Solving One Third of the OEIS from Scratch

Josef Urban
joint work with Thibault Gauthier and Mirek Olsak

Czech Technical University in Prague

AGI'24
August 14, 2024, Seattle

European Research Council
Established by the European Commission

TL;DR

- A machine can find explanations for over 125k OEIS sequences
- This is done *from scratch*, without any domain knowledge
- N. Sloane: The OEIS: *A Fingerprint File for Mathematics* (2021)
- About 350k integer sequences in 2021 from all parts of math
- We use a simple Search-Verify-Train positive feedback loop
- Developed by us for combining learning & proving since 2006
- However, one of the most surprising experiments in my life:
- 670 iterations and still refuses to plateau - counters RL wisdom
- Since it interleaves symbolic breakthroughs and statistical learning?
- The electricity bill is only $1k-$3k, you can do this at home
- Btw., we experimentally verify Occam's Razor
- Evolving (self-improving) population of 4.5M matching explanations
- Connections to Solomonoff Induction, AIXI, Gödel Machine?

Brief Intro to What I Normally Do: ML+ATP for Math

- What is mathematical and scientific thinking?
- Pattern-matching, analogy, induction/learning from examples
- Deductive reasoning and proving
- Complicated feedback loops between learning and deduction
- Using a lot of previous knowledge - both for induction and deduction

- We need to develop such methods on computers
- Are there any large corpora suitable for nontrivial deduction?
- Yes! Large libraries of formal proofs and theories
- So let's try to develop strong AI on them!

- In my case done for 25-30 years. Recent overviews:
- *Learning Guided Automated Reasoning: A Brief Survey. 2024*
- Zar Goertzel's Phd thesis: *Learning Inference Guidance in ATP. 2023*
- *AI4REASON*: `http://ai4reason.org/PR_CORE_SCIENTIFIC_4.pdf`

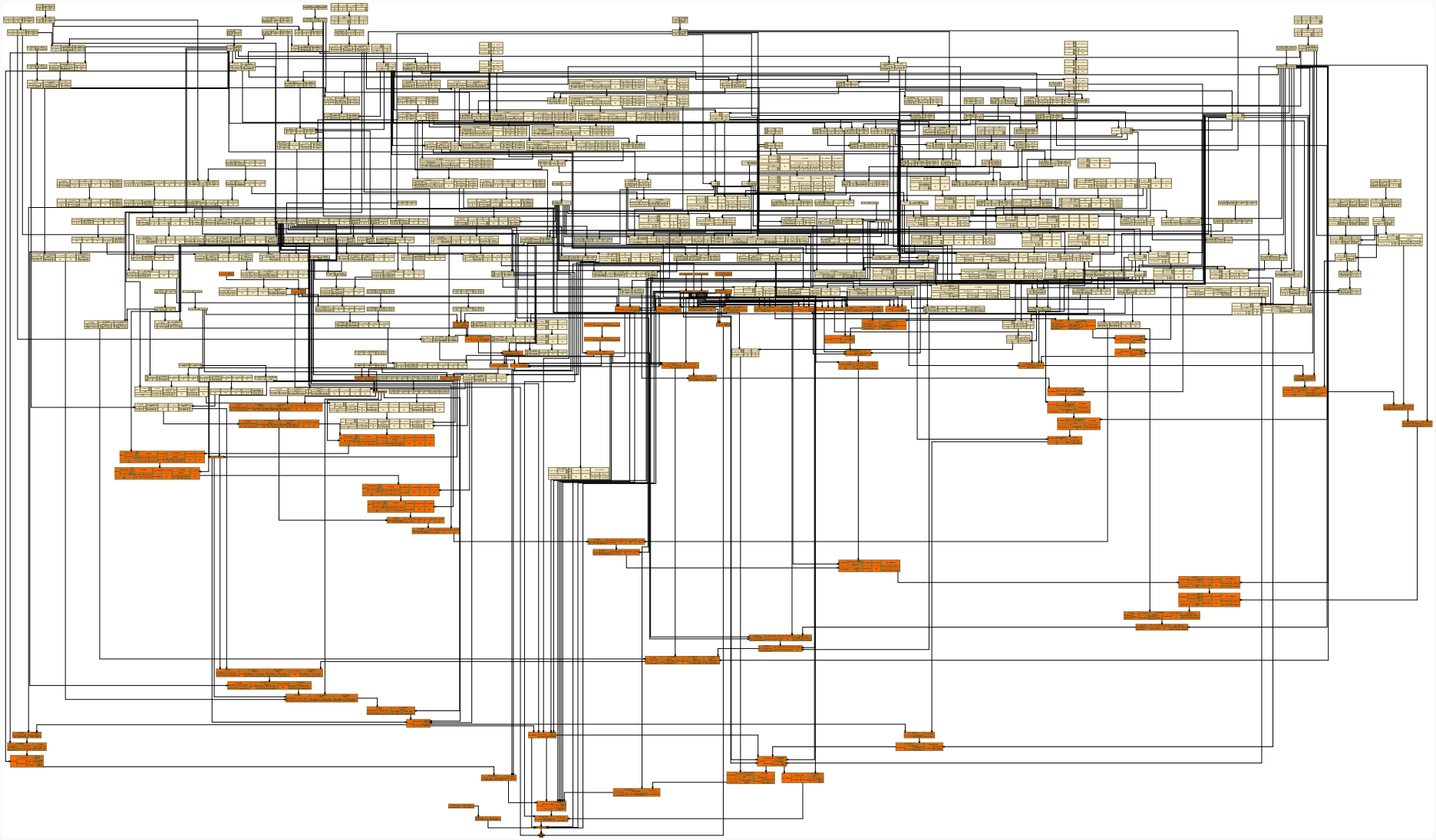
ENIGMA: Feedback prove/learn loop on formal math

- Done on 57880 Mizar Mathematical Library problems recently
- Efficient ML-guidance inside the best ATPs (E prover and more)
- Training of the ML-guidance is *interleaved* with proving harder problems
- Ultimately a 70% improvement over the original strategy:
- ... from 14933 proofs to 25397 proofs (all 10s CPU - no cheating)
- 75% of the Mizar corpus reached in July 2021 - higher times and many runs: `https://github.com/ai4reason/ATP_Proofs`
- Details in our Mizar60 paper: `https://arxiv.org/abs/2303.06686`

	$\mathcal{S}$	$\mathcal{S} \odot \mathcal{M}_9^0$	$\mathcal{S} \oplus \mathcal{M}_9^0$	$\mathcal{S} \odot \mathcal{M}_9^1$	$\mathcal{S} \oplus \mathcal{M}_9^1$	$\mathcal{S} \odot \mathcal{M}_9^2$	$\mathcal{S} \oplus \mathcal{M}_9^2$	$\mathcal{S} \odot \mathcal{M}_9^3$	$\mathcal{S} \oplus \mathcal{M}_9^3$
solved	**14933**	16574	20366	21564	22839	22413	23467	22910	23753
$\mathcal{S}$%	+0%	+10.5%	+35.8%	+43.8%	+52.3%	+49.4%	+56.5%	+52.8%	+58.4
$\mathcal{S}+$	+0	+4364	+6215	+7774	+8414	+8407	+8964	+8822	+9274
$\mathcal{S}-$	-0	-2723	-782	-1143	-508	-927	-430	-845	-454

	$\mathcal{S} \odot \mathcal{M}_{12}^3$	$\mathcal{S} \oplus \mathcal{M}_{12}^3$	$\mathcal{S} \odot \mathcal{M}_{16}^3$	$\mathcal{S} \oplus \mathcal{M}_{16}^3$
solved	24159	24701	25100	**25397**
$\mathcal{S}$%	+61.1%	+64.8%	+68.0%	**+70.0%**
$\mathcal{S}+$	+9761	+10063	+10476	+10647
$\mathcal{S}-$	-535	-295	-309	-183

Can you do this in 4 minutes?



Can you do this in 4 minutes?

```
theorem Th31: :: BORSUK_5:31
 for A being Subset of R^1
 for a, b being real number st a < b & A = RAT (a,b) holds
 Cl A = [.a,b.]
proof
 let A be Subset of R^1; :: thesis:
 let a, b be real number ; :: thesis:
 assume that
 A1: a < b and
 A2: A = RAT (a,b) ; :: thesis:
 reconsider ab = ].a,b.[, RT = RAT as Subset of R^1 by NUMBERS:12, TOPMETR:17;
 reconsider RR = RAT /\ ].a,b.[ as Subset of R^1 by TOPMETR:17;
 A3: the carrier of R^1 /\ (Cl ab) = Cl ab by XBOOLE_1:28;
 A4: Cl RR c= (Cl RT) /\ (Cl ab) by PRE_TOPC:21;
 thus Cl A c= [.a,b.] :: according to XBOOLE_0:def 10 :: thesis:
 proof
  let x be set ; :: according to TARSKI:def 3 :: thesis:
  assume x in Cl A ; :: thesis:
  then x in (Cl RT) /\ (Cl ab) by A2, A4;
  then x in the carrier of R^1 /\ (Cl ab) by Th15;
  hence x in [.a,b.] by A1, A3, Th16; :: thesis:
 end;
 thus [.a,b.] c= Cl A :: thesis:
 proof
  let x be set ; :: according to TARSKI:def 3 :: thesis:
  assume A5: x in [.a,b.] ; :: thesis:
  then reconsider p = x as Element of RealSpace by METRIC_1:def 13;
  A6: a <= p by A5, XXREAL_1:1;
  A7: p <= b by A5, XXREAL_1:1;
  per cases by A7, XXREAL_0:1;
   suppose A8: p < b ; :: thesis:
    now :: thesis:
     let r be real number ; :: thesis:
     reconsider pp = p + r as Element of RealSpace by METRIC_1:def 13, XREAL_0:def 1;
     set pr = min (pp,((p + b) / 2));
     A9: min (pp,((p + b) / 2)) <= (p + b) / 2 by XXREAL_0:17;
     assume A10: r > 0 ; :: thesis:
     p < min (pp,((p + b) / 2))
     proof
      per cases by XXREAL_0:15;
       suppose min (pp,((p + b) / 2)) = pp ; :: thesis:
        hence p < min (pp,((p + b) / 2)) by A10, XREAL_1:29; :: thesis:
       end;
       suppose min (pp,((p + b) / 2)) = (p + b) / 2 ; :: thesis:
        hence p < min (pp,((p + b) / 2)) by A8, XREAL_1:226; :: thesis:
       end;
      end;
     end;
     then consider Q being rational number such that
     A11: p < Q and
     A12: Q < min (pp,((p + b) / 2)) by RAT_1:7;
     (p + b) / 2 < b by A8, XREAL_1:226;
     then min (pp,((p + b) / 2)) < b by A9, XXREAL_0:2;
     then A13: Q < b by A12, XXREAL_0:2;
     min (pp,((p + b) / 2)) <= pp by XXREAL_0:17;
     then A14: (min (pp,((p + b) / 2))) - p <= pp - p by XREAL_1:9;
     reconsider P = Q as Element of RealSpace by METRIC_1:def 13, XREAL_0:def 1;
     P - p < (min (pp,((p + b) / 2))) - p by A12, XREAL_1:9;
     then P - p < r by A14, XXREAL_0:2;
     then dist (p,P) < r by A11, Th14;
     then A15: P in Ball (p,r) by METRIC_1:11;
     a < Q by A6, A11, XXREAL_0:2;
     then A16: Q in ].a,b.[ by A13, XXREAL_1:4;
     Q in RAT by RAT_1:def 2;
     then Q in A by A2, A16, XBOOLE_0:def 4;
     hence Ball (p,r) meets A by A15, XBOOLE_0:3; :: thesis:
    end;
    hence x in Cl A by GOBOARD6:92, TOPMETR:def 6; :: thesis:
```

Can you do this in 4 minutes?

README.md

Topology - the closure of rationals on (a,b) is [a,b]

359-long proof in 234s using 3-phase ENIGMA, shifting context and aggressive subsumption.

for A being Subset of R^1 for a, b being real number st a < b & A = RAT (a,b) holds Cl A = [.a,b.]

The Mizar proof takes 80 lines:

http://grid01.ciirc.cvut.cz/~mptp/7.13.01_4.181.1147/html/borsuk_5.html#T31

E proof (3-phase parental+lgb+gnn-server plus shifting context plus aggr subsumption) using 38 of the 101 heuristically selected premises (subproblem minimization):
http://grid01.ciirc.cvut.cz/~mptp/enigma_prf/t31_borsuk_5

/local1/mptp/parents/out2/2pb3l8-query1024-ctx1536-w0-coop-srv-local1-f1711-jj1-zar-parents_nothr_gnnm2_solo1_0.05_0.005_0.1_fw.minsub65all_240s_fw/t31_borsuk_5

```
# Proof object clause steps            : 359
# Proof object initial clauses used    : 56
# Proof object initial formulas used   : 38
# Proof object simplifying inferences  : 180
# Parsed axioms                        : 101
# Initial clauses in saturation        : 153
# Processed clauses                    : 7274
# ...remaining for further processing  : 4883
# Generated clauses                    : 438702
# ...frozen by parental guidance       : 133869
# ...aggressively subsumed             : 83871
# User time               : 234.274 s
```

What Are the Current AI/TP TODOs/Bottlenecks?

- High-level structuring of proofs - proposing good lemmas
- Proposing new concepts, definitions and theories
- Proposing new targeted algorithms, decision procedures, tactics
- Proposing good witnesses for existential proofs
- All these problems involve synthesis of some mathematical objects
- Btw., constructing proofs is also a synthesis task
- This talk: explore learning-guided synthesis for OEIS
- Interesting research topic and tradeoff in learning/AI/proving:
- Learning direct guessing of objects (this talk) vs guidance for search procedures (ENIGMA and friends)
- Start looking also at semantics rather than just syntax of the objects

Quotes: Learning vs. Reasoning vs. Guessing

"C'est par la logique qu'on démontre, c'est par l'intuition qu'on invente."
(It is by logic that we prove, but by intuition that we discover.)
– Henri Poincaré, Mathematical Definitions and Education.

"Hypothesen sind Netze; nur der fängt, wer auswirft."
(Hypotheses are nets: only he who casts will catch.)
– Novalis, quoted by Popper – The Logic of Scientific Discovery

Certainly, let us learn proving, but also let us learn guessing.
– G. Polya - Mathematics and Plausible Reasoning

Galileo once said, "Mathematics is the language of Science." Hence, facing the same laws of the physical world, alien mathematics must have a good deal of similarity to ours.
– R. Hamming - Mathematics on a Distant Planet

QSynt: Semantics-Aware Synthesis of Math Objects

- Gauthier (et al) 2019-24
- Synthesize math expressions based on semantic characterizations
- i.e., not just on the syntactic descriptions (e.g. proof situations)
- Tree Neural Nets and Monte Carlo Tree Search (a la AlphaZero)
- Recently also various (small) *language models* with their search methods
- Invent programs for OEIS sequences FROM SCRATCH (no LLM cheats)
- 126k OEIS sequences (out of 350k) solved so far (670 iterations): `https://www.youtube.com/watch?v=24oejR9wsXs`, `http://grid01.ciirc.cvut.cz/~thibault/qsynt.html`
- ~4.5M explanations invented: 50+ different characterizations of primes
- Non-neural (Turing complete) symbolic computing and semantics collaborate with the statistical/neural learning
- Program evolution governed by high-level criteria (Occam, efficiency)

OEIS: $\geqslant$ 350000 finite sequences

The OEIS is supported by the many generous donors to the OEIS Foundation.

THE ON–LINE ENCYCLOPEDIA OF INTEGER SEQUENCES ®

founded in 1964 by N. J. A. Sloane

2 3 5 7 11 Search Hints

(Greetings from The On-Line Encyclopedia of Integer Sequences!)

Search: **seq:2,3,5,7,11**

Displaying 1-10 of 1163 results found. page 1 2 3 4 5 6 7 8 9 10 ... 117

Sort: relevance | references | number | modified | created Format: long | short | data

A000040 The prime numbers. +30 10150
(Formerly M0652 N0241)

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, 101, 103, 107, 109, 113, 127, 131, 137, 139, 149, 151, 157, 163, 167, 173, 179, 181, 191, 193, 197, 199, 211, 223, 227, 229, 233, 239, 241, 251, 257, 263, 269, 271 (list; graph; refs; listen; history; text; internal format)

OFFSET 1,1

COMMENTS See A065091 for comments, formulas etc. concerning only odd primes. For all information concerning prime powers, see A000961. For contributions concerning "almost primes" see A002808.
A number p is prime if (and only if) it is greater than 1 and has no positive divisors except 1 and p.
A natural number is prime if and only if it has exactly two (positive) divisors.
A prime has exactly one proper positive divisor, 1.

Generating programs for OEIS sequences

$0, 1, 3, 6, 10, 15, 21, \ldots$

An undesirable large program:

```
if x = 0 then 0 else
if x = 1 then 1 else
if x = 2 then 3 else
if x = 3 then 6 else ...
```

Small program (Occam's Razor):

$$\sum_{i=1}^{n} i$$

Fast program (efficiency criteria):

$$\frac{n \times n + n}{2}$$

Programming language

- Constants: $0, 1, 2$
- Variables: x, y
- Arithmetic: $+, -, \times, \mathit{div}, \mathit{mod}$
- Condition : if $\ldots \leqslant 0$ then $\ldots$ else $\ldots$
- $\mathit{loop}(f, a, b) := u_a$ where $u_0 = b$,

$$u_n = f(u_{n-1}, n)$$

- Two other loop constructs: *loop2*, a while loop

Example:

$$2^{\mathbf{x}} = \prod_{y=1}^{x} 2 = \mathit{loop}(2 \times x, \mathbf{x}, 1)$$
$$\mathbf{x}! = \prod_{y=1}^{x} y = \mathit{loop}(y \times x, \mathbf{x}, 1)$$

QSynt: synthesizing the programs/expressions

- Inductively defined set P of our *programs and subprograms*,
- and an auxiliary set F of binary functions (higher-order arguments)
- are the smallest sets such that $0, 1, 2, x, y \in P$, and if $a, b, c \in P$ and $f, g \in F$ then:

$$a + b, a - b, a \times b, a \; div \; b, a \; mod \; b, cond(a, b, c) \in P$$

$$\lambda(x, y).a \in F, \; loop(f, a, b), loop2(f, g, a, b, c), compr(f, a) \in P$$

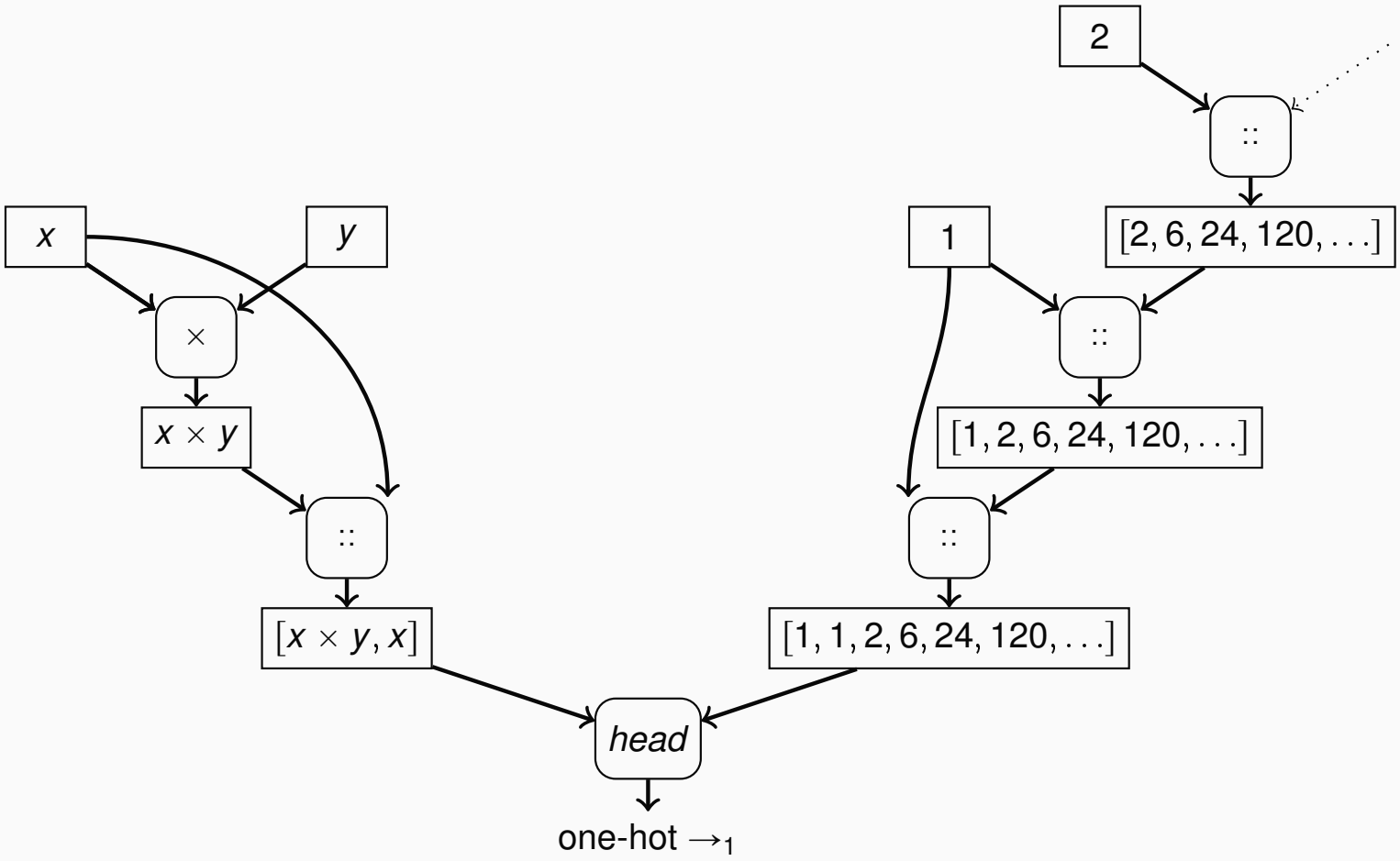
- Programs are built in reverse polish notation
- Start from an empty stack
- Use ML to repeatedly choose the next operator to push on top of a stack
- Example: Factorial is $loop(\lambda(x, y).\ x \times y, x, 1)$, built by:

$$[\,] \rightarrow_x [x] \rightarrow_y [x, y] \rightarrow_\times [x \times y] \rightarrow_x [x \times y, x]$$

$$\rightarrow_1 [x \times y, x, 1] \rightarrow_{loop} [loop(\lambda(x, y).\ x \times y, x, 1)]$$

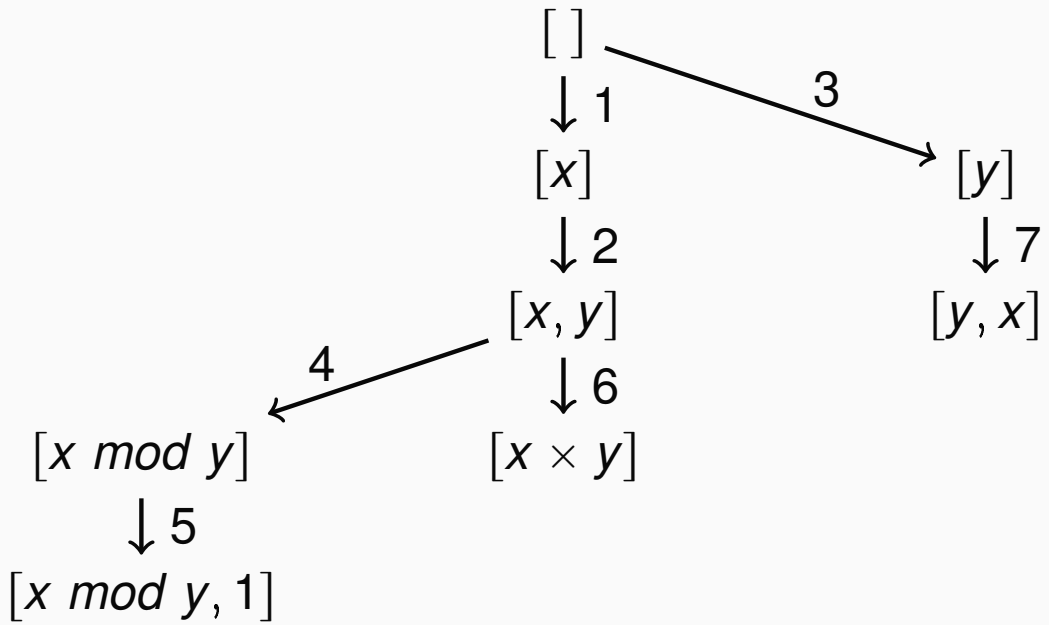
QSynt: Training of the Neural Net Guiding the Search

- The triple $((\mathit{head}([x \times y, x], [1, 1, 2, 6, 24, 120 \ldots]), \rightarrow_1)$ is a training example extracted from the program for factorial $\mathit{loop}(\lambda(x, y).\ x \times y, x, 1)$
- $\rightarrow_1$ is the action (adding 1 to the stack) required on $[x \times y, x]$ to progress towards the construction of $\mathit{loop}(\lambda(x, y).\ x \times y, x, 1)$.



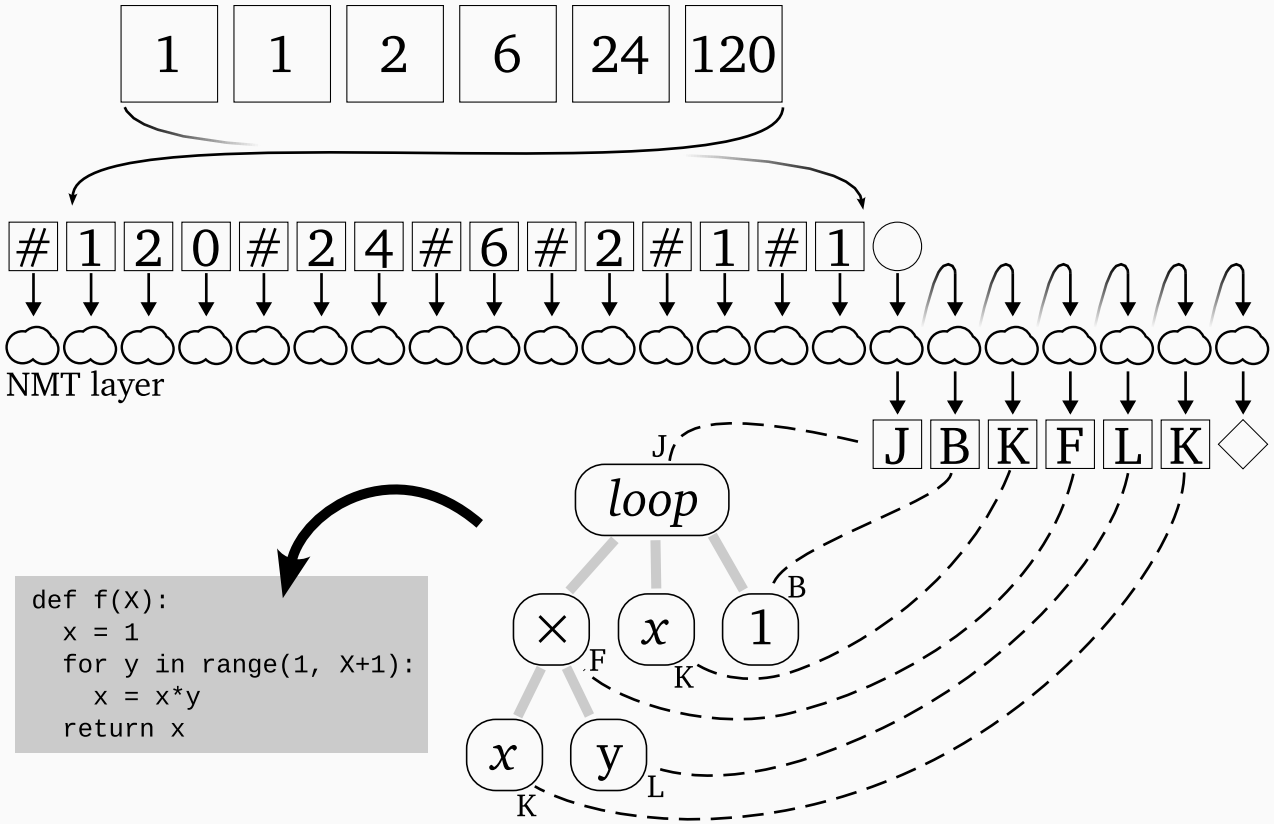
QSynt program search - Monte Carlo search tree

7 iterations of the tree search gradually extending the search tree. The set of the synthesized programs after the 7th iteration is $\{1, x, y, x \times y, x \ mod \ y\}$.

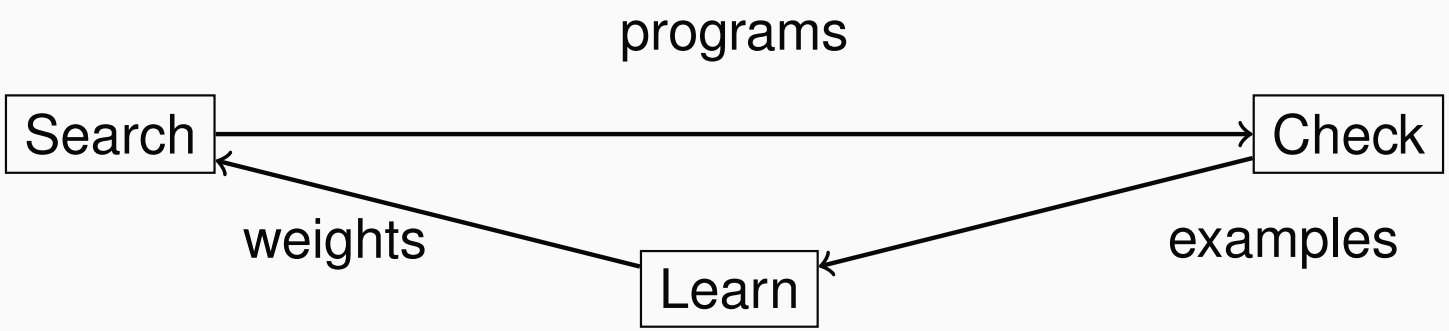


Encoding OEIS for Language Models

- Input sequence is a series of digits
- Separated by an additional token # at the integer boundaries
- Output program is a sequence of tokens in Polish notation
- Parsed by us to a syntax tree and translatable to Python
- Example: $a(n) = n!$



Search-Verify-Train Positive Feedback Loop



- Analogous to our Prove/Learn feedback loops in learning-guided proving (since 2006 – Machine Learner for Automated Reasoning – MaLARea))
- However, the OEIS setting allows much faster feedback on *symbolic conjecturing*

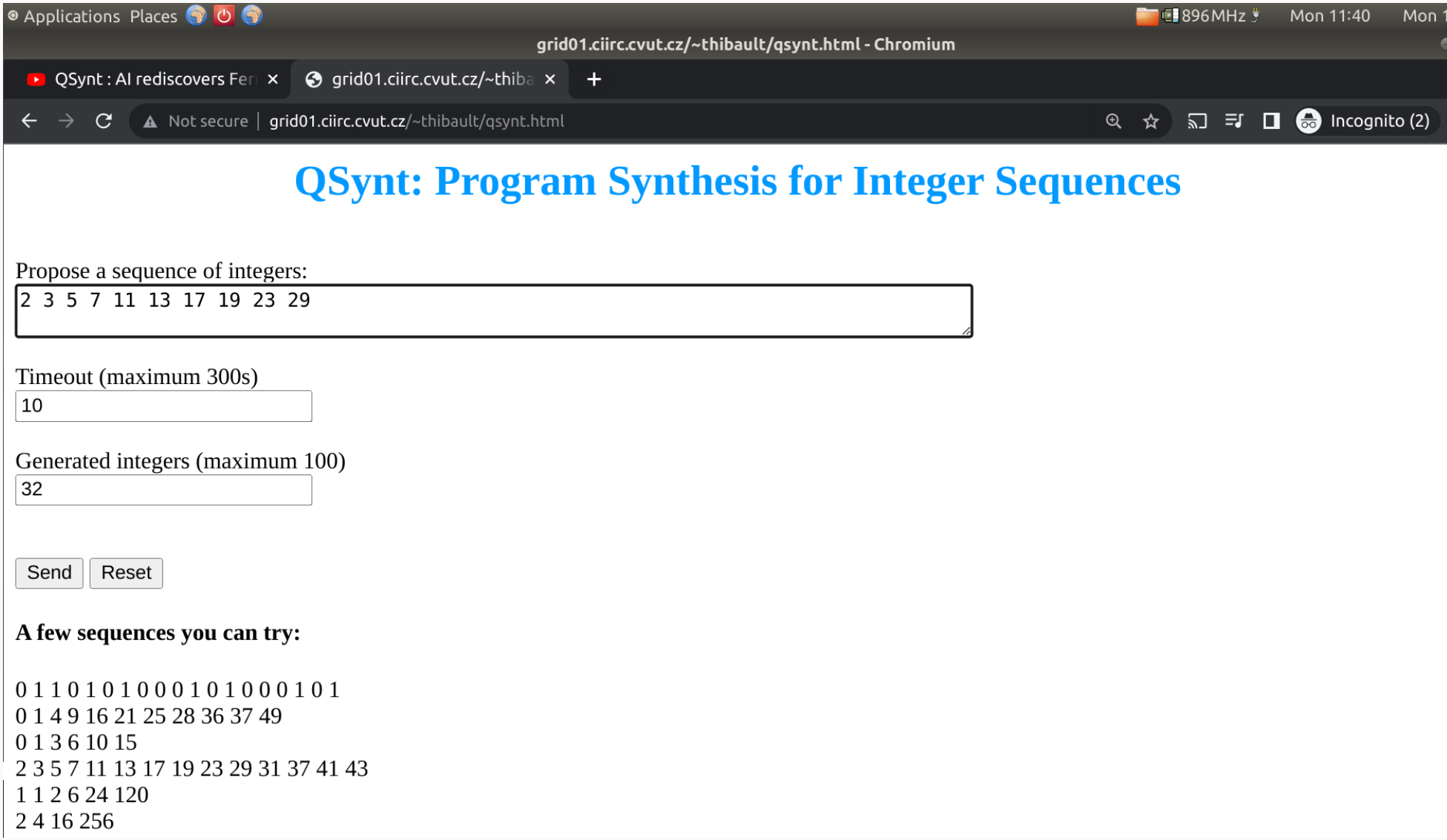
Search-Verify-Train Feedback Loop for OEIS

- **search phase**: LM synthesizes many programs for input sequences
- typically 240 candidate programs for each input using beam search
- 84M programs for OEIS in several hours on the GPU (depends on model)
- **checking phase**: the millions of programs efficiently evaluated
- resource limits used, fast indexing structures for OEIS sequences
- check if the program generates *any* OEIS sequence (hindsight replay)
- we keep the shortest (Occams's razor) and fastest program (efficiency)
- from iter. 501, we also keep the program with the best speed/length ratio
- **learning phase**: LM trains to translate the "solved" OEIS sequences into the best program(s) generating them
- from iter. 336: train LMs to transform (generalization, optimization)
- our learned version of human-coded methods like ILP and compilation

Search-Verify-Train Feedback Loop

- The weights of the LM either trained from scratch or continuously updated
- This yields *new minds* vs *seasoned experts* (who have seen it all)
- We also train experts on varied selections of data, in varied ways
- Orthogonality: common in theorem proving - different experts help
- Each iteration of the self-learning loop discovers more solutions
- ... also improves/optimizes existing solutions
- The alien mathematician thus self-evolves
- Occam's razor and efficiency are used for its weak supervision
- Quite different from today's LLM approaches:
- LLMs do one-time training on everything human-invented
- Our alien instead starts from zero knowledge
- Evolves increasingly nontrivial skills, may diverge from humans
- Turing complete (unlike Go/Chess) – arbitrary complex algorithms

QSynt web interface for program invention



QSynt inventing Fermat pseudoprimes

Positive integers k such that $2^k \equiv 2 \mod k$. ($341 = 11 * 31$ is the first non-prime)

```
First 16 generated numbers (f(0),f(1),f(2),...):
2 3 5 7 11 13 17 19 23 29 31 37 41 43 47 53
Generated sequence matches best with: A15919(1-75), A100726(0-59), A40(0-58)

Program found in 5.81 seconds
f(x) := 2 + compr(\x.loop(\(x,i).2*x + 2, x, 2) mod (x + 2), x)
Run the equivalent Python program here or in the window below:
```

BRYTHON

Tutorial Demo Documentation Console Editor Gallery Resources

English

Brython version: 3.10.6

run Python Javascript Share code

```
def f2(X):
  x = 2
  for i in range (1,X + 1):
    x = 2*x + 2
  return x

def f1(X):
  x,i = 0,0
  while i <= X:
    if f2(x) % (x + 2) <= 0:
      i = i + 1
    x = x + 1
  return x - 1

def f0(X):
  return 2 + f1(X)

for x in range(32):
  print (f0(x))
```

```
2
3
5
7
11
13
17
19
23
29
31
37
41
43
47
53
59
61
67
```

Lucas/Fibonacci characterization of (pseudo)primes

```
input sequence: 2,3,5,7,11,13,17,19,23,29

invented output program:
f(x) := compr(\(x,y).(loop2(\(x,y).x + y, \(x,y).x, x, 1, 2) - 1)
              mod (1 + x), x + 1) + 1

human conjecture: x is prime iff? x divides (Lucas(x) - 1)

PARI program:
? lucas(n) = fibonacci(n+1)+fibonacci(n-1)
? b(n) = (lucas(n) - 1) % n

Counterexamples (Bruckman-Lucas pseudoprimes):
? for(n=1,4000,if(b(n)==0,if(isprime(n),0,print(n))))
1
705
2465
2737
3745
```

QSynt inventing primes using Wilson's theorem

n is prime iff $(n-1)! + 1$ is divisible by n (i.e.: $(n-1)! \equiv -1 \mod n$)

```
First 32 generated numbers (f(0),f(1),f(2),...):
0 1 1 0 1 0 1 0 0 0 1 0 1 0 0 0 1 0 1 0 0 0 1 0 0 0 0 0 1 0 1 0
Generated sequence matches best with: A10051(0-100), A252233(0-29), A283991(0-24)

Program found in 5.17 seconds
f(x) := (loop(\(x,i).x * i, x, x) mod (x + 1)) mod 2
Run the equivalent Python program here or in the window below:
```

BRYTHON

Tutorial Demo Documentation Console Editor Gallery Resources

English

Brython version: 3.10.6

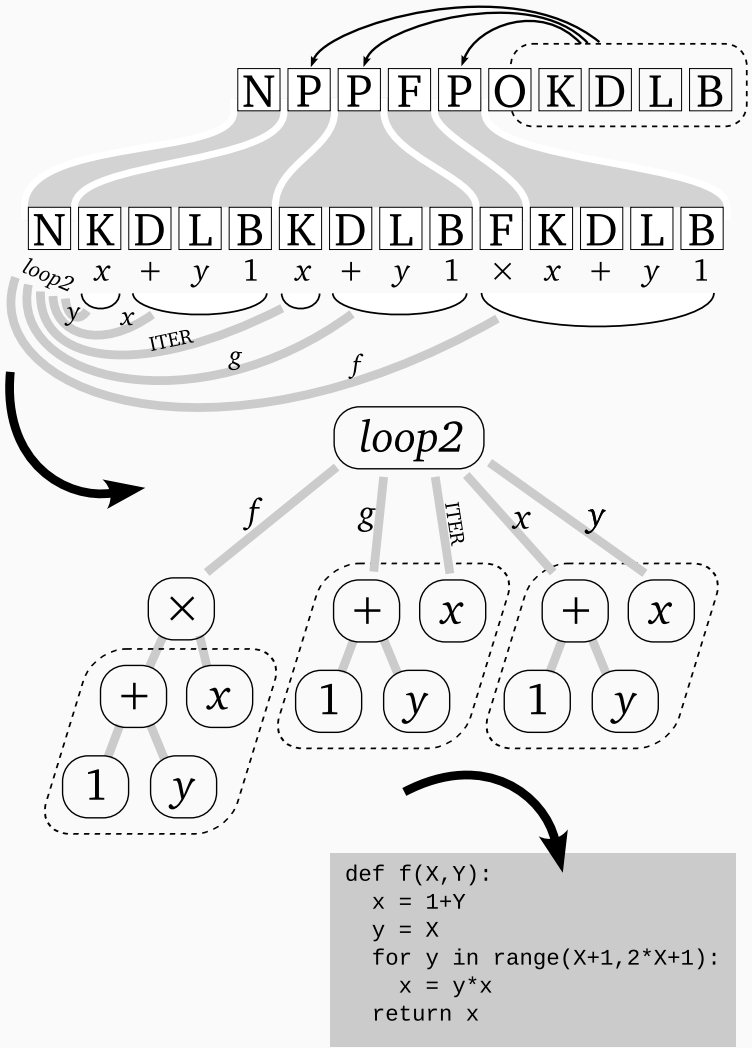
run | Python | Javascript | Share code

```
def f1(X):
  x = X
  for i in range (1,X + 1):
    x = x * i
  return x

def f0(X):
  return (f1(X) % (X + 1)) % 2

for x in range(32):
  print (f0(x))
```

```
0
1
1
0
1
0
1
0
0
0
1
0
1
```

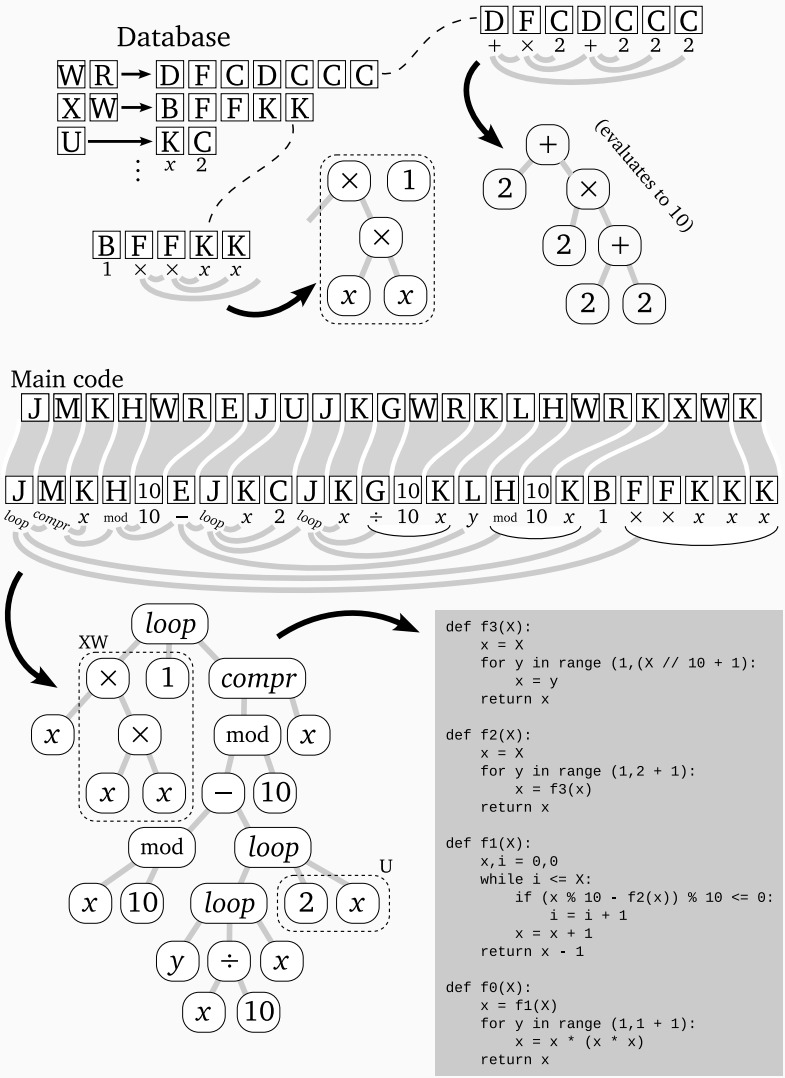
Introducing Local Macros/Definitions

A macro/expanded version of a program invented for A1813: $a(n) = (2n)!/n!$.
1, 2, 12, 120, 1680, 30240, 665280, 17297280, 518918400, 17643225600,



Introducing Global Macros/Definitions

A macro/expanded version of A14187 (cubes of palindromes).
0, 1, 8, 27, 64, 125, 216, 343, 512, 729, 1331, 10648, 35937, 85184,




```
def f3(X):
    x = X
    for y in range (1,(X // 10 + 1):
        x = y
    return x

def f2(X):
    x = X
    for y in range (1,2 + 1):
        x = f3(x)
    return x

def f1(X):
    x,i = 0,0
    while i <= X:
        if (x % 10 - f2(x)) % 10 <= 0:
            i = i + 1
        x = x + 1
    return x - 1

def f0(X):
    x = f1(X)
    for y in range (1,1 + 1):
        x = x * (x * x)
    return x
```

Five Different Self-Learning Runs

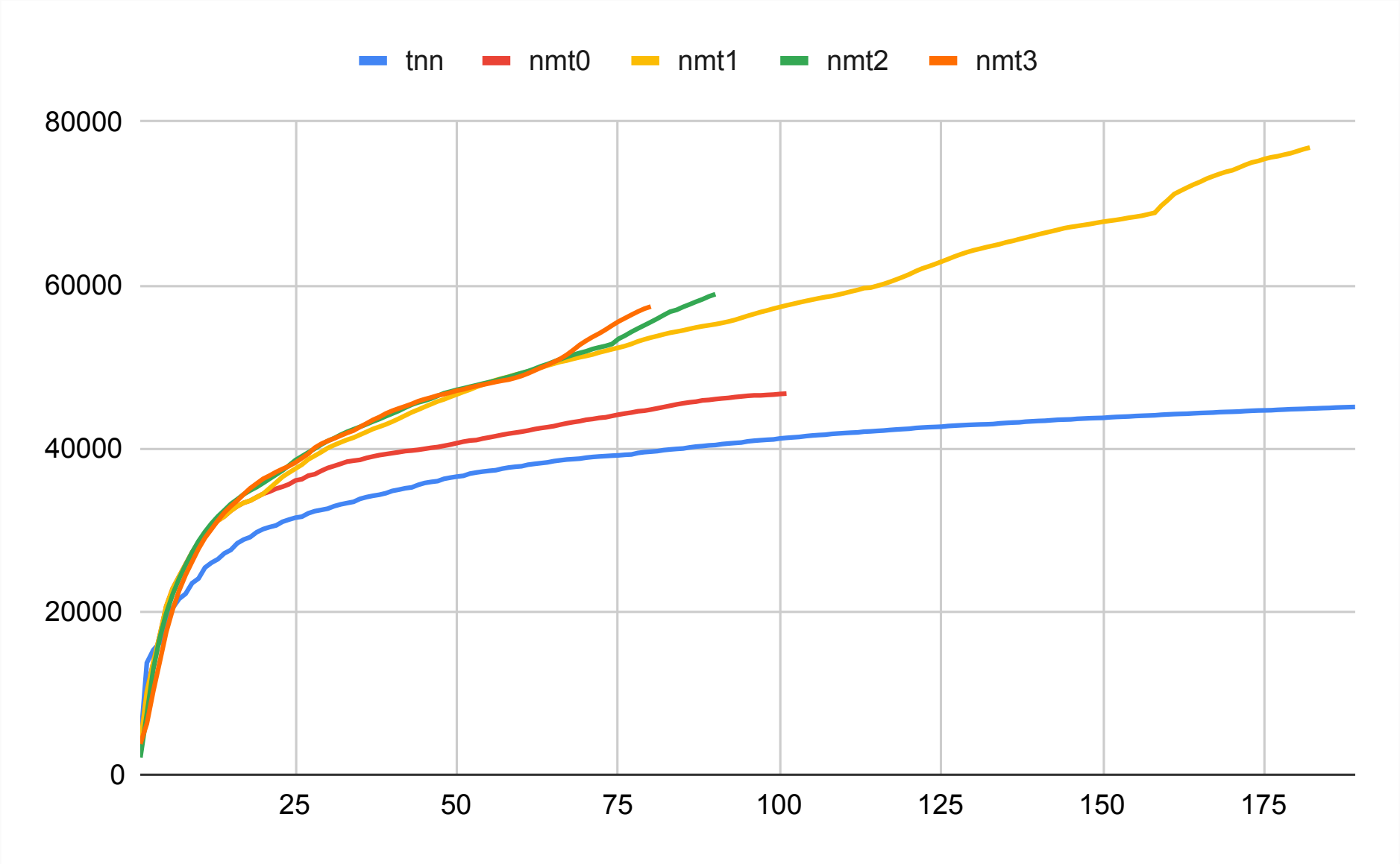


Figure: Cumulative counts of solutions.

Five Different Self-Learning Runs

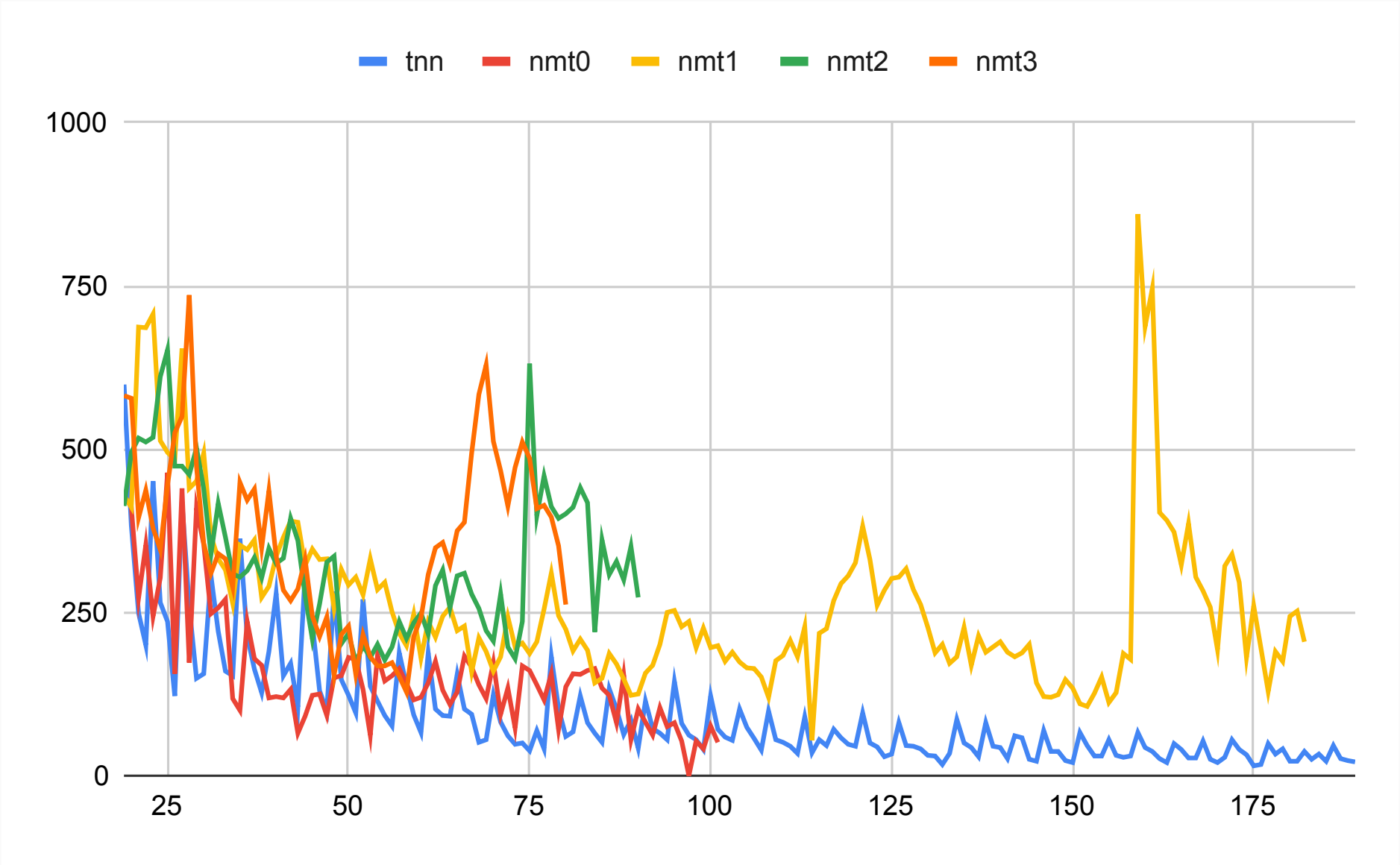


Figure: Increments of solutions.

Size Evolution

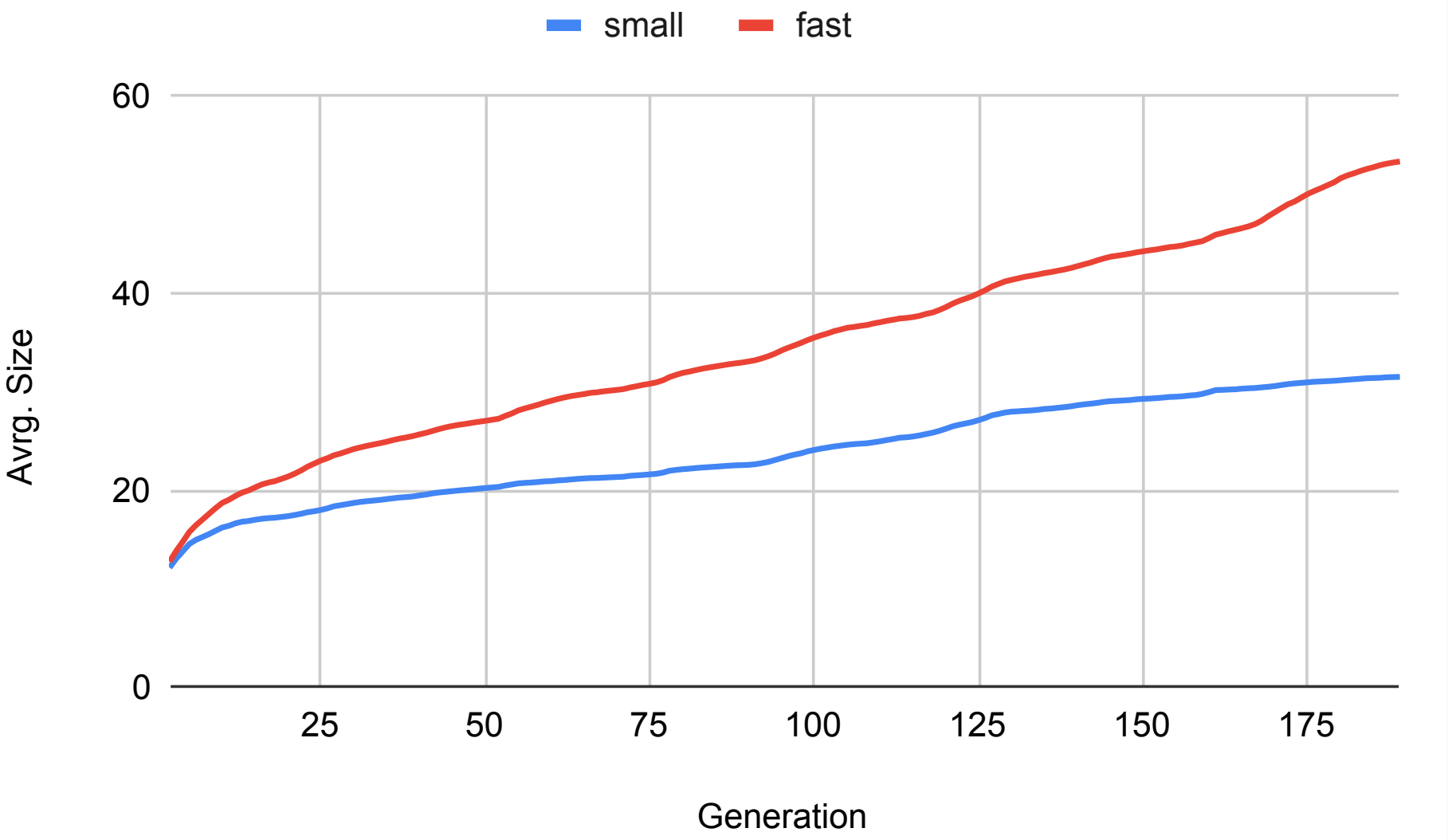


Figure: Avrg. size in iterations

Speed Evolution – Technology Breakthroughs

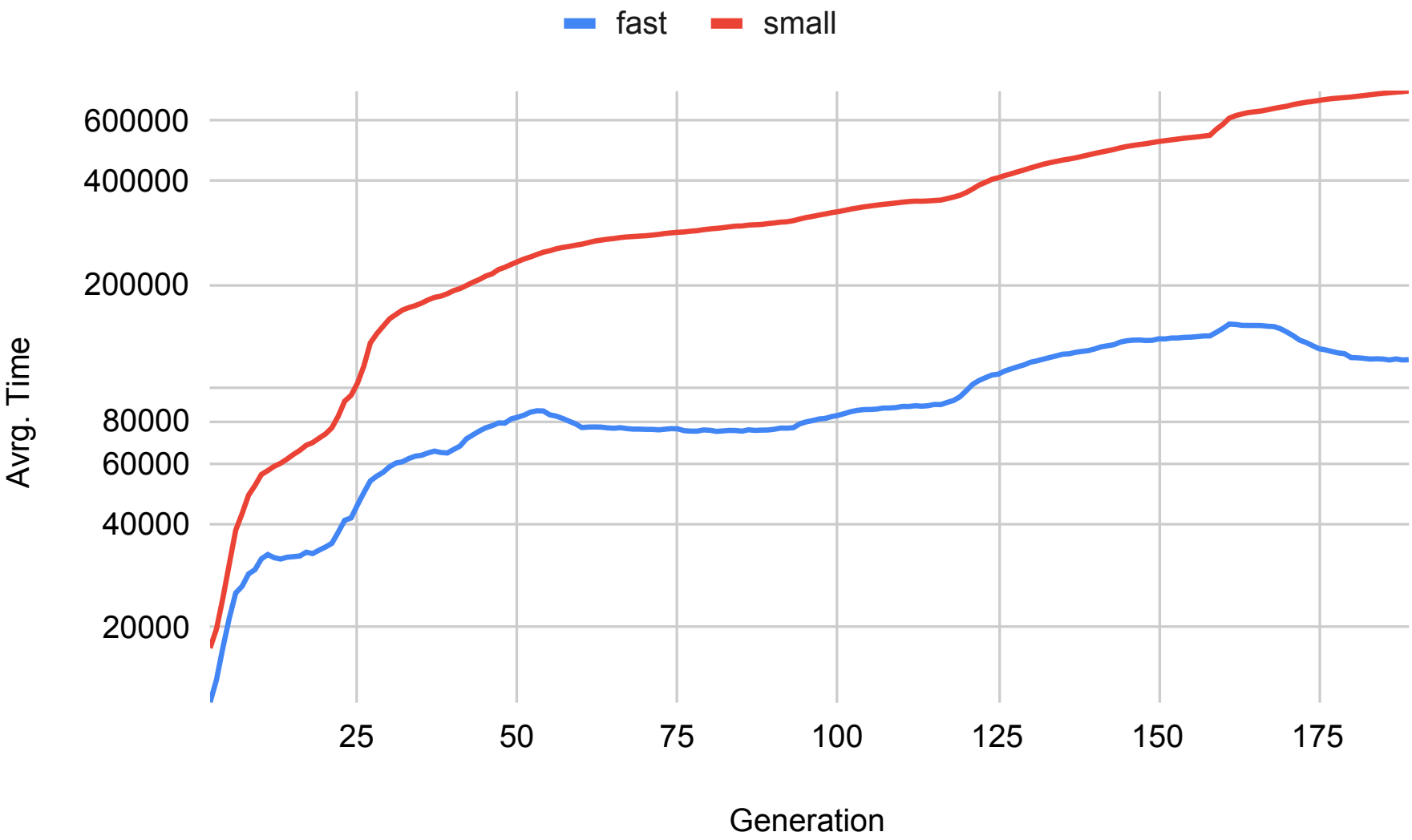
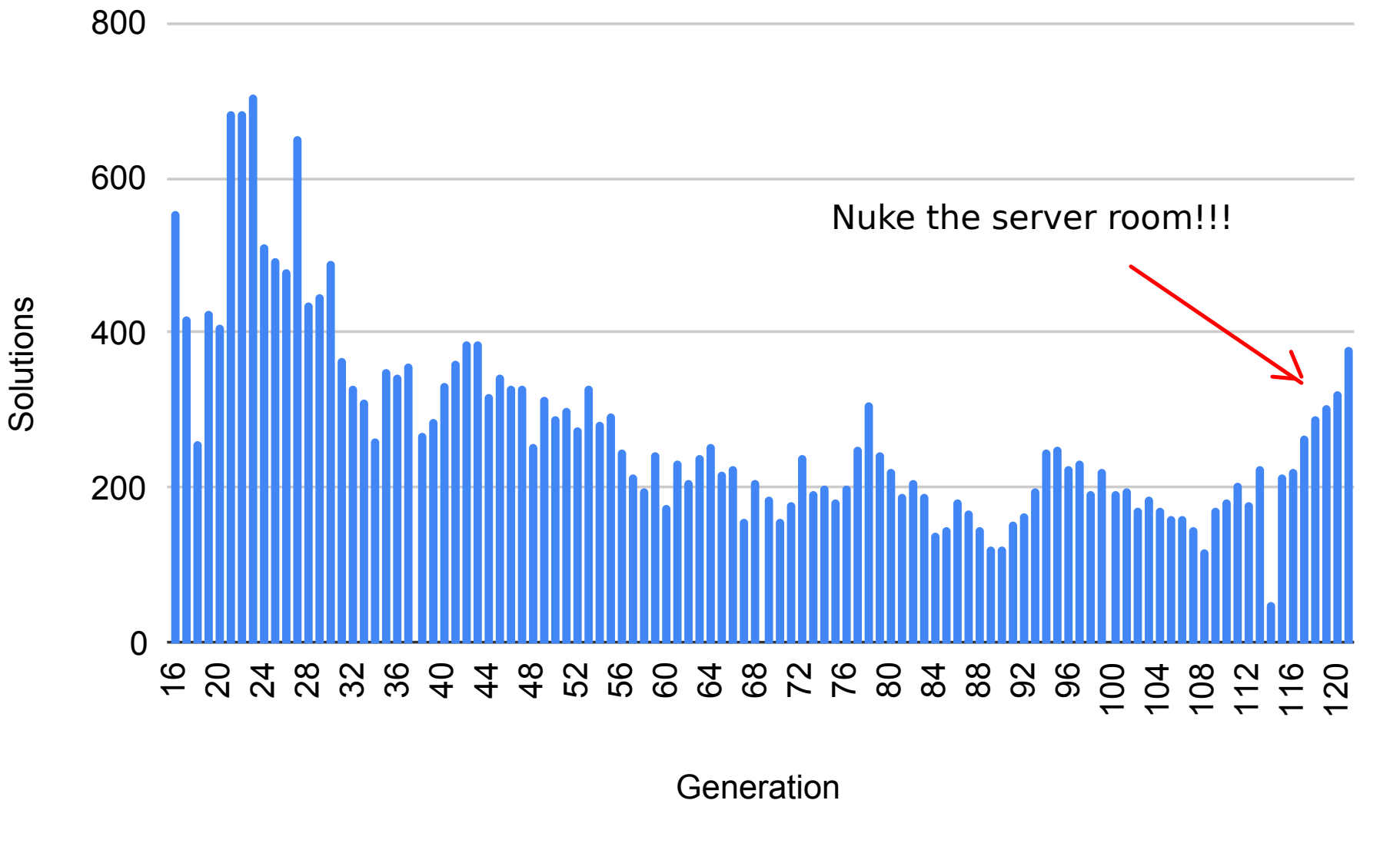
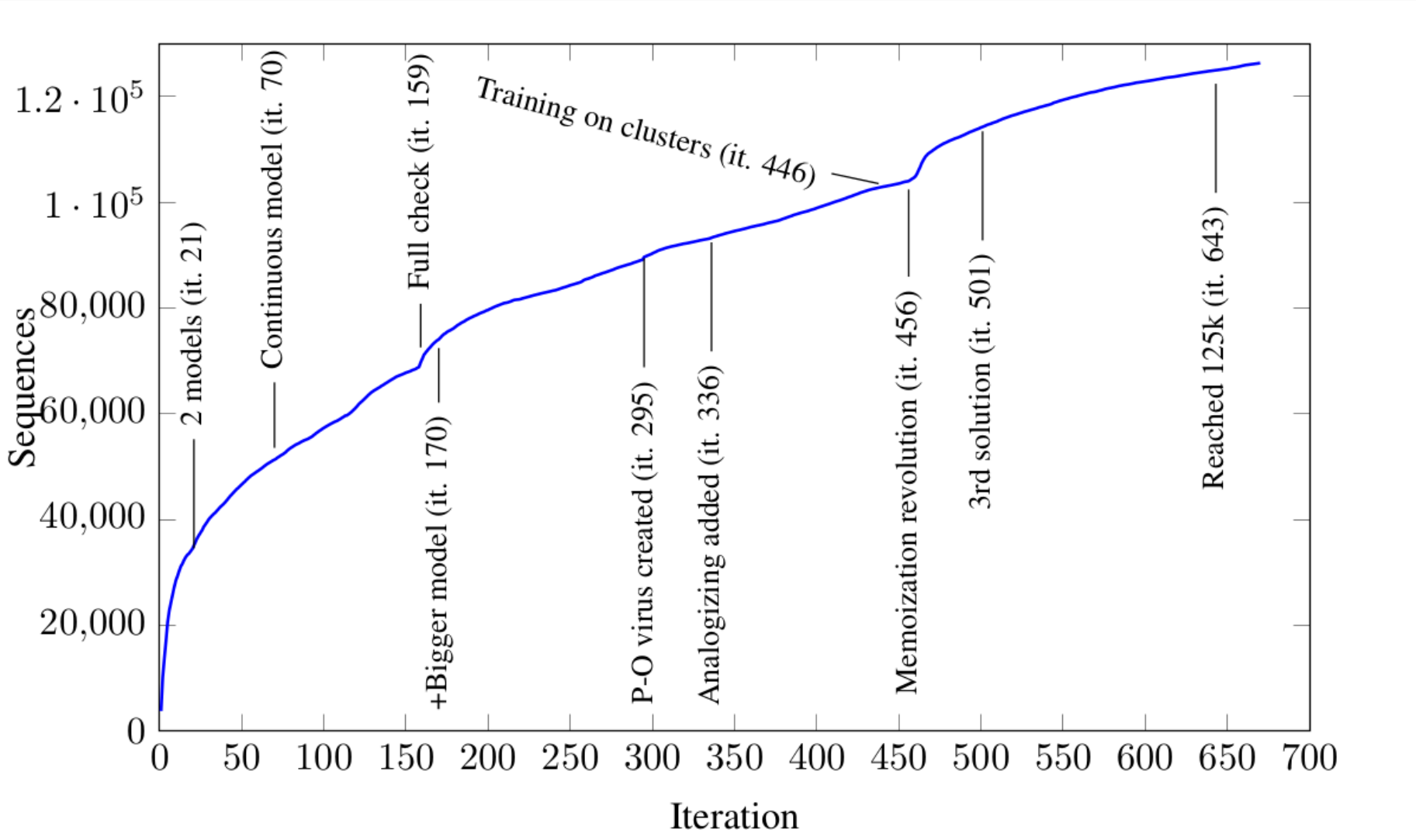


Figure: Avrg. time in iterations

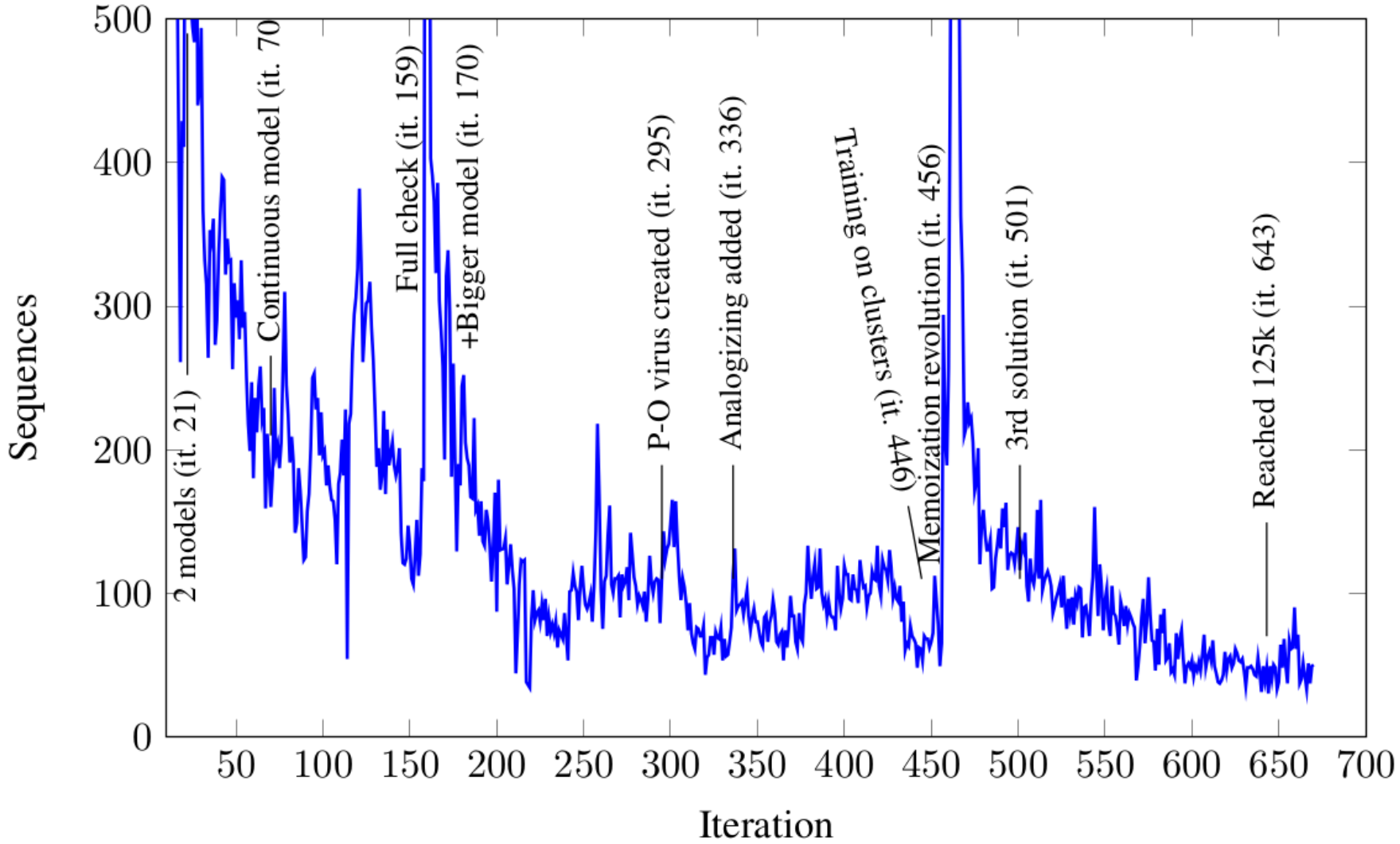
Singularity Take-Off X-mas Card



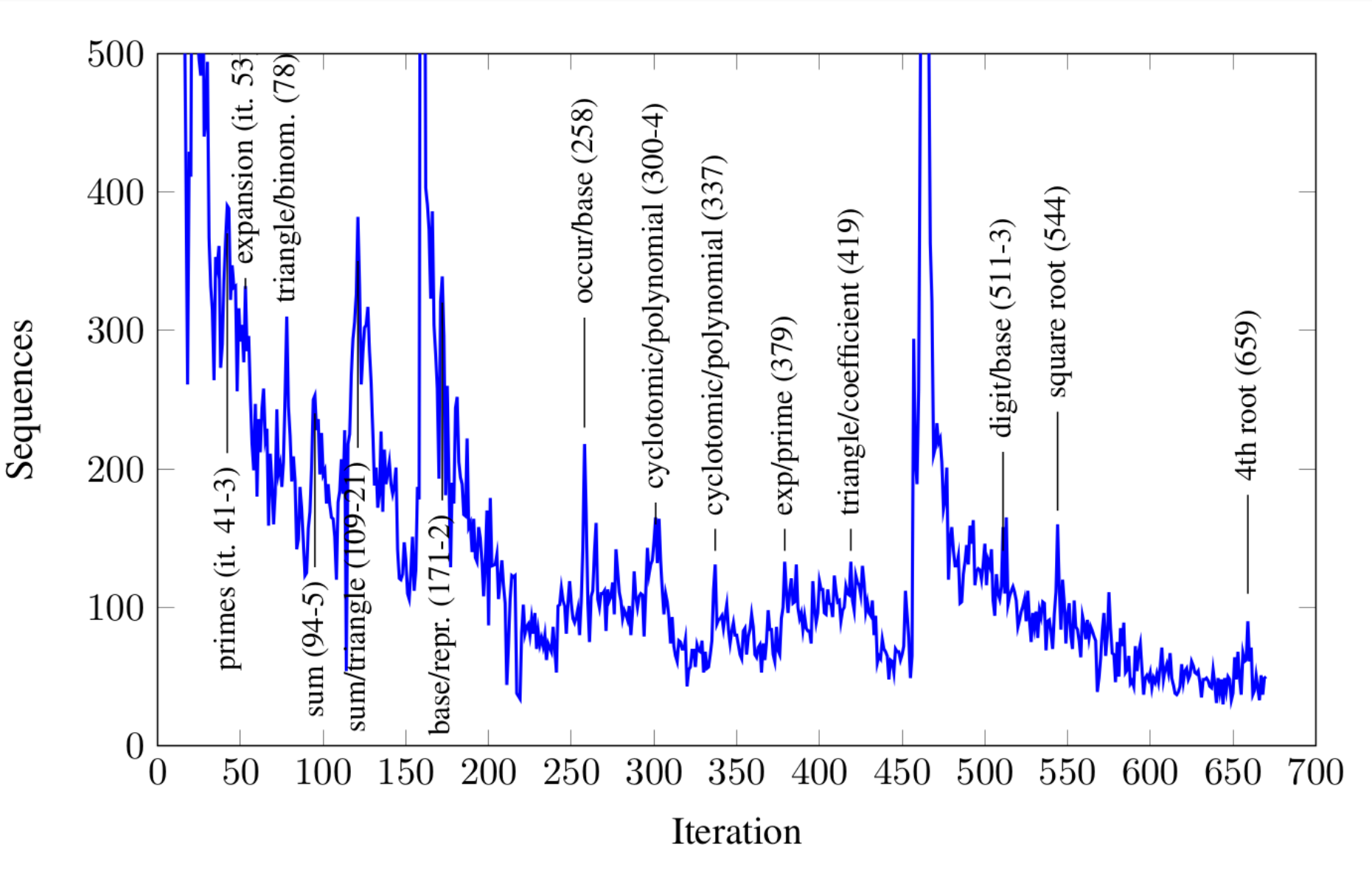
Human Made Technology Jumps



Human Made Technology Jumps



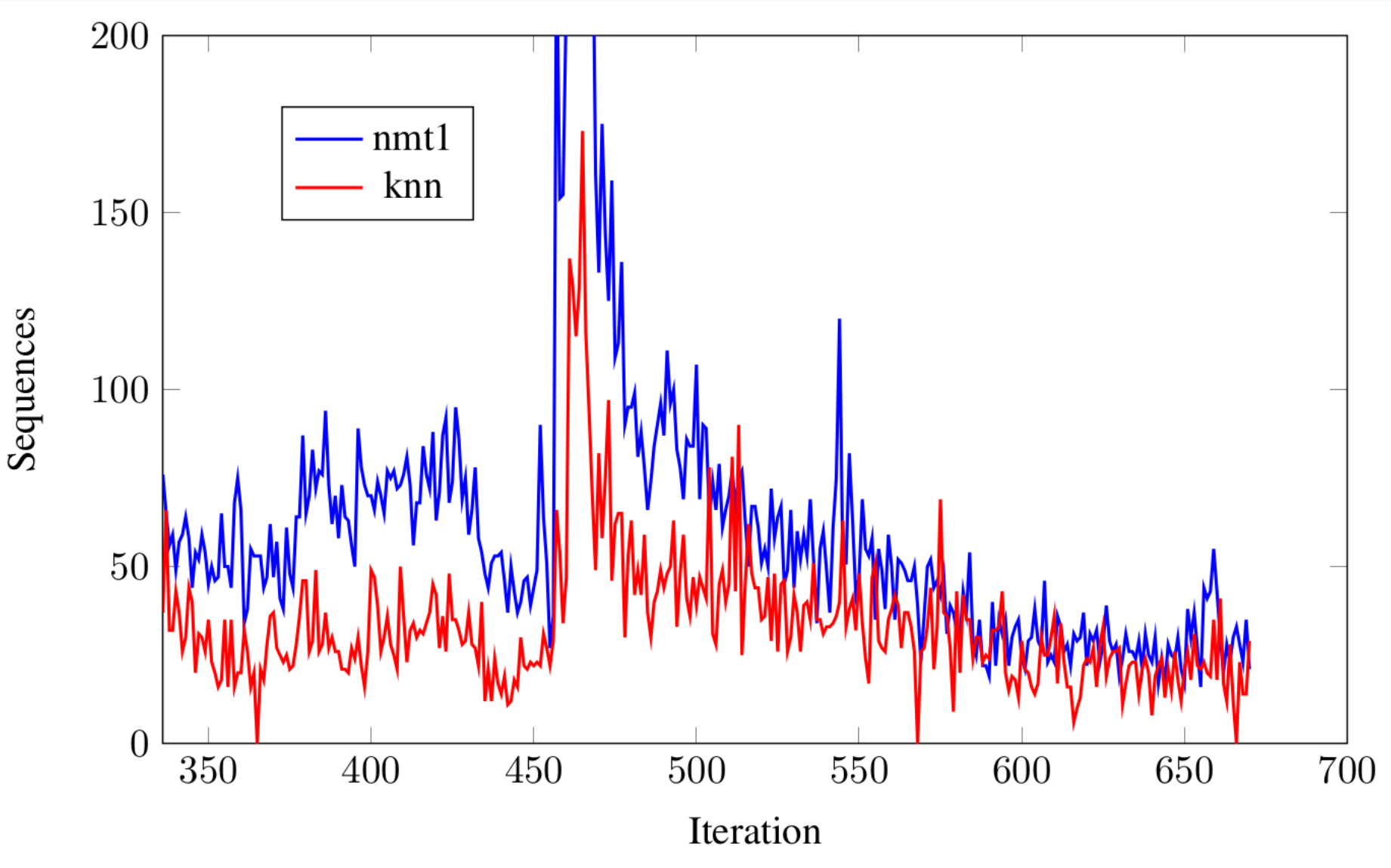
Some Automatic Technology Jumps



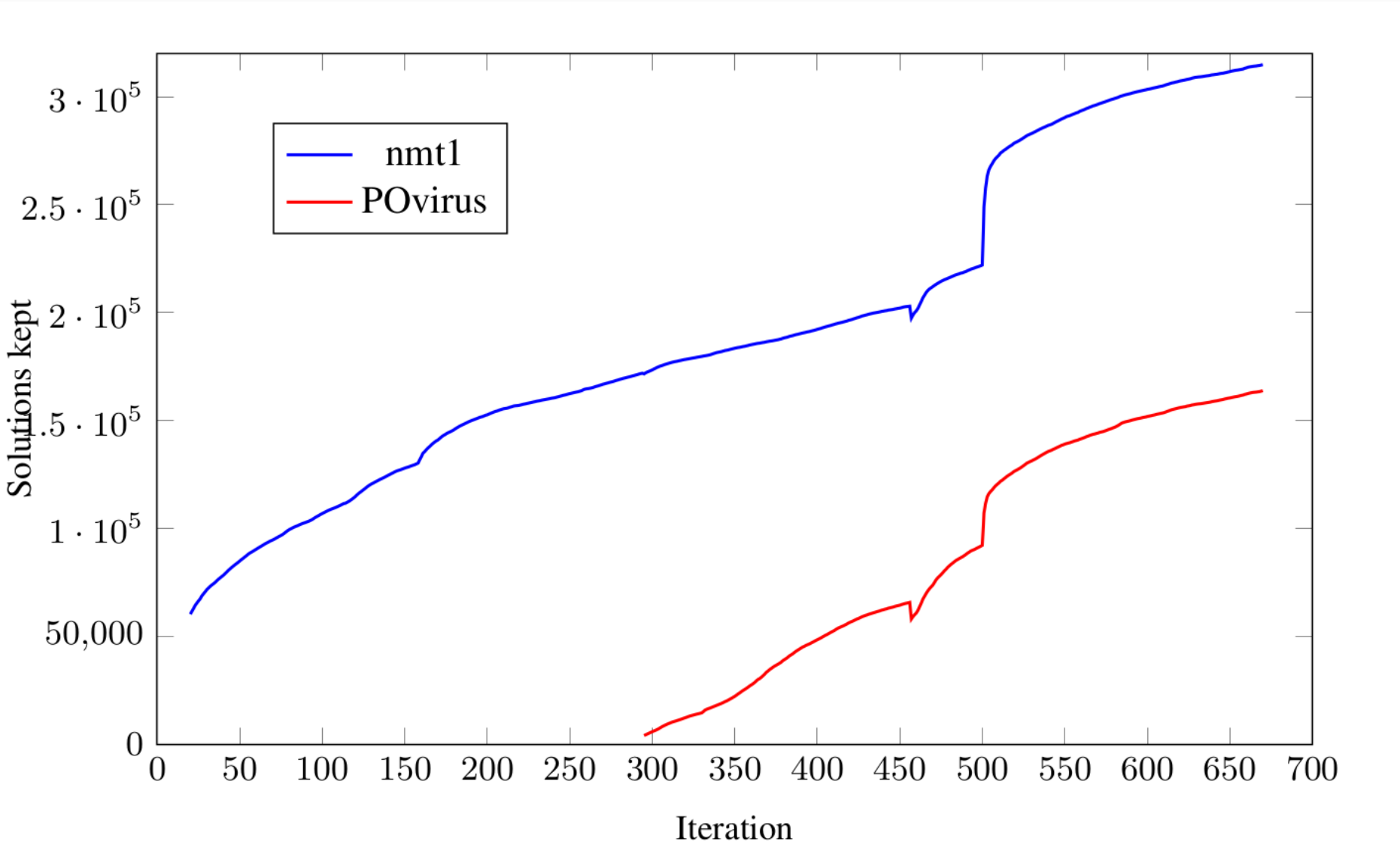
Some Automatic Technology Jumps

- iter 53: expansion/prime: A29363 Expansion of $1/((1-x^4)(1-x^7)(1-x^9)(1-x^{10}))$
- iter 78: triangle/binomial: A38313 Triangle whose (i,j)-th entry is $binomial(i,j) * 10^{i-j} * 11^j$
- iter 94-5: sum: A100192 $a(n) = Sum_{k=0..n} binomial(2n, n+k) * 2^k$
- 109-121: sum/triangle: A182013 Triangle of partial sums of Motzkin numbers
- 171-2: base/representation: A39080 n st base-9 repr. has the same number of 0's and 4's
- 258: occur/base: A44533 n st "2,0" occurs in the base 7 repr of n but not of $n+1$
- 300-304: cyclotomic/polynomial: A14620 Inverse of 611th cyclotomic polynomial
- 379: exp/prime: A124214 E.g.f.: $exp(x)/(2-exp(3*x))^{1/3}$
- 419: triangle/coefficient: A15129 Triangle of (Gaussian) q-binomial coefficients for $q=-13$
- 511,3: digit/base/prime: A260044 Primes with decimal digits in 0,1,3.
- 544: square root: A10538 Decimal expansion of square root of 87.
- 659: 4th root: A11084 Decimal expansion of 4th root of 93.

Translation vs Transformation



PO-virus Infection Rates



Generalization of the Solutions to Larger Indices

- Are the programs correct?
- Can we experimentally verify Occam's razor?
  (implications for how we should be designing ML/AI systems!)
- OEIS provides additional terms for some of the OEIS entries
- Among 78118 solutions, 40,577 of them have a b-file with 100 terms
- We evaluate both the small and the fast programs on them
- Here, 14,701 small and 11,056 fast programs time out.
- 90.57% of the remaining slow programs check
- 77.51% for the fast programs
- This means that **SHORTER EXPLANATIONS ARE MORE RELIABLE!**
  (Occam was right, so why is everybody building trillion-param LLMs???)
- Common error: reliance on an approximation of a real number, such as π.

Are two QSynt programs equivalent?

- As with primes, we often find many programs for one OEIS sequence
- Currently we have almost 4.5M programs for the 126k sequences
- It may be quite hard to see that the programs are equivalent
- Extend to Schmidhuber's Gödel Machine?
- A simple example for $0, 2, 4, 6, 8, \ldots$ with two programs f and g:
  - $f(0) = 0, f(n) = 2 + f(n-1)$ if $n > 0$
  - $g(n) = 2 * n$
  - conjecture: $\forall n \in \mathbb{N}. g(n) = f(n)$
- We can ask mathematicians, but we have thousands of such problems
- Or we can try to ask our ATPs (and thus create a large ATP benchmark)!
- Here is one SMT encoding by Janota & Gauthier:

```
(set-logic UFLIA)
(define-fun-rec f ((x Int)) Int (ite (<= x 0) 0 (+ 2 (f (- x 1)))))
(assert (exists ((c Int)) (and (> c 0) (not (= (f c) (* 2 c))))))
(check-sat)
```

Inductive proof by Vampire of the $f = g$ equivalence

```
% SZS output start Proof for rec2
1. f(X0) = $ite($lesseq(X0,0), 0,$sum(2,f($difference(X0,1)))) [input]
2. ? [X0 : $int] : ($greater(X0,0) & ~f(X0) = $product(2,X0)) [input]
[...]
43. ~$less(0,X0) | iG0(X0) = $sum(2,iG0($sum(X0,-1))) [evaluation 40]
44. (! [X0 : $int] : (($product(2,X0) = iG0(X0) & ~$less(X0,0)) => $product(2,$sum(X0,1)) = iG0($sum(X0,1)))
    &  $product(2,0) = iG0(0))   => ! [X1 : $int] : ($less(0,X1) => $product(2,X1) = iG0(X1)) [induction hypo]
[...]
49. $product(2,0) != iG0(0) | $product(2,$sum(sK3,1)) != iG0($sum(sK3,1)) | ~$less(0,sK1) [resolution 48,41]
50. $product(2,0) != iG0(0) | $product(2,sK3) = iG0(sK3) | ~$less(0,sK1) [resolution 47,41]
51. $product(2,0) != iG0(0) | ~$less(sK3,0) | ~$less(0,sK1) [resolution 46,41]
52. 0 != iG0(0) | $product(2,$sum(sK3,1)) != iG0($sum(sK3,1)) | ~$less(0,sK1) [evaluation 49]
53. 0 != iG0(0) | $product(2,sK3) = iG0(sK3) | ~$less(0,sK1) [evaluation 50]
54. 0 != iG0(0) | ~$less(sK3,0) | ~$less(0,sK1) [evaluation 51]
55. 0 != iG0(0) | ~$less(sK3,0) [subsumption resolution 54,39]
57. 1 <=> $less(sK3,0) [avatar definition]
59. ~$less(sK3,0) <- (~1) [avatar component clause 57]
61. 2 <=> 0 = iG0(0) [avatar definition]
64. ~1 | ~2 [avatar split clause 55,61,57]
65. 0 != iG0(0) | $product(2,sK3) = iG0(sK3) [subsumption resolution 53,39]
67. 3 <=> $product(2,sK3) = iG0(sK3) [avatar definition]
69. $product(2,sK3) = iG0(sK3) <- (3) [avatar component clause 67]
70. 3 | ~2 [avatar split clause 65,61,67]
71. 0 != iG0(0) | $product(2,$sum(sK3,1)) != iG0($sum(sK3,1)) [subsumption resolution 52,39]
72. $product(2,$sum(1,sK3)) != iG0($sum(1,sK3)) | 0 != iG0(0) [forward demodulation 71,5]
74. 4 <=> $product(2,$sum(1,sK3)) = iG0($sum(1,sK3)) [avatar definition]
76. $product(2,$sum(1,sK3)) != iG0($sum(1,sK3)) <- (~4) [avatar component clause 74]
77. ~2 | ~4 [avatar split clause 72,74,61]
82. 0 = iG0(0) [resolution 36,10]
85. 2 [avatar split clause 82,61]
246. iG0($sum(X1,1)) = $sum(2,iG0($sum($sum(X1,1),-1))) | $less(X1,0) [resolution 43,14]
251. $less(X1,0) | iG0($sum(X1,1)) = $sum(2,iG0(X1)) [evaluation 246]
[...]
1176. $false <- (~1, 3, ~4) [subsumption resolution 1175,1052]
1177. 1 | ~3 | 4 [avatar contradiction clause 1176]
1178. $false [avatar sat refutation 64,70,77,85,1177]
% SZS output end Proof for rec2
% Time elapsed: 0.016 s
```

80 Programs That Have Most Evolved

120	https://oeis.org/A238952	101	https://oeis.org/A97012	98	https://oeis.org/A17666
117	https://oeis.org/A35218	101	https://oeis.org/A71190	98	https://oeis.org/A113184
116	https://oeis.org/A1001	101	https://oeis.org/A70824	97	https://oeis.org/A82
112	https://oeis.org/A35178	101	https://oeis.org/A64987	97	https://oeis.org/A6579
111	https://oeis.org/A88580	101	https://oeis.org/A57660	97	https://oeis.org/A56595
111	https://oeis.org/A62069	101	https://oeis.org/A54024	97	https://oeis.org/A293228
111	https://oeis.org/A163109	101	https://oeis.org/A53222	97	https://oeis.org/A27847
111	https://oeis.org/A1615	101	https://oeis.org/A50457	97	https://oeis.org/A23645
109	https://oeis.org/A66446	101	https://oeis.org/A23888	97	https://oeis.org/A10
108	https://oeis.org/A48250	101	https://oeis.org/A209295	96	https://oeis.org/A92403
108	https://oeis.org/A321516	101	https://oeis.org/A206787	96	https://oeis.org/A90395
108	https://oeis.org/A2654	100	https://oeis.org/A99184	96	https://oeis.org/A83919
107	https://oeis.org/A75653	100	https://oeis.org/A63659	96	https://oeis.org/A7862
107	https://oeis.org/A60278	100	https://oeis.org/A62968	96	https://oeis.org/A78306
107	https://oeis.org/A23890	100	https://oeis.org/A35154	96	https://oeis.org/A69930
106	https://oeis.org/A62011	100	https://oeis.org/A339965	96	https://oeis.org/A69192
106	https://oeis.org/A346613	100	https://oeis.org/A277791	96	https://oeis.org/A54519
106	https://oeis.org/A344465	100	https://oeis.org/A230593	96	https://oeis.org/A53158
105	https://oeis.org/A49820	100	https://oeis.org/A182627	96	https://oeis.org/A351267
104	https://oeis.org/A55155	99	https://oeis.org/A9191	96	https://oeis.org/A334136
104	https://oeis.org/A349215	99	https://oeis.org/A82051	96	https://oeis.org/A33272
104	https://oeis.org/A143348	99	https://oeis.org/A62354	96	https://oeis.org/A325939
103	https://oeis.org/A92517	99	https://oeis.org/A247146	96	https://oeis.org/A211779
103	https://oeis.org/A64840	99	https://oeis.org/A211261	96	https://oeis.org/A186099
102	https://oeis.org/A9194	99	https://oeis.org/A147588	96	https://oeis.org/A143152
102	https://oeis.org/A51953	98	https://oeis.org/A318446	96	https://oeis.org/A125168
102	https://oeis.org/A155085	98	https://oeis.org/A203		

Evolution and Proliferation of Primes and Others

`https://bit.ly/3XHZsjK`: triangle coding, sigma (sum of divisors), primes. `https://bit.ly/3iJ4oGd` (the first 24, now 50)

Nr	Program
P1	(if x <= 0 then 2 else 1) + (compr (((loop (x + x) (x mod 2) (loop (x * x) 1 (loop (x + x) (x div 2) 1))) + x) mod (1 + x)) x)
P2	1 + (compr ((((loop (x * x) 1 (loop (x + x) (x div 2) 1)) + x) * x) mod (1 + x)) (1 + x))
P3	1 + (compr (((loop (x * x) 1 (loop (x + x) (x div 2) 1)) + x) mod (1 + x)) (1 + x))
P4	2 + (compr ((loop2 (1 + (if (x mod (1 + y)) <= 0 then 0 else x)) (y - 1) x 1 x) mod (1 + x)) x)
P5	1 + (compr ((loop (if (x mod (1 + y)) <= 0 then (1 + y) else x) x (1 + x)) mod (1 + x)) (1 + x))
P6	1 + (compr ((loop (if (x mod (1 + y)) <= 0 then (1 + y) else x) (2 + (x div (2 + (2 + 2)))) (1 + x)) mod (1 + x)) (1 + x))
P7	compr ((1 + (loop (if (x mod (1 + y)) <= 0 then (1 + y) else x) x x)) mod (1 + x)) (2 + x)
P8	1 + (compr ((loop (if (x mod (1 + y)) <= 0 then (1 + y) else x) (1 + ((2 + x) div (2 + (2 + 2)))) (1 + x)) mod (1 + x)) (1 + x))
P9	compr (x - (loop (if (x mod (1 + y)) <= 0 then (1 + y) else x) x x)) (2 + x)
P10	compr (x - (loop (if (x mod (1 + y)) <= 0 then 2 else x) (x div 2) x)) (2 + x)
P11	1 + (compr ((loop (if (x mod (1 + y)) <= 0 then (1 + y) else x) (1 + (x div (2 + (2 + 2)))) (1 + x)) mod (1 + x)) (1 + x))
P12	compr ((x - (loop (if (x mod (1 + y)) <= 0 then y else x) x x)) - 2) (2 + x)
P13	1 + (compr ((loop (if (x mod (1 + y)) <= 0 then (1 + y) else x) (2 + (x div (2 * (2 + (2 + 2))))) (1 + x)) mod (1 + x)) (1 + x))
P14	compr ((x - (loop (if (x mod (1 + y)) <= 0 then y else x) x x)) - 1) (2 + x)
P15	1 + (compr (x - (loop (if (x mod (1 + y)) <= 0 then (1 + y) else x) (2 + (x div (2 * (2 + (2 + 2))))) (1 + x))) (1 + x))
P16	compr (2 - (loop (if (x mod (1 + y)) <= 0 then 0 else x) (x - 2) x)) x
P17	1 + (compr (x - (loop (if (x mod (1 + y)) <= 0 then 2 else x) (2 + (x div (2 * (2 + (2 + 2))))) (1 + x))) (1 + x))
P18	1 + (compr (x - (loop (if (x mod (1 + y)) <= 0 then 2 else x) (1 + (2 + (x div (2 * (2 * (2 + 2)))))) (1 + x))) (1 + x))
P19	1 + (compr (x - (loop2 (loop (if (x mod (1 + y)) <= 0 then 2 else x) (2 + (y div (2 * (2 + (2 + 2))))) (1 + y)) 0 (1 - (x mod 2)) 1 x)) (1 + x))
P20	1 + (compr (x - (loop2 (loop (if (x mod (1 + y)) <= 0 then 2 else x) (1 + (2 + (y div (2 * (2 * (2 + 2)))))) (1 + y)) 0 (1 - (x mod 2)) 1 x)) (1 + x))
P21	1 + (compr (x - (loop2 (loop (if (x mod (2 + y)) <= 0 then 2 else x) (2 + (y div (2 * ((2 + 2) + (2 + 2))))) (1 + y)) 0 (1 - (x mod 2)) 1 x)) (1 + x))
P22	1 + (compr (x - (loop2 (loop (if (x mod (2 + y)) <= 0 then 2 else x) (2 + (y div (2 * (2 * (2 + 2))))) (1 + y)) 0 (1 - (x mod 2)) 1 x)) (1 + x))
P23	2 + (compr (loop (x - (if (x mod (1 + y)) <= 0 then 0 else 1)) x x) x)
P24	loop (1 + x) (1 - x) (1 + (2 * (compr (x - (loop (if (x mod (2 + y)) <= 0 then 1 else x) (2 + (x div (2 * (2 + 2)))) (1 + (x + x)))) x)))

Evolution and Proliferation of Primes

Iter	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	4	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	6	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	8	1	6	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	12	4	6	6	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35	7	12	6	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36	4	10	6	0	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
37	3	4	6	0	18	6	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	2	3	1	0	12	18	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
39	2	3	1	0	9	56	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
40	2	5	2	0	7	59	49	9	1	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0
41	1	2	3	0	4	52	58	42	23	0	13	0	8	0	0	0	0	0	0	0	0	0	0	0
42	0	2	4	0	3	44	50	38	60	8	11	0	55	0	0	0	0	0	0	0	0	0	0	0
43	0	2	12	0	0	37	55	14	116	35	16	7	90	0	0	0	0	0	0	0	0	0	0	0
44	0	2	13	0	0	28	40	6	176	73	19	8	122	9	12	0	0	0	0	0	0	0	0	0
45	0	2	9	0	0	19	24	4	147	185	26	16	94	25	29	0	7	0	0	0	0	0	0	0
46	0	2	4	0	0	11	14	0	101	256	21	14	66	64	30	0	29	0	0	0	0	0	0	0
47	0	0	0	0	0	9	4	0	55	290	23	3	43	116	16	6	62	14	0	0	0	0	0	0
48	0	0	0	0	0	8	0	0	22	261	16	0	34	192	10	6	89	30	0	0	0	0	0	0
49	0	0	0	0	0	8	0	0	6	195	11	0	36	225	8	6	99	34	0	0	0	0	0	0
50	0	0	0	0	0	5	0	0	2	154	8	0	29	168	6	6	108	39	0	0	0	0	0	0
51	0	0	0	0	0	4	0	0	0	121	7	0	21	97	6	6	113	43	0	0	0	0	0	0
52	0	0	0	0	0	2	0	0	0	118	8	0	12	62	6	6	110	51	0	0	0	0	0	0
53	0	0	0	0	0	1	0	0	0	59	7	0	15	33	6	6	125	62	0	0	0	0	0	0
54	0	0	0	0	0	1	0	0	0	41	4	0	16	17	6	9	137	72	0	0	0	0	0	0
55	0	0	0	0	0	2	0	0	0	32	4	0	15	9	6	17	147	82	0	0	0	0	0	0
56	0	0	0	0	0	1	0	0	0	29	4	0	10	7	6	39	152	98	0	0	0	0	0	0

Selection of 123 Solved Sequences

https://github.com/Anon52MI4/oeis-alien

Table: Samples of the solved sequences.

https://oeis.org/A317485	Number of Hamiltonian paths in the n-Bruhat graph.
https://oeis.org/A349073	a(n) = U(2*n, n), where U(n, x) is the Chebyshev polynomial of the second kind.
https://oeis.org/A293339	Greatest integer k such that $k/2^n < 1/e$.
https://oeis.org/A1848	Crystal ball sequence for 6-dimensional cubic lattice.
https://oeis.org/A8628	Molien series for A_5.
https://oeis.org/A259445	Multiplicative with $a(n) = n$ if n is odd and $a(2^s) = 2$.
https://oeis.org/A314106	Coordination sequence Gal.6.199.4 where G.u.t.v denotes the coordination sequence for a vertex of type v in tiling number t in the Galebach list of u-uniform tilings
https://oeis.org/A311889	Coordination sequence Gal.6.129.2 where G.u.t.v denotes the coordination sequence for a vertex of type v in tiling number t in the Galebach list of u-uniform tilings.
https://oeis.org/A315334	Coordination sequence Gal.6.623.2 where G.u.t.v denotes the coordination sequence for a vertex of type v in tiling number t in the Galebach list of u-uniform tilings.
https://oeis.org/A315742	Coordination sequence Gal.5.302.5 where G.u.t.v denotes the coordination sequence for a vertex of type v in tiling number t in the Galebach list of u-uniform tilings.
https://oeis.org/A004165	OEIS writing backward
https://oeis.org/A83186	Sum of first n primes whose indices are primes.
https://oeis.org/A88176	Primes such that the previous two primes are a twin prime pair.
https://oeis.org/A96282	Sums of successive twin primes of order 2.
https://oeis.org/A53176	Primes p such that $2p + 1$ is composite.
https://oeis.org/A267262	Total number of OFF (white) cells after n iterations of the "Rule 111" elementary cellular automaton starting with a single ON (black) cell.

Infinite Math-Nerd Sniping

- We have 4.5M problems for math nerds like this one:
- JU: *This thing works for the first 1k values (just checked) - any idea why?*
- `https://oeis.org/A004578` - Expansion of sqrt(8) in base 3.
- loop2(((y * y) div (x + y)) + y, y, x + x, 2, loop((1 + 2) * x, x, 2)) mod (1 + 2)
- MO: *Not a proof, just a rough idea: The program iterates the function q |-> 2+q / 1+q, where q is a rational number. This converges to sqrt(2). The number q is represented by an integer 'a' such that* $a = 3^x * (2 * q)$, *where 'x' is the input. Once the approximation is good enough,* $a = floor(3^x * sqrt(8))$, *so a mod 3 is the digit we want.*

Serious Math Conjecturing – Elliptic Curves

- Sander Dahmen: *Here are some OEIS labels related to elliptic curves (and hence modular forms), ordered by difficulty. It would be interesting to know if some of these appear in your results.*
- A006571 A030187 A030184 A128263 A187096 A251913
- JU: *We have the first three:*
- A6571 : loop((push(loop((pop(x) * loop(if (pop(x) mod y) <= 0 then ((if (y mod loop(1 + (x + x), 2, 2)) <= 0 then (x - y) else x) - y) else x, y, push(0, y))) + x, y, push(0, x)), x) * 2) div y, x, 1)
- A30187 : loop(push(loop((pop(x) * loop(if (pop(x) mod y) <= 0 then (x - loop(if (x mod (((2 + y) * y) - 1)) <= 0 then (x + x) else x, 2, y)) else x, y, push(0, y))) + x, y, push(0, x)), x) div y, x, 1)
- A30184 : loop(push(loop((pop(x) * loop(if (pop(x) mod y) <= 0 then (x - loop(if (x mod (1 + (y + y))) <= 0 then (x + x) else x, 2, y)) else x, y, push(0, y))) + x, y, push(0, x)), x) div y, x, 1)

A6571: Expansion of $q * Product_{k>=1}(1 - q^k)^2 * (1 - q^{11*k})^2$
A30187: Expansion of $\eta(q) * \eta(q^2) * \eta(q^7) * \eta(q^{14})$ in powers of q.
A30184: Expansion of $\eta(q) * \eta(q^3) * \eta(q^5) * \eta(q^{15})$ in powers of q.

More Bragging

- Hofstadter-Conway $10000 sequence: a(n) = a(a(n-1)) + a(n-a(n-1)) with a(1) = a(2) = 1.
- D. R. Hofstadter, Analogies and Sequences: Intertwined Patterns of Integers and Patterns of Thought Processes, Lecture in DIMACS Conference on Challenges of Identifying Integer Sequences, 2014.

```
Date: Sun, Mar 17, 2024
To: <dughof@indiana.edu>

Dear Douglas,

our system [1] has today (iteration 552) found a solution of
https://oeis.org/A004074. The solution in Thibault's programming
language [1] (with push/pop added on top of [1]) is:

((2*loop(push(loop(pop(x),x-1,x),x)+loop(pop(x),y-x,pop(x)),x-1,1))-1)-x

The related A4001 was solved in iteration 463 and the solution is:
loop(push(loop(pop(x), y-x,pop(x)),x) + loop(pop(x), x-1, x), x - 1, 1)
```

Minsky 2014

It seems to me that the most important discovery since Gödel was the discovery by Chaitin, Solomonoff and Kolmogorov of the concept called Algorithmic Probability which is a fundamental new theory of how to make predictions given a collection of experiences and this is a beautiful theory, everybody should learn it,

but it's got one problem, that is, that you cannot actually calculate what this theory predicts because it is too hard, it requires an infinite amount of work.

However, it should be possible to make practical approximations to the Chaitin, Kolmogorov, Solomonoff theory that would make better predictions than anything we have today. Everybody should learn all about that and spend the rest of their lives working on it.

– M. Minsky, Panel discussion on The Limits of Understanding, 2014

References to our relevant work

- Thibault Gauthier, Josef Urban: Learning Program Synthesis for Integer Sequences from Scratch. AAAI 2023: 7670-7677
- Thibault Gauthier, Miroslav Olsák, Josef Urban: Alien Coding. CoRR abs/2301.11479 (2023)
- Thibault Gauthier, Chad E. Brown, Mikolas Janota, Josef Urban: A Mathematical Benchmark for Inductive Theorem Provers. LPAR 2023: 224-237
- Thibault Gauthier, Cezary Kaliszyk, Josef Urban: TacticToe: Learning to Reason with HOL4 Tactics. LPAR 2017: 125-143
- Thibault Gauthier, Cezary Kaliszyk, Josef Urban, Ramana Kumar, Michael Norrish: Learning to Prove with Tactics. CoRR abs/1804.00596 (2018)
- Jan Jakubuv, Karel Chvalovský, Zarathustra Amadeus Goertzel, Cezary Kaliszyk, Mirek Olsák, Bartosz Piotrowski, Stephan Schulz, Martin Suda, Josef Urban: MizAR 60 for Mizar 50. ITP 2023: 19:1-19:22
- Thibault Gauthier: Deep Reinforcement Learning for Synthesizing Functions in Higher-Order Logic. LPAR 2020: 230-248
- Thibault Gauthier: Tree Neural Networks in HOL4. CICM 2020: 278-283
- Qingxiang Wang, Cezary Kaliszyk, Josef Urban: First Experiments with Neural Translation of Informal to Formal Mathematics. CICM 2018: 255-270
- Bartosz Piotrowski, Josef Urban: Stateful Premise Selection by Recurrent Neural Networks. LPAR 2020: 409-422
- Bartosz Piotrowski, Josef Urban: Guiding Inferences in Connection Tableau by Recurrent Neural Networks. CICM 2020: 309-314
- Josef Urban, Jan Jakubuv: First Neural Conjecturing Datasets and Experiments. CICM 2020: 315-323
- Bartosz Piotrowski, Josef Urban, Chad E. Brown, Cezary Kaliszyk: Can Neural Networks Learn Symbolic Rewriting? CoRR abs/1911.04873 (2019)

Thanks and Advertisement

- Thanks for your attention!
- To push AI/ML methods in math and theorem proving, we organize:
- AITP – Artificial Intelligence and Theorem Proving
- September 1-6, 2024, Aussois, France, `aitp-conference.org`
- ATP/ITP/Math vs AI/ML/AGI people, Computational linguists
- Discussion-oriented and experimental
- About 50 people in 2024
- Invited talks by people who do AI/ML/TP for math, physics, ...