CayleyPy Growth: group (artificial) intelligence

Sergey Galkin (PUC-Rio) and Igor Kiselev (Waterloo)
on behalf of

CayleyPy collaboration

What is CaylePy?

Crowd-sourced project and collaboration, more than 100 people involved,
featuring
A. Chervov, D. Fedoriaka, E. Konstantinova, A. Soibelman, A. Smolensky, F. Petrov, F. Levkovich-Maslyuk, H. Isambert

An open-source python library
https://github.com/cayleypy/cayleypy

CayleyPy-1: NeurIPS 2025 (spotlight) Mexico

CayleyPy-2: Accepted in ATMP

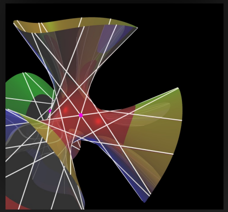
CayleyPy-3: Growth (49 authors from 10+ countries): this workshop and talk
5th MATH-AI workshop @ NeurIPS'2025,
Oral Presentation on Dec 6, 14:50-15:10, San Diego.

CayleyPy-3: arXiv:2509.19162 (a brief version, just 46 pages),
Draft of a "complete" version is more than 500 pages long so far. Multiple spin-offs.

Definitions and examples

Def. (A. Renyi) "a mathematician is a machine for turning coffee into theorems"

Example ☕ $\ldots\ [Y^{[2]}] = [\mathbf{P}^d \times Y] + \mathbf{L}^2[F_Y]\ \ldots$ ☕ $\ldots\ 27$



Pipeline: observe patterns, pose conjectures, prove theorems,
conceptualize definitions, find examples

Inductive reasoning: key examples & critical compute/scale.

CayleyPy is… a collaboration of people and machines for turning coffee and compute into conjectures and theorems.

A new and, so far, unconventional way of doing mathematics.
Math shall be a common task, not a vanity fair..

Unconventional collaboration
(Cf. Yejin Choi)

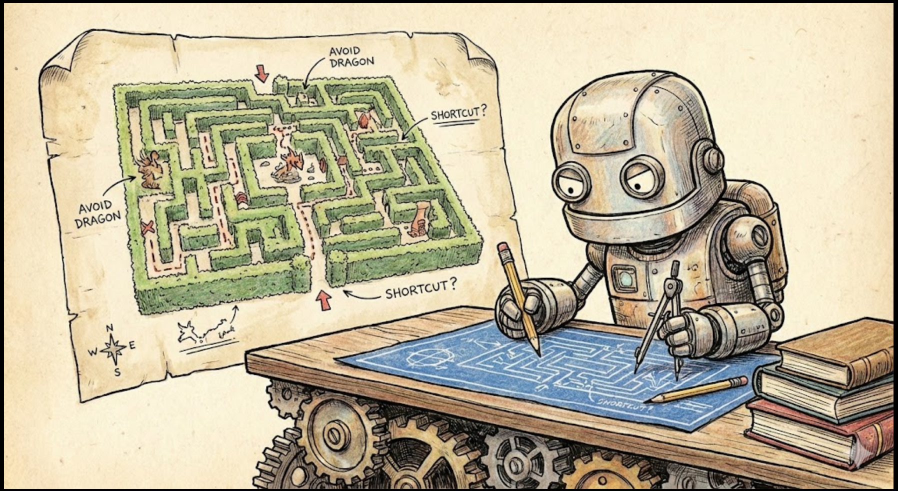
Pipeline:

open-source library development

Kaggle competitions

Scientists in the loop: mathematicians, physicists, bioinformaticians,…

Reasoners put effort to:



- reaching goals
- finding paths
- proving theorems
- formalizing statements
- coding
- planning tasks (shopping, navigation,…)

googol := 10^{100}

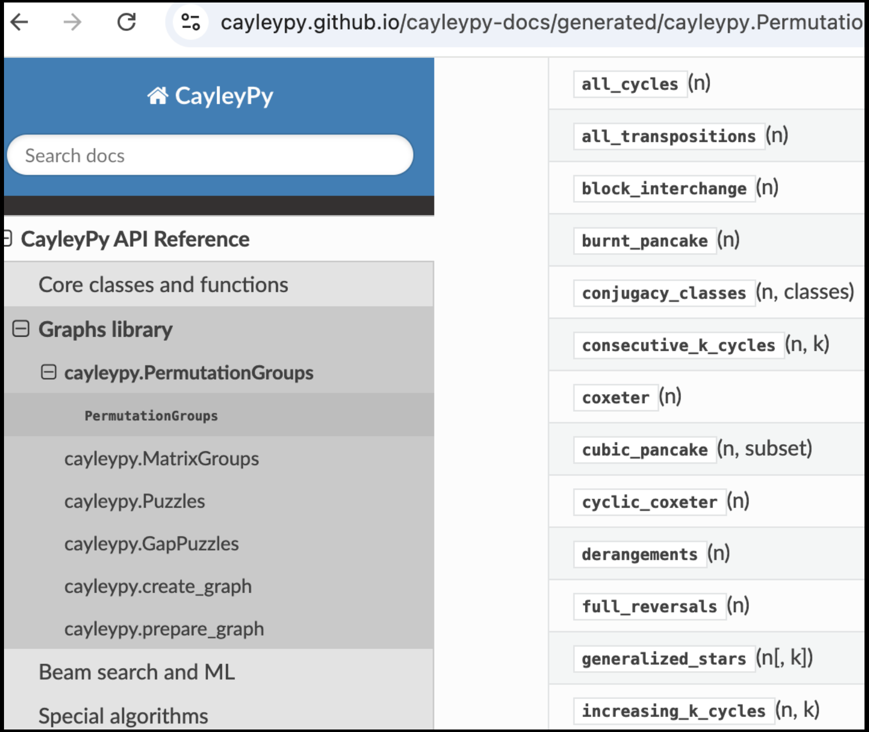
- !ALERT! Beware of: combinatorial e x p l o S I O N 💥🎇🤯

- We scale up graph search to googol-size graphs, for graphs regular enough
- (Cayley graphs and Schreier coset graphs, so far; 15-puzzle in beta)

CayleyPy: AI-based open-source Python library to handle GOOGOL size graphs

400+ stars on Github; Documentation; Tutorials; 300 notebooks with experiments

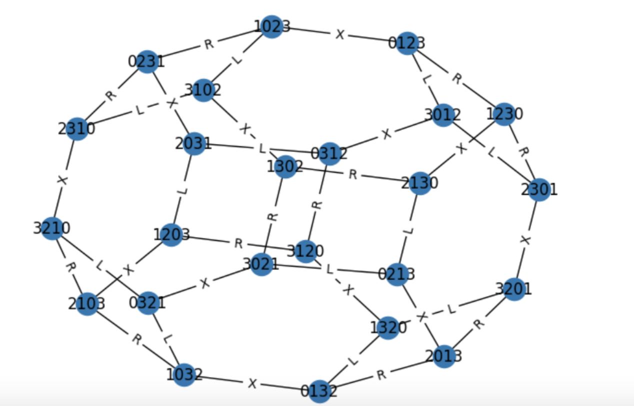
https://cayleypy.github.io/cayleypy-docs/index.html




```python
import networkx as nx
from cayleypy import CayleyGraph, PermutationGroups

graph = CayleyGraph(PermutationGroups.lrx(4))
result = graph.bfs( return_all_edges=True, return_all_hashes=True)
G = result.to_networkx_graph()

pos = nx.kamada_kawai_layout(G)
nx.draw(G, pos, with_labels=True)
edge_labels = nx.get_edge_attributes(G, 'label')
nx.draw_networkx_edge_labels(G, pos, edge_labels=edge_labels)
None
```



One task — many ways to explain:

Solve a puzzle (fun!)

e.g. Rubik's cube

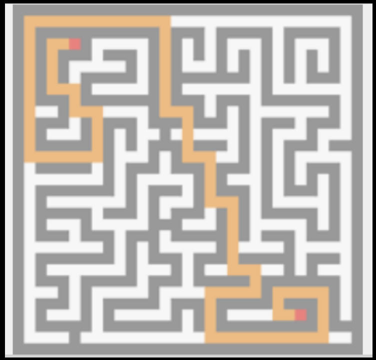
The God's number = ?

Math: factorize a matrix,
decompose a permutation
into a product of generators

$$(13) = (12)\,(23)\,(12)$$

$$\begin{pmatrix} 3 & 4 \\ 5 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Complexity = ?



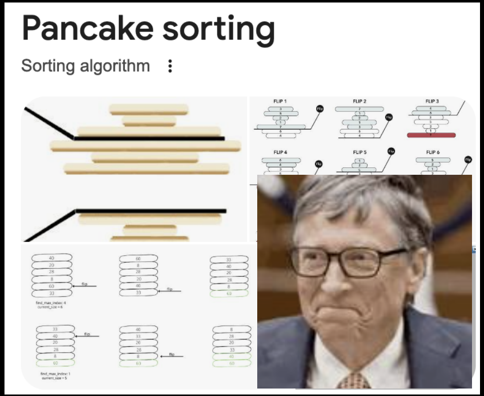
Find a path

on a (Cayley) graph

Diameter = ?

CS: sorting problems

Worst case performance = ?



Results and SOTA

1. C. Sims (1970): Schreier–Sims algorithm for permutation groups
2. Knuth, Donald E. (1991): randomized version $(10^{20} - 10^{30})$
3. CayleyPy beats them by orders and orders of magnitude $(10^{50} - 10^{120})$

Complexity:

1. 1981 NP-hard: optimal solutions and diameters
2. 1998-...: optimal solution NP-complete for Rubik's cube, etc

n	GAP Length	Ideal Length n(n-1)/2	AI Length
9	41	36	36
10	51	45	45
11	65	55	55
12	78	66	66
13	99	78	78
14	111	91	91
15	268	105	105
16	2454	120	120
17	380	136	136
18	20441	153	153
19	3187	171	171
20	217944	190	190
21	-	210	210

Conjecture of L. Babai

1. For any generators of S_n, diameter is less than $O(n^2)$ (any polynomial bound - open problem)
2. CayleyPy 100-1000 times FASTER than GAP, so it enables extensive computational experiments
3. New: explicit generators which conjecturally maximize the diameter.
   A simple pattern: "a square with whiskers" + involutions
4. New conjecture: the diameter is less than $n^2/2 + 4n$

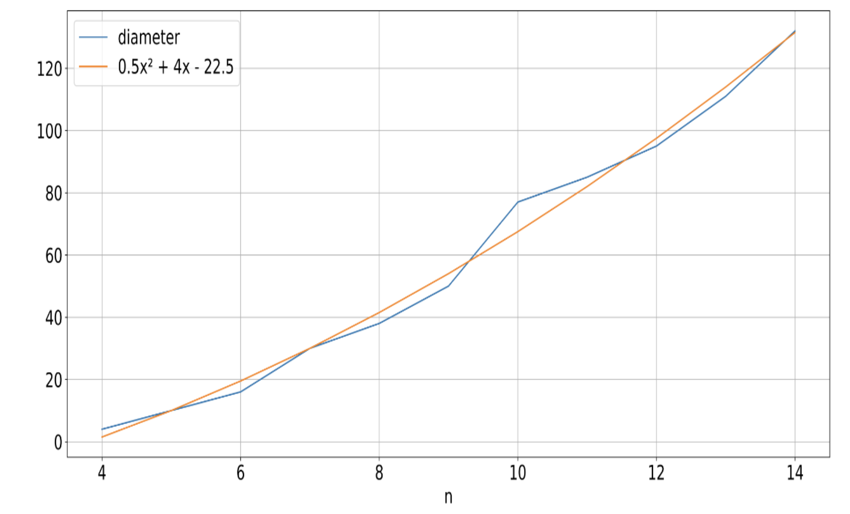
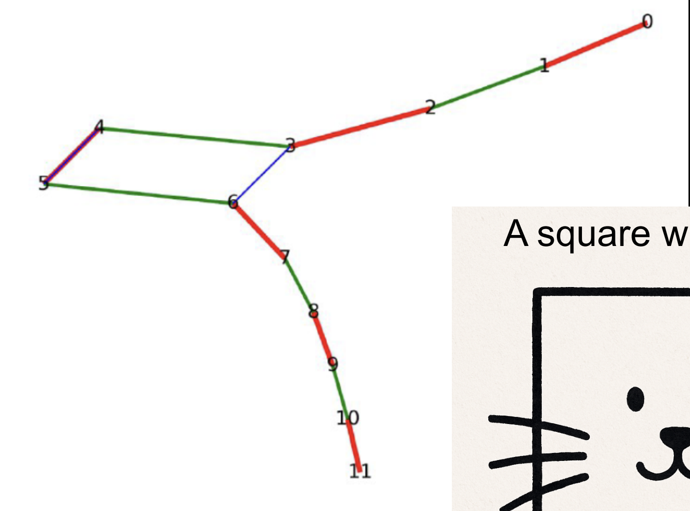
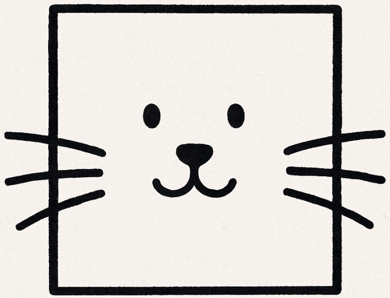


FIGURE 10. Maximal diameters found up to $n \leq 14$ and quadratic fit: $n^2/2 + 4n - 22.5$.



A square with whiskers



Sou ótimo. 对不对?

Surprise: Diameters are “findable” (?!) quasi-polynomials

Example of a quasi-polynomial: $n^2/3$ **if** $n = 3k$, **and** $(n^2 - 1)/3$ **otherwise.**

Conjectures:

1. **Diameters of S_n are quasi-polynomials (for constructive / “analytic” generators — open problem)**
2. **Length (word metric) of ANY “analytic” element is quasi-polynomial for any “analytic” generators**

Example — https://oeis.org/A039745 — Глушков

Two generators: L = (1 2 … n), the left cyclic shift,

X = (1 2), a transposition

1968 problem by “a father of Soviet cybernetics” V. M. Glushkov

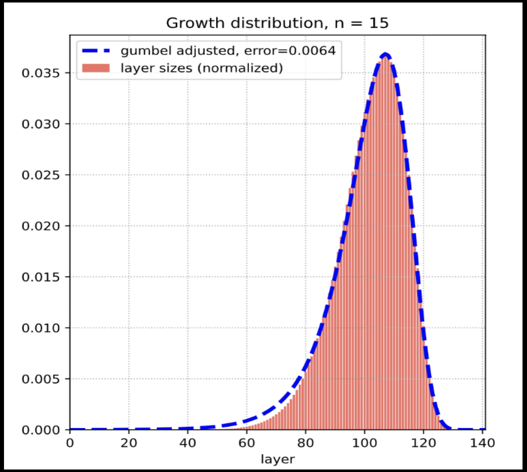
New conjecture for $n > 4$: diameter $(3n^2 - 8n + 9)/4$, n odd

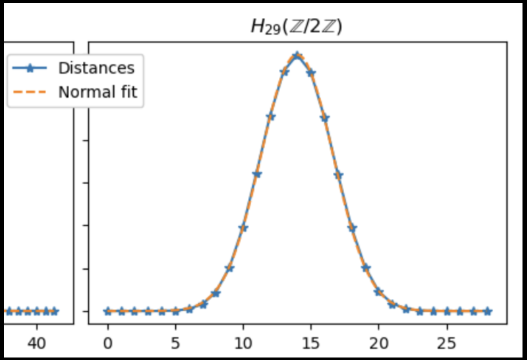
$(3n^2 - 8n + 12)/4$**, n even. Checked for $n < 16$.**

Hundreds of conjectures ("life exists beyond diameters")

Black diamond - new open conjecture, blue diamond - we can prove it

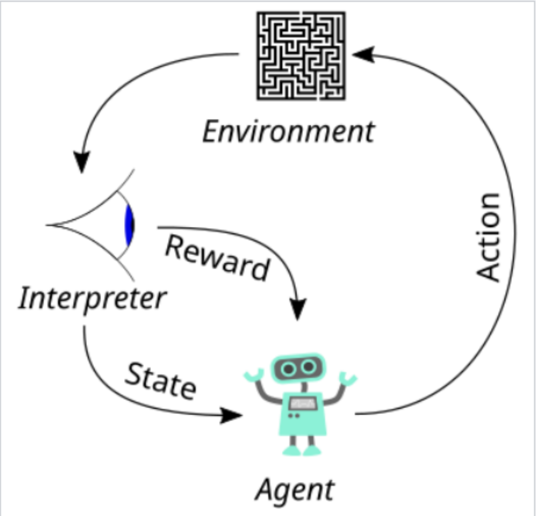
Generators	Group	Diameter	PDF	F-la	Mean	Mode	Var	Skew	Kurt	Antipodes	Algorithm	Metric	Spectrum	Mixing Time
Coxeter	S_n	$\frac{n(n-1)}{2}$	Gauss	+	$\frac{n(n-1)}{4}$	$\frac{n(n-1)}{4}$	$+$	$\to 0$	$\to 0$	1	Bubble	$+$	?	?
Cyclic Coxeter	S_n	$\left\lfloor \frac{n^2}{4} \right\rfloor$	Gauss◆	?	$0.17(n^2-n+1)$◆	≈Mean◆	?	$\to 0$◆	$\to 0$◆	1|2◆	?	+	Wig◆	?
LRX	S_n	$\frac{n(n-1)}{2}*$	Gumbel◆	?	$\approx 0.38n^2-n$◆	$\approx 0.39n^2-n$◆	?	$\to -0.7$◆	$\to 3.3$◆	◆	◆ (blue)	?	Uni◆	$> n^3$◆
LX-Glushkov	S_n	$\frac{3n^2-8n+9\|12}{4}$◆	Gumbel◆	Fib/?	$\approx 0.57n^2-2n$◆	$\approx 0.57n^2-1.6n$◆	?	$\to -0.7$◆	$\to 0.5$◆	$*$	◆	?	?	?
LARX	S_n	$\frac{n^2-2\|5}{2}$◆	?	?	$0.4n^2-0.7n$◆	≈Mean◆	?	?	?	+◆|?	?	?	?	?
LARX+I	S_n	$\frac{n(n+6)-12\|19}{4}$◆	?	?	$\approx \frac{n(n+1)}{4}$◆	≈Mean◆	?	?	?	+◆	?	?	?	?
LSL	S_n	$\frac{n(n-3)}{2}+3$◆	?	?	$\approx 0.4n^2-1.5n$◆	≈Mean◆	?	?	?	+◆	?	?	?	?
LSL+I	S_n	$\frac{n(n+4)}{4}-3\|\frac{17}{4}$◆	?	?	$\approx 0.2n^2$	$\approx 0.2n^2$	?	?	?	?	?	?	?	?
3-cyc	A_n	$\left\lfloor \frac{n}{2} \right\rfloor$	?	+−	$\approx D-0.5\|1$◆	$D-0\|1$◆	?	?	?	+	+	+	Int	?
(0ij)	A_n	$\left\lfloor \frac{3(n-1)}{4} \right\rfloor$	?	?	$\approx 0.55n$◆	$\approx Mean$◆	?	?	?	?	?	?	?	?
(01i)	A_n	$\frac{3n-5}{2}+\frac{i^n+(-i)^n}{4}$◆	?	?	$\approx n-2$◆	$n-1$◆	?	?	?	?	+◆ (blue)	?	?	?
(01i)I	A_n	$\left\lfloor \frac{3n-6}{2} \right\rfloor$ ◆ (blue)	?	?	$\approx n+1.25\ln n+\ldots$◆	≈Mean◆	◆	◆	◆	◆	◆ (blue)^{O}	◆ (blue)	Int◆ (blue)	nlog(n)◆





Reinforcement Learning as Graph Pathfinding

Standard RL (forget it):



The typical framing of a reinforcement learning (RL) scenario: an agent takes actions in an environment, which is interpreted into a reward and a state representation, which are fed back to the agent.

The dictionary (cf. Richard Sutton's table):

States	=	**Nodes of the graph**
Actions	=	**Edges**
Reward	=	**Weights of edges**
Cumulative reward	=	**Length of the path**

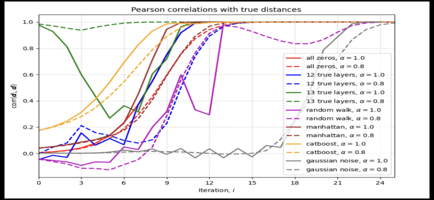
d(g) = distance from g to any e, satisfy:

$$d(g) = 1 + \min_{n:\ neigbours\ of\ g} d(n) \text{ Bellman equation}$$

$$d(e) = 0 \text{ boundary condition}$$

A "tropical" equation.

Solving Bellman = simple iteration



CayleyPy-1 approach is based on *diffusion distance*

1) **Train set via random walks**: Ran many random walks trajectories starting from identity "id" and store pairs: (g,k): g - node of graph, k - in how many steps it was achieved by random walk.

Train set - pairs: (permutation , n_steps)

2) **Model**: Train machine learning model - to predict "k" from "permutation"

3) **Pathfinder (graph search, beam search)**:

Start from node, take all its neighbors, compute neural net predictions, choose 10 best (min predictions) nodes, repeat until "id" is found

https://neurips.cc/virtual/2025/loc/mexico-city/poster/120075

Neural nets make mistakes: Beam Search to the Rescue

Beam Search as a metaphorical model of a research community:

A single researcher is exponentially less effective than a group, because alone it “cannot break the wall” and stuck in local minima.

Pipeline of all RL-like systems (e.g. AlfaGo):

1 Neural Net says where to move

(trained by "playing with itself")

2 Graph search compensates its mistakes (Monte-Carlo Tree Search, Beam Search, A*, many other global optimization algorithms)

Beam search:

Has a single parameter - size of the beam (size of research group) say 10:

1 Look on all neighbors - choose 10 best

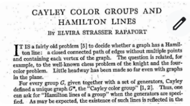
2 Iterate

Surprise: 10 cannot be compensated by 10 hours of a single “researcher”, because you stuck in local minima and only other researchers can provide breakthroughs to come out of it

Polymath, crowd-sourcing — attract 1000+ AI experts to math problem.
Easy — Kaggle just needs to “press one button”.

CayleyPy RapaportM2 - solve math mystery from 1959

Help mathematicians to find the God's number and algorithm - develop method to solve/sort permutations using E.Rapaport transformations



CAYLEY COLOR GROUPS AND HAMILTON LINES

By ELVIRA STRASSER RAPAPORT

IT IS a fairly old problem [5] to decide whether a graph has a Hamilton line: a closed connected path of edges without multiple points and containing each vertex of the graph. The question is related, for example, to the well-known chess problem of the knight and the four-color problem. Little headway has been made so far even with graphs in the plane.

For every group G, given together with a set of generators, Cayley defined a unique graph G^*, the "Cayley color group" [1, 2]. Thus, one can ask for "Hamilton lines of a group" when the generators are specified. As may be expected, the existence of such lines is reflected in the

#	Team	Members	Score	Entries	Last
1	Ruslan Grimov		120228357	3	8d
2	Alim Bijiev		153013709	3	13d

CayleyPy Pancake sorting - outperform B.Gates

Prefix (pancake) sorting algorithm was proposed by B.Gates in 1979 - can you do better ?

CayleyPy V.M.Glushkov permutations problem

Solve the problem by "father of Soviet cybernetics" V.M.Glushkov (1968)

CayleyPy Reversals - Estimate Evolution Distance

70 million years from human to mice - help to check that is true 😄 Develop reversal sorting algorithms as optimal as possible

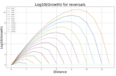


FIGURE 18. Growth for reversals

FIGURE 19. The reversal graph, $n = 4$

Kaggle is a “world championship” for AI challenges. (It belongs to Google).

Math problems: optimize 1 number + easy to check correctness are perfect for Kaggle

Challenges attract THOUSANDS participants. Competition are extremely tough! Is not about money (no chance to get them, Sasha got twice 😜).

It is about:

- Hormones — adrenaline,
- Fame / glory / honor / “Reputation” — be “Kaggle grandmaster” — job will find you by itself [vanity for good]
- … and respect

Create a challenge. If Kaggle decides to “assign reputation” then you got 1000 AI-experts…

A pancake sorting theorem of William Gates

- Gates, W.; Papadimitriou, C. (1979). "Bounds for Sorting by Prefix Reversal". *Discrete Mathematics*. **27**: 47–57. doi:10.1016/0012-365X(79)90068-2.

In 1979, Bill Gates and Christos Papadimitriou[11] gave an upper bound of 5/3 n

Pancake graph: Cayley graph for S_n, with generators: take first k pancakes, swap them and put back.

Example (k=4): 🐶🐩🌭🐕‍🦺🦮🐕‍🦺🌭🐶🐶🐶

🐕‍🦺🌭🐩🐶🦮🐕‍🦺🌭🐶🐶🐶

$$a\ b\ c\ d\ e\ d\ c\ a\ a\ a \mapsto d\ c\ b\ a\ e\ d\ c\ a\ a\ a$$

Pancake Diameter is UNKNOWN - open problem for n>20.

Diameter of Rubik's cube UNKNOWN for n >3. Etc…

A biological application: Cayley distance ≃ evolutionary distance

Transforming cabbage into turnip: polynomial algorithm for sorting signed permutations by reversals

1999 cited: 1215

Sridhar Hannenhalli, Pavel A. Pevzner

Authors: Sridhar Hannenhalli, Pavel A. Pevzner Authors Info & Claims

Transforming men into mice (polynomial algorithm for genomic distance

5.1.5. *Transposons (block transpositions).*

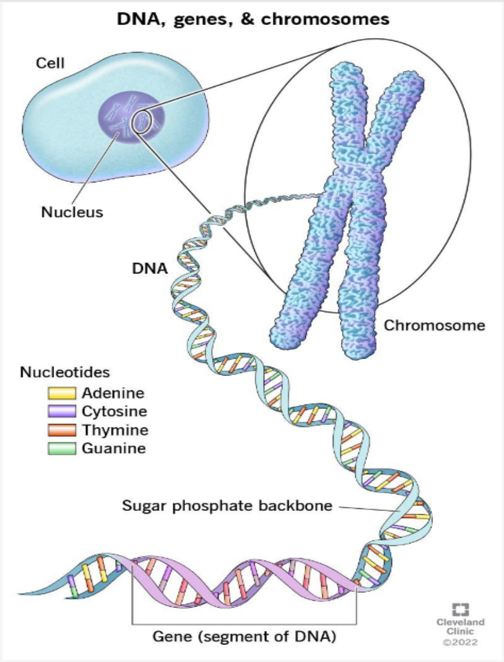
The generators called somewhat misleadingly in biological literature ”transpositions” not to be confused with standard mathematical transpositions, better called ”block transpositions”, or we will call them just ”transposons” - since their biological motivation is to describe ”transposons” (movig fragments of genome). They can be described as follows: cutting out a segment and reinserting it elsewhere on the same chromosome (that segment is abstraction of biological transposon):

$$\dots u \underbrace{v \dots w}_{\Delta} x \dots y \dots \longmapsto \dots u\, x \dots y \underbrace{v \dots w}_{\Delta} \dots .$$

They are supported by CayleyPy under the name ”transposons”.

Conjecture.(OEIS-A065603) Diameter of transposons is $\lceil \frac{n+1}{2} \rceil$ for $n \neq 13, 15$.

The conjecture formulated in [ErikssonErikssonKarlanderSvenssonWästlund01] (section 3 end), based on earlier results in [BafnaPevzner98].



[Submitted on 27 Aug 2024 (v1), last revised 11 Feb 2025 (this version, v2)]

What makes math problems hard for reinforcement learning: a case study

Ali Shehper, Anibal M. Medina-Mardones, Lucas Fagan, Bartłomiej Lewandowski, Angus Gruen, Yang Qiu, Piotr Kucharski, Zhenghan Wang, Sergei Gukov

Selman Akbulut and Robion Kirby. "A potential smooth counterexample in dimension 4 to the Poincare conjecture, the Schoenflies conjecture, and the Andrews–Curtis conjecture". *Topology* 24.4 (1985) (cit. on pp. 4, 7).

Pathfinding on an infinite graph

2. Andrews–Curtis conjecture

The Andrews–Curtis Conjecture concerns the study of *balanced presentations* of the trivial group, i.e. presentations of the trivial group with an equal number of generators and relators. The conjecture proposes that any balanced presentation of the trivial group

$$\langle x_1, \ldots, x_n \mid r_1, \ldots, r_n \rangle$$

can be converted to the trivial presentation

$$\langle x_1, \ldots, x_n \mid x_1, \ldots, x_n \rangle$$

through a series of the following operations known as *AC-moves* [AC65]:

(AC1) Substitute some r_i by $r_i r_j$ for $i \neq j$.
(AC2) Replace some r_i by r_i^{-1}.
(AC3) Change some r_i to $g r_i g^{-1}$ where g is a generator or its inverse.

A notable family of potential counterexamples, denoted $\mathrm{AK}(n)$, due to Akbulut and Kirby,

$$\mathrm{AK}(n) = \langle x, y \mid x^n = y^{n+1}, xyx = yxy \rangle, \quad n \geq 3$$

has been open in mathematics for more than four decades [AK85].[3] The length of the presentation $\mathrm{AK}(n)$ is $2n+7$, and until now, it was unknown whether this length can even be reduced using AC moves. We find that for every fixed n, $\mathrm{AK}(n)$ is AC-equivalent to each element of the following 1-parameter family of presentations parameterized by $k \in \mathbb{Z}$,

$$P(n,k) = \langle x, y \mid y^{n-k-1} x^{-1} yx = xyx^{-1} y^{n-k} \ , \ x = y^{-k} x^{-1} yxy \rangle.$$

When $n \geq 5$, this family contains presentations of length less than $2n+7$. The shortest presentation occurs in the case $k = n-1$, which gives us the following result, proven as Theorem 4 in the main text.

Theorem A. *For every $n \geq 2$, $AK(n)$ is AC-equivalent to the presentation*

$$\langle x, y \mid x^{-1} yx = xyx^{-1} y \ , \ xy^{n-1} x = yxy \rangle,$$

of length $n + 11$. This gives a reduction in length of $AK(n)$ for all $n \geq 5$.

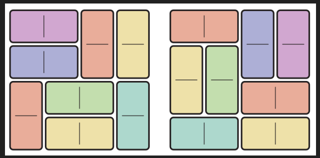
Another important family of potential counterexamples, due to Miller and Schupp, is

$$\mathrm{MS}(n,w) = \langle x, y \mid x^{-1} y^n x = y^{n+1}, x = w \rangle,$$

where $n > 0$ and w is a word with exponent sum zero on x. It has been open for more than 25 years [MS99].

E AÍ ? What's next? Quo vadis, CayleyPy-4? Am, а дальше как?

scale, Scale, SCALE, Scale, SCALE — more compute is most welcome!!!



: ICLR'26 @ Rio de Janeiro

Graphs from *groupoids* — 15-puzzle, dimer models, dominoes, perfect matchings, flip graphs, … regular graphs,… finite pieces of infinite groups, pathfinding on more general graphs…

Scale collaboration (bottlenecks — interfaces).

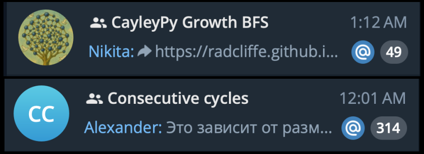
Model for new collaborations.

If you can just get your mind together, then come on across to me.

What simple game

Could be a faithful representation space

For (mathematical) research?



Then which simple game could be a faithful representation for (mathematical) research?

MATH IS NOT THE PROBLEM TO BE SOLVED BUT THE REALITY TO BE EXPERIENCED.

'Great. we can drink sake and talk maths.'
'You mean talk maths and murder.'
Yukawa shrugged and wrinkled his brow. 'Maybe so. Though I did come up with a new problem for you .... Which is harder: devising an unsoluble problem, or solving that problem? And it's not an empty question. I guarantee this puzzle has an answer. Interesting, no?'

Have you ever been experienced?
Well, I have
And let me prove this to you

```
import Mathlib
```

CayleyPy 3 collaboration: Alexey Aparnev, Alim Bidzhiev, Simo Alami Chehboune, Lyudmila Cheldieva, Alexander Chervov, Artem Chevychelov, Orianne Debeaupuis, Stanislav Diner, Antonina Dolgorukova, Mikhail Evseev, Dmytro Fedoriaka, Sergey Galkin, Elizaveta Geraseva, Ruslan Grimov, Liliya Grunwald, Artem Isakov, Herve Isambert, Igor Kiselev, Zakhar Kogan, Eduard Koldunov, Ivan Koltsov, Elena Konstantinova, Anton Kostin, Dmitrii Kovalenko, Artem Krasnyi, Alexey Kravchenko, Evelina Kudasheva, Sergey Kudashev, Fedor Levkovich-Maslyuk, Michael Litvinov, Sergei Lytkin, Dariya Mamayeva, Altana Natyrova, Anton Naumov, Olga Nikitina, Sergei Nikolenko, Mark Obozov, Oleg Papulov, Anastasia Sheveleva, Dmitry Shiltsov, Lidia Shishina, Andrei Smolensky, Alexander Soibelman, Arsenii Sychev, Anton Titarenko, Rustem Turtayev, Nick Vilkin-Krom, Dmitry Volovich, Vladislav Zamkovoy

not necessarily stoned, but beautiful