[\mathbf{m}] = \frac{\left(\sum_k k\,m_k\right)!}{\left(1+\sum_k (k-1)\,m_k\right)!\,\prod_k m_k!}$$

The 10^{12}-term table of weights is not a table at all; it is a product (binomial theorem):

$$\sum_{i,W} T[i,W]\,u^i y^W = \prod_{k=1}^{d} \left(1+u\,y^k\right)^M$$

And the remaining factorial ratio is a Beta integral, so the sum over i becomes evaluation at u = −x:

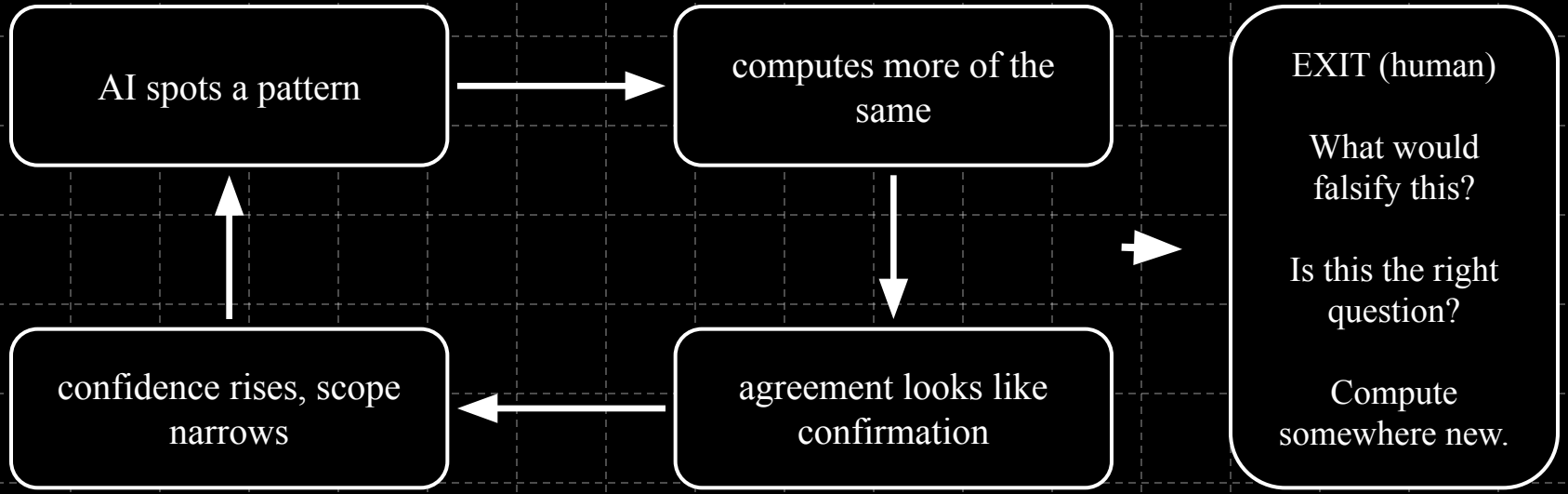
$$\frac{i!}{(M+1+i)!} = \frac{1}{M!}\int_0^1 x^i (1-x)^M\,dx$$

Exact quadrature at DM + 1 nodes, an order-10 recurrence per node, 2.64 s per 30-bit prime, about a thousand primes and CRT, with a proven bound fixing how many.

Where The AI Went In Circles

- $a_{2,2}$: reformulation after reformulation, each announced as a breakthrough, each re-hitting the same obstruction in new notation.
- That loop ended only when a pole-order count produced a number: about 70,000 unknowns, four orders of magnitude out of reach.
- KZ: an order-6 Richardson value, stable for several orders, was accepted as converged. It was wrong in the eighth digit.
- Holonomicity was filed as “research-grade” for the whole campaign. The computation that settled it took seconds.
- Cost: it proposed computations without estimating runtime. A construction needing six coefficients was written to compute twelve, and never finished.

The Loop, And The Way Out



The model never left this loop on its own. Every escape in the paper started with a human question that changed what was being computed, not how much.

When I Made The Call

- "Those odds look high." The recount gave $E \approx 0.2$, not 0.5.
- "Do something structural." The box lottery became the track decomposition and the Staircase Theorem.
- "Audit before we claim." It found Sheiner (2026) and Garg (2026) on Chomp, and the note was reframed around Gale's 1974 question.
- "Go past the target." Geode was extended from 5D to every $D \leq 10$.
- "Do it ourselves." For KZ we took the self-reliant route instead of outsourcing the probabilistic step.
- "Referee this." Nothing went to Dr. Zeilberger until the note was written, exhaustively refereed, and the code was public.

What It Is Genuinely Superb At

- Writing a GPU retrograde solver, a CUDA kernel and a CRT driver in an hour, then debugging them at 2 a.m. without complaint.
- Turning a vague idea into a machine-certifiable derivation.
- Typesetting, refactoring, re-running, and re-running again.
- Never getting bored of the fortieth variant of the same computation.
- Being a tireless second pair of hands for someone who lacks the specialty needed to solve such problems.

What I'd Tell Someone Trying This

- Speed came from the machine. Correctness came from structure we imposed on it.
- Two engines that must agree catch bugs. One engine that agrees with you does not.
- Five confirming cases is a hypothesis, not a law. We were burned twice in one problem.
- The decisive moves (go structural, audit before claiming, stop extrapolating) were human. The model would happily have kept going.
- Keep an error ledger and publish it. It is the most useful section of the paper.

Part 2

(a glitch in the matrix)

Thank You

- Paper: arXiv:2608.11290
- Code and data: github.com/jaideepsaipadhi/zeilberger-challenges
- Thanks to Doron Zeilberger, Manuel Kauers, Tewodros Amdeberhan, and Shalosh B. Ekhad, whose tables calibrated the Chomp work
- Questions?