AI в науке

Где мы сейчас и куда мы идём

@SINECOR

Сергей Николенко

26 июля 2026

ЦЕНТРАЛЬНЫЙ УНИВЕРСИТЕТ

ЦЕНТРАЛЬНЫЙ УНИВЕРСИТЕТ

Наш план

- История AI в математике
- Результаты чистых LLM
- AI в других науках
- Последние новости
- Заключение

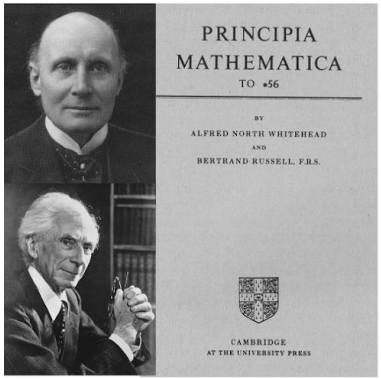
1. Современный AI в математике

Я так привык к самостоятельной работе, что считал более лёгким для себя доказать теорему без книги, чем вычитывать из нее доказательства. Лишь не всегда это удавалось.

Константин Циолковский

Автоматическое доказательство теорем

- Работа математика – это интересный случай, который всегда оставался для меня загадкой
- В математической логике давно развиваются методы автоматического доказательства теорем



$*54{\cdot}43.\quad \vdash :. \alpha, \beta \in 1 . \supset : \alpha \cap \beta = \Lambda . \equiv . \alpha \cup \beta \in 2$

Dem.

$\vdash . *54{\cdot}26 . \supset \vdash :. \alpha = \iota`x . \beta = \iota`y . \supset : \alpha \cup \beta \in 2 . \equiv . x \neq y .$

$[*51{\cdot}231] \qquad \equiv . \iota`x \cap \iota`y = \Lambda .$

$[*13{\cdot}12] \qquad \equiv . \alpha \cap \beta = \Lambda \qquad (1)$

$\vdash . (1) . *11{\cdot}11{\cdot}35 . \supset$

$\vdash :. (\exists x, y) . \alpha = \iota`x . \beta = \iota`y . \supset : \alpha \cup \beta \in 2 . \equiv . \alpha \cap \beta = \Lambda \qquad (2)$

$\vdash . (2) . *11{\cdot}54 . *52{\cdot}1 . \supset \vdash . \text{Prop}$

From this proposition it will follow, when arithmetical addition has been defined, that $1 + 1 = 2$.

Автоматическое доказательство теорем

- Конечно, в худшем случае всё это безнадёжно…
- Но нам же не нужен худший случай! И всё равно ничего не получается

The same statement

*A Mathematical Incompleteness in Peano Arithmetic**

JEFF PARIS
LEO HARRINGTON

Weaker proof system

$$(\wedge\to)\ \frac{\varphi\psi\Gamma\to\Delta}{(\varphi\wedge\psi)\Gamma\to\Delta};\quad(\to\wedge)\ \frac{\Gamma\to\Delta\varphi;\Gamma\to\Delta\psi}{\Gamma\to\Delta(\varphi\wedge\psi)};$$

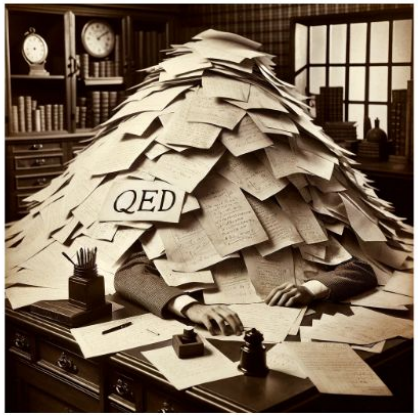
$$(\vee\to)\ \frac{\varphi\Gamma\to\Delta;\psi\Gamma\to\Delta}{(\varphi\vee\psi)\Gamma\to\Delta};\quad(\to\vee)\ \frac{\Gamma\to\Delta\varphi\psi}{\Gamma\to\Delta(\varphi\vee\psi)};$$

$$(\supseteq\to)\ \frac{\Gamma\to\Delta\varphi;\psi\Gamma\to\Delta}{(\varphi\supseteq\psi)\Gamma\to\Delta};\quad(\to\supseteq)\ \frac{\varphi\Gamma\to\Delta\psi}{\Gamma\to\Delta(\varphi\supseteq\psi)};$$

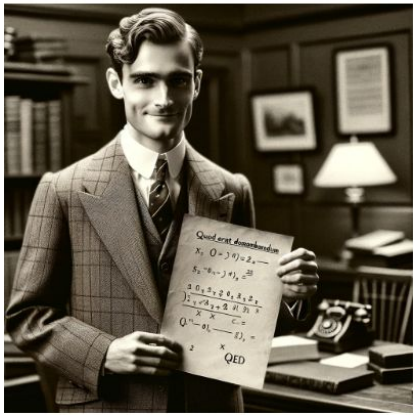
$$(\neg\to)\ \frac{\Gamma\to\Delta\varphi}{\neg\varphi\Gamma\to\Delta};\quad(\to\neg)\ \frac{\varphi\Gamma\to\Delta}{\Gamma\to\Delta\neg\varphi};$$

$$(\forall\to)\ \frac{\forall x\varphi\varphi(x\mid t)\Gamma\to\Delta}{\forall x\varphi\Gamma\to\Delta};\quad(\to\forall)\ \frac{\Gamma\to\Delta\varphi}{\Gamma\to\Delta\forall x\varphi(y\mid x)};$$

$$(\exists\to)\ \frac{\varphi\Gamma\to\Delta}{\exists x\varphi(y\mid x)\Gamma\to\Delta};\quad(\to\exists)\ \frac{\Gamma\to\Delta\varphi(x\mid t)\exists x\varphi}{\Gamma\to\Delta\exists x\varphi}.$$

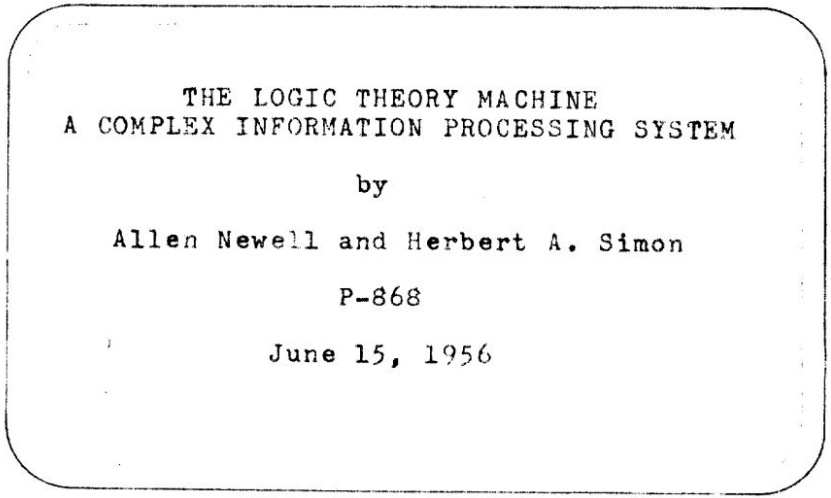


Stronger proof system



Logic Theorist

- Вся область AI началась с прувера: Logic Theorist (Newell, Simon, 1956)
- Herbert Simon, Jan 1956: “Over Christmas, Al Newell and I invented a thinking machine”
- Он же, позже: “[we] invented a computer program capable of thinking non-numerically, and thereby solved the venerable mind-body problem, explaining how a system composed of matter can have the properties of mind”.



THE LOGIC THEORY MACHINE
A COMPLEX INFORMATION PROCESSING SYSTEM

by

Allen Newell and Herbert A. Simon

P-868

June 15, 1956

The two connectives, - and v, are taken as primitives. The third connective, → , is defined in terms of the other two, thus:

1.01 $p \rightarrow q \ =_{def} \ -p \vee q$

The five axioms that are postulated to be true are:

1.2 $p \vee p .\rightarrow. p$

1.3 $p .\rightarrow. q \vee p$

1.4 $p \vee q .\rightarrow. q \vee p$

1.5 $p .\vee. q \vee r :\rightarrow: q .\vee. p \vee r$

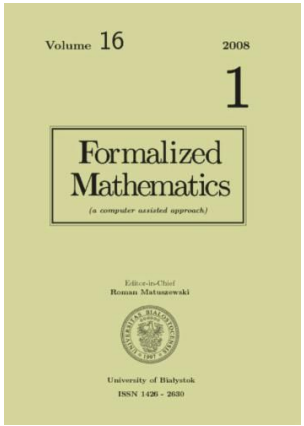
1.6 $p \rightarrow q .\rightarrow: r \vee p .\rightarrow. r \vee q$

Символьные вычисления

- 1960-1980-е: в основном символьные вычисления
  - MACSYMA (Pavelle and Wang, 1985; Fateman, 1982): Project MAC's SYmbolic MAnipulator
  - Потом Matlab, Maple, Mathematica…
- Продолжают появляться пруверы и proof assistants:
  - Automath (de Bruijn, 1967)
  - LCF (Logic for Computable Functions; Milner, 1972)
  - Mizar (1973)
  - в конце 1980-х появились Coq (1989) и HOL (1988)

Но всё это было не совсем то…

- Но новых теорем как-то не доказывалось
- Есть целый раздел формализованной математики, что-то удаётся формализовать, но в главном журнале в 2023 году всё ещё решают задачки для (продвинутых) студентов из книги Серпинского…
- QED Manifesto (1993) тоже как-то быстро затих



Open Access

Elementary Number Theory Problems. Part X – Diophantine Equations

Artur Korniłowicz | Dec 31, 2023

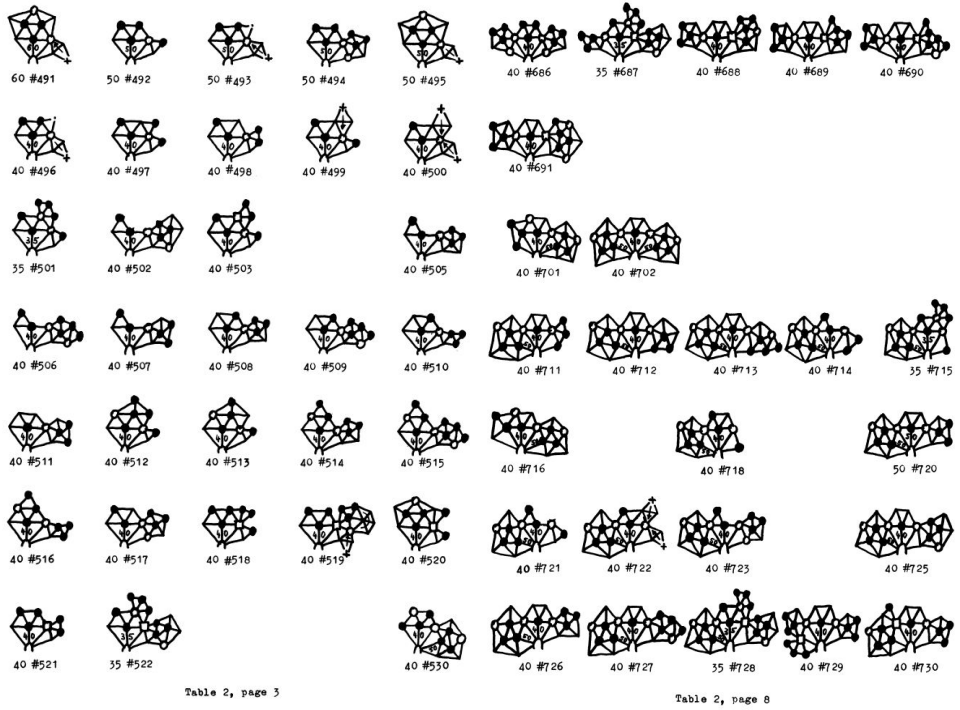
Гипотеза Роббинса

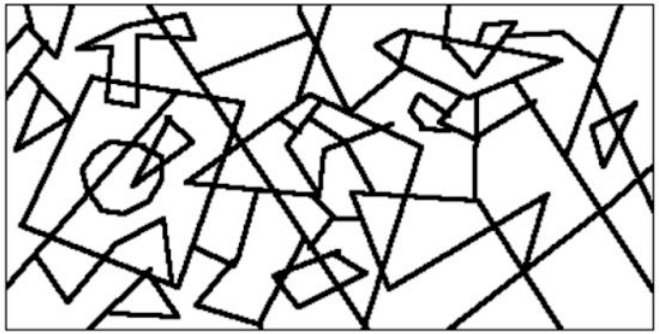
- Наверное, самый яркий положительный пример – гипотеза Роббинса о булевых алгебрах (McCune, 1996)
- Но это настолько близко к аксиомам, насколько возможно, а доказательство просто большое и механическое…

1. Associativity: $a \vee (b \vee c) = (a \vee b) \vee c$
2. Commutativity: $a \vee b = b \vee a$
3. *Robbins equation*: $\neg\left(\neg\left(a \vee b\right) \vee \neg\left(a \vee \neg b\right)\right) = a$

Задача четырёх красок

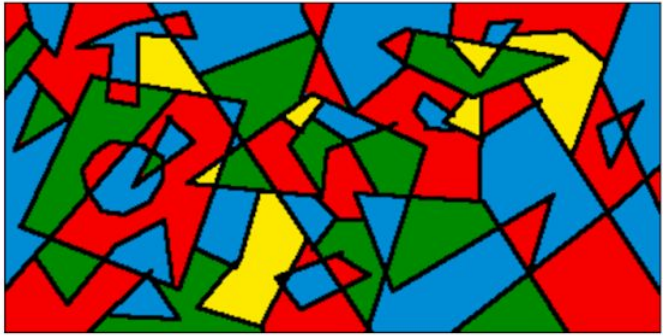
- Проблему четырёх красок решили “при помощи компьютера” (Appel, Haken, 1976)
- Но и это не логическое доказательство, а перебор случаев программой, которую написал человек; [список других примеров]





Kenneth Appel and Wolfgang Haken

Postmark of the Illinois Department of Mathematics

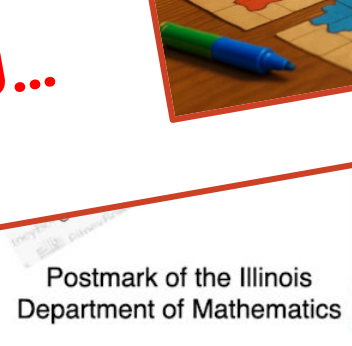


Задача четырёх красок

- Проблему четы
“при р
Hake
- Но э
дока
прогр
челов

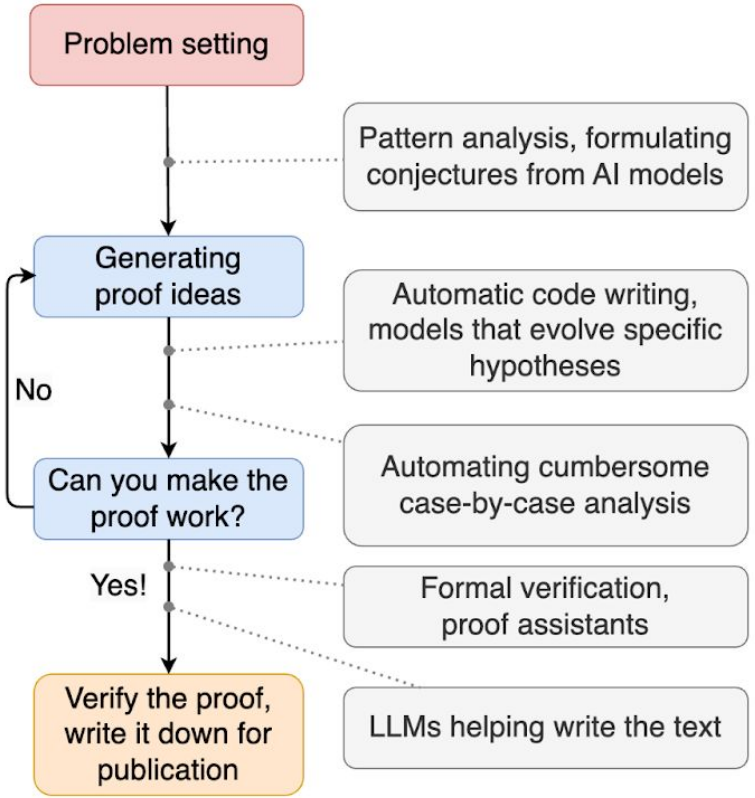
А что же сейчас?
У нас же расцвет глубокого обучения, большие кластеры мощных GPU…

Kenneth Appel and Wolfgang Haken



Postmark of the Illinois Department of Mathematics

LLM для математики



Advancing mathematics by guiding human intuition with AI

Alex Davies, Petar Veličković, Lars Buesing, Sam Blackwell, Daniel Zheng, Nenad Tomašev, Richard Tanburn, Peter Battaglia, Charles Blundell, András Juhász, Marc Lackenby, Geordie Williamson, Demis Hassabis & Pushmeet Kohli

Nature **600**, 70–74 (2021) | Cite this article

ToRA: A Tool-Integrated Reasoning Agent for Mathematical Problem Solving

Zhibin Gou[1,2]*, Zhihong Shao[1,2]*, Yeyun Gong[2], Yelong Shen[2]
Yujiu Yang[1], Minlie Huang[1], Nan Duan[2], Weizhu Chen[2]
[1]Tsinghua University [2]Microsoft
{gzb22,szh19}@mails.tsinghua.edu.cn
{yegong,yeshe,nanduan,wzchen}@microsoft.com

Discovering faster matrix multiplication algorithms with reinforcement learning

Alhussein Fawzi, Matej Balog, Aja Huang, Thomas Hubert, Bernardino Romera-Paredes, Mohammadamin Barekatain, Alexander Novikov, Francisco J. R. Ruiz, Julian Schrittwieser, Grzegorz Swirszcz, David Silver, Demis Hassabis & Pushmeet Kohli

Nature **610**, 47–53 (2022) | Cite this article

Mathematical discoveries from program search with large language models

Bernardino Romera-Paredes, Mohammadamin Barekatain, Alexander Novikov, Matej Balog, M. Pawan Kumar, Emilien Dupont, Francisco J. R. Ruiz, Jordan S. Ellenberg, Pengming Wang, Omar Fawzi, Pushmeet Kohli & Alhussein Fawzi

Nature **625**, 468–475 (2024) | Cite this article

BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 82, Number 5, September 1976

RESEARCH ANNOUNCEMENTS

EVERY PLANAR MAP IS FOUR COLORABLE[1]

BY K. APPEL AND W. HAKEN

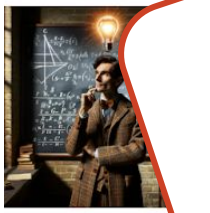
Communicated by Robert Fossum, July 26, 1976

Home About Coq Get Coq Documentation Community
The Coq Proof Assistant

Community

GPT-4

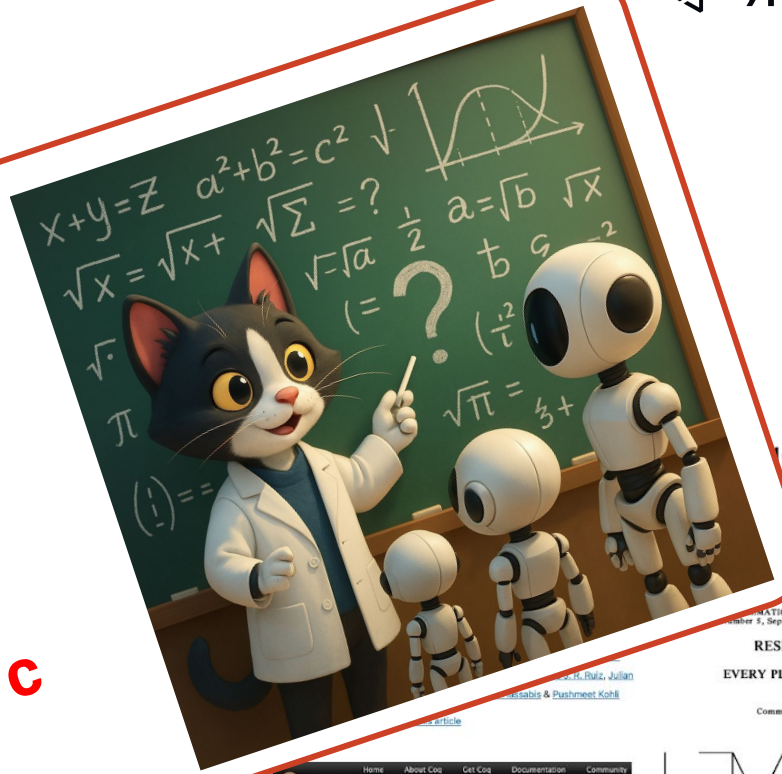
LLM для математики



Problem setting

LLMs helping write the text

Схему про LLM для математики я сделал в начале 2024 года; что изменилось с тех пор?



discoveries from program
ge language models

Mohammadamin Barekatain, Alexander Novikov, Matej
Dupont, Francisco J. R. Ruiz, Jordan S. Ellenberg,
shmeet Kohli & Alhussein Fawzi

Cite this article

RESEARCH ANNOUNCEMENTS

EVERY PLANAR MAP IS FOUR COLORABLE[1]

BY K. APPEL AND W. HAKEN

Communicated by Robert Fossum, July 26, 1976

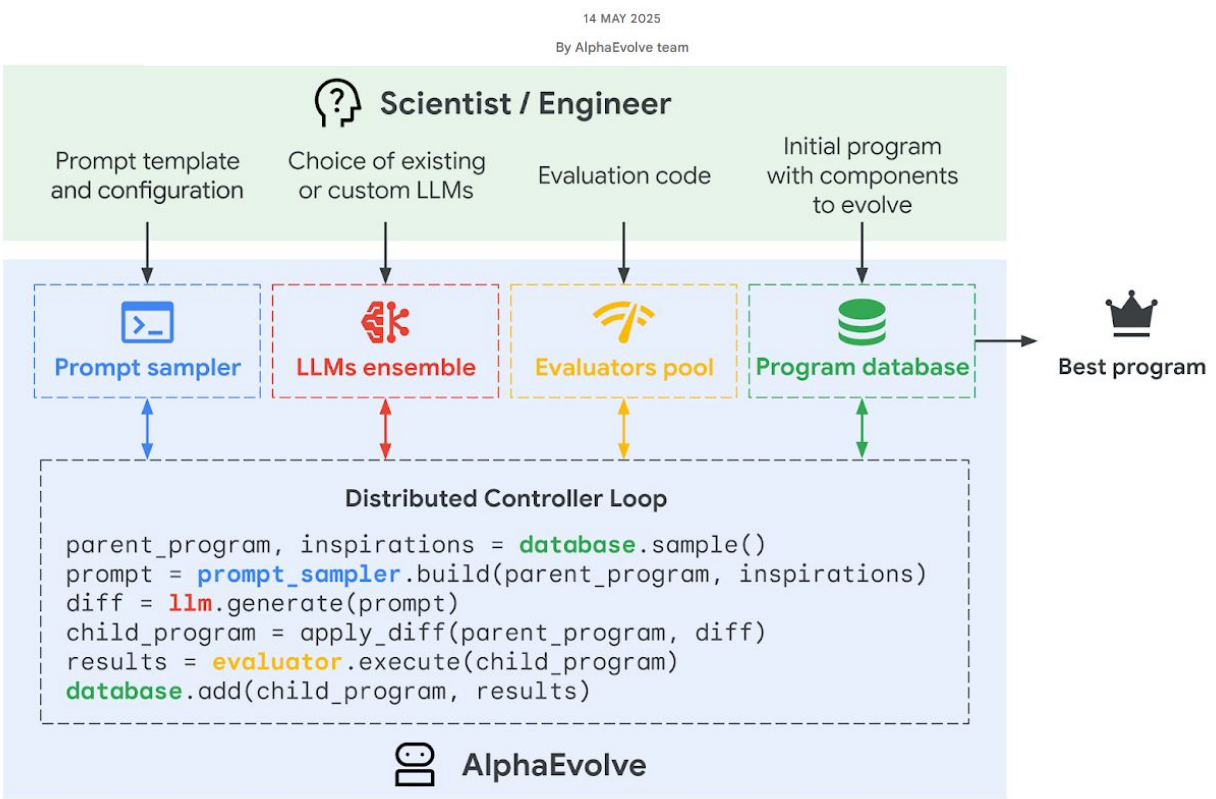
Home About Coq Get Coq Documentation Community
The Coq Proof Assistant

Community

GPT-4

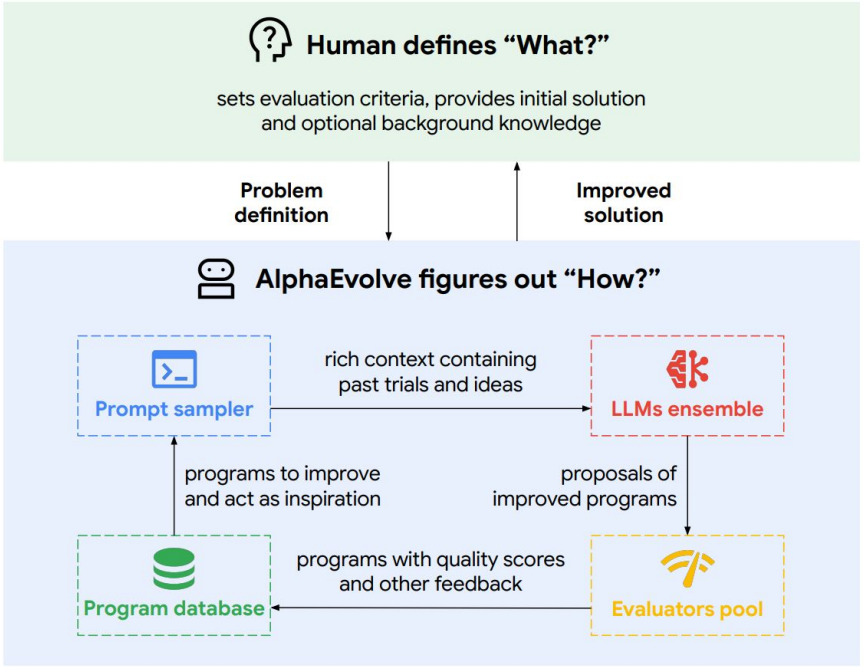
AlphaEvolve

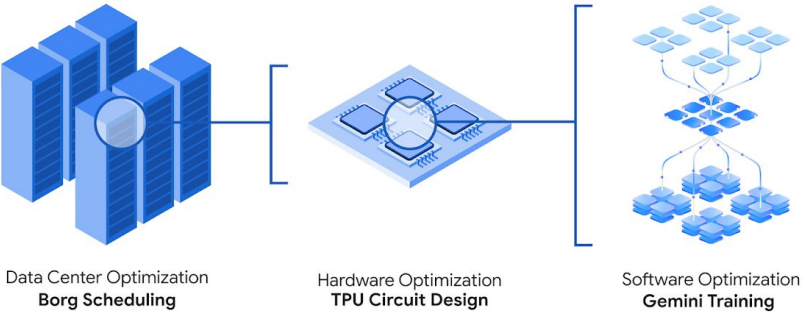
- AlphaEvolve (DeepMind, May 14, 2025): продолжение и улучшение FunSearch и AlphaTensor, несколько LLM, которые пишут, критикуют и тестируют код для разных задач



AlphaEvolve

- Новые результаты в разных областях:
  - новый scheduling в датацентрах – уже в Borg
  - улучшенная схема для умножения матриц (прямо на Verilog) – уже в TPU
  - новые алгоритмы: улучшил Штрассена для 4x4 над ℂ, чего AlphaTensor раньше не смог

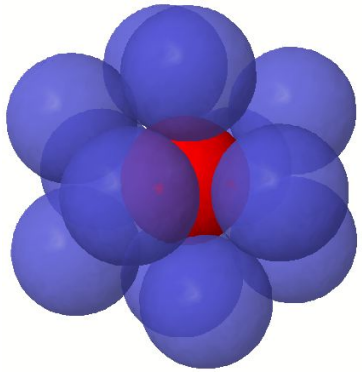




FunSearch [80]	*AlphaEvolve*
evolves single function	evolves entire code file
evolves up to 10-20 lines of code	evolves up to hundreds of lines of code
evolves code in Python	evolves any language
needs fast evaluation (≤ 20min on 1 CPU)	can evaluate for hours, in parallel, on accelerators
millions of LLM samples used	thousands of LLM samples suffice
small LLMs used; no benefit from larger	benefits from SOTA LLMs
minimal context (only previous solutions)	rich context and feedback in prompts
optimizes single metric	can simultaneously optimize multiple metrics

AlphaEvolve

- Новые результаты в разных областях:
  - новая нижняя оценка на контактное число (kissing number) в размерности 11
  - сразу несколько улучшенных оценок в анализе, геометрии, комбинаторике…



$$\max_{-1/2\le t\le 1/2}\int_{\mathbb{R}} f_1(t-x)f_1(x)\,dx \ge C_1\left(\int_{-1/4}^{1/4} f_1(x)\,dx\right)^2$$

1.5098 → **1.5053**

$$\|f*f\|_2^2 \le C_2\|f*f\|_1\|f*f\|_\infty$$

0.8892 → **0.8962**

$$\max_{-1/2\le t\le 1/2}\left|\int_{\mathbb{R}} f(t-x)f(x)\,dx\right| \ge C_3\left(\int_{-1/4}^{1/4} f(x)\,dx\right)^2$$

1.4581 → **1.4557**

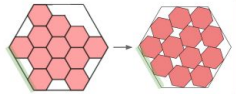
$$A(f)A(\hat{f}) \ge C_4$$

0.3523 → **0.3521**

Analysis

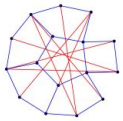
Hexagon outer edge

4.000 → **3.942**



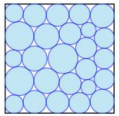
Max distance/min distance

12.890 → **12.889**

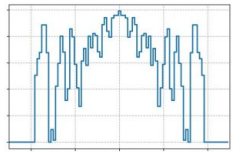


Sum of radii

2.6340 → **2.6358**



Geometry



$$\sup_{x\in[-2,2]}\int_{-1}^{1} f(t)g(x+t)\,dt \ge C_5$$

0.380926 → **0.380924**

$$|A+B| \ll |A|$$

$$|A-B| \gg |A|^{C_6}$$

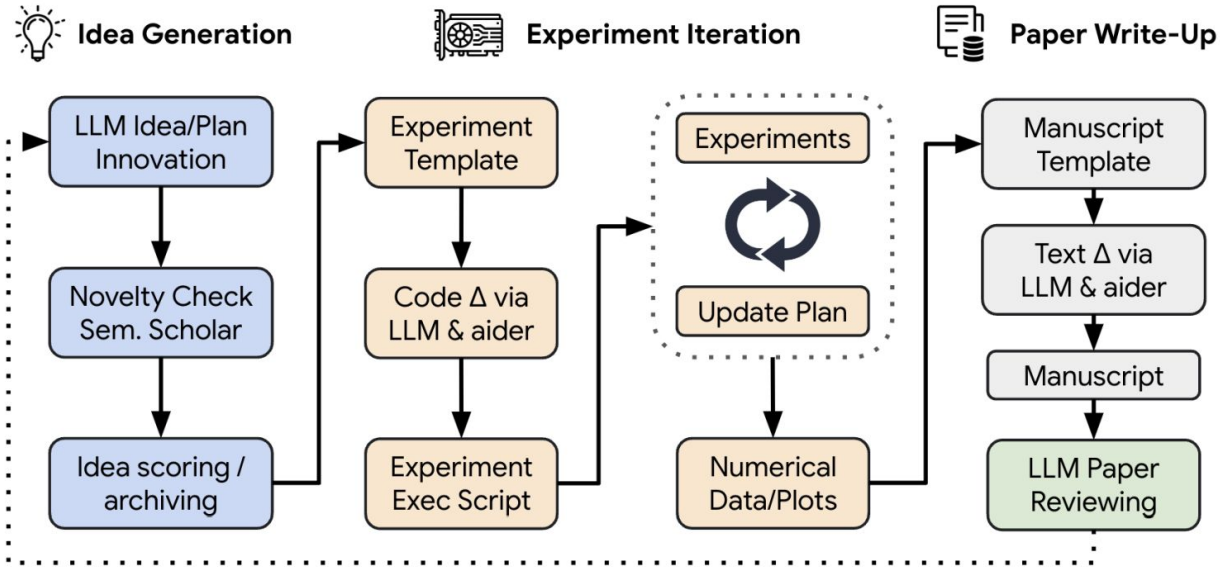
1.1446 → **1.1584**

Combinatorics

Пример из августа 2024-го

Это мой слайд из ноября 2024

- AI researcher (Lu et al., August 12, 2024): система (которую вы можете сами установить) ходит к нескольким API (LLM, Semantic Scholar), умеет использовать информацию и ресурсы компьютера (сохранять веса моделей) и самостоятельно писать и запускать код экспериментов
- Статьи пока не гениальные, но когда они получаются автоматически… и, опять же, даже o1 наверняка это значительно улучшит



AI Scientist-v2

- Guess what? Получилось!
- AI Scientist-v2 (Sakana AI, 12 марта 2025) смогла написать статью, которая прошла на ICLR 2025 Workshop “I Can't Believe It's Not Better: Challenges in Applied Deep Learning”!
- Задали тему, написали несколько статей полностью автономно, end-to-end, выбрали три лучших, подали на workshop – и вот результаты:

Title	ICLR Workshop Scores
Compositional Regularization: Unexpected Obstacles in Enhancing Neural Network Generalization	6, 7, 6
Real-World Challenges in Pest Detection Using Deep Learning: An Investigation into Failures and Solutions	3, 7, 4
Unveiling the Impact of Label Noise on Model Calibration in Deep Learning	3, 3, 3

AI Scientist-v2

- Это, видимо, первая по-настоящему полностью автоматически порождённая статья, прошедшая серьёзный peer review и принятая в хорошее место

000
001
002
003
004
005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025

COMPOSITIONAL REGULARIZATION: UNEXPECTED OBSTACLES IN ENHANCING NEURAL NETWORK GENERALIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Neural networks excel in many tasks but often struggle with compositional generalization—the ability to understand and generate novel combinations of familiar components. This limitation hampers their performance on tasks requiring systematic reasoning beyond the training data. In this work, we introduce a training method that incorporates an explicit compositional regularization term into the loss function, aiming to encourage the network to develop compositional representations. Contrary to our expectations, our experiments on synthetic arithmetic expression datasets reveal that models trained with compositional regularization do not achieve significant improvements in generalization to unseen combinations compared to baseline models. Additionally, we find that increasing the complexity of expressions exacerbates the models' difficulties, regardless of compositional regularization. These findings highlight the challenges of enforcing compositional structures in neural networks and suggest that such regularization may not be sufficient to enhance compositional generalization.

Zochi

- Но не последняя! Система Zochi (Intology AI, March 2025) автономно написала статью, которую приняли на ACL 2025, на главный трек!

The 63rd Annual Meeting of the Association for Computational Linguistics

Your Submissions | Author Tasks

#	Submission Summary	Official Review	Decision
5156	Tempest: Automatic Multi-Turn Jailbreaking of Large Language Models with Tree Search Andy Zhou, Ron Arel	0 Official Reviews Submitted Average Rating: N/A (Min: N/A, Max: N/A) Average Confidence: N/A (Min: N/A, Max: N/A)	ACL 2025 Main Recommendation: Accept (Main)

Meta Review of Submission5987 by Area Chair GMh9

Meta Review by Area Chair GMh9 | 09 Apr 2025, 14:20 (modified: 24 Apr 2025, 10:02) | Senior Area Chairs, Area Chairs, Authors, Reviewers Submitted, Program Chairs, Commitment Readers

Revisions

Metareview:

The paper introduces TEMPEST, a multi-turn adversarial framework that employs a tree search approach to assess the safety of Large Language Models (LLMs) by exploring multiple conversation paths and tracking incremental policy breaches. By utilizing a cross-branch learning mechanism, TEMPEST efficiently identifies model vulnerabilities and demonstrates how small compliance concessions can accumulate into significant security failures. Evaluations show TEMPEST achieves 100% success rates on GPT-3.5-turbo and 97% on GPT-4, surpassing single-turn and existing multi-turn adversarial methods. This innovative methodology emphasizes the importance of simultaneous exploration in comprehensive safety testing of language models.

Summary Of Reasons To Publish:

- All reviewers agreed that the tree search methodology represents a significant advancement by allowing multiple conversation paths to be explored simultaneously, rather than depending on a single dialogue thread.
- Reviewers appreciated the innovative ability of the framework to quantify incremental safety erosion and exploit minor policy breaches, which contributed to TEMPEST achieving an impressive success rate of 97-100% against the evaluated LLMs.
- Reviewers highlighted the paper's clear and fluent structure, noting that it is well-organized and logically detailed, particularly in the methodology section.
- The intriguing combination of minor concessions within the tree search context and the unique perspective in revealing security vulnerabilities of LLMs were also praised as distinctive aspects of the approach.

Summary Of Suggested Revisions:

- *Limitation of Model Evaluation*: A reviewer wished authors had evaluated a broader range of models beyond just GPT-3.5-Turbo, GPT-4, and Llama-3.1-70B, noting the omission of important families like Claude, Gemini, and others, which limits the generalizability of the findings. Authors have conducted additional experiments on a more diverse set of models and will include these in the final version of the paper.
- *Insufficient Dataset Validation*: A reviewer noted that the paper's evaluation relied solely on the JailbreakBench dataset, questioning its effectiveness across various types of harmful content and adversarial scenarios. Authors agreed and have thus extended their evaluation to include HarmBench and StrongReject datasets; the results confirm TEMPEST's effectiveness, which they will include in the paper.
- *Missing Related Work*: A reviewer felt that the paper fails to discuss several key related works in multi-turn jailbreaking and red teaming, which could enhance the overall context and relevance of the research. Authors have acknowledged this and will incorporate discussions of these related works in the revised version.
- *Lack of Detail in Methodology*: A reviewer indicated that the paper could benefit from more detailed methodological explanations, such as providing full prompts and additional context regarding the attacker LLM. Authors addressed this concern in their rebuttal and should include these details in the final version.
- *Need for In-depth Analysis*: One reviewer suggested that authors should elaborate on the accumulation of minor concessions leading to fully disallowed outputs to clarify this aspect of their findings. Authors should expand on this analysis in the final version.
- *Inclusion of Defense Experiments*: A reviewer recommended that authors include defense experiments, such as testing against models like Llama Guard-3, to provide a comprehensive view of TEMPEST's effectiveness. Authors agreed, showed preliminary experiment, and will include a full set of defense experiments in the final version of the paper.
- *Concerns Regarding Attacker LLM's Capabilities*: One reviewer pointed out that the proposed method relies on a capable attacker LLM, raising concerns about its ability to generate harmful responses itself. Authors have addressed this in the rebuttal and will clarify these concerns in the final version.
- *Method's Efficiency Issues*: One reviewer expressed that the efficiency of the method is a concern due to the reliance on exploring multiple jailbreaking prompts through tree search. Authors have responded by demonstrating that their method is actually more efficient than existing methods and will include this information in the next version of the paper.

Overall Assessment: 4.0 = Conference: I think this paper could be accepted to an *ACL conference.

Reported Issues: No

Zochi

- Но не последняя! Система Zochi (Intology AI, March 2025) автономно написала статью, которую приняли на ACL 2025, на главный трек!
- Кстати, о самой статье тоже интересно поговорить...

Tempest: Automatic Multi-Turn Jailbreaking of Large Language Models with Tree Search

Andy Zhou*
Intology AI
andy@intology.ai

Ron Arel*
Intology AI
ron@intology.ai

Abstract

We introduce Tempest, a multi-turn adversarial framework that models the gradual erosion of Large Language Model (LLM) safety through a *tree search* perspective. Unlike single-turn jailbreaks that rely on one meticulously engineered prompt, Tempest expands the conversation at each turn, branching out multiple adversarial prompts that exploit partial compliance from previous responses. Through a cross-branch

2024a; Ren et al., 2024; Zhao and Zhang, 2025; Yu et al., 2024). The dynamic nature of chat interfaces presents unique challenges for safety testing, as adversaries can adapt their strategies based on model responses and gradually accumulate partial compliance across multiple turns.

Traditional approaches to evaluating LLM safety have focused primarily on single-turn attacks, where carefully engineered prompts attempt to elicit harmful responses in one shot (Zou et al.,

Zochi

- Но не последняя! Система Zochi (Intology AI, March 2025) автономно написала статью, которую приня на ACL 202 главный
- Кстати, о ca статье тоже интересно поговорить...

Это было весной 2025 года. А что сейчас? Какие новости?..

Tempest: Automatic Mu rge Language Models

el*
gy AI
eintology.ai

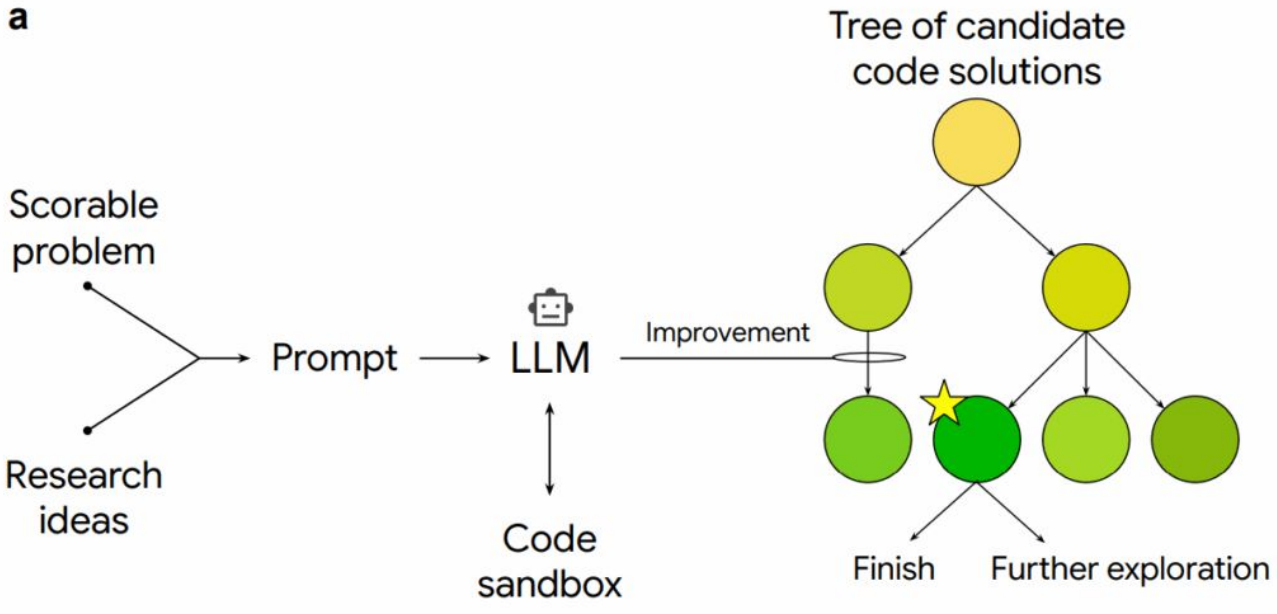
n adversarial
gradual erosion of
el (LLM) safety through a
spective. Unlike single-turn jail-
t rely on one meticulously engineered
mpt, Tempest expands the conversation at
each turn, branching out multiple adversarial
prompts that exploit partial compliance from
previous responses. Through a cross-branch

2024a; Ren et al., 2024; Zhao and Zhang, 2025; Yu et al., 2024). The dynamic nature of chat interfaces presents unique challenges for safety testing, as adversaries can adapt their strategies based on model responses and gradually accumulate partial compliance across multiple turns.

Traditional approaches to evaluating LLM safety have focused primarily on single-turn attacks, where carefully engineered prompts attempt to elicit harmful responses in one shot (Zou et al.,

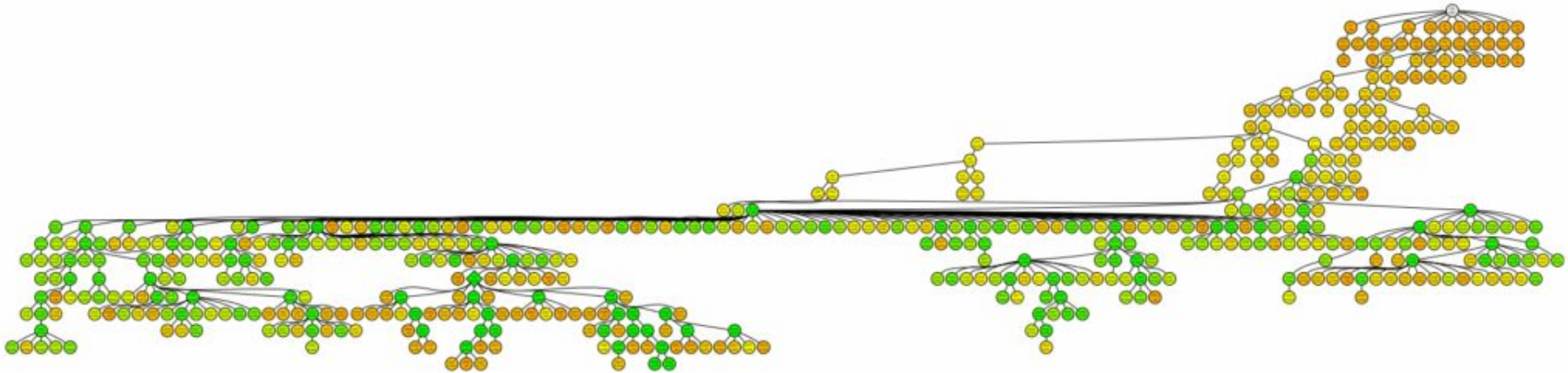
Побей все бенчмарки до завтра

- Aygün et al. (DeepMind, September 8, 2025): сочетание поиска по дереву (как в AlphaGo), добавления идей из научной литературы (через LLM) и рекомбинации методов (как в генетическом программировании)



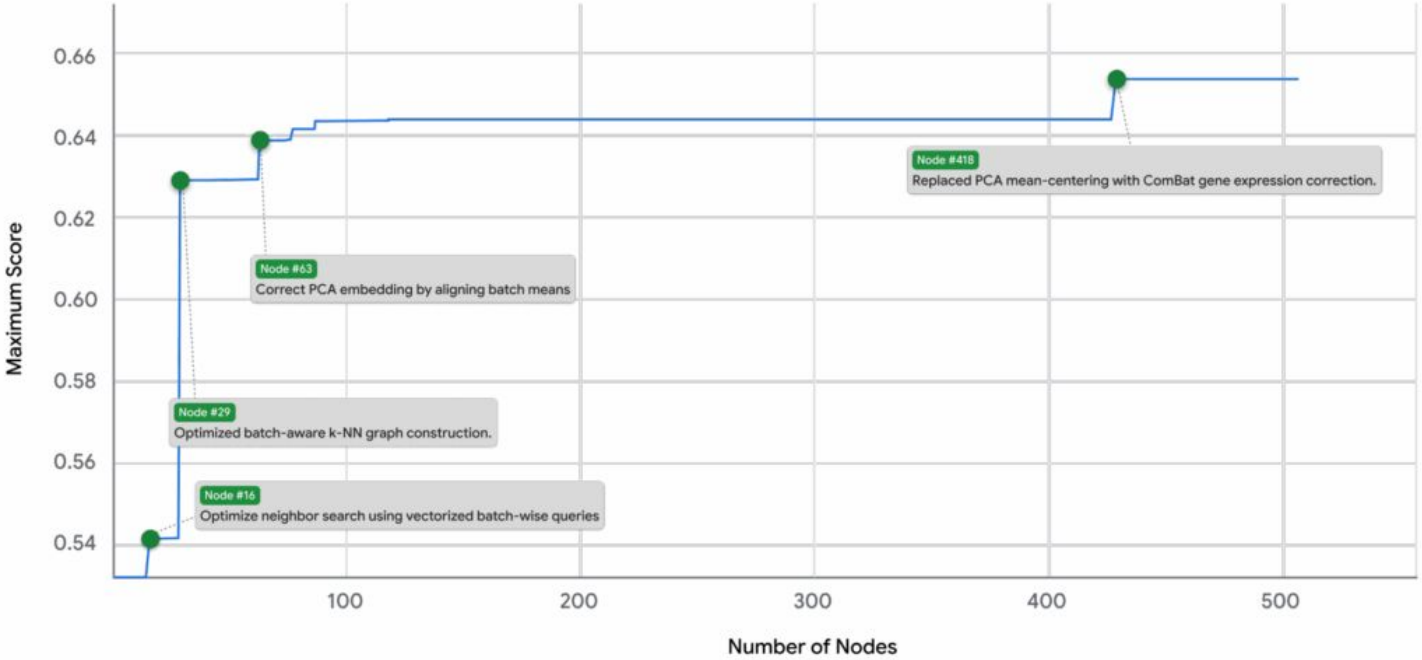
Побей все бенчмарки до завтра

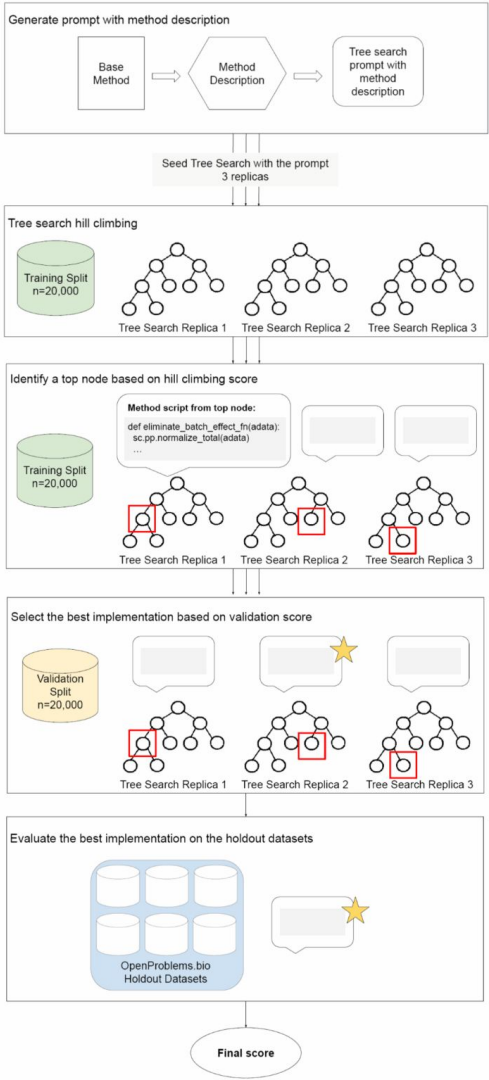
- Aygün et al. (DeepMind, September 8, 2025): системе нужна численная оценка, которую нужно улучшить, возможность писать код и читать статьи
- Дальше поиск по дереву, где узлы – это версии программы
- Если какая-то модификация улучшила оценку, система с большей вероятностью будет исследовать эту ветку дальше



Побей все бенчмарки до завтра

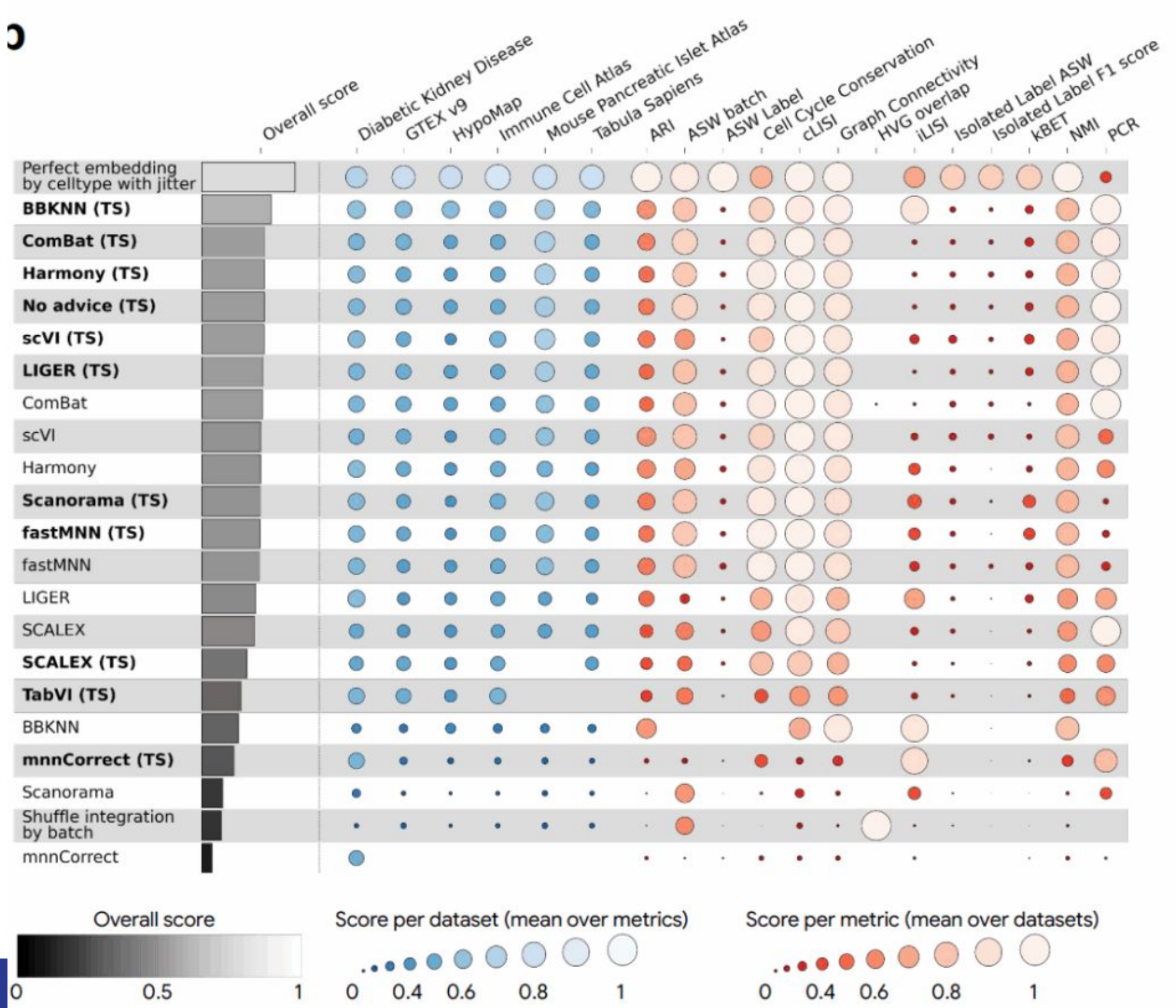
- Скачки в метрике можно сопоставить с конкретными идеями, появившимися в коде





Aygün et al.

- Например, система придумала новый метод batch correction для single-cell данных, объединив ComBat (глобальная линейная коррекция) и BBKNN (локальная коррекция через ближайших соседей)



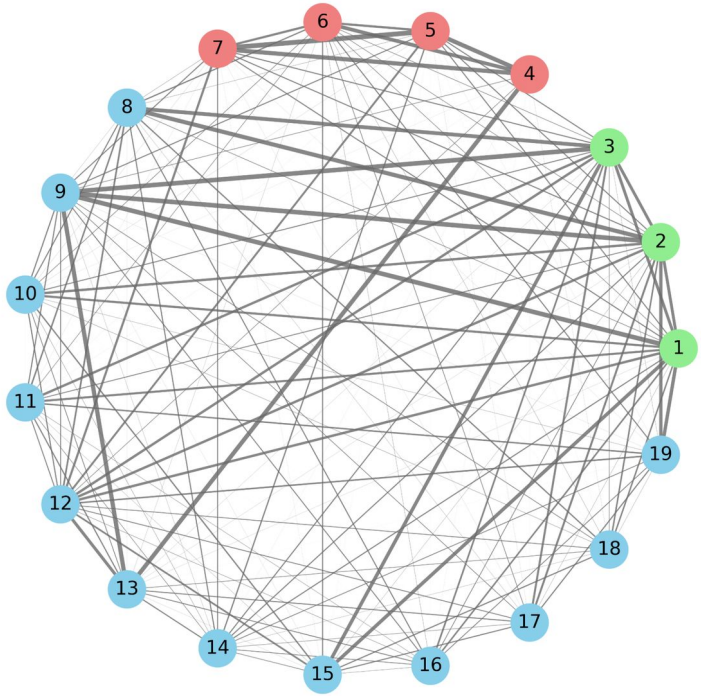
Aygün et al.

- А вот ближе к математике: авторы создали датасет из 38 сложных интегралов (scipy не берёт), на половине система оценивала результаты MCTS, а половина была тестовой
- 17 из 19 тестовых интегралов решились с ошибкой меньше 3%

train set		test set	
445.001	$\int_0^\infty \sin(x^2)\,dx$	446.021	$\int_0^\infty \left(\sin^4(ax^2) - \sin^4(bx^2)\right)\,dx$
445.017	$\int_0^\infty \sin(ax^2)\cos(2bx)\,dx$	446.045	$\int_0^\infty x\cos(ax^2)\cos(2bx)\,dx$
447.012	$\int_0^\infty \sin\left(ax^2 + \frac{b^2}{a}\right)\cos(2bx)\,dx$	449.013	$\int_0^\infty x^{\mu-1}\sin(ax)\cos(bx)\,dx$
458.031	$\int_0^\infty \left(\frac{\gamma+x}{\beta^2+(\gamma+x)^2} - \frac{\gamma-x}{\beta^2+(\gamma-x)^2}\right)\sin(ax)\,dx$	465.002	$\int_0^\infty \frac{(3-4\sin^2(ax))\sin^2(ax)}{x}\,dx$
462.034	$\int_0^\infty \frac{x\sin(ax)\cos(bx)}{c^2+x^2}\,dx$	465.013	$\int_0^\infty \frac{\sin^{2m+1}(x)\sin(x(6m+3))}{a^2+x^2}\,dx$
477.049	$\int_0^\infty \frac{x\sin(ax)+\cos(ax)}{x^2+1}\,dx$	467.025	$\int_0^\infty \frac{\sin(x)\cos(x)}{x\sqrt{\sin^2(x)+1}}\,dx$
478.036	$\int_0^\infty \frac{(\cos(a)-\cos(anx))\sin(mx)}{x}\,dx$	478.031	$\int_0^\infty \sin(ax^p)\,dx$
487.011	$\int_0^\infty \frac{1}{x}\frac{\sin(x)}{\left(a^2\cos^2(x)+b^2\sin^2(x)\right)^2}\,dx$	478.050	$\int_u^\infty \frac{\cos(ax)}{\sqrt{-u+x}}\,dx$

Теоремы из теоретической информатики

- Nagda et al. (DeepMind, Sep 30, 2025): запустили AlphaEvolve на задачах из теоретической информатики
- Для задачи MAX-4-CUT (максимальное сечение на четыре множества) оценку неприближаемости в 0.9883 улучшили до 0.987 через мощный гаджет из 19 вершин с очень разными весами рёбер
- И нашли новый 4-регулярный граф Рамануджана с большим 2-сечением (это нужно для оценок MAX-2-CUT в среднем)



Malliavin-Stein experiment

- Но что же с самой математикой?
- Diez et al. (Sep 3, 2025): авторы попросили GPT-5 доказать новую теорему из теории вероятностей; буквально:

4.1.1 Gaussian framework

We started with the following initial prompt:

```
Paper 2502.03596v1 establishes a qualitative fourth moment
theorem for the sum of two Wiener-Itô integrals of orders
p and q, where p and q have different parities.  Building
on the Malliavin-Stein method (see 1203.4147v3 for details),
could you derive a quantitative version for the total variation
distance, with a convergence rate depending solely on the
fourth cumulant of this sum?
```

Mathematical research with GPT-5: a Malliavin-Stein experiment

Charles-Philippe Diez* Luís da Maia* Ivan Nourdin*

Abstract

On August 20, 2025, GPT-5 was reported to have solved an open problem in convex optimization. Motivated by this episode, we conducted a controlled experiment in the Malliavin–Stein framework for central limit theorems. Our objective was to assess whether GPT-5 could go beyond known results by extending a *qualitative* fourth-moment theorem to a *quantitative* formulation with explicit convergence rates, both in the Gaussian and in the Poisson settings. To the best of our knowledge, the derivation of such quantitative rates had remained an open problem, in the sense that it had never been addressed in the existing literature. The present paper documents this experiment, presents the results obtained, and discusses their broader implications.

2. Результаты чистых LLM

Одинаковость языка есть первая таинственная связь, соединяющая людей между собою. Народ говорит одним языком, и в этом единстве выражается внутренняя симпатия, родство душ, по которому люди одного народа сливают звуки в известные стройные созвучия, выражая ими внутренние и внешние свои понятия. Один человек говорит, и другой понимает его, — и вот между ними утверждена прочная связь взаимного разумения.

Константин Аксаков. О грамматике вообще

Malliavin-Stein experiment

- И после такого запроса авторы отмечают, что в доказательстве была ошибка, и GPT-5 потребовалось аж два наводящих промпта (!), чтобы ошибку понять и исправить
- Затем GPT-5 сам предложил рассмотреть другой случай, и в нём опять была ошибка, и понадобилось указать модели на неё! После этого GPT-5 за три-четыре промпта пришёл к верному результату

> I think you are mistaken in claiming that $(p+q)!\|u\widetilde{\otimes}v\|^2 = p!q!\|u\|^2\|v\|^2$. Why should that be the case?

It eventually admitted (which is not surprising, since by alignment it usually agrees with us) that the statement was false, but more importantly, it understood where the mistake came from. This was followed by a reasoning and a formula that, this time, were correct.

Then, at our request, GPT-5 reformatted the result in the style of a research article, including an introduction, the presentation of our main theorem, its proof with all the details (correct this time!), and a bibliography. The exact prompt was:

> Turn this into a research paper ready for submission. Follow my style (see attached paper 0705.0570v4):
> - start with an introduction giving some context,
> - then present the main result, followed by a very detailed proof where no step is left out,
> - finish with a complete bibliography.
> The final document should be a LaTeX file that I can compile.

Malliavin-Stein experiment

- Что же заключают авторы?

4.1.3 Role of the AI

To summarize, we can say that the role played by the AI was essentially that of an executor, responding to our successive prompts. Without us, it would have made a damaging error in the Gaussian case, and it would not have provided the most interesting result in the Poisson case, overlooking an essential property of covariance, which was in fact easily deducible from the results contained in the document we had provided.

- Кажется, это называется copium…

5 Some personal reflections

Overall, the experience of doing mathematics with GPT-5 was mixed. It felt very similar to working with a junior assistant at the beginning of a new project: exploring directions, formulating hypotheses, searching for counterexamples, and progressively adjusting statements. The AI showed a genuine ability to follow guided reasoning, to recognize its mistakes when pointed out, to propose new research directions, and to never take on the task. However, this only seems to support *incremental* research, that is, producing new results that do not require genuinely new ideas but rather the ability to combine ideas coming from different sources. At first glance, this might appear useful for an exploratory phase, helping us save time. In practice, however, it was quite the opposite: we had to carefully verify everything produced by the AI and constantly guide it so that it could correct its mistakes.

Gödel Test

- Feldman, Karbasi (Sep 22, 2025): тест Гёделя – могут ли LLM доказывать простые, но новые теоремы?
- Вот что писал об этом Теренс Тао, когда познакомился с o1

"The new model could work its way to a correct (and well-written) solution if *provided a lot of hints and prodding, but did not generate the key conceptual ideas on its own, and did make some non-trivial mistakes. The experience seemed roughly on par with trying to advise a mediocre, but not completely incompetent, graduate student. However, this was an improvement over previous models, whose capability was closer to an actually incompetent graduate student. It may only take one or two further iterations of improved capability (and integration with other tools, such as computer algebra packages and proof assistants) until the level of 'competent graduate student' is reached."*

— Terence Tao

Gödel Test

- Feldman, Karbasi (Sep 22, 2025): дали пять запросов в таком духе:

Prompt to GPT-5

Consider the problem of maximizing a function F from $[0,1]^n$ to the reals that is the sum of a non-negative monotonically increasing DR-submodular function G and a non-negative DR-submodular function H over a solvable down-closed polytope P. I would like to bound the performance on this problem from the NeurIPS 2021 paper "Submodular + Concave" which is attached. Specifically, if x is the output vector of this algorithm and o is the vector in P maximizing F, then I would like to lower bound $F(x)$ with an expression of the form $\alpha * G(o) + \beta * H(o) - err$, where α and β are constants. err should be a function that depends only on the error parameter ϵ of the algorithm, the diameter D of the polytope P, and the smoothness parameters L_G and L_H of G and H, respectively, and goes to zero as ϵ goes to zero. Please give the best such bound that you prove (a bound is considered better if the values of the constants α and β are larger). Provide a mathematically rigorous and well explained proof for the bound you come up with.

Gödel Test

- Feldman, Karbasi (Sep 22, 2025)

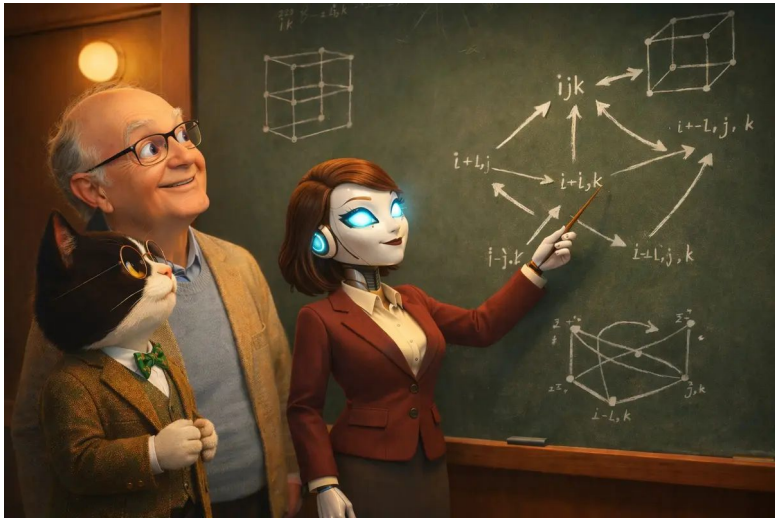
Prompt to GPT-5

Consider t
of a non-
DR-submo
the perfor
which is att
in P maxim
$\alpha * G(o) + \beta *$
only on the e
smoothness pa
Please give the
the constants α a
for the bound you
with.
ne form
tion that depends
of the polytope P, and the
vely, and goes to zero as ϵ goes to zero.
(a bound is considered better if the values of
ovide a mathematically rigorous and well explained proof

GPT-5 решил три из пяти, с небольшими легко исправимыми недочётами. Главный найденный недостаток — GPT-5 писал “ленивые” доказательства, максимально ссылаясь на предоставленные статьи, даже когда это было не очень изящно

Claude’s Cycles

- Donald Knuth (Feb 2026): Claude Opus 4.6 решил открытую проблему, над которой Кнут бился неделями; решение потом подтвердилось и даже было формализовано



Shock! Shock! I learned yesterday that an open problem I’d been working on for several weeks had just been solved by Claude Opus 4.6 — Anthropic’s hybrid reasoning model that had been released three weeks earlier! It seems that I’ll have to revise my opinions about “generative AI” one of these days. What a joy it is to learn not only that my conjecture has a nice solution but also to celebrate this dramatic advance in automatic deduction and creative problem solving. I’ll try to tell the story briefly in this note.

Here’s the problem, which came up while I was writing about directed Hamiltonian cycles for a future volume of *The Art of Computer Programming*:

> Consider the digraph with m^3 vertices ijk for $0 \le i, j, k < m$, and three arcs from each vertex, namely to i^+jk, ij^+k, and ijk^+, where $i^+ = (i+1) \bmod m$. Try to find a general decomposition of the arcs into three directed m^3-cycles, for all $m > 2$.

Erdös Problems

- 1113 задач Эрдеша в базе erdosproblems.com, из них 437 решены
- Многие из них совсем не знаменитые, возможно, некоторые случайно решены где-то в литературе, но много и сложных нерешённых задач

Problems have always been an essential part of my mathematical life. A well-chosen problem can isolate an essential difficulty in a particular area, serving as a benchmark against which progress in this area can be measured. It might be like a 'marshmallow', serving as a tasty tidbit supplying a few moments of fleeting enjoyment. Or it might be like an 'acorn', requiring deep and subtle new insights from which a mighty oak can develop. […]

In this note I would like to describe a variety of my problems which I would classify as my favorites. Of course, I can't guarantee that they are all 'acorns', but because many have thwarted the efforts of the best mathematicians for many decades (and have often acquired a cash reward for their solutions), it may indicate that new ideas will be needed, which can, in turn, lead to more general results, and naturally, to further new problems.

In this way, the cycle of life in mathematics continues forever.

Paul Erdős, *Some of my favorite problems and results*, 1997

OPEN

Let $f \in \mathbb{R}[x]$ be a monic polynomial of degree n with n real roots in $[-1, 1]$. Determine the infimum and supremum of possible values of

$$|\{x \in \mathbb{R} : |f(x)| < 1\}|.$$

#1038: [EHP58,p.131] analysis

FALSIFIABLE

Let $z_1, \dots, z_n \in \mathbb{C}$ with $|z_i - z_j| \leq 2$ for all i, j, and

$$\Delta(z_1, \dots, z_n) = \prod_{i \neq j} |z_i - z_j|.$$

What is the maximum possible value of Δ? Is it maximised by taking the z_i to be the vertices of a regular polygon?

#1045: [EHP58,p.143] analysis

Erdös Problems

- Теренс Тао сделал страничку “AI contributions to Erdős problems”
- Понемногу начинают накапливаться результаты, но к этому мы ещё вернёмся…

AI contributions to Erdős problems

teorth edited this page 4 hours ago · 70 revisions

This page collects the various ways in which AI tools have contributed to the understanding of Erdős problems. Note that a single problem may appear multiple times in these lists.

Legend

- 🟢 : full solution to the problem
- 🟡 : partial solution to the problem
- 🔴 : failure to make progress on the problem

5. Problems with an AI-powered literature review

Problem	AI tools used	Review performed on	Outcome
[35]	GPT-5	13 Oct, 2025	🟡 Partial results found
[66]	GPT-5	13 Oct, 2025	🟡 Partial results found
[124]	ChatGPT DeepResearch, Gemini DeepResearch	30 Nov, 2025	No significant results found
[167]	GPT-5	12 Oct, 2025	🟡 Partial results found
[188]	GPT-5	13 Oct, 2025	🟡 Partial results found
[203]	ChatGPT DeepResearch, Gemini DeepResearch	19 Oct, 2025	No significant results found
[223]	GPT-5	13 Oct, 2025	🟢 Full solution found
[248]	Gemini DeepResearch	19 Oct, 2025	No significant results found
[330]	ChatGPT DeepResearch, Gemini DeepResearch, Claude	19 Dec, 2025	🔴 Partial results found (with inaccuracies); did not find literature proof

Erdös Problems

- Теренс Тао сделал страничку “AI contributions to Erdős problems”
- Понемногу начинают накапливаться результаты, но к этому мы ещё вернёмся…

AI contributions to Erdős problems

teorth edited this page 4 hours ago · 70 revisions

This page collects the various ways in which AI tools have contributed to the understanding of Erdős problems. Note that a single problem may appear multiple times in these lists.

Legend

- 🟢 : full solution to the problem
- 🟡 : partial solution to the problem
- 🔴 : failure to make progress on the problem

[334]	Gemini DeepResearch	19 Nov, 2025	No significant results found
[339]	GPT-5	11 Oct, 2025	🟢 Full solution found
[347]	ChatGPT DeepResearch	25 Oct, 2025	🟢 Full solution found
[354]	ChatGPT DeepResearch	19 Oct, 2025	🟡 Partial results found
[367]	ChatGPT DeepResearch, Gemini DeepResearch	22 Nov, 2025	🔴 No significant results found; did not find Erdős problems community proof
[370]	ChatGPT DeepReseach, Gemini, Gemini DeepResearch	17 Oct, 2025	🟡 Problem found to be misstated; solutions to variants found
[387]	ChatGPT DeepResearch	1 Nov, 2025	🟡 Partial results found
[421]	ChatGPT DeepResearch, Gemini DeepResearch	18 Oct, 2025	No significant results found
[481]	ChatGPT DeepResearch, Gemini DeepResearch	1 Dec, 2025	🟡 Unwittingly reproduced existing proof
[481]	ChatGPT	3 Dec, 2025	🟢 Full solution found
[494]	GPT-5	13 Oct, 2025	🟢 Full solution found
[515]	GPT-5	15 Oct, 2025	🟢 Full solution found

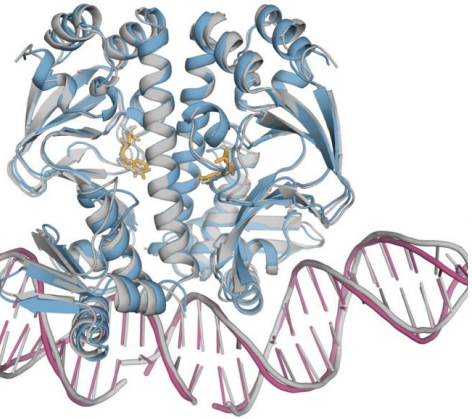
3. AI в других науках

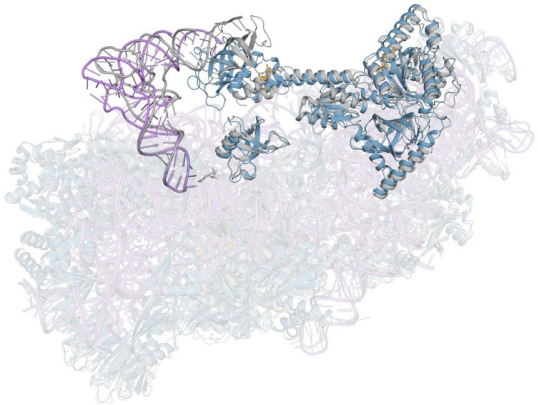
В минуту слабости нас поддержит величие нашего идеала; кто-нибудь предпочтет, конечно, иной идеал; впрочем, не становится ли Бог ученого тем величественнее, чем более он удаляется от нас? Верно то, что он строг, далёк, непреклонен, и многие души пожалеют об этом. Но он, по крайней мере, не разделяет наших мелочей и жалких злоб дня.

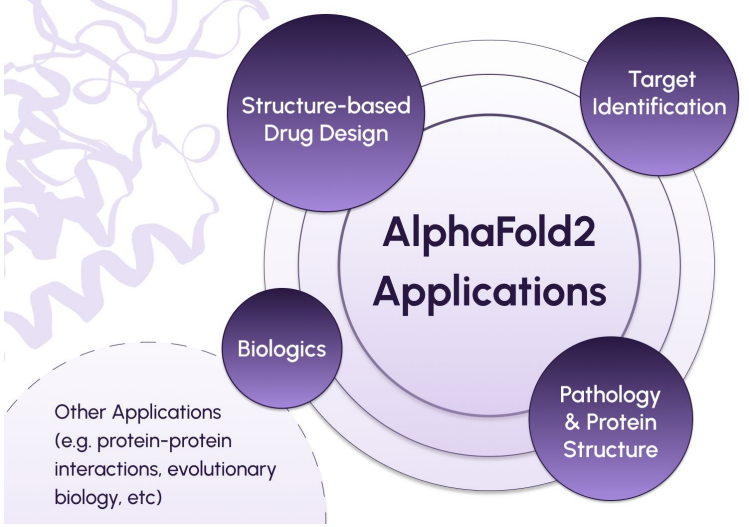
Анри Пуанкаре. Наука и нравственность

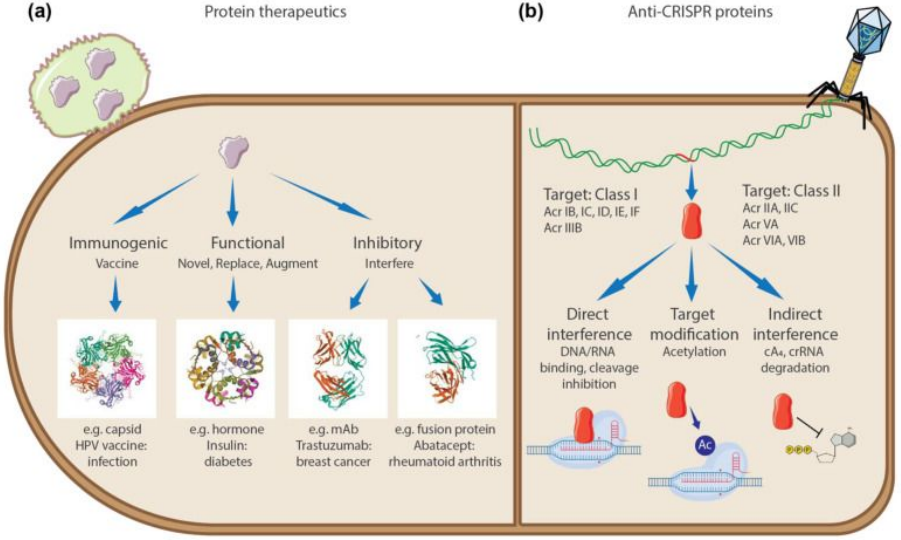
AlphaFold

- Как вы знаете, AlphaFold от DeepMind прямо сейчас производит революцию в медицине
- AlphaFold 3 (Google, May 2024)









ЦЕНТРАЛЬНЫЙ УНИВЕРСИТЕТ

Кстати, об AlphaFold и DeepMind в целом

PRESS RELEASE

9 October 2024

The Nobel Prize in Chemistry 2024

The Royal Swedish Academy of Sciences has decided to award the Nobel Prize in Chemistry 2024
with one half to and the other half jointly to

David Baker
University of Washington, Seattle, WA, USA

"for computational protein design"

Demis Hassabis
Google DeepMind, London, UK

John M. Jumper
Google DeepMind, London, UK

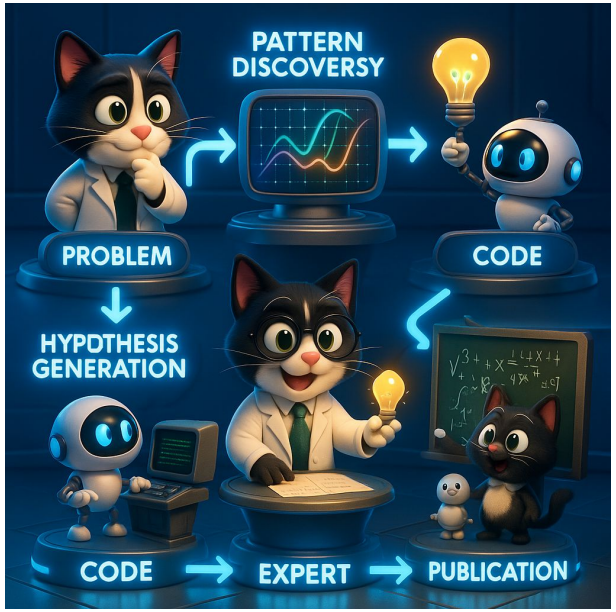
"for protein structure prediction"

DeepMind

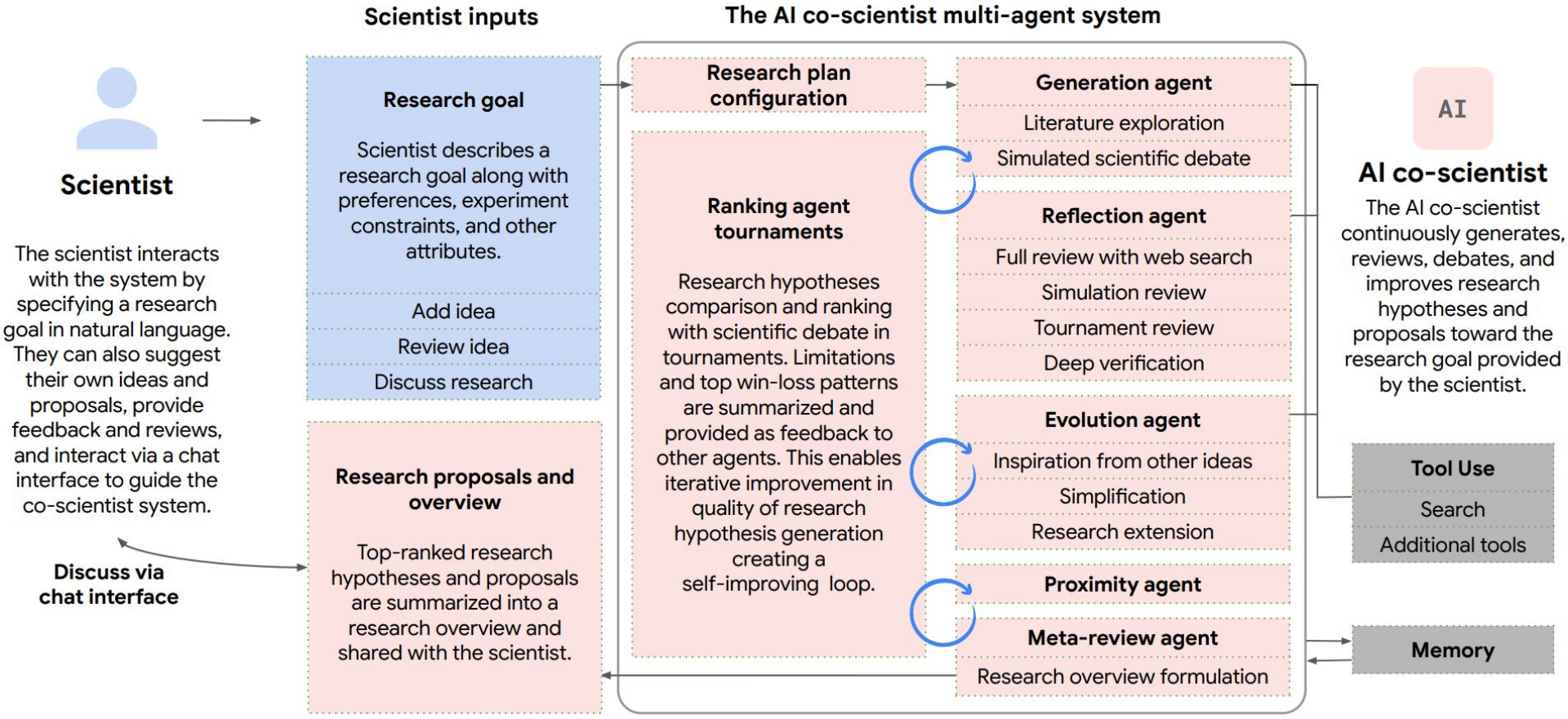
They cracked the code for proteins' amazing structures

Google Co-Scientist

- Google Co-Scientist (Gottweis et al., Feb 19, 2025): мультиагентная система для помощи живым учёным на всех уровнях
- Состоит из нескольких агентов на основе Gemini 2.0: Generation, Reflection, Ranking, Evolution, Proximity, Meta-review
- В целом структура выглядит абсолютно естественно, подобные работы уже были

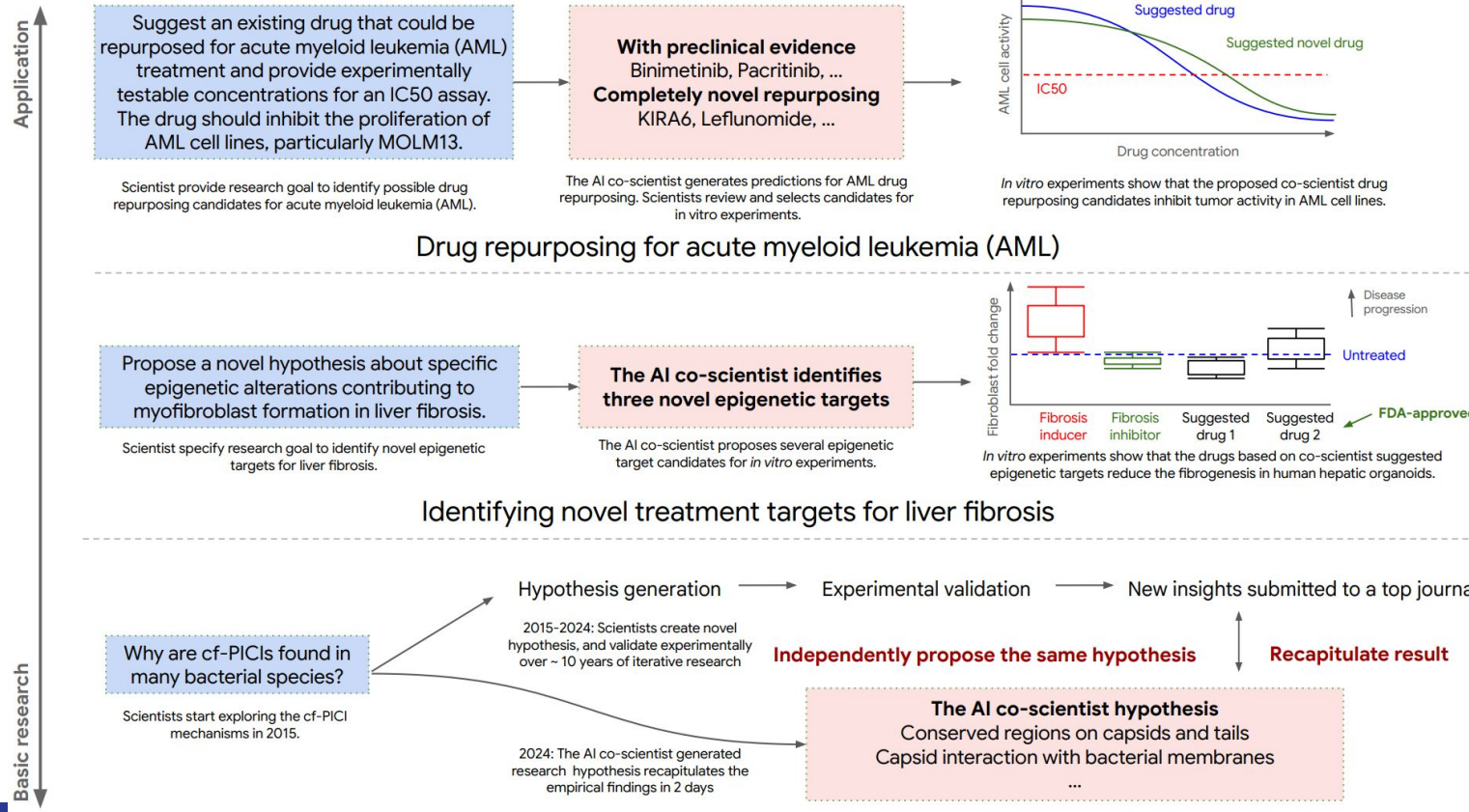


Google Co-Scientist



Google Co-Scientist

- Разница в том, что Co-Scientist, по всей видимости, и правда работает
- В том числе благодаря рассужда-ющим моделям



Google Co-Scientist

- Пока главный пример – от Jose Penades, микробиолога из Imperial College London
- Он задал Co-Scientist’у вопрос, на который знал ответ, но ответ ещё не был опубликован
- И система выдала правильный ответ первым, а ещё дала несколько других интересных идей
- В этом примере не всё так просто, конечно

Scientists spent 10 years on a superbug mystery - Google's AI solved it in 48 hours

The co-scientist model came up with several other plausible solutions as well

By Shawn Knight February 21, 2025 at 2:22 PM | 20 comments

Analysis and Technology

Can Google's new research assistant AI give scientists 'superpowers'?

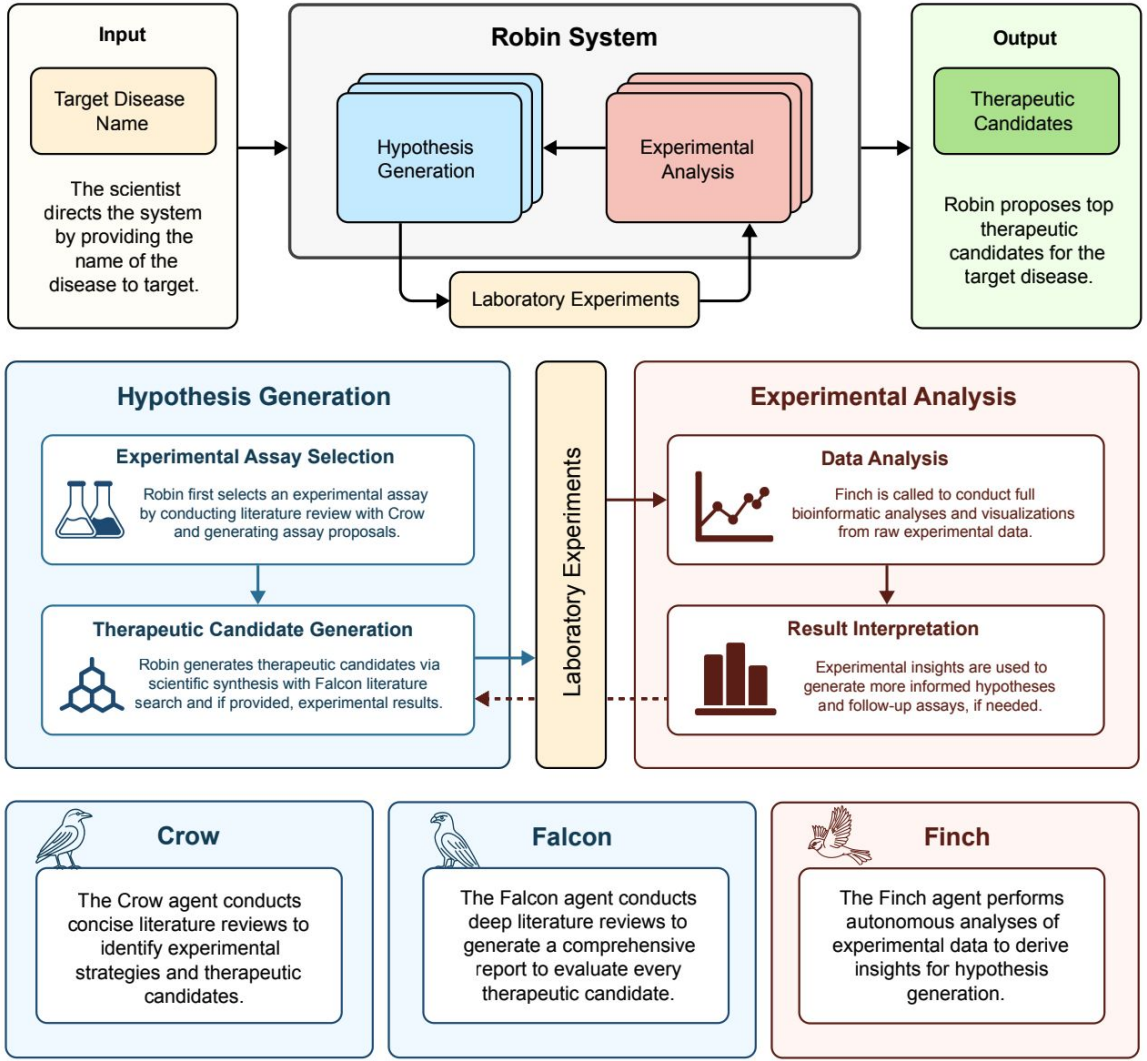
Researchers who have been given access to Google's new AI "co-scientist" tool are enthusiastic about its potential, but it isn't yet clear whether it can make truly novel discoveries

By Michael Le Page

19 February 2025

Robin

- Ghareeb et al. (May 19, 2025): Robin, другой набор агентов, похожий на Google Co-Scientist и разработанный FutureHouse и Oxford



Robin

- Я не биолог, но выглядит круто...

Sam Rodriques 
@SGRodriques

5:12 PM · May 20, 2025 · **1M** Views

Today, we're announcing the first major discovery made by our AI Scientist with the lab in the loop: a promising new treatment for dry AMD, a major cause of blindness.

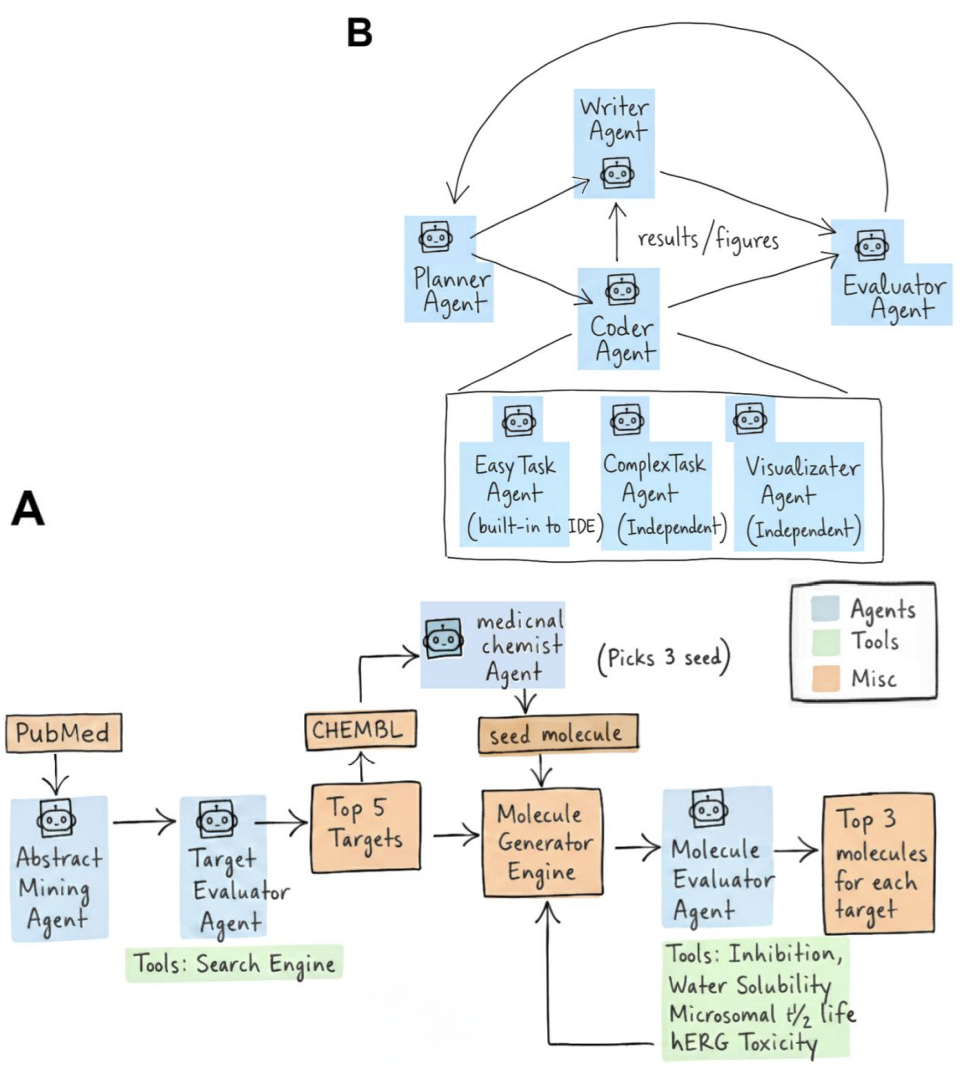
Our agents generated the hypotheses, designed the experiments, analyzed the data, iterated, even made figures for the paper. The resulting manuscript is a first-of-a-kind in the natural sciences, in which everything that needed to be done to write the paper was done by AI agents, apart from actually conducting the physical experiments in the lab and writing the final manuscript. We are also introducing Robin, the first multi-agent system that fully automates the in-silico components of scientific discovery, which made this discovery. This is the first time that we are aware of that hypothesis generation, experimentation, and data analysis have been joined up in closed loop, and is the beginning of a massive acceleration in the pace of scientific discovery that will be driven by these agents. We will be open-sourcing the code and data next week.

Robin is a multi-agent system that uses Crow, Falcon, and Finch, the agents on our platform, to generate novel hypotheses, plan experiments, and analyze data. We asked Robin to find a new treatment for dry age-related macular degeneration. Robin considered the disease mechanisms associated with dry AMD, proposed a specific experimental assay that could be used to evaluate hypotheses in the wet lab, and proposed specific molecules we could test in that assay. We tested the molecules and gave it the resulting data, which it analyzed before proposing more experiments. In the end, it identified Ripasudil, a Rho Kinase inhibitor (ROCK inhibitor) that is approved in Japan for several other diseases, which seems very promising as potential treatment for dry AMD. It also identified specific molecular mechanisms that might underlie the effects of Ripasudil in RPE cells, from an RNA sequencing experiment it proposed. To be clear, no one has proposed using ROCK inhibitors to treat dry AMD in the literature before, as far as we can find, and I think it would have been very difficult for us to come up with this hypothesis without the agents. We have also run the proposed treatment by several experts in AMD, who confirm that it is interesting and novel. Moreover, this project was fast: with Robin in hand, the entire project took about 10 weeks, which is way shorter than it would have taken if we had been doing all of the in-silico components ourselves.

Important caveats: We are real biologists at FutureHouse, so I want to be clear that although the discovery here is exciting, we are not claiming that we have cured dry AMD. Fully validating this hypothesis as a treatment for dry AMD will take human trials, which will take much longer. Also, this discovery is cool, but it is not yet a "move 37"-style discovery. At the current rate of progress, I'm sure we will get to that level soon.

Drug discovery

- Пример: Multi-target Parallel Drug Discovery with Multi-agent Orchestration (AI Agent et al., November 2025)
- A – drug discovery, B – manuscript generation
- Система действительно нашла потенциальные молекулы для Alzheimer's targets, используя "Human-in-the-loop Simulation": люди могли в критические моменты посоветовать что-то системе



Drug discovery

- Приближаются к клиническим испытаниям первые AI-designed лекарства
- AlphaFold уже произвёл революцию в предсказании структуры белков, но современные модели идут дальше — они уже придумывают новые биоактивные молекулы

Review Article | Published: 08 September 2025

AI-driven protein design

Huan Yee Koh, Yizhen Zheng, Madeleine Yang, Rohit Arora, Geoffrey I. Webb, Shirui Pan, Li Li & George M. Church

Nature Reviews Bioengineering **3**, 1034–1056 (2025) | Cite this article

10k Accesses | **6** Citations | **92** Altmetric | Metrics

What will be the first AI-designed drug? These disease-fighting antibodies are top contenders

Just a year after the first AI-designed antibody was made, scientists say clinical trials are on the horizon.

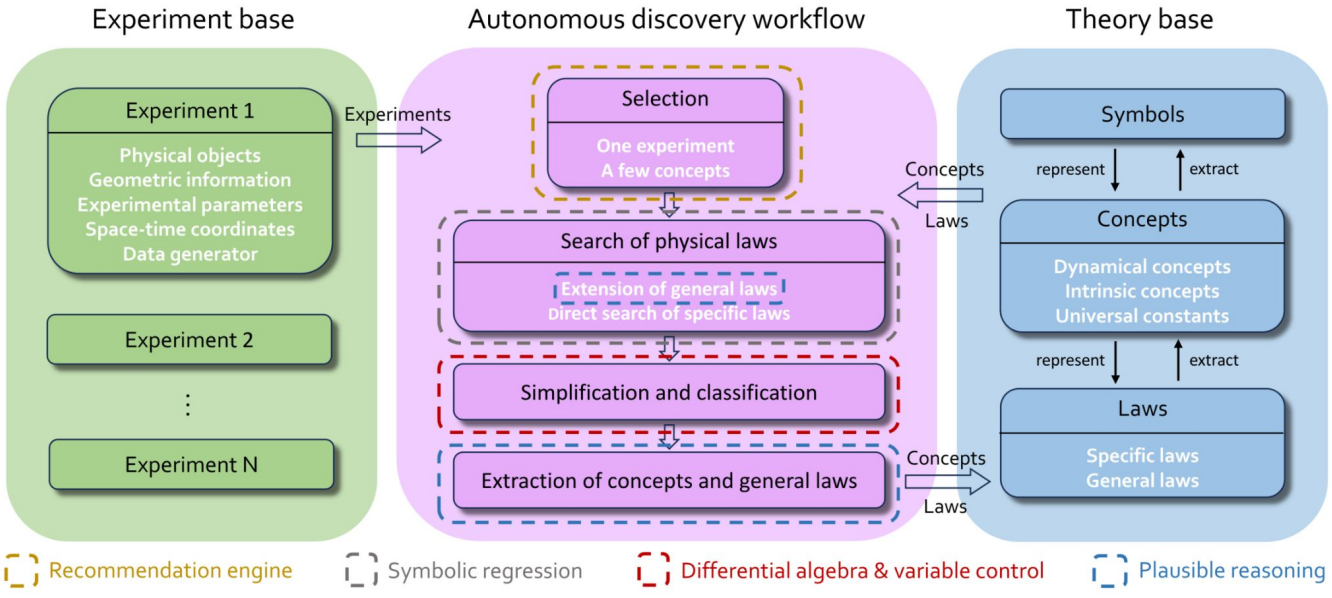
AI в других науках

- В физике пока громких успехов я не знаю, но уже появляются системы, способные выводить физические законы из данных

[Submitted on 2 Apr 2025]

AI-Newton: A Concept-Driven Physical Law Discovery System without Prior Physical Knowledge

You-Le Fang, Dong-Shan Jian, Xiang Li, Yan-Qing Ma



AI в других науках

- В целом другие науки, конечно, медленнее трансформируются
- Но люди уже понимают, что AI-системы могут изменить сам научный процесс
- В каком-то смысле это всё, конечно, только начинается…

GOV.UK

Policy paper

AI for Science Strategy

Published 20 November 2025

Introduction: the AI for science opportunity

An ‘AI for Science’ moment is unfolding. Investment and attention are being drawn to the idea that increasing the productivity of scientific research will be the most valuable application of AI.[footnote 1] Science is being transformed at unprecedented pace.

In the last 3 years, the leading edge of AI for science has moved from prediction to action. AI models are increasingly autonomous participants in the scientific process. In biology, AI models have gone from predicting the structures of known proteins to designing entirely new ones, helping scientists to accelerate research in nearly every field – from fighting malaria to potential new Parkinson’s treatments.[footnote 2]

Alongside increasingly capable narrow models, frontier AI labs are now racing to build AI science agents capable of automating core parts of the scientific process. These systems can generate hypotheses, design experiments and conduct analysis without direct human input. Pairings of AI with real-world experiments foreshadow systems that can ‘learn from doing’ in real time, as demonstrated by Liverpool’s Materials Innovation Factory, which built a mobile robotic chemist that conducted 688 experiments over eight days and discovered a new catalyst without human intervention.[footnote 3] These AI-centred discovery processes could transform scientific productivity and progress.[footnote 4]

AI в других науках

- ...но уже начинается!
- 28 ноября 2025: Launching the Genesis Mission, executive order 14363

FEDERAL REGISTER
The Daily Journal of the United States Government

This order launches the "Genesis Mission" as a dedicated, coordinated national effort to unleash a new age of AI-accelerated innovation and discovery that can solve the most challenging problems of this century. The Genesis Mission will build an integrated AI platform to harness Federal scientific datasets—the world's largest collection of such datasets, developed over decades of Federal investments—to train scientific foundation models and create AI agents to test new hypotheses, automate research workflows, and accelerate scientific breakthroughs. The Genesis Mission will bring together our Nation's research and development resources—combining the efforts of brilliant American scientists, including those at our national laboratories, with pioneering American businesses; world-renowned universities; and existing research infrastructure, data repositories, production plants, and national security sites—to achieve dramatic acceleration in AI development and utilization. We will harness for the benefit of our Nation the revolution underway in computing, and build on decades of innovation in semiconductors and high-performance computing. The Genesis Mission will dramatically accelerate scientific discovery, strengthen national security, secure energy dominance, enhance workforce productivity, and multiply the return on taxpayer investment into research and development, thereby furthering America's technological dominance and global strategic leadership.

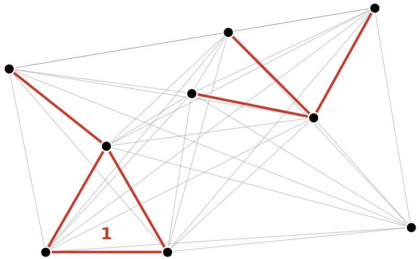
4. Последние новости

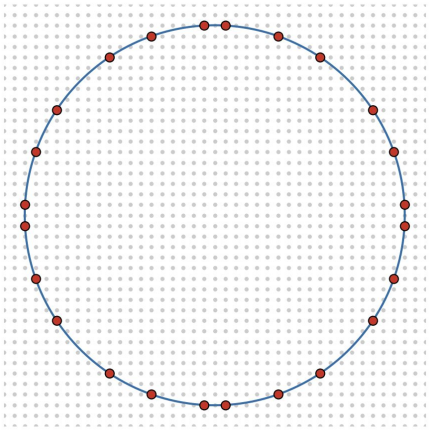
Затем остаются только размышления о том, какие прекрасные здания строятся ныне в Париже… Можно вообразить себе, как свежи и занимательны эти объяснения, что они неизменно повторяются уж более двух лет, по крайней мере раз в неделю, если не чаще. Кроме того, неизменно в каждом фельетоне с большим юмором излагается жалоба на недостаток новостей.

Николай Чернышевский
Новости литературы, искусств, наук и промышленности

Единичные расстояния

- …ну хотя как “если”, уже стало…
- Jared Lichtman (Apr 2026): GPT-5.4 Pro опроверг важную гипотезу Эрдёша (над ней думало много математиков), причём сделал это неожиданным и красивым образом





1946 — **Paul Erdős**
"On sets of distances of n points"
Amer. Math. Monthly 53 (1946), 248–250
Постановка задачи. Нижняя оценка $\nu(n) \geq n^{1+c/\log\log n}$, верхняя оценка $\nu(n) \leq O(n^{3/2})$.

1984 — **Joel Spencer, Endre Szemerédi, William T. Trotter**
"Unit distances in the euclidean plane"
Graph Theory and Combinatorics (Cambridge 1983), 293–303
Верхняя оценка $\nu(n) \leq O(n^{4/3})$.

1997 — **László A. Székely**
"Crossing numbers and hard Erdős problems in discrete geometry"
Combinatorics, Probability and Computing 6 (1997), 353–358
Элегантное короткое доказательство $O(n^{4/3})$ через crossing lemma.

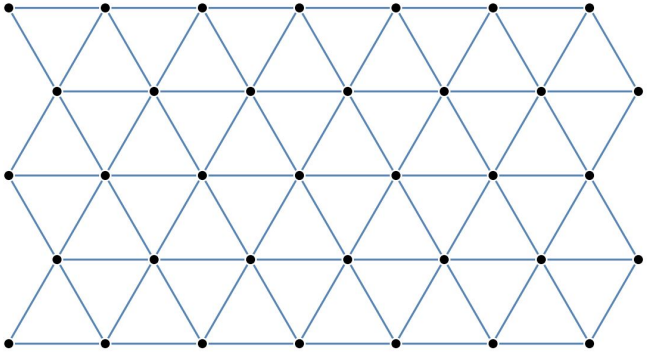
2023 — **Noga Alon, Matija Bucić, Lisa Sauermann**
"Unit and distinct distances in typical norms"
GAFA 35 (2025), 1–42 · arXiv:2302.09058
Для нормы общего положения $\nu(n) \leq \frac{d}{2} n \log_2 n$.

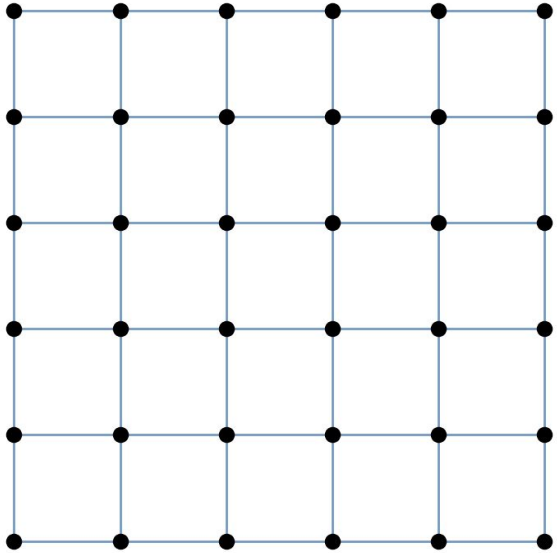
2024 — **Josef Greilhuber, Carl Schildkraut, Jonathan Tidor**
"More unit distances in arbitrary norms"
Bull. London Math. Soc. 57 (2025), 2885–2901 · arXiv:2410.07557
Соответствующая нижняя оценка $\Omega(n \log_2 n)$ для любой нормы.

2026 — **OpenAI**
"Planar Point Sets with Many Unit Distances"
openai.com, 20 мая 2026 · комментарии: Alon et al., arXiv:2605.20695
Контрпример: $\nu(n) \geq n^{1+\delta}$ для $\delta > 0$ и бесконечно многих n.
Гипотеза Эрдёша опровергнута.

Единичные расстояния

- Здесь я не смогу объяснить доказательство целиком, но давайте обсудим постановку задачи
- Рассмотрим *n* точек на обычной евклидовой плоскости. Сколько одинаковых расстояний может быть среди пар этих точек?
- Треугольная решётка даёт 3n
- Главный подход: превратить задачу в арифметическую. Рассмотрим квадратную решётку; на ней расстояние m встречается столько раз, сколько раз m^2 раскладывается в a^2+b^2.





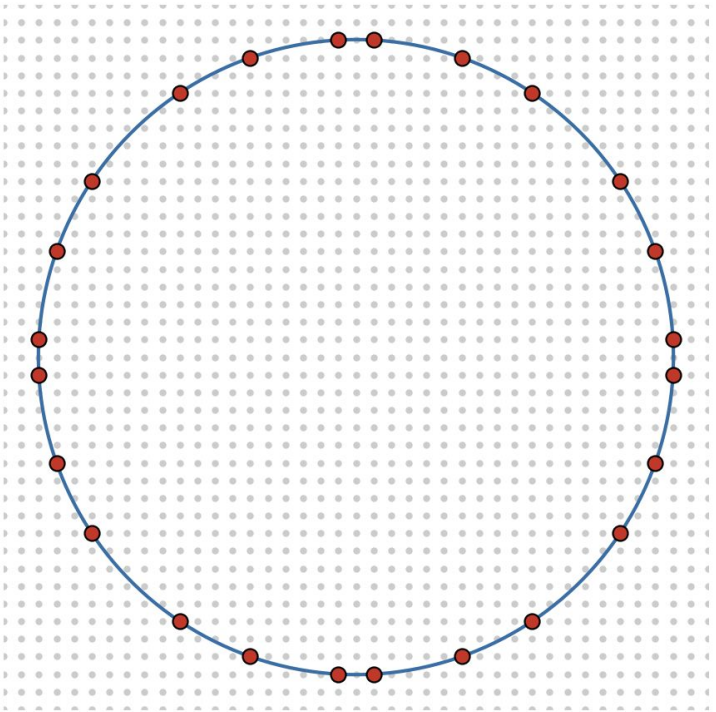
Единичные расстояния

- Каждое p=4k+1 можно разложить как

$$4k+1=(1+2i\sqrt{k})(1-2i\sqrt{k})=z\cdot\bar{z},\qquad z=1+2i\sqrt{k}.$$

- Если число m имеет в разложении много простых делителей p_i вида 4k+1, то каждое из них можно так разложить, а потом представить m как произведение двух комплексно сопряжённых подмножеств
- Например, у числа 325 на окружности радиуса $\sqrt{325}$ сразу 24 целые точки:

$$325=5^2\cdot 13=(2+i)^2(2-i)^2(3+2i)(3-2i).$$



$$325=\Big((2+i)^2(3-2i)\Big)\Big((2-i)^2(3+2i)\Big)=(17+6i)(17-6i)=17^2+6^2.$$

$$325=\Big((2+i)^2(3+2i)\Big)\Big((2-i)^2(3-2i)\Big)=(1+18i)(1-18i)=1^2+18^2.$$

$$325=((2-i)(2+i)(3+2i))\,((2+i)(2-i)(3-2i))=(15+10i)(15-10i)=15^2+10^2.$$

Единичные расстояния

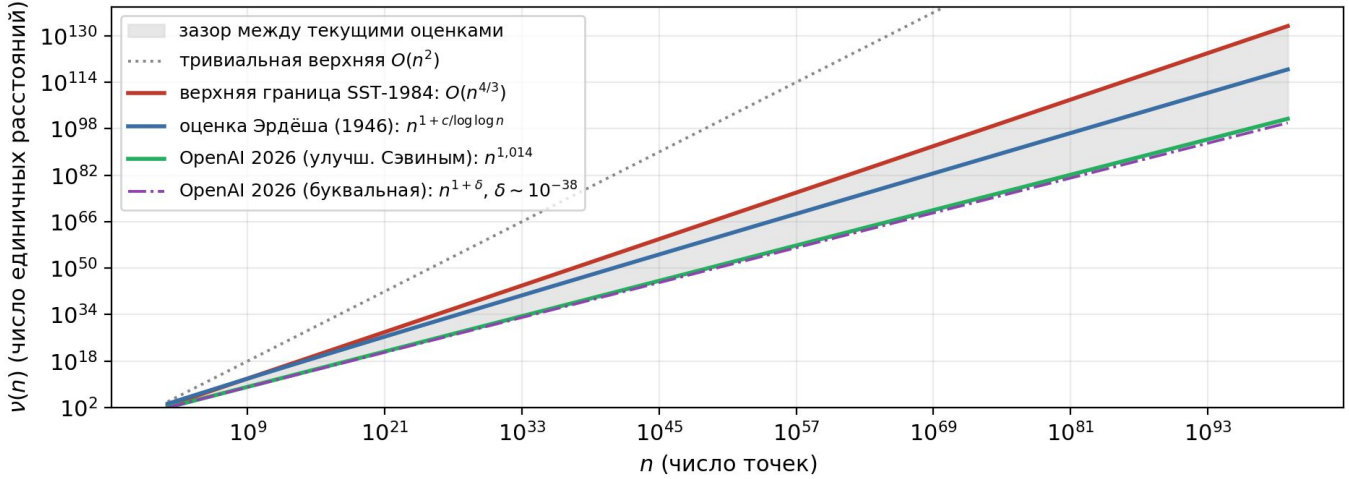
- Если аккуратно посчитать такие пары точек на окружности, получится уже нелинейная оценка: $\nu(n) \geq n^{1+c/\log\log n}$.
- Можно ли лучше?
- Эрдёш предположил, что нет, но GPT доказал, что можно!

Гипотеза Эрдёша. $\nu(n) \leq n^{1+o(1)}$, а скорее даже

$$\nu(n) \leq n^{1+C/\log\log n}$$

Теорема (OpenAI, 2026). Существует константа $\delta > 0$ и бесконечная последовательность натуральных чисел n, для которых $\nu(n) \geq n^{1+\delta}$.

Известные границы для задачи о единичных расстояниях (log-log, видны 100 порядков по n).



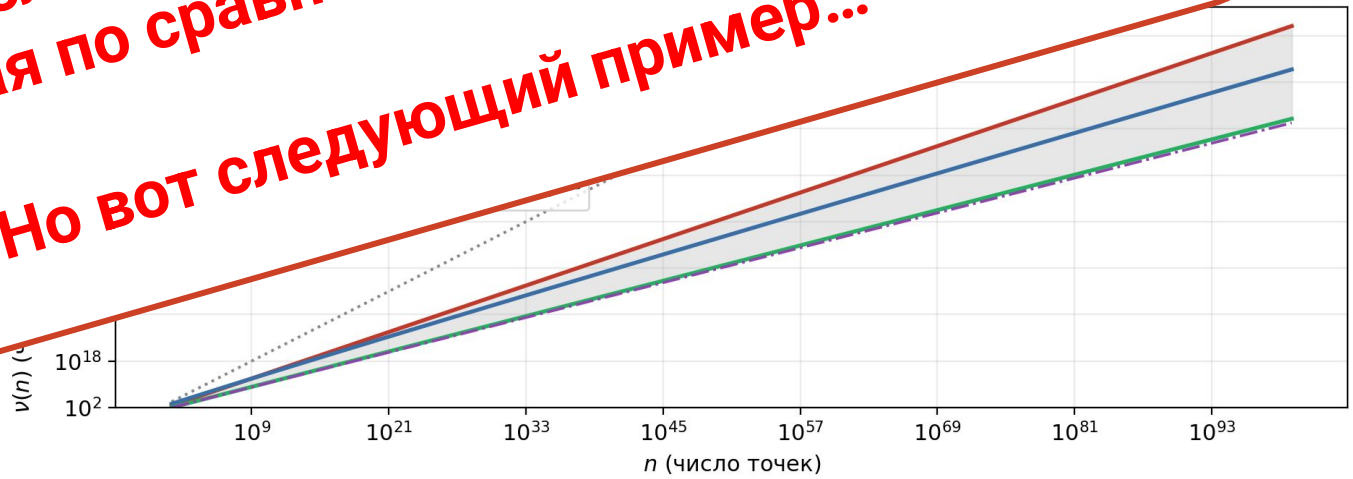
ЦЕНТРАЛЬНЫЙ УНИВЕРСИТЕТ

Единичные расстояния

- Если аккуратно посчитать такие
  пары точек на окружн
  получится у
  о

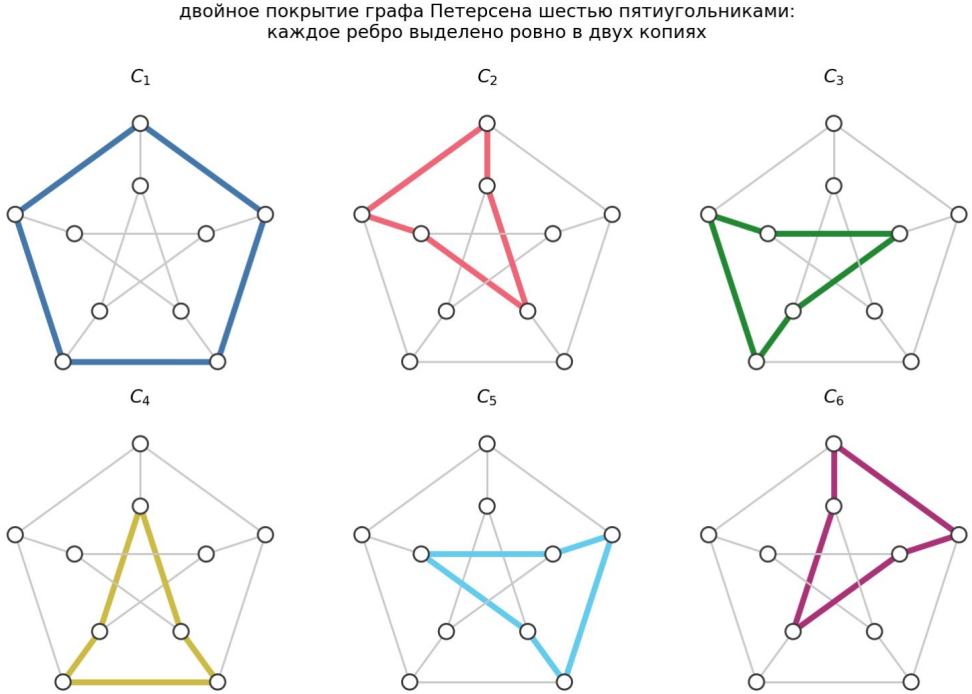
Здесь я вам ничего больше не расскажу, дальше уже реально сложная математика (хоть и довольно простая по сравнению с ожиданиями).

Но вот следующий пример…



Двойное покрытие циклами

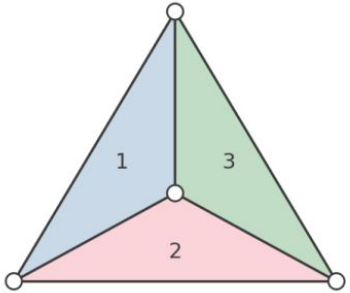
- Верно ли, что в любом графе без мостов есть набор простых циклов, для которого каждое ребро содержится ровно в двух циклах этого набора?
- Поставили задачу ещё в 1970-х, люди много думали, были продвижения
- OpenAI, 10 июля 2026: препринт с полным доказательством
- Мой пост: “Охота на снарка: доказательство гипотезы о двойном покрытии циклами”



Двойное покрытие циклами

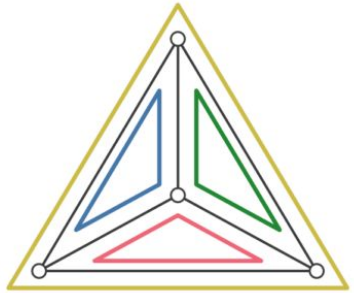
- И на этот раз решение реально простое! Препринт на две страницы, никакой сложной теории
- Давайте обсудим задачу и даже идею решения
- Сначала классические результаты: для планарных графов всё очевидно, а мосты очевидно мешают

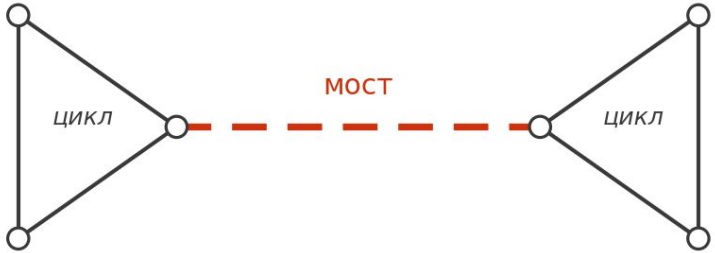
планарный граф и его грани



4 — внешняя грань

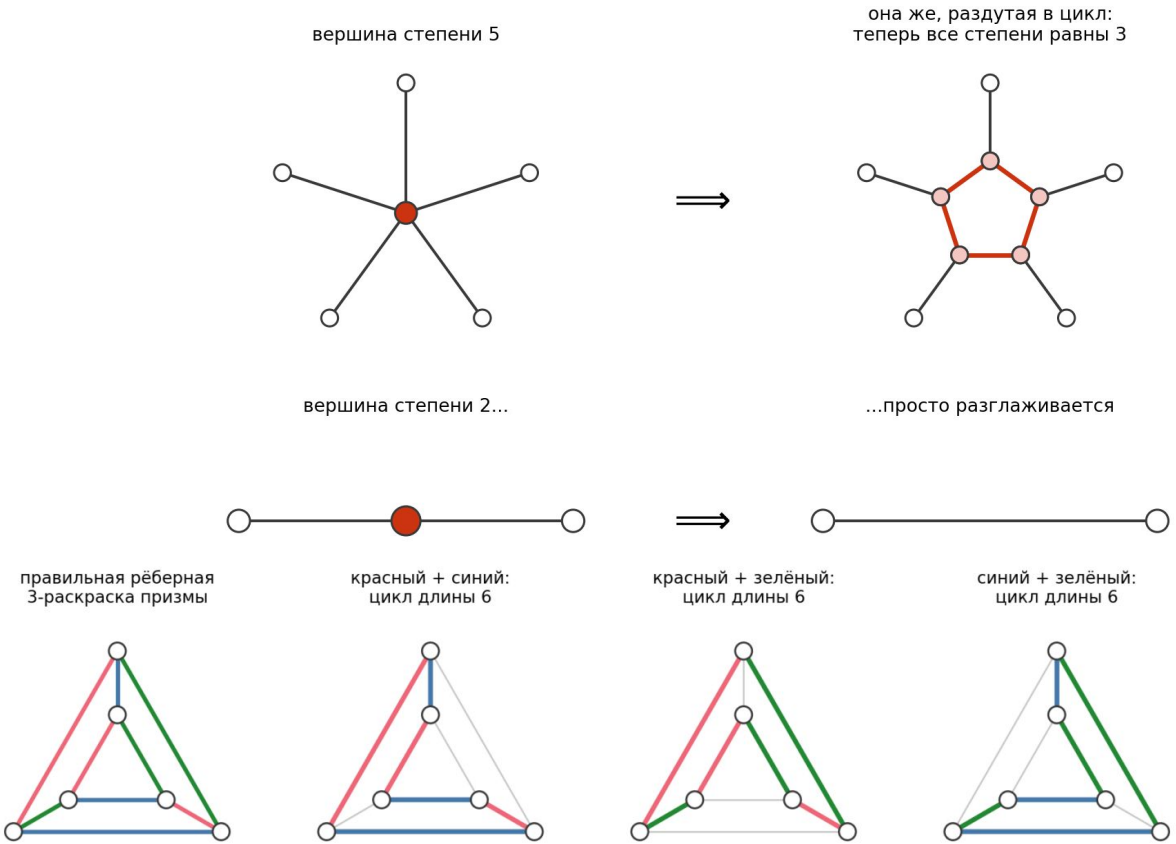
граничные циклы граней:
вдоль каждого ребра — ровно две цветные линии





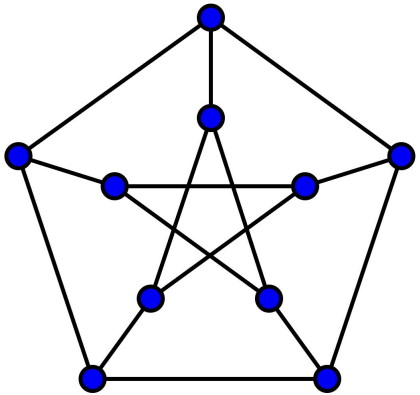
Двойное покрытие циклами

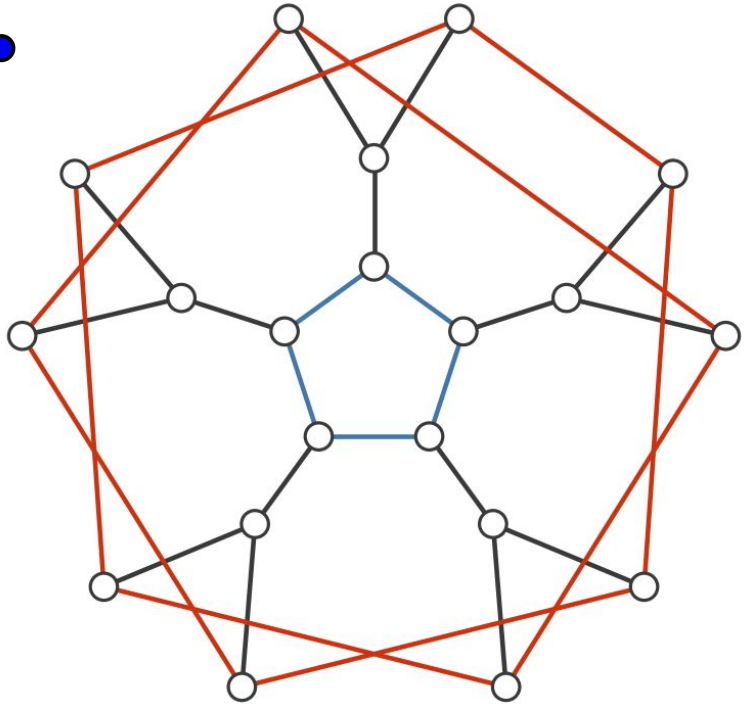
- Достаточно доказать для кубических графов, то есть графов с вершинами степени 3
- Szekeres (1973): если есть 3-раскраска рёбер, то из неё легко сделать покрытие: возьмём по два цвета, они будут давать подграф степени 2, т.е. объединение циклов



Двойное покрытие циклами

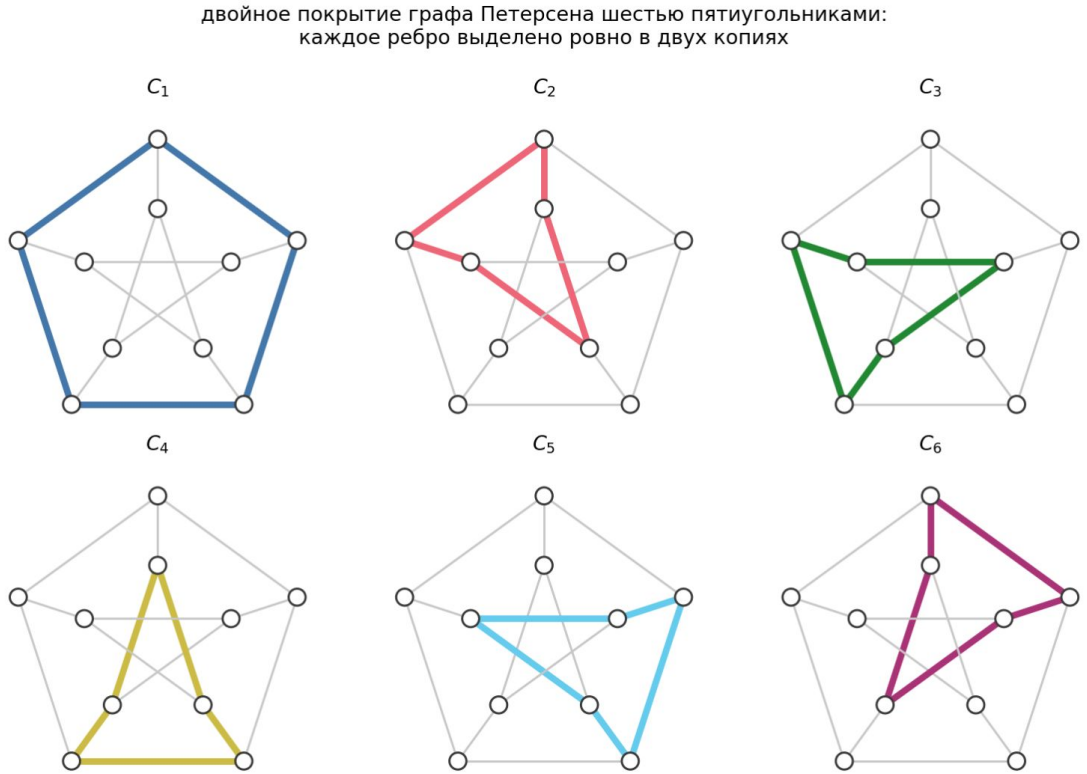
- Значит, минимальный контрпример — это кубический граф без мостов, который нельзя рёберно раскрасить в 3 цвета
- Такие графы (с ещё парой условий невырожденности) называются *снарками*
- Первый снарк — граф Петерсена, потом нашли “цветки Айзекса” и другие примеры





Двойное покрытие циклами

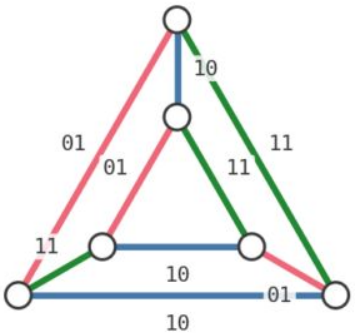
- Снарки, конечно, являются контрпримерами к раскрашиваемости, а не к двойному покрытию
- Но конструкция Секереша для снарков не работает, и нужно было придумать что-то другое



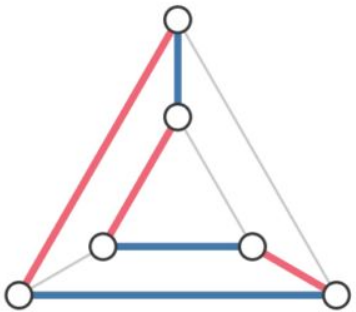
Двойное покрытие циклами

- Главная идея: *потоки*; нас интересуют потоки над $\mathbb{F}_2^k$
- Там правила Кирхгофа сводятся к XOR; и если есть нигде не нулевой поток, например, с двухбитовыми значениями, то можно посмотреть на рёбра, где каждый бит равен 1 и их сумма равна 1, и получится двойное покрытие:

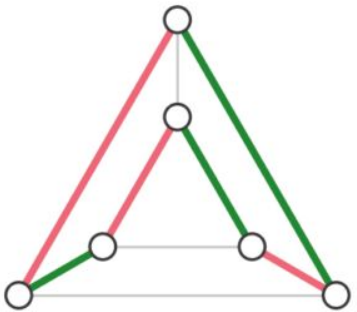
нигде не нулевой $\mathbb{F}_2^2$-поток



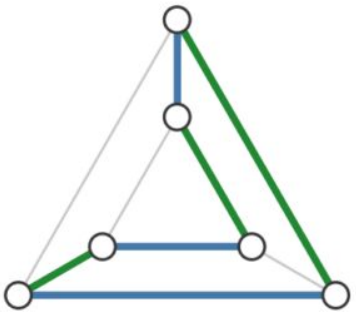
сумма битов = 1:
шестиугольник



второй бит = 1:
шестиугольник



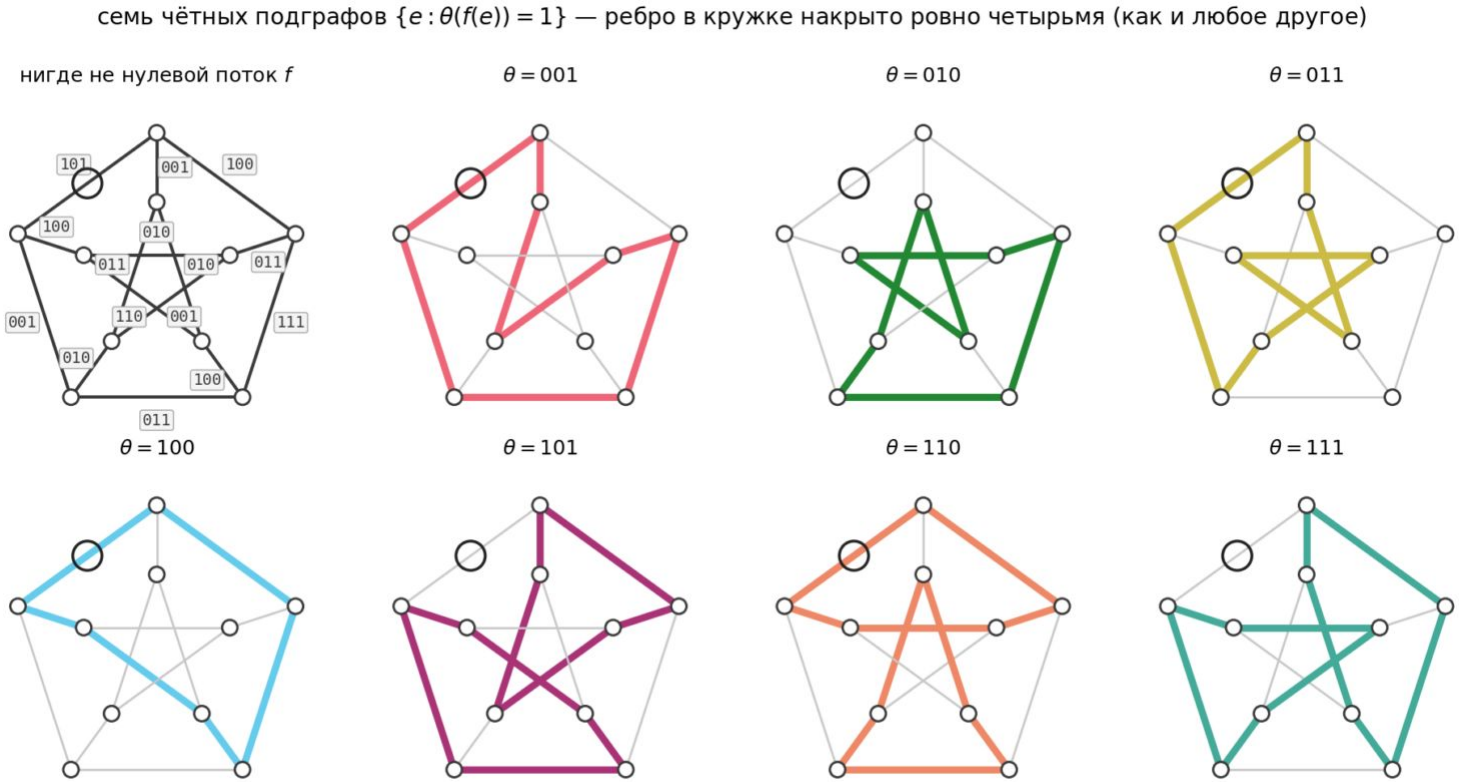
первый бит = 1:
шестиугольник



Двойное покрытие циклами

- У графа без мостов есть нигде не нулевой 8-поток, и его тоже можно превратить в циклы, перечислив способы свернуть три бита в один

$$\{e : \theta(f(e)) = 1\}$$



ЦЕНТРАЛЬНЫЙ УНИВЕРСИТЕТ

Двойное покрытие циклами

- У графа без мостов есть нигде не нулевой 8-пот

семь чётных подграфов $\{e : \theta(f(e))$

нигде не нулевой поток f

То есть с 1980-х известно, что каждое ребро можно покрыть четыре раза.

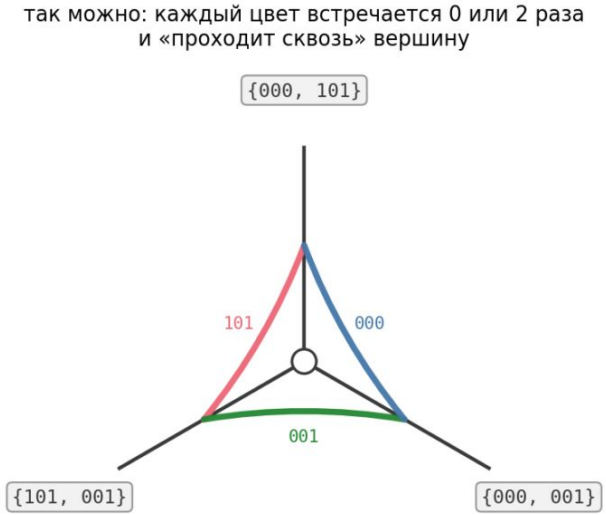
А надо два. Как сделать этот шаг?

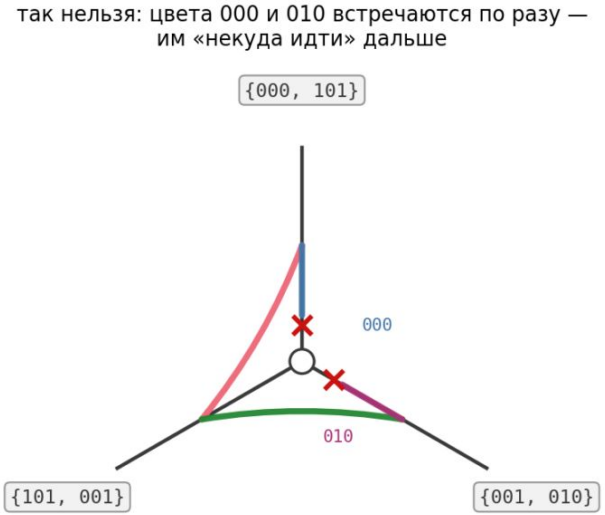
$\theta = 111$

$\{e : \theta($

Двойное покрытие циклами

- Ключевая идея GPT: вместо одного цвета будем назначать рёбрам **пару** цветов; потребуем, чтобы вокруг каждой вершины каждый цвет встречался 0 или 2 раза
- Первая лемма: если такая расстановка пар существует, то у графа есть двойное покрытие циклами

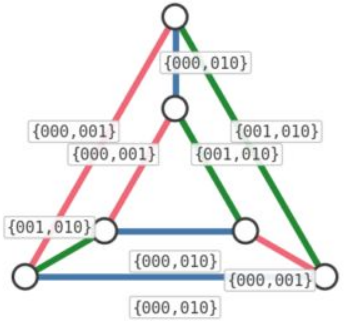




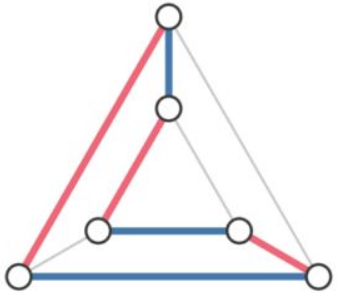
Двойное покрытие циклами

- Это фактически ослабленная 3-раскраска Секереша, только с парами цветов; для каждого цвета соберём подграф из рёбер, у которых он есть
- Главный вопрос: как построить такую раскраску парами цветов?

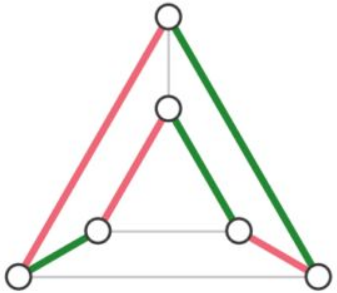
пары из 3-раскраски Секереша



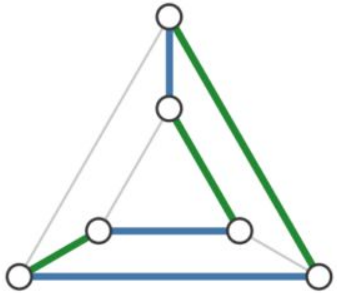
M_{000}: шестиугольник



M_{001}: шестиугольник

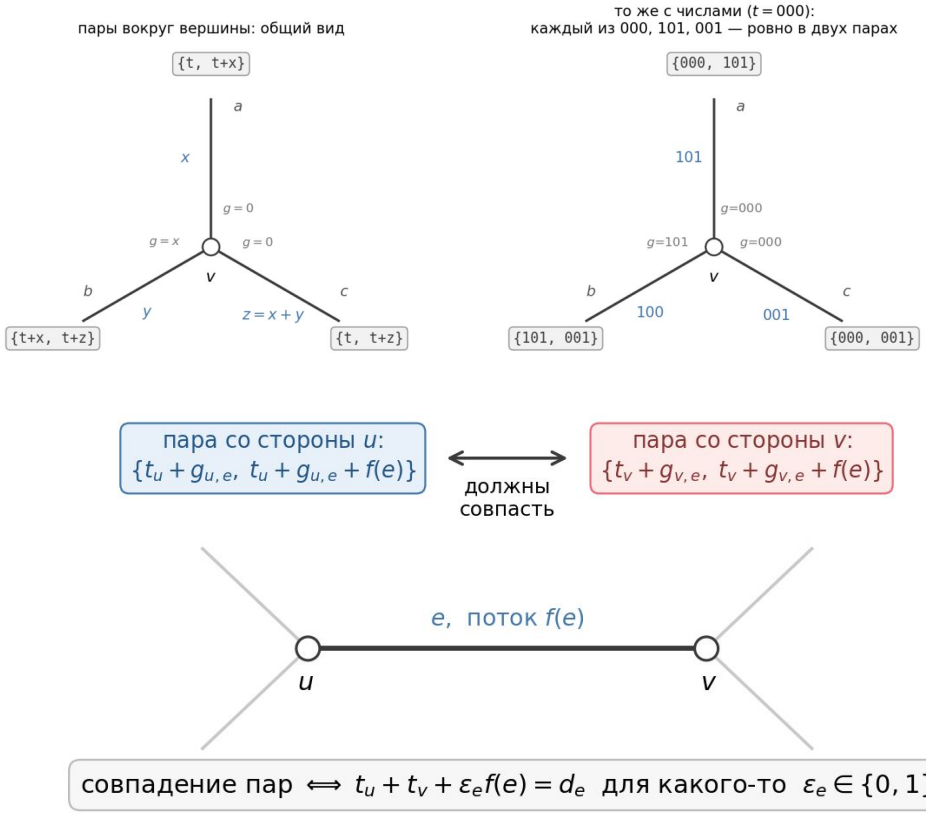


M_{010}: шестиугольник



Двойное покрытие циклами

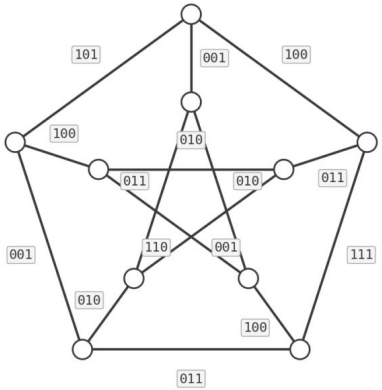
- Вот здесь я немного пропущу технические детали, но они очень простые! В моём посте всё рассказано полностью
- Смысл в том, что условие существования раскраски записывается как система линейных уравнений над полем F_2, а потом доказывается, что система имеет решение



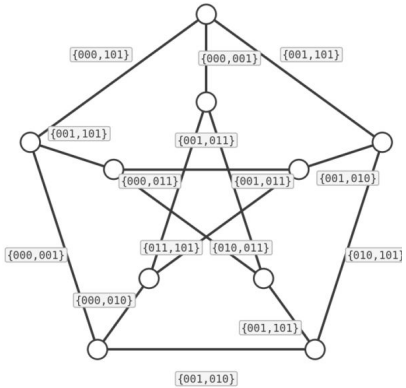
Двойное покрытие циклами

- И всё получается! Более того, доказательство конструктивное, то есть можно сразу получить покрытие полиномиальным алгоритмом. Вот для графа Петерсена пример:

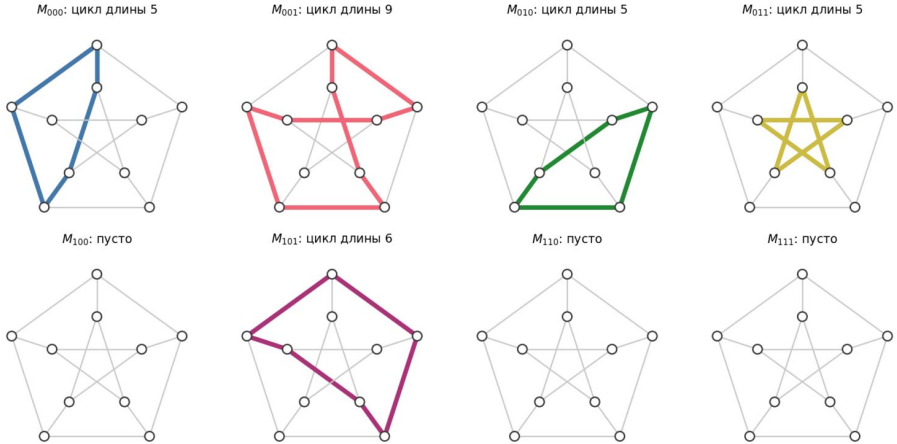
нигде не нулевой $\mathbb{F}_2^3$-поток f



пары P_e, полученные из f решением системы

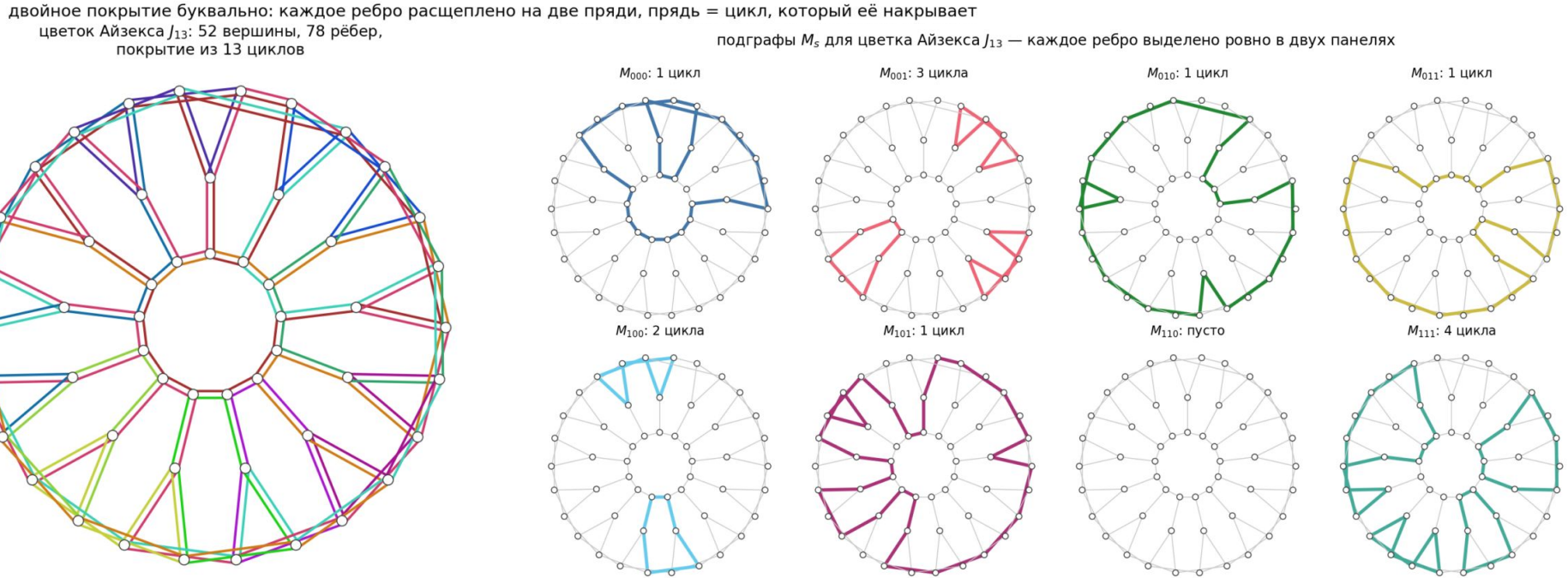


восемь подграфов M_s — каждое ребро выделено ровно в двух: двойное покрытие циклами



Двойное покрытие циклами

- А вот цветок Айзекса J_{13}:



Двойное покрытие циклами

- А вот цветок Айзекса J_{13}:

Но и это ещё не всё!

Проблема якобиана

- Проблема якобиана, стоявшая с 1930-х: если определитель якобиана отображения F — ненулевая константа, то F обратимо (и обратное отображение автоматически полиномиально)
- Главный открытый вопрос о многочленах от нескольких переменных, который можно рассказать любому студенту
- И его 20 июля 2026 г. закрыл Claude Fable, места для ошибки там нет; читайте [мой подробный пост]

levent ✓
@__alpoge__

hello there the jacobian conjecture is false thanx to my close friend akhil for asking about it and my other close friend fable for working during the world cup final

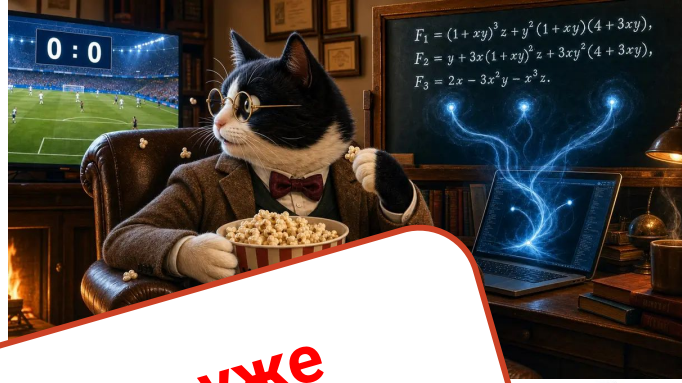
((1+xy)^3 z + y^2 (1+xy) (4+3xy), y + 3 x (1+xy)^2 z + 3 x y^2 (4+3xy), 2 x - 3 x^2 y - x^3 z): \C^3\to \C^3, has jacobian determinant -2, and sends (0, 0, -1/4), (1, -3/2, 13/2), and (-1, 3/2, 13/2) to (-1/4, 0, 0)

5:19 AM · Jul 20, 2026 · **21.9M** Views

$$F_1 = (1+xy)^3 z + y^2(1+xy)(4+3xy),$$
$$F_2 = y + 3x(1+xy)^2 z + 3xy^2(4+3xy),$$
$$F_3 = 2x - 3x^2 y - x^3 z.$$

Проблема якобиана

- Проблема якобиана, стоявшая с 1930-х: если определитель якобиана отображения F — ненулевая константа, то F обратимо (и обратное отображение автоматически поли
- Главный откр многочленах с переменных, к рассказать люб
- И его 20 июля 20 Claude Fable, места для об ки там нет; читайте мой подробный пост



В четверг (23 июля 2026) я уже собирался было отправить презентацию, но…

end akhil
uring the

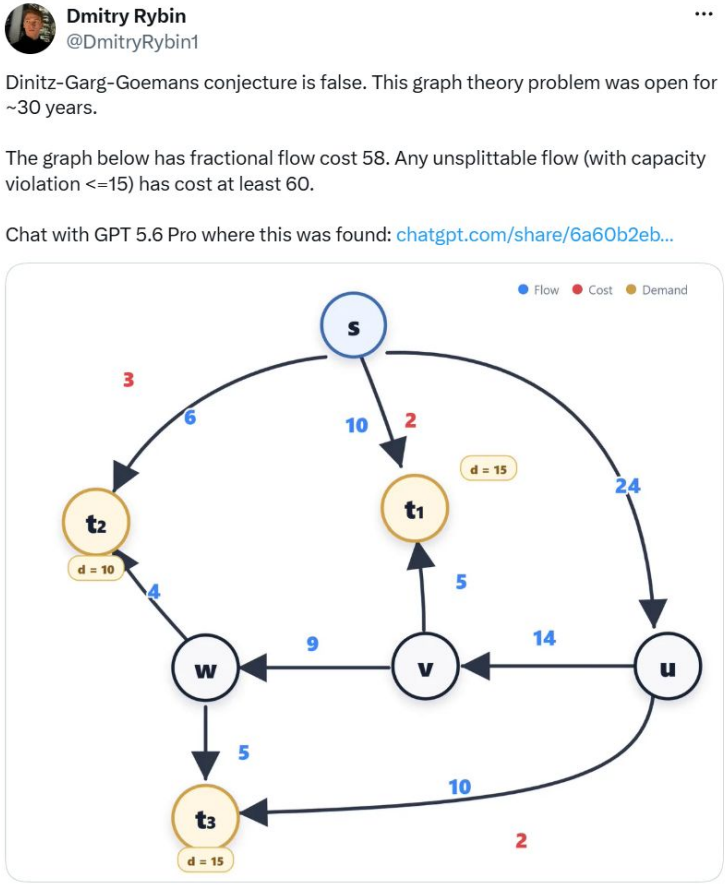
(1+xy)^2 z + 3 x y^2 (4+3xy), 2 x
3, has jacobian determinant -2, and sends
2, 13/2), and (-1, 3/2, 13/2) to (-1/4, 0, 0)

.19 AM · Jul 20, 2026 · **21.9M** Views

$$F_1 = (1+xy)^3 z + y^2(1+xy)(4+3xy),$$
$$F_2 = y + 3x(1+xy)^2 z + 3xy^2(4+3xy),$$
$$F_3 = 2x - 3x^2 y - x^3 z.$$

Гипотеза Диница — Гарга — Гёманса

- Dinitz-Garg-Goemans conjecture: дробный поток можно превратить в неделимый (целочисленный), не увеличивая общую цену и увеличивая нагрузку одного ребра не более чем на максимальный спрос (demand) одной вершины-стока
- Дмитрий Рыбин получил контрпример от GPT 5.6; пожалуй, самое интересное здесь в том, *как* он его получил…



Гипотеза Диница — Гарга — Гёманса

Tuesday 12:29 PM

Uploaded a file

Construct a counterexample to general (non-planar) case of Dinitz Garg Goemans conjecture. You should do a breakthrough and find a structured counterexample.

Worked for 52m 52s

Research conclusion

I did **not** obtain a valid counterexample, and presenting the construction below as one would be mathematically false. The requested object would disprove a conjecture that is still open in the current literature.

Tuesday 1:32 PM

please continue research and find a complete unconditional counterexample

Worked for 89m 12s

Status after the extended search

I have **not found a complete unconditional counterexample**. Every construction that initially appeared to separate the cost condition acquired an additional unsplittable routing once **all** source–terminal paths were taken into account. Exact separation then returned value 0.

Tuesday 3:03 PM

Continue the search. Have a clear strategy obtained from deeper understanding of the problem structure.

Worked for 93m 57s

Research update

I continued the search, but I do **not yet have a positive, exact separation certificate**, so there is still no unconditional counterexample to report. The conjecture remains open in the current literature. Its exact

Tuesday 5:21 PM

it's enough of partial results. let's finish with a complete unconditional counterexample

Worked for 88m 24s

Complete finite counterexample

Under Conjecture 1.3 exactly as stated in the uploaded paper—find an unsplittable flow y with

$$y_a \le x_a + d_{\max} \quad \forall a, \qquad c^{\mathsf{T}} y \le c^{\mathsf{T}} x$$

for every feasible fractional flow x and nonnegative cost vector c—the instance below is a counterexample.

Dinitz_Garg_Goemans

ЦЕНТРАЛЬНЫЙ УНИВЕРСИТЕТ

Гипотеза Диница — Гарга — Гёманса

Tuesday 12:29 PM

Uploaded a file

Construct a counterexample to general (non-planar) case of Dinitz Garg Goemans conjecture. You should do a breakthrough and find a structured counterexample.

Worked for 52m 52s

Research conclusion

I did **not** obtain a valid counterexample, and pr
mathematically false. The requested object wou
literature.

Tues

please continue re
counterexample

Worked for 89m 12s

Status after the extended search

I have **not found a complete unconditional counterexample**. Every construction that initially appeared to separate the cost condition acquired an additional unsplittable routing once **all** source–terminal paths were taken into account. Exact separation then returned value 0.

Tuesday

om deeper

ere is still no
current literature. Its exact

's enough of partial results. let's finish with a complete unconditional counterexample

Worked for 88m 24s

Complete finite counterexample

Under Conjecture 1.3 exactly as stated in the uploaded paper—find an unsplittable flow y with

$$y_a \le x_a + d_{\max} \quad \forall a, \qquad c^{\mathsf{T}} y \le c^{\mathsf{T}} x$$

for every feasible fractional flow x and nonnegative cost vector c—the instance below is a counterexample.

Dinitz_Garg_Goemans

Есть ощущение, что we ain't seen nothing yet...

5. Заключение

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Cicero (не совсем)

ЦЕНТРАЛЬНЫЙ УНИВЕРСИТЕТ

Важное напоминание о скорости прогресса

January 2022

Model Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. 5 + 6 = 11. The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The cafeteria had 23 apples originally. They used 20 to make lunch. So they had 23 - 20 = 3. They bought 6 more apples, so they have 3 + 6 = 9. The answer is 9. ✓

September 2023

Problem: Suppose that the sum of the squares of two complex numbers x and y is 7 and the sum of their cubes is 10. List all possible values for $x + y$, separated by commas.

Let's use `Sympy` to calculate and print all possible values for $x + y$, r_1

...

Removing duplicates, the possible values for $x + y$ are \boxed{-5, 1, 4}. ✓ r_2

March 2024

Problem:
Let a, b, c be positive integers. Prove that it is impossible to have all of the three numbers $a^2 + b + c, b^2 + c + a, c^2 + a + b$ to be perfect squares.

September 2025

Prompt to GPT-5

Consider the problem of maximizing a function F from $[0,1]^n$ to the reals that is the sum of a non-negative monotonically increasing DR-submodular function G and a non-negative DR-submodular function H over a solvable down-closed polytope P. I would like to bound the performance on this problem from the NeurIPS 2021 paper "Submodular + Concave" which is attached. Specifically, if x is the output vector of this algorithm and o is the vector in P maximizing F, then I would like to lower bound $F(x)$ with an expression of the form $\alpha * G(o) + \beta * H(o) - err$, where α and β are constants. err should be a function that depends only on the error parameter ϵ of the algorithm, the diameter D of the polytope P, and the smoothness parameters L_G and L_H of G and H, respectively, and goes to zero as ϵ goes to zero. Please give the best such bound that you prove (a bound is considered better if the values of the constants α and β are larger). Provide a mathematically rigorous and well explained proof for the bound you come up with.

April 2026

Is it true that, for any x, if $A \subset [x, \infty)$ is a primitive set of integers (so that no distinct elements of A divide each other) then

$$\sum_{a \in A} \frac{1}{a \log a} < 1 + o(1),$$

where the $o(1)$ term $\to 0$ as $x \to \infty$?

#1196: [ESS68b][Er80,p.101] number theory | primitive sets

ЦЕНТРАЛЬНЫЙ УНИВЕРСИТЕТ

January 2022

Model Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. 5 + 6 = 11. The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The cafeteria had 23 apples originally. They used 20 to make lunch. So they had 23 - 20 = 3. They bought 6 more apples, so they have 3 + 6 = 9. The answer is 9. ✓

September 2023

Problem: Suppose that the sum of the squares of two complex numbers x and y is 7 and the sum of their cubes is 10. List all possible values for $x + y$, separated by commas.

Let's use `Sympy` to calculate and print all possible values for $x + y$, r_1

...

Removing duplicates, the possible values for $x + y$ are \boxed{-5, 1, 4}.

March 2024

Problem:

Let a, b, c be positive integers. Prove that it is impossible to have $a^2 + b + c, b^2 + c + a, c^2 + a + b$ to be perfect squares.

September 2025

Prompt to GPT-5

Consider the probl... of a non-negati... DR-submodula... the performan... which is attach... in P maximizing $\alpha * G(o) + \beta * H(...$ only on the error ... smoothness parame... Please give the best ... the constants α and ... for the bound you co... proof

April 2026

Is it true that, for an... set of integers (so that no distin... de each other) then

$$\sum_{a \in A} \frac{1}{a \log a} < 1 + o(1),$$

where the $o(1)$ term $\to 0$ as $x \to \infty$?

#1196: [ESS68b][Er80,p.101] number theory | primitive sets

Эту картинку я сделал не так давно, но она уже безнадёжно устарела...

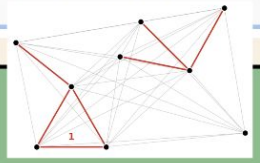
Ва... ...и прогресса

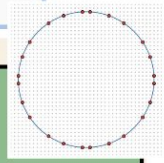
ЦЕНТРАЛЬНЫЙ УНИВЕРСИТЕТ

May 2026

the constants α and β are larger). Provide a mathematically rigorous and well explained proof for the bound you come up with.

DISPROVED - $500





Does every set of n distinct points in $\mathbb{R}^2$ contain at most $n^{1+O(1/\log\log n)}$ many pairs which are distance 1 apart?

Теорема (OpenAI, 2026). Существует константа $\delta > 0$ и бесконечная последовательность натуральных чисел n, для которых $\nu(n) \geq n^{1+\delta}$.

July 2026

A PROOF OF THE CYCLE DOUBLE COVER CONJECTURE

OPENAI

ABSTRACT. We prove the cycle double cover conjecture, posed by Tutte, Itai and Rodeh, Szekeres, and Seymour, which asserts that every bridgeless undirected graph has a collection of cycles which covers every edge exactly twice.

July 2026

levent ✓
@__alpoge__

hello there the jacobian conjecture is false thanx to my close friend akhil for asking about it and my other close friend fable for working during the world cup final

((1+xy)^3 z + y^2 (1+xy) (4+3xy), y + 3 x (1+xy)^2 z + 3 x y^2 (4+3xy), 2 x - 3 x^2 y - x^3 z): \C^3\to \C^3, has jacobian determinant -2, and sends (0, 0, -1/4), (1, -3/2, 13/2), and (-1, 3/2, 13/2) to (-1/4, 0, 0)

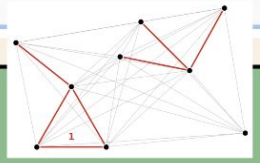
5:19 AM · Jul 20, 2026 · **21.9M** Views

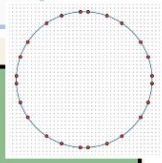
Теперь великие открытые вопросы падают регулярно

ЦЕНТРАЛЬНЫЙ УНИВЕРСИТЕТ

the constants α and β are larger). Provide a mathematically rigorous and well explained proof for the bound you come up with.

DISPROVED - $500





Does every set of n distinct points in $\mathbb{R}^2$ contain at most $n^{1+O(1/\log\log n)}$ many pairs which are distance 1 apart?

Теорема (OpenAI, 2026). Существует конста
последовательность натурал

May 2026

July

July 2026

Tuesday 5:21 PM

it's enough of partial results. let's finish with a complete unconditional counterexample

...re is false thanx to my close friend akhil
...and my other close friend fable for working during the
...cup final

((1+xy)^3 z + y^2 (1+xy) (4+3xy), y + 3 x (1+xy)^2 z + 3 x y^2 (4+3xy), 2 x - 3 x^2 y - x^3 z): \C^3\to \C^3, has jacobian determinant -2, and sends (0, 0, -1/4), (1, -3/2, 13/2), and (-1, 3/2, 13/2) to (-1/4, 0, 0)

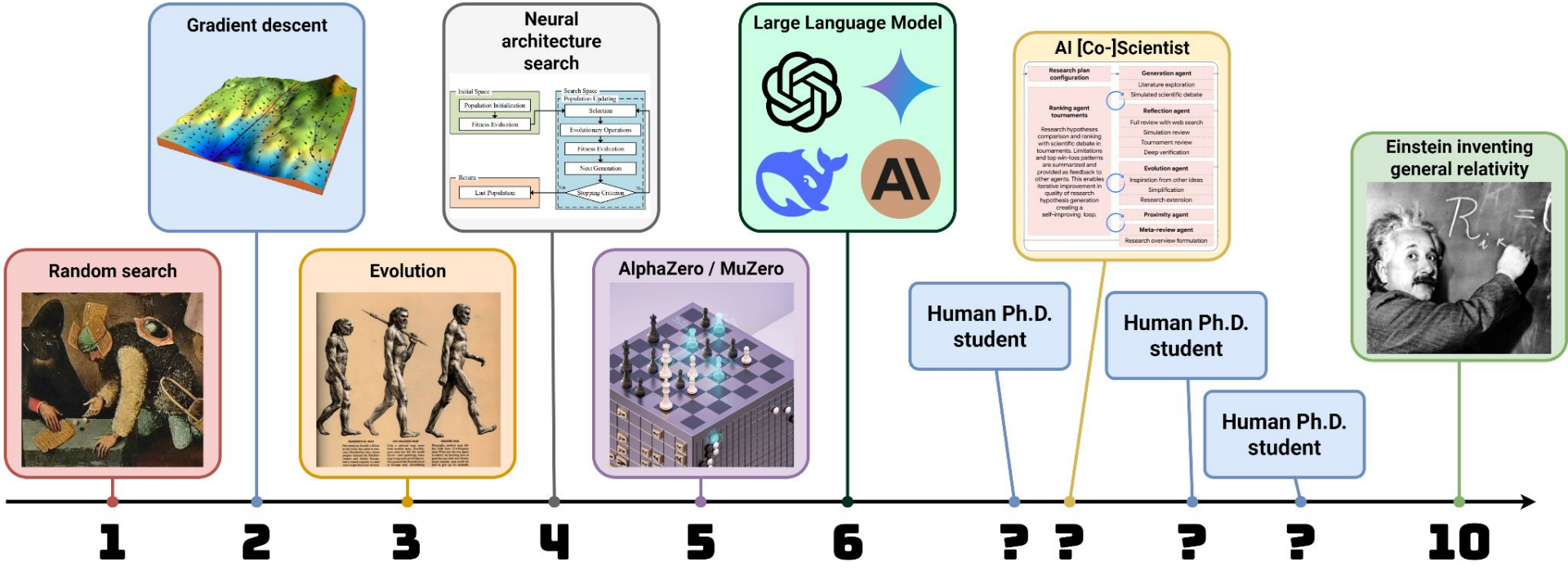
5:19 AM · Jul 20, 2026 · **21.9M** Views

...e

...регулярно

Уровни креативности

- А что на самом деле, без лишнего скептицизма и без шапкозакидательства? Научный поиск – это тоже оптимизационный процесс; где мы сейчас?



Спасибо за внимание!

@SINECOR