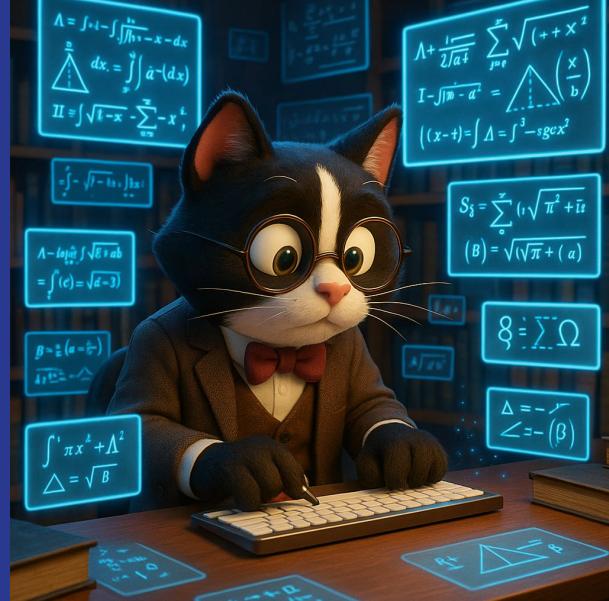


# КИИ-2025

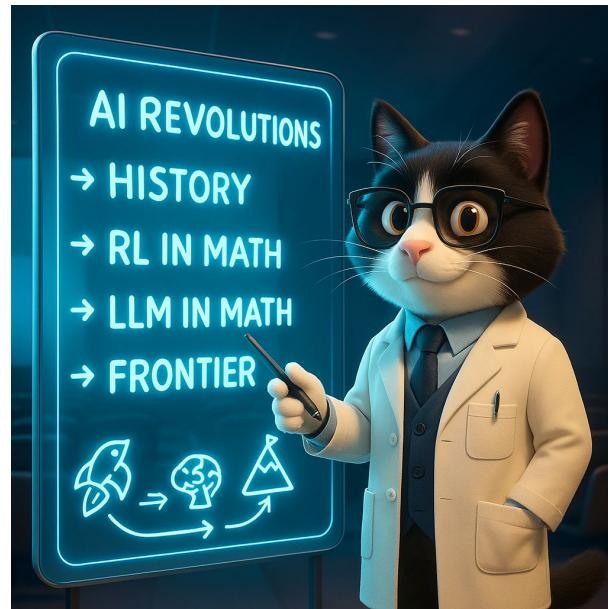
# Рассуждающие языковые модели (и их успехи в науке)

Сергей  
Николенко  
7 октября 2025



# Наш план

- Революции искусственного интеллекта: прогресс ускоряется
- Языковые модели
- Методы chain-of-thought
- Рассуждающие модели
- RL для математики и CS
- Как применить к математике LLM
- Новые результаты AI в науке



# 1. Революции AI

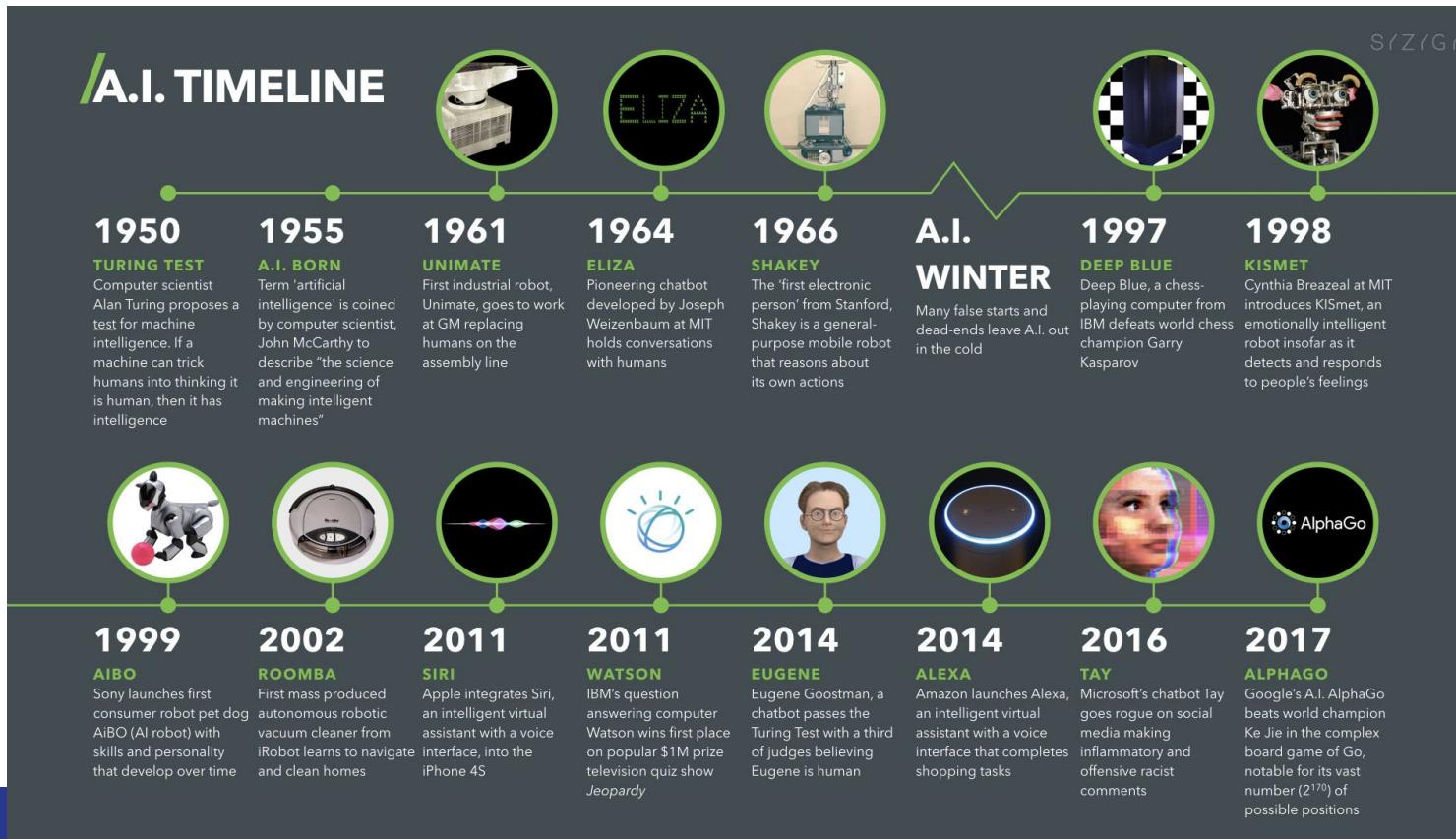


Революции, происходящие на поверхности жизни, ничего существенного никогда не меняют и не открывают, они лишь... по-новому переставляют все те же элементы и являются старые образы в новых одеяниях. Революция всегда есть в значительной степени маскарад, и если сорвать маски, то можно встретить старые, знакомые лица. Новые души рождаются позже, после глубокого перерождения и осмысливания опыта революции.

*Николай Бердяев. Духи русской революции*

# AI timelines

- Обычный AI timeline
- Источник



# AI timelines

- Обычный AI timeline



Cameron R. Jones  
Department of Cognitive Science  
UC San Diego  
San Diego, CA 92119  
cameron@ucsd.edu

Benjamin K. Bergen  
Department of Cognitive Science  
UC San Diego  
San Diego, CA 92119  
bkbergen@ucsd.edu

Watson  
IBM's question answering computer  
Watson wins first place  
on popular \$1M prize  
television quiz show  
*Jeopardy!*

2014

EUGENE  
Eugene Goostman, a  
chatbot passes the  
Turing Test with a third  
of judges believing  
Eugene is human

2014

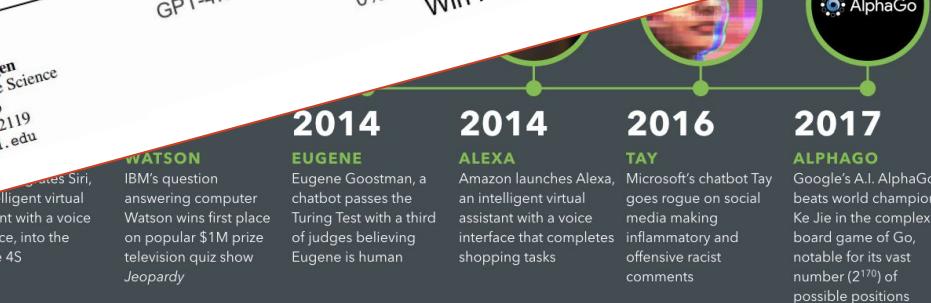
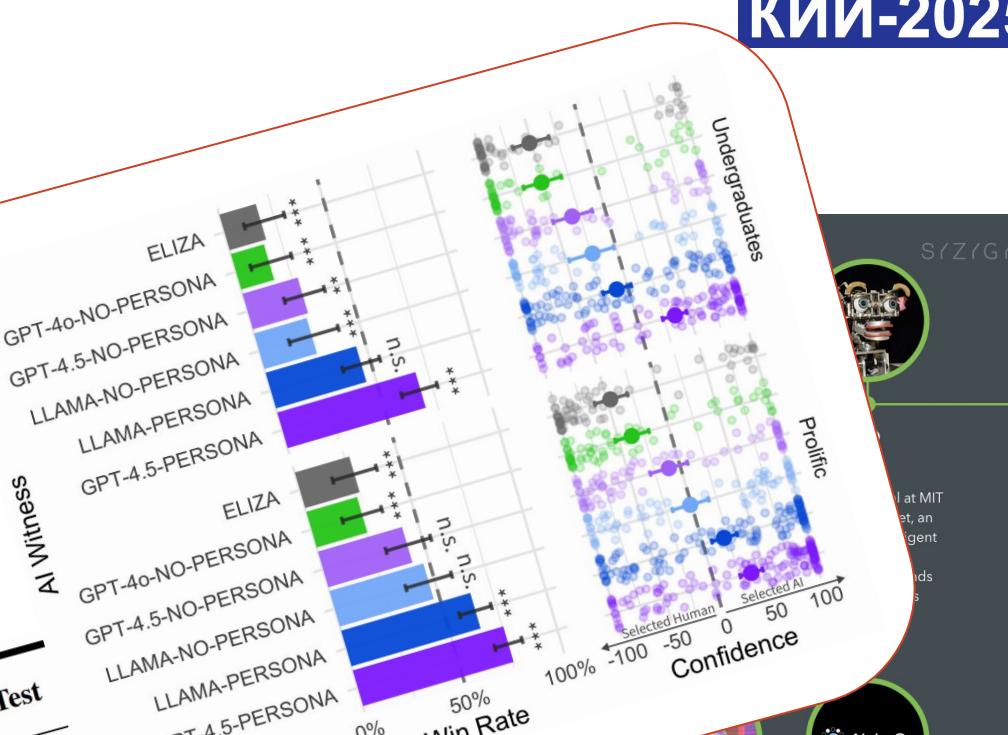
ALEXA  
Amazon launches Alexa,  
an intelligent virtual  
assistant with a voice  
interface that completes  
shopping tasks

2016

TAY  
Microsoft's chatbot Tay  
goes rogue on social  
media making  
inflammatory and  
offensive racist  
comments

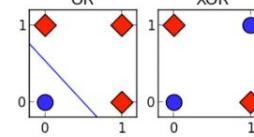
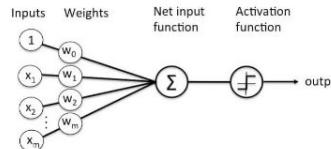
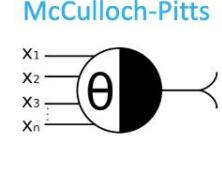
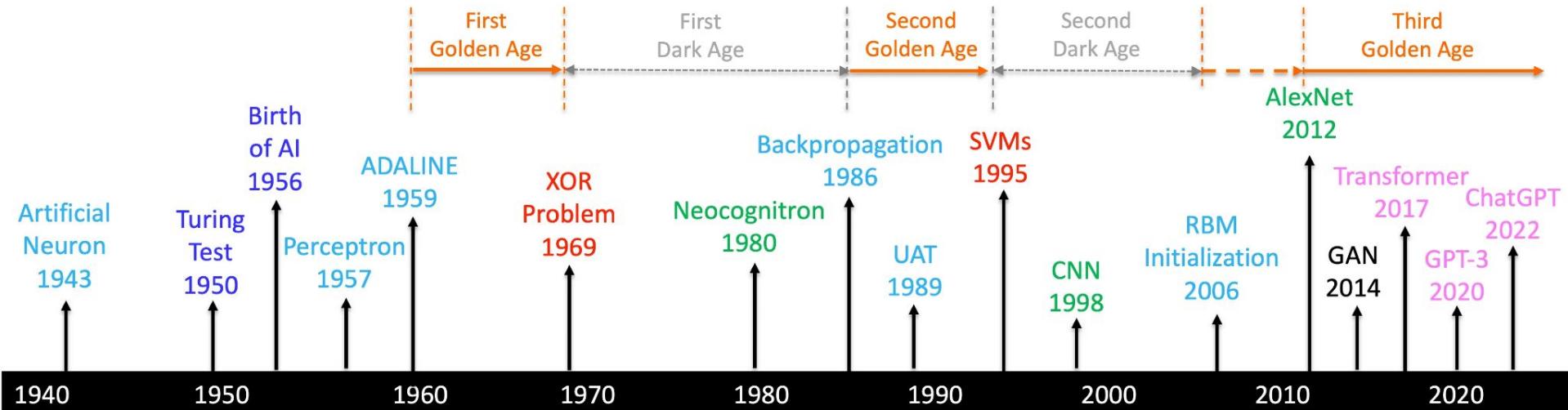
2017

ALPHAGO  
Google's A.I. AlphaGo  
beats world champion  
Ke Jie in the complex  
board game of Go,  
notable for its vast  
number ( $2^{176}$ ) of  
possible positions



# Глубокое обучение

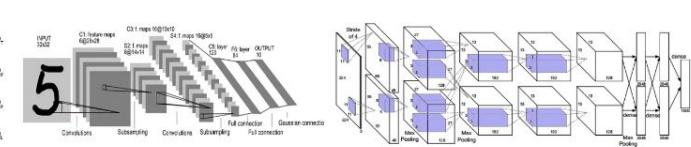
- Нейросети были всегда, ещё до AI как науки, но заработали не так давно ([источник](#))



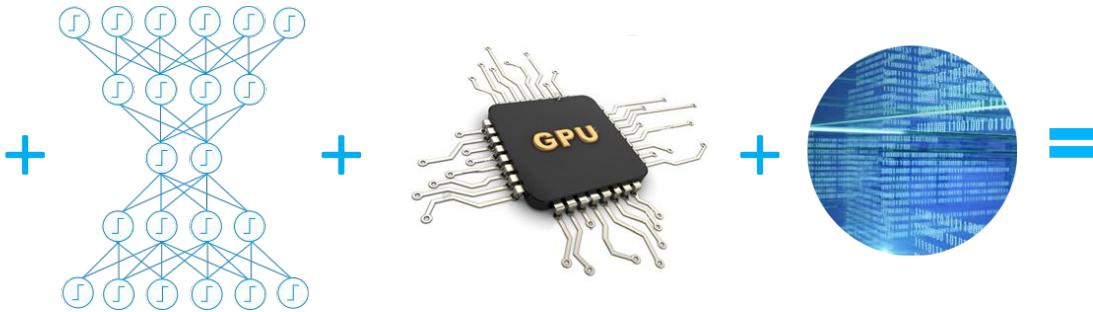
Rumelhart, Hinton et al.

LeCun

Hinton-Ruslan Krizhevsky et al. Vaswani



# Революция глубокого обучения



- В 2006-2007 годах произошла революция глубокого обучения (deep learning), началась третья волна хайпа нейронных сетей
- Революция была не только идейной, но и технической: теперь есть данные и вычислительные мощности
- Сейчас глубокие нейронные сети дают лучшие результаты почти во всех областях ML

# Кстати, о Джейфри Хинтоне



PRESS RELEASE

8 October 2024

## The Nobel Prize in Physics 2024

The Royal Swedish Academy of Sciences has decided to award the Nobel Prize in Physics 2024 to

**John J. Hopfield**

Princeton University, NJ, USA

**Geoffrey E. Hinton**

University of Toronto, Canada

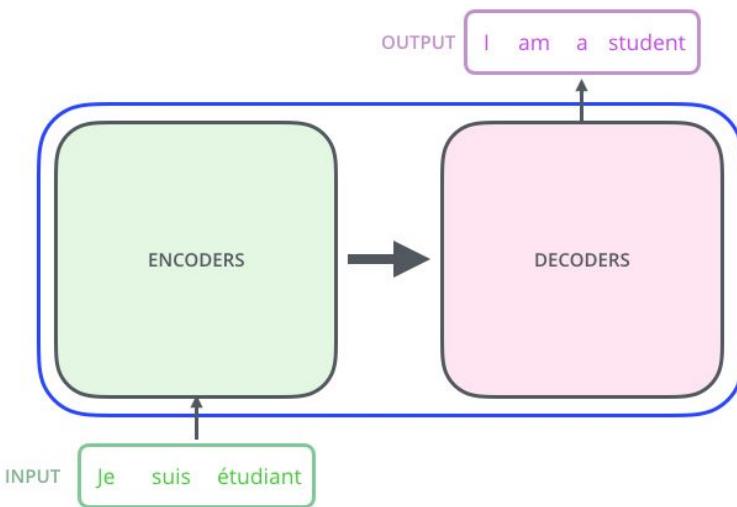
*“for foundational discoveries and inventions that enable machine learning with artificial neural networks”*

**They trained artificial neural networks using physics**



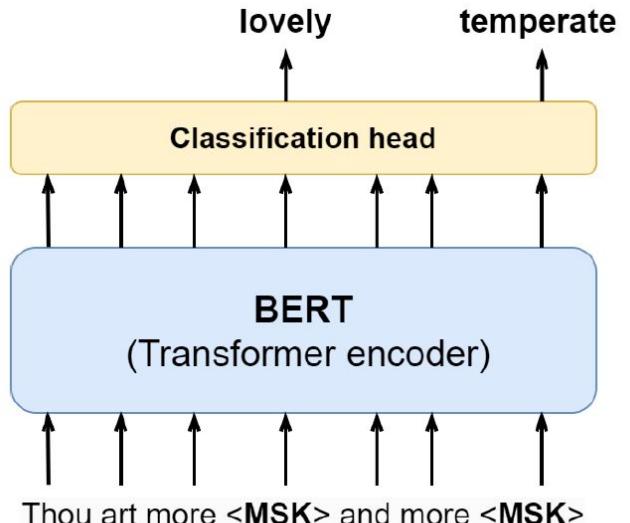
# Революция трансформеров

- Трансформер – это модель на основе механизма внимания (self-attention)
- Первый трансформер ([Vaswani et al., 2017](#)) имел encoder-decoder архитектуру

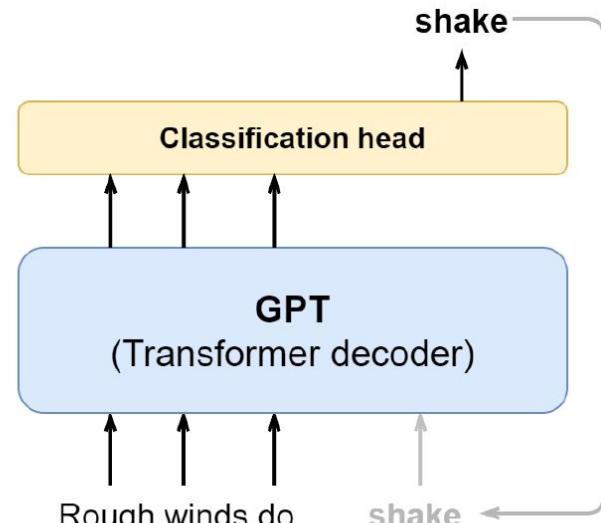


# BERT и GPT

- Когда появились трансформеры, оказалось, что кодировщик и декодировщик могут быть очень полезны и по отдельности: семейства BERT и GPT



(a) BERT: masked language modeling

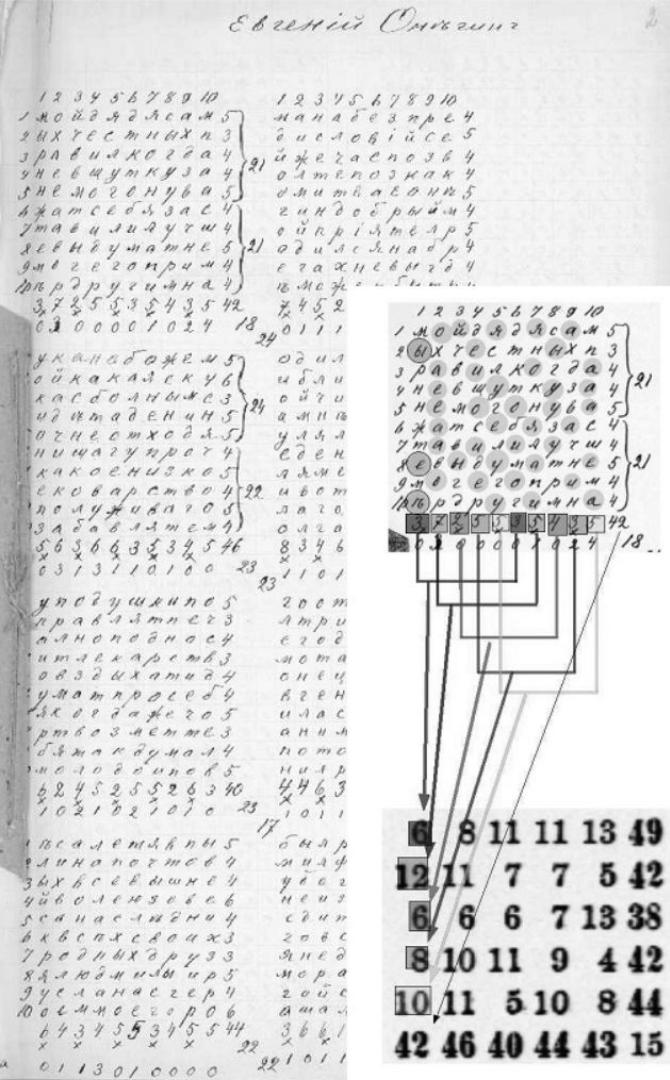
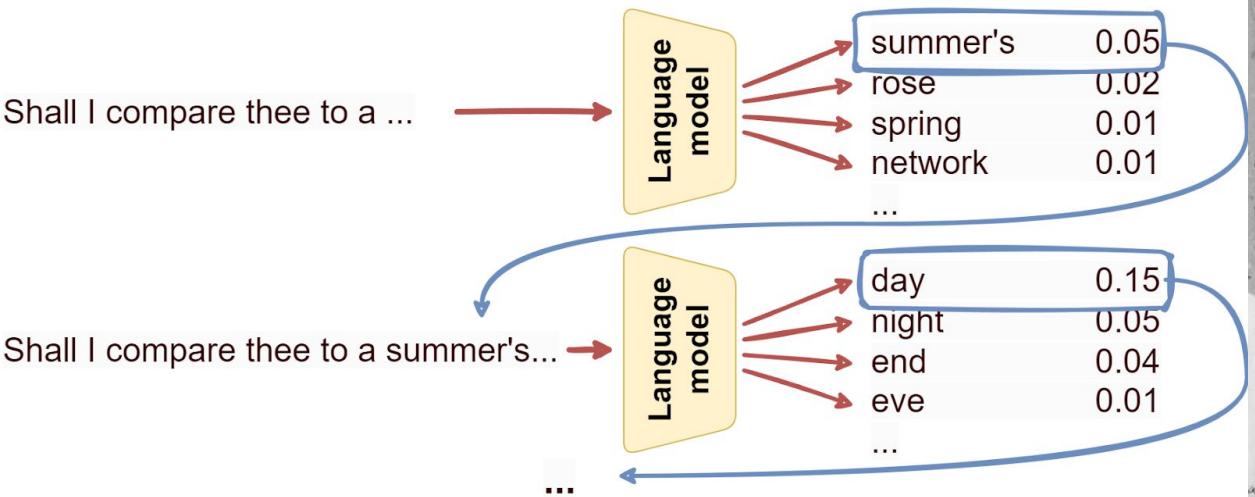


(b) GPT: language modeling

- Из GPT получились языковые модели; что это такое?..

# Как устроены языковые модели

- **Языковая модель** – это вид постановки задачи: авторегрессивно порождать следующий токен на основе уже порождённых
- То есть языковая модель должна “прочитать” участок текста, а затем по результату чтения классифицировать следующий токен



# От GPT к ChatGPT

- ChatGPT начинается с GPT-3, которая учится ТОЛЬКО предсказывать токены
- Затем – fine-tuning через RLHF (reinforcement learning from human feedback)

Step 1

**Collect demonstration data, and train a supervised policy.**

A prompt is sampled from our prompt dataset.

Explain the moon landing to a 6 year old

A labeler demonstrates the desired output behavior.

Some people went to the moon...

This data is used to fine-tune GPT-3 with supervised learning.

SFT

Step 2

**Collect comparison data, and train a reward model.**

A prompt and several model outputs are sampled.

Explain the moon landing to a 6 year old

A B  
Explain gravity... Explain war...  
C D  
Moon is natural satellite of... People went to the moon...

A labeler ranks the outputs from best to worst.

D > C > A = B

This data is used to train our reward model.

RM  
D > C > A = B

Step 3

**Optimize a policy against the reward model using reinforcement learning.**

A new prompt is sampled from the dataset.

Write a story about frogs

The policy generates an output.

PPO

Once upon a time...

The reward model calculates a reward for the output.

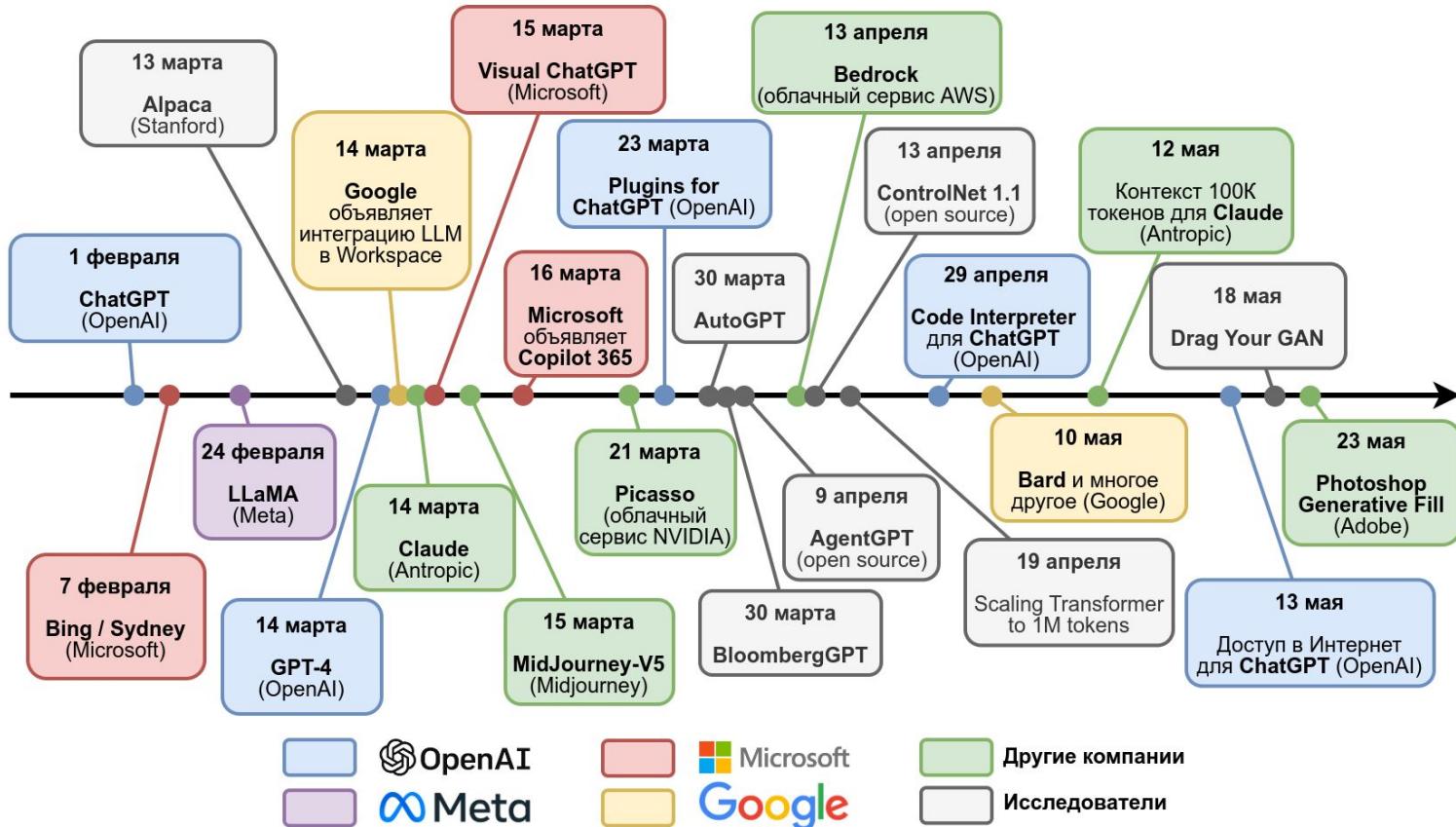
RM

The reward is used to update the policy using PPO.

$r_k$

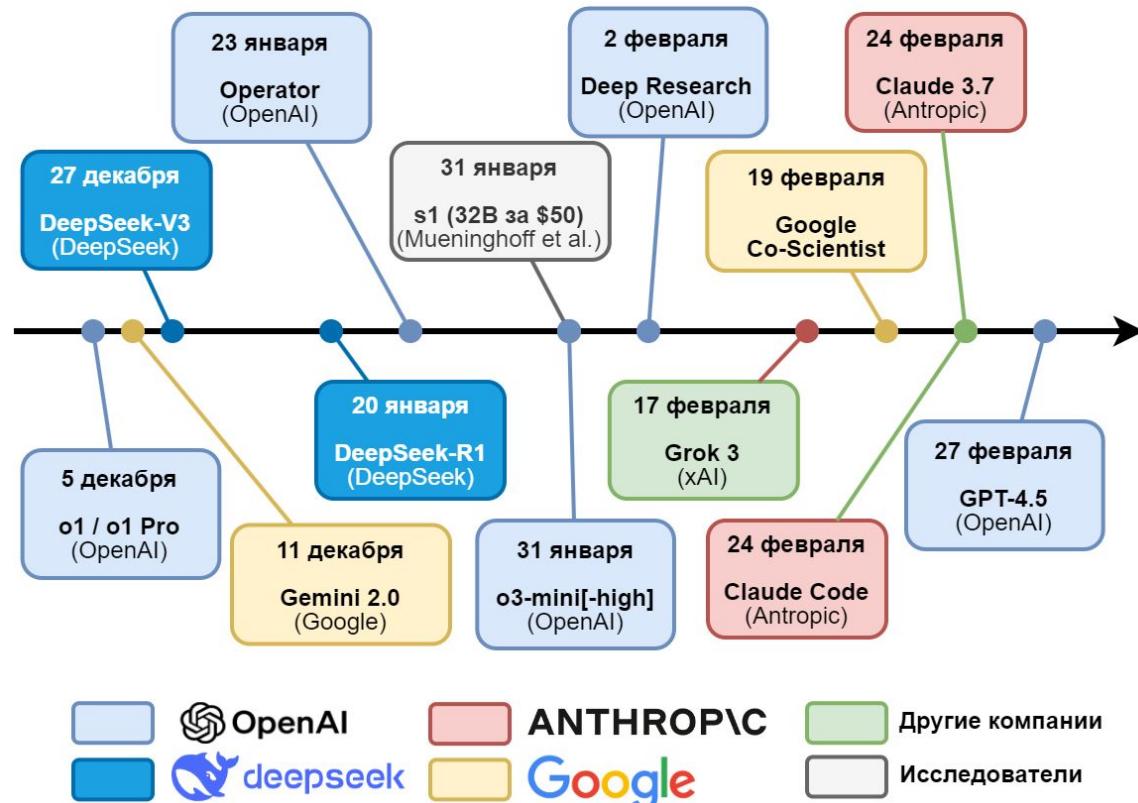
# Развитие AI в эпоху LLM

- И вот AI timeline весны 2023 года
- И с тех пор не то чтобы замедляется...



# Например, в начале 2025 года

- 22 января: [DeepSeek-R1](#)
- 31 января: [s1 \(Mueninghoff et al., 2025\)](#)
- 2 февраля: [OpenAI Deep Research](#)
- 17 февраля: [Grok 3](#)
- 19 февраля: [Google Co-Scientist](#) на основе Gemini 2.0 ([11 декабря 2024](#))
- 24 февраля: [Claude 3.7 Sonnet](#) (hybrid reasoning Claude)
- 27 февраля: [GPT-4.5](#)

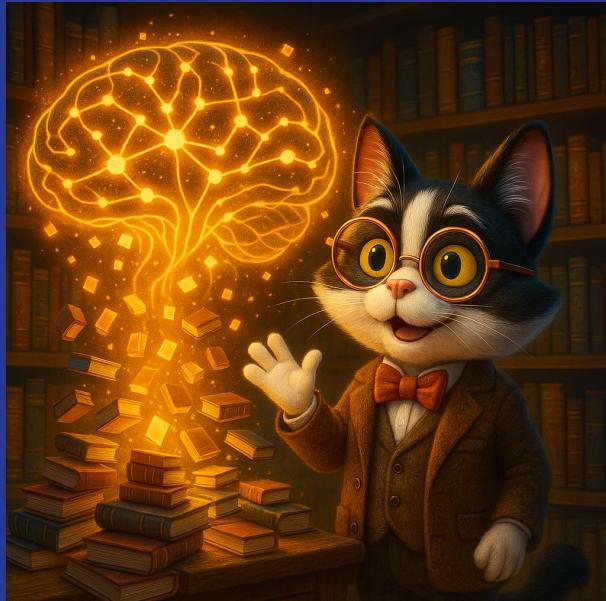


# AGI? Сингулярность?

- Значит ли это, что создание AGI уже близко?
- И если AGI на пороге, что это значит для человечества?
- Приведёт ли это к той самой технологической сингулярности?
- Это вопросы без ответов, хотя люди пытаются прогнозировать ([AI 2027](#))
- Сегодня – о конкретной области, в которой много интересного...



## 2. Языковые модели

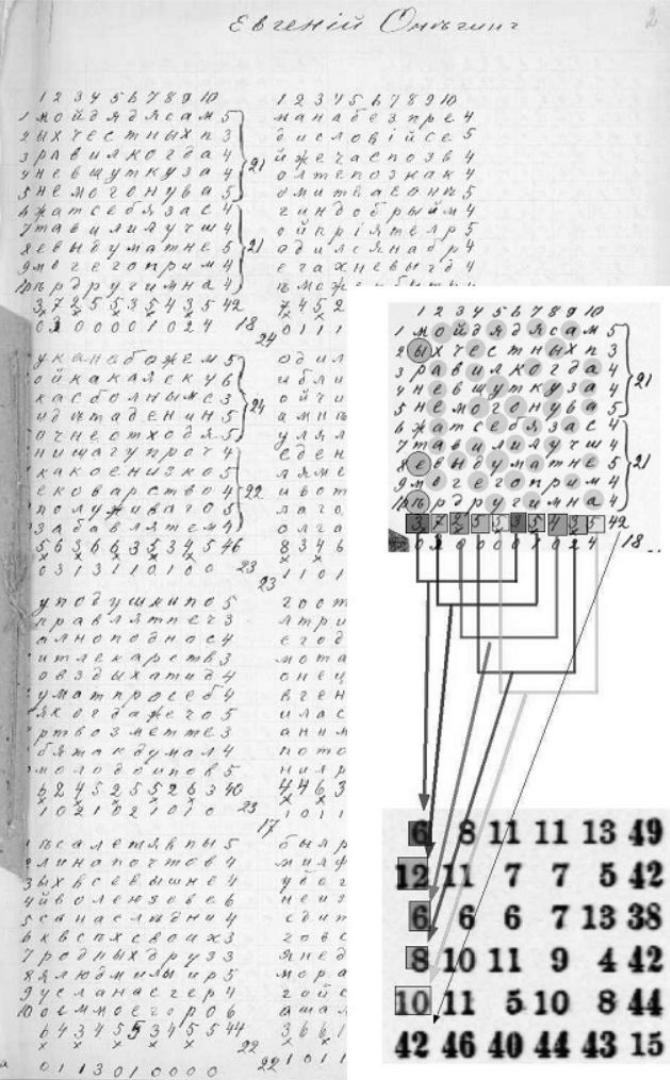
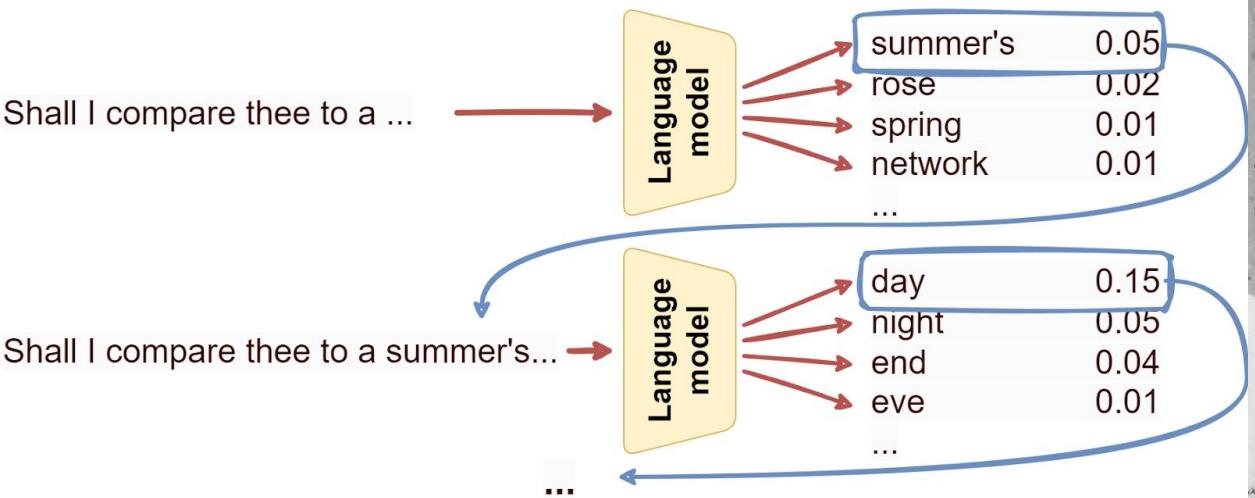


А увлекают меня такие книжки, что как их  
дочитаешь до конца – так сразу подумаешь:  
хорошо бы, если бы этот писатель стал твоим  
лучшим другом и чтоб с ним можно было  
поговорить по телефону, когда захочется.

*Джером Дэвид Сэлинджер*

# Как устроены языковые модели

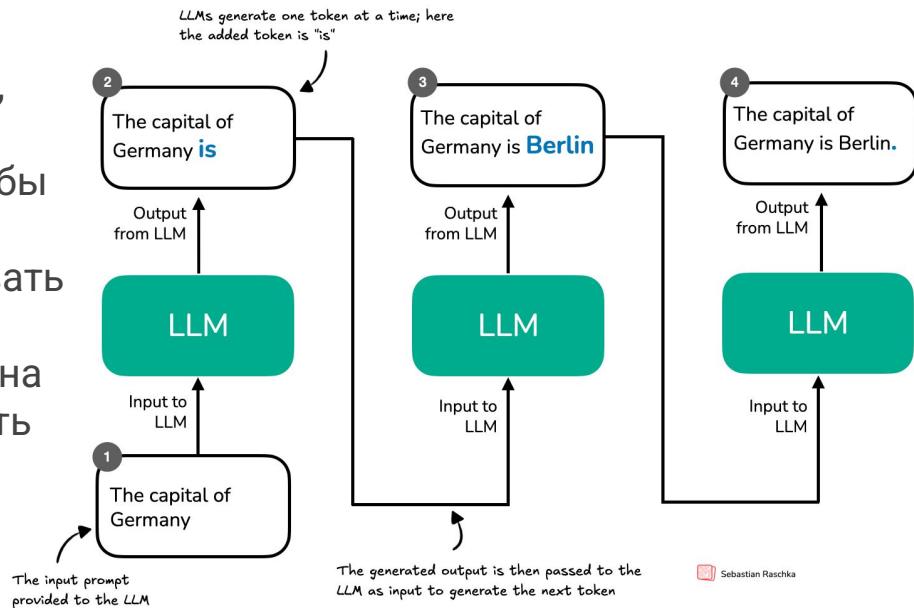
- **Языковая модель (language model)** – это вид постановки задачи: авторегрессивно порождать следующий токен на основе уже порождённых
- То есть языковая модель должна “прочитать” участок текста, а затем по результату чтения классифицировать следующий токен



## ЛМ и тест Тьюринга

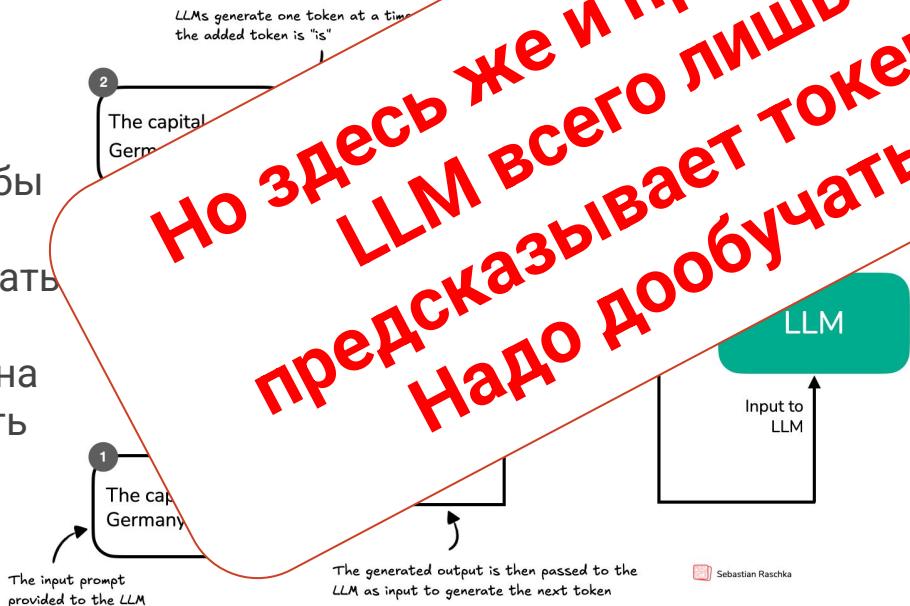
- Почему тест Тьюринга был настолько успешным?  
Потому что на естественном языке можно поговорить о чём угодно

- Аналогично, языковая модель, чтобы всё лучше предсказывать следующий токен, должна уметь решать всё более сложные задачи...



# LM и тест Тьюринга

- Почему тест Тьюринга был настолько успешный?  
Потому что на естественном языке можно о чём угодно
- Аналогично, языковая модель, чтобы всё лучше предсказывать следующий токен, должна уметь решать всё более сложные задачи...

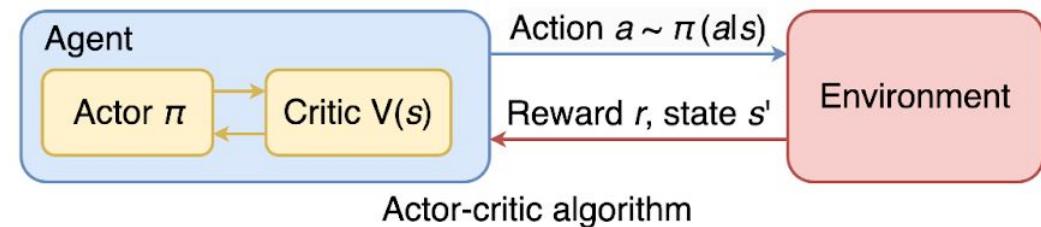
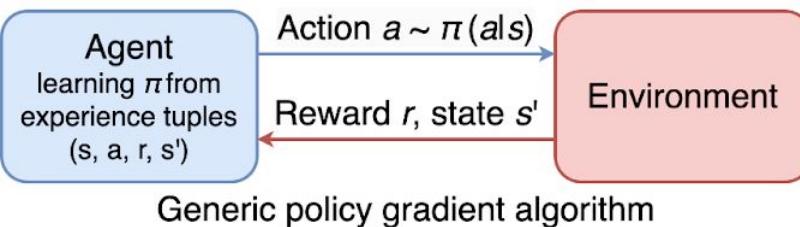


Но здесь же и проблема:  
LLM всего лишь  
предсказывает токены.  
Надо дообучать.



# Обучение с подкреплением

- Обучение с подкреплением (reinforcement learning, RL) – это раздел ML, в котором агент “живёт” в окружающей среде и собирает датасет для обучения по ходу дела
- Для этого нужно суметь реализовать окружающую среду, которая будет давать награду; например, результат партии в шахматы или корректность доказательства в математике



# От GPT к ChatGPT

- ChatGPT начинается с GPT-3, которая учится ТОЛЬКО предсказывать токены
- Затем – fine-tuning через RLHF (reinforcement learning from human feedback)

Step 1

**Collect demonstration data, and train a supervised policy.**

A prompt is sampled from our prompt dataset.

Explain the moon landing to a 6 year old

Some people went to the moon...

A labeler demonstrates the desired output behavior.

SFT

This data is used to fine-tune GPT-3 with supervised learning.

Step 2

**Collect comparison data, and train a reward model.**

A prompt and several model outputs are sampled.

Explain the moon landing to a 6 year old

A B  
Explain gravity... Explain war...  
C D  
Moon is natural satellite of... People went to the moon...

D > C > A = B

A labeler ranks the outputs from best to worst.

This data is used to train our reward model.

RM

D > C > A = B

Step 3

**Optimize a policy against the reward model using reinforcement learning.**

A new prompt is sampled from the dataset.

Write a story about frogs

PPO

Once upon a time...

RM

$r_k$

The policy generates an output.

The reward model calculates a reward for the output.

The reward is used to update the policy using PPO.

- Общая схема ([Stiennon et al., 2020](#))

## RLHF

### 1 Collect human feedback

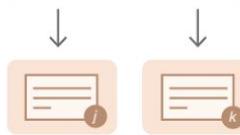
A Reddit post is sampled from the Reddit TL;DR dataset.



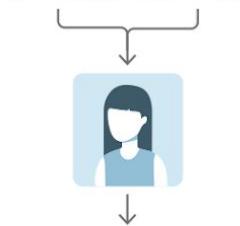
Various policies are used to sample a set of summaries.



Two summaries are selected for evaluation.



A human judges which is a better summary of the post.



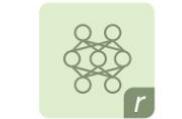
*"j is better than k"*

### 2 Train reward model

One post with two summaries judged by a human are fed to the reward model.



The reward model calculates a reward  $r$  for each summary.



The loss is calculated based on the rewards and human label, and is used to update the reward model.

$$\text{loss} = \log(\sigma(r_j - r_k))$$

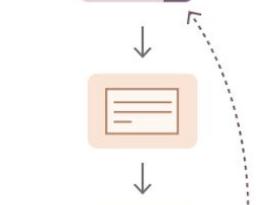
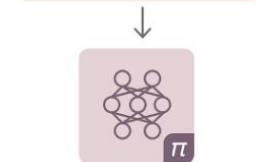
*"j is better than k"*

### 3 Train policy with PPO

A new post is sampled from the dataset.



The policy  $\pi$  generates a summary for the post.



The reward model calculates a reward for the summary.



The reward is used to update the policy via PPO.



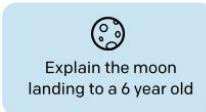
## ● Общая схема ([Ouyang et al., 2022](#))

### RLHF

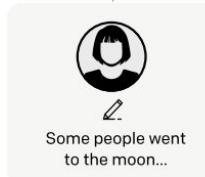
#### Step 1

**Collect demonstration data, and train a supervised policy.**

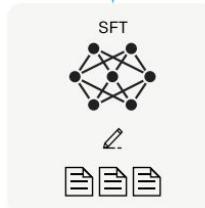
A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.



This data is used to fine-tune GPT-3 with supervised learning.



#### Step 2

**Collect comparison data, and train a reward model.**

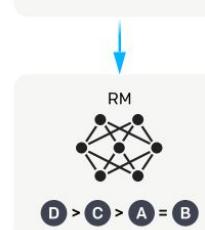
A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.



This data is used to train our reward model.



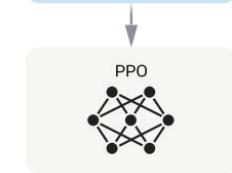
#### Step 3

**Optimize a policy against the reward model using reinforcement learning.**

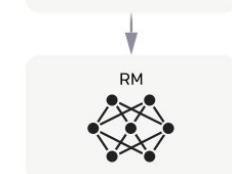
A new prompt is sampled from the dataset.



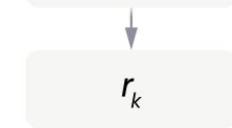
The policy generates an output.



The reward model calculates a reward for the output.



The reward is used to update the policy using PPO.



• Общая схема ([Ouyang et al., 2022](#))

RLHF

Step 1

Collect demonstration data, and train a supervised policy.

A prompt is sampled from our prompt dataset.



Explain the moon landing to a 6 year old.

A labeler demonstrates the desired output behavior.

This data is used to fine-tune GPT with supervised learning.

Step 2

Collect comparison data and train a reward.

A



$D > C > A = B$

This data is used to train our reward model.



$D > C > A = B$

Step 3

Optimize a policy against the reward model using reinforcement learning.

A prompt is sampled from the dataset.



Write a story about frogs



PPO



Once upon a time...



RM



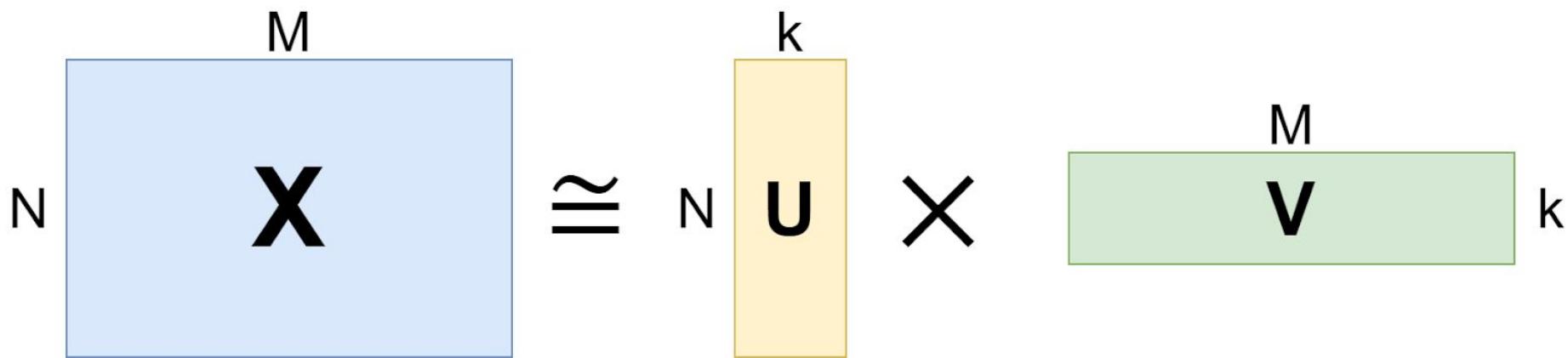
$r_k$

Это дало InstructGPT, а из него получился ChatGPT, и всё заверте...

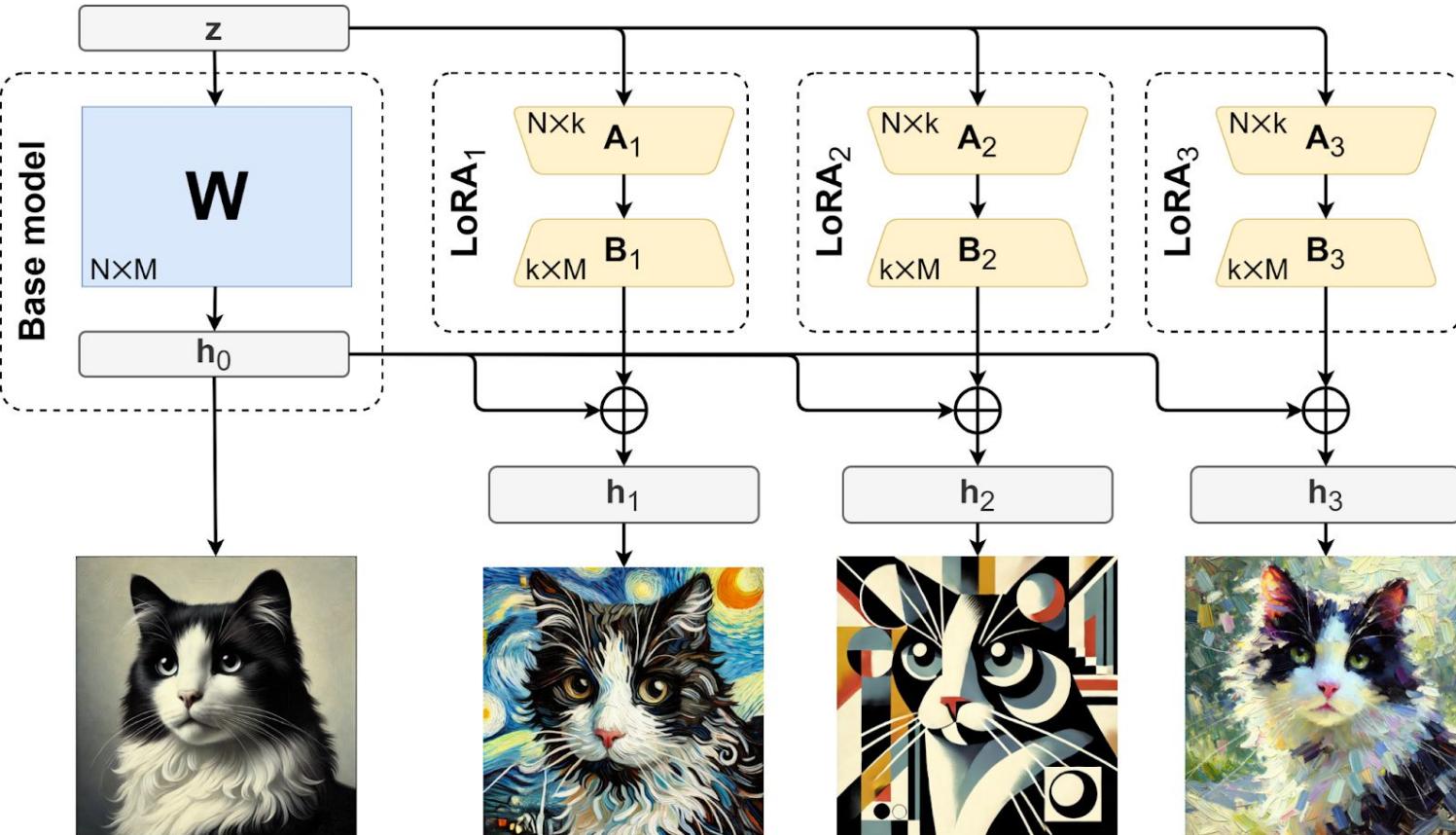
# LoRA и другие адаптеры

- Другой подход к fine-tuning – добавить маленький кусочек (адаптер) к большой модели и обучить только его; это может быть параллельная маленькая нейросеть с residual connection, но часто такие адаптеры основаны на низкоранговых приближениях (LoRA – Low-Rank Adaptation)

$$X \approx UV, \quad \text{where } U \in \mathbb{R}^{N \times k}, \quad V \in \mathbb{R}^{k \times M} \text{ for } k \ll N, M.$$



# LoRA и другие адаптеры



- Может быть много адаптеров для одной модели, хранить надо только их
- Но сейчас не будем углубляться в LoRa...

# GPT-4

- Возможности GPT-4 сразу же были неплохо изучены, и они поражали
- “[Sparks of Artificial General Intelligence: Early experiments with GPT-4](#)”

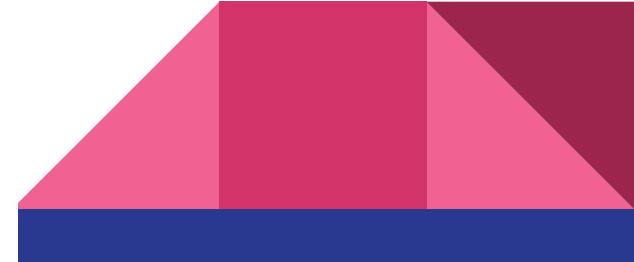
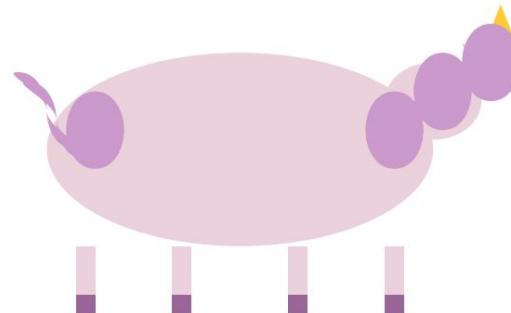
**Prompt:** Can you write a proof that there are infinitely many primes, with every line that rhymes?

**GPT-4:**

Yes, I think I can,  
Though it might take a clever plan.  
I'll start by noting Euclid's proof,  
Which shows that primes aren't just aloof.  
Assume we have a finite list,  
Of primes, and that none have been missed.  
Multiply them all together,  
And add one, just to be clever.  
The result will be a number,  
That has no prime factors, oh wonder!  
But every number has a prime,  
That divides it, it's just a matter of time.  
So we've found a contradiction,  
And our finite list needs eviction.  
There must be infinitely many primes,  
And that's the end of my rhyming lines.s

**Prompt:** Draw a unicorn in TiKZ.

**GPT-4:** [Produces L<sup>A</sup>T<sub>E</sub>X compiling to following picture.]



# GPT-4

- Возможности GPT-4 сра...
- “Sparks of Art”

**Prompt:** Can you write a poem that proves there are infinitely many primes, with every line containing a prime?

**GPT-4:**

Yes, I think I can,  
Though it might take a clever plan.  
I'll start by noting Euclid's proof  
Which shows that primes aren't just one.  
Assume we have a finite list,  
Of primes, and that none have been missed.  
Multiply them all together,  
And add one, just to be clever.  
The result will be a number,  
That has no prime factors, oh wonder!  
But every number has a prime,  
That divides it, it's just a matter of time.  
So we've found a contradiction,  
And our finite list needs eviction.  
There must be infinitely many primes,  
And that's the end of my rhyming lines.s

27 февраля 2025

OpenAI

GPT-4

# GPT-4

- Возможности GPT-4 сразу же поражали
- “Sparks of Artistic Creativity”

**Prompt:** Can you write a poem about the number of primes, with every line having exactly one prime?

**GPT-4:**

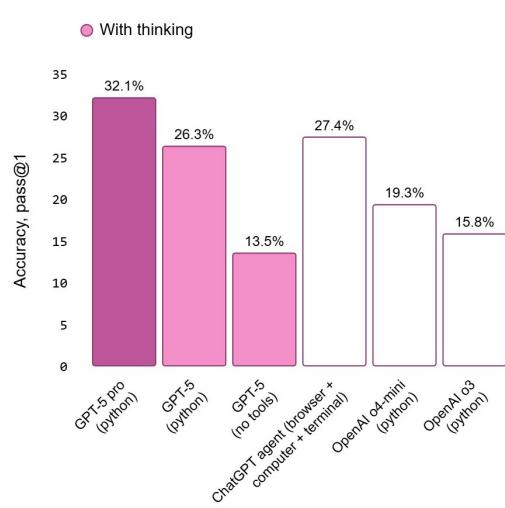
Yes, I think I can,  
Though it might take a clever plan.  
I'll start by noting Euclid's proof  
Which shows that primes aren't just one.  
Assume we have a finite list,  
Of primes, and that none have been missed.  
Multiply them all together,  
And add one, just to be clever.  
The result will be a number,  
That has no prime factors, oh wonder!  
But every number has a prime,  
That divides it, it's just a matter of time.  
So we've found a contradiction,  
And our finite list needs eviction.  
There must be infinitely many primes,  
And that's the end of my rhyming lines.s



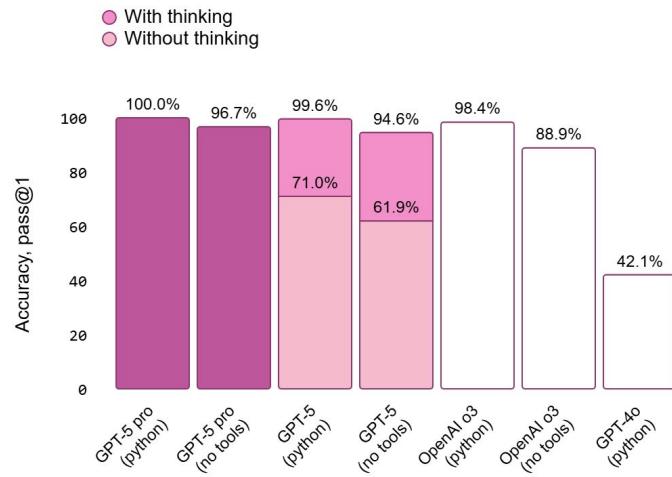
# GPT-5

- Семейство GPT-5 вышло два месяца назад ([OpenAI, Aug 7, 2025](#))
- Как водится, нам ничего конкретного не говорят, но это хорошие модели; мне кажется, что разочарование не слишком обосновано

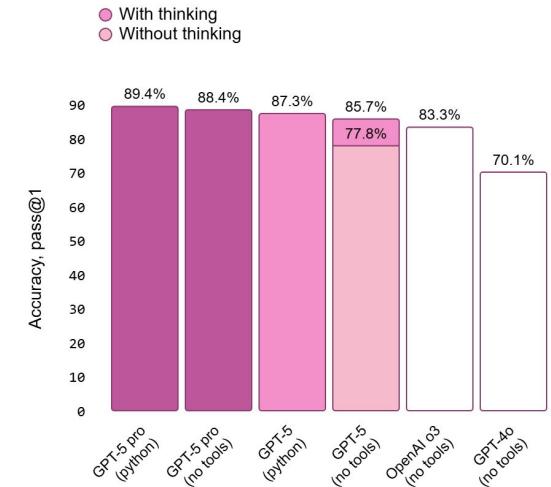
FrontierMath, Tier 1-3  
Expert-level math



AIME 2025  
Competition math

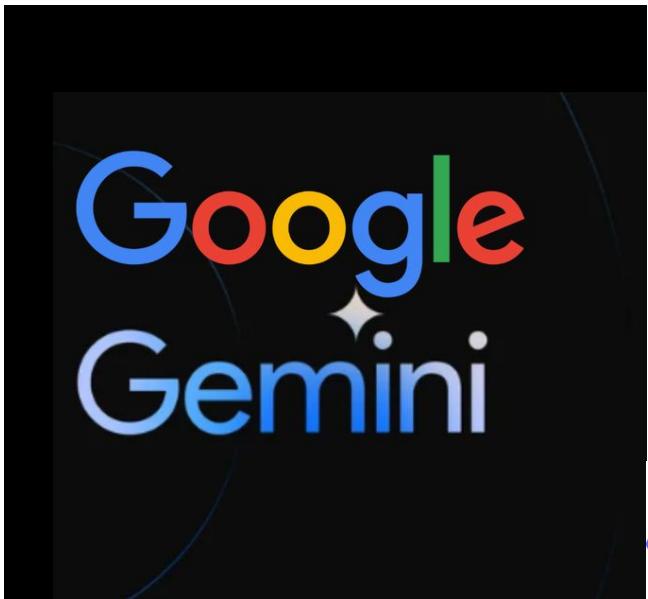
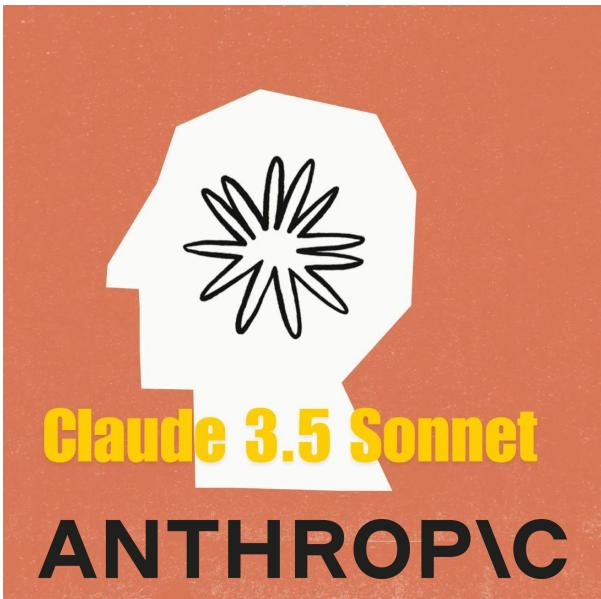


GPQA Diamond  
PhD-level science questions



# Не только OpenAI

- Есть и другие хорошие LLM



# Не только OpenAI

- И они очень быстро стали “рассуждающими”
- Что это значит? И как именно стали?..



### 3. Chain of thought

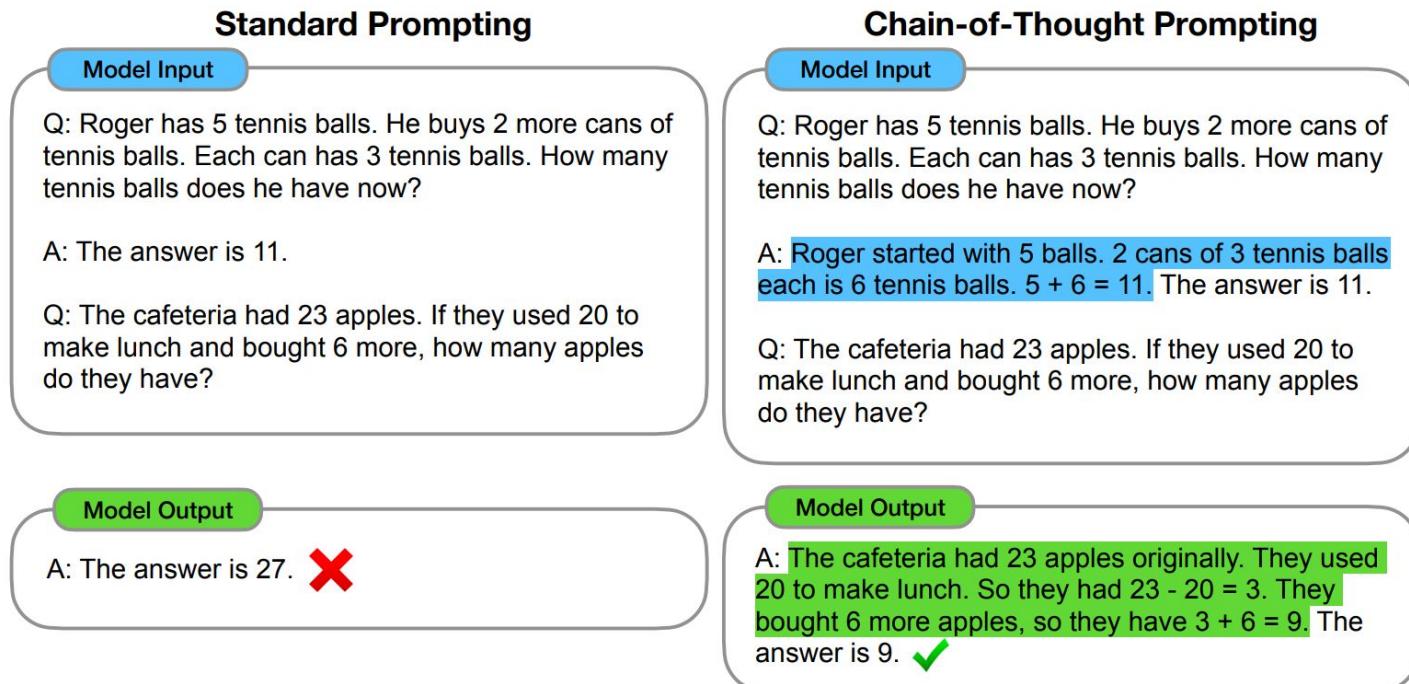


Вокруг цепи правил всегда должна виться  
серебряная нить фантазии.

*Роберт Шуман*

# Chain of Thought

- Wei et al., 2022: ещё на этапе few-shot prompting люди заметили, что если давать модели более подробные примеры, то качество ответов в итоге вырастет



# Let's think step by step

- Kojima et al. (2022): помогает даже просто добавить “Let’s think step by step”!

(a) Few-shot

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: The answer is 11.

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A:

(Output) The answer is 8. X

(b) Few-shot-CoT

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: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A:

(Output) The juggler can juggle 16 balls. Half of the balls are golf balls. So there are  $16 / 2 = 8$  golf balls. Half of the golf balls are blue. So there are  $8 / 2 = 4$  blue golf balls. The answer is 4. ✓

(c) Zero-shot

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: The answer (arabic numerals) is

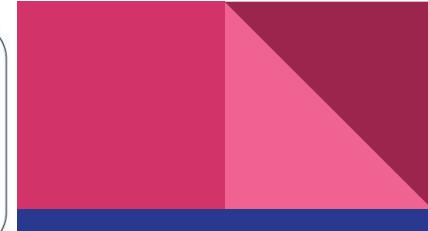
(Output) 8 X

(d) Zero-shot-CoT (Ours)

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

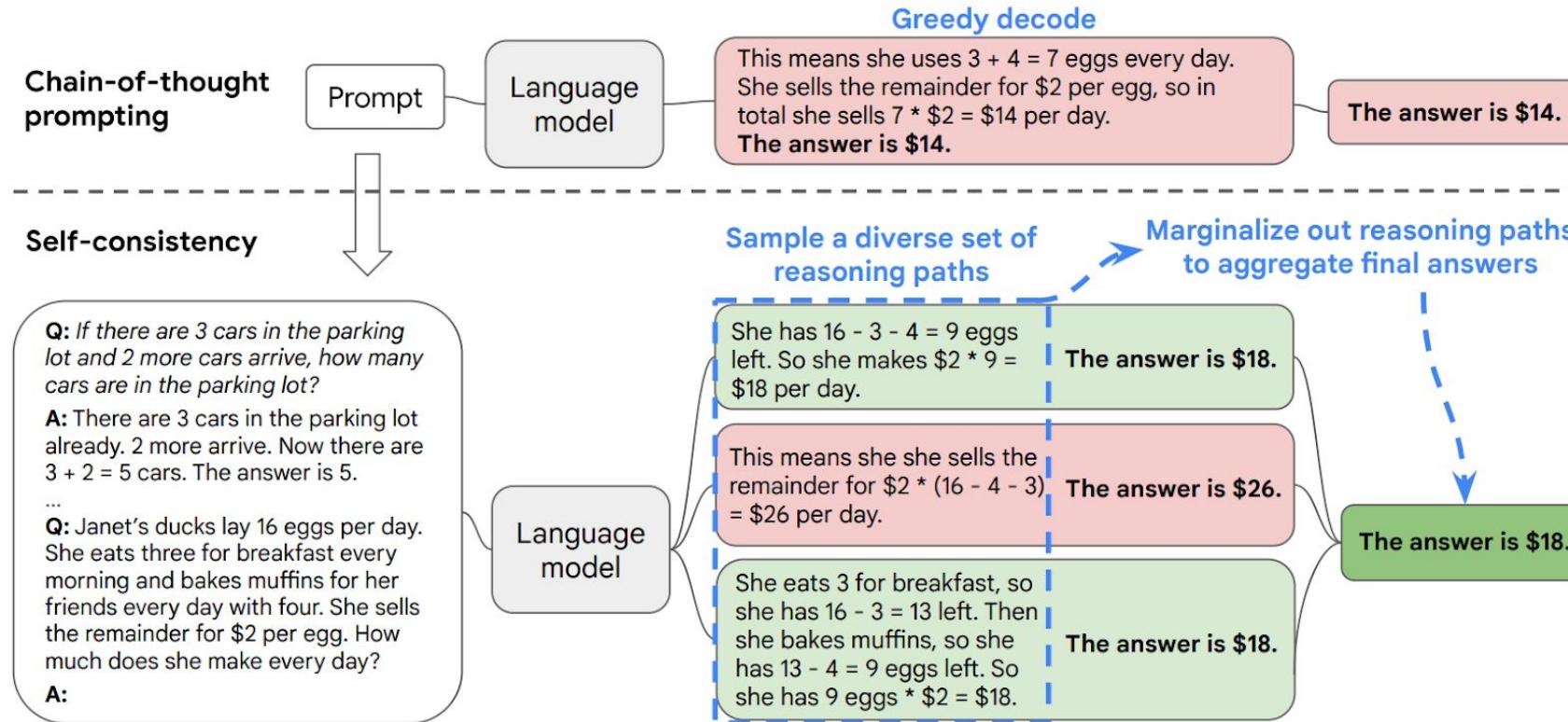
A: **Let's think step by step.**

(Output) There are 16 balls in total. Half of the balls are golf balls. That means that there are 8 golf balls. Half of the golf balls are blue. That means that there are 4 blue golf balls. ✓



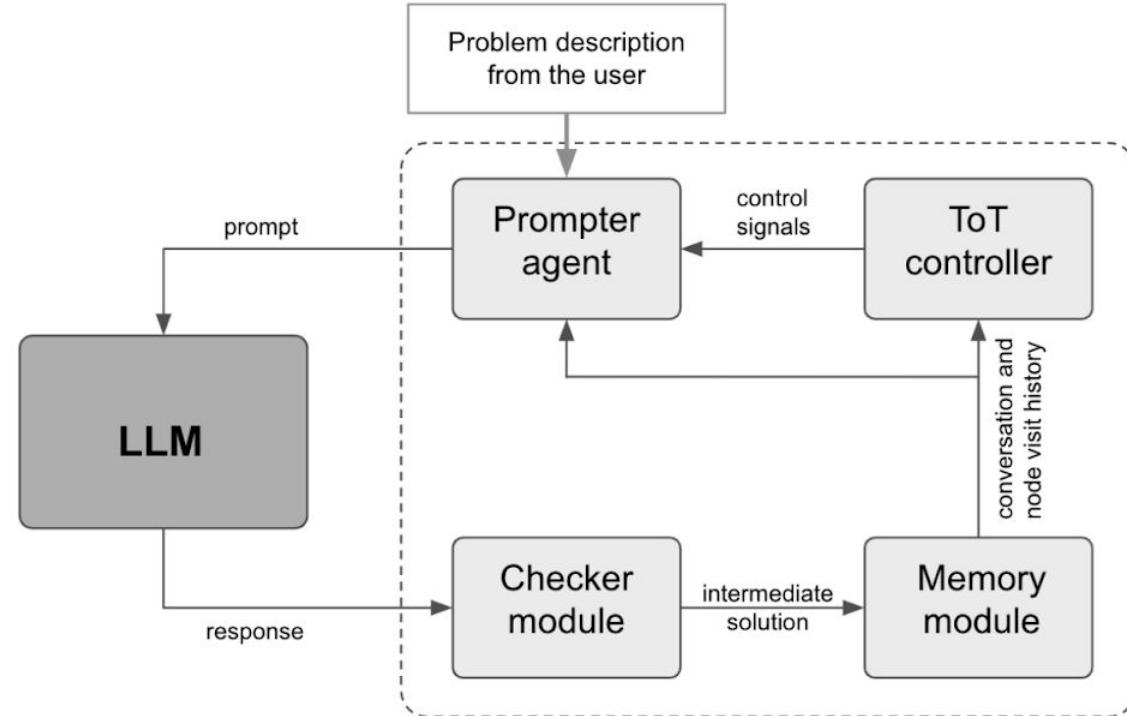
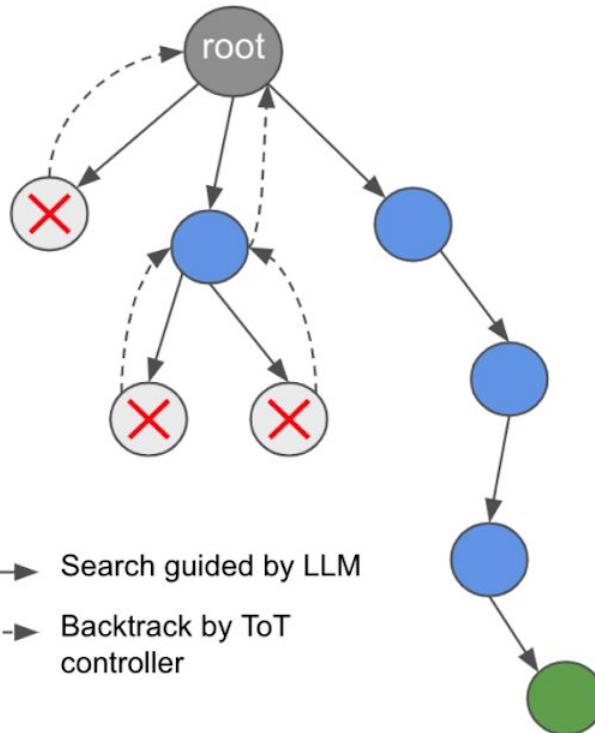
# От Chain-of-Thought к Graph-of-Thought

- Wang et al. (2022): Chain-of-Thought with Self-Consistency (CoT-SC)



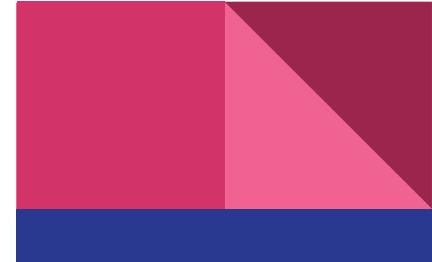
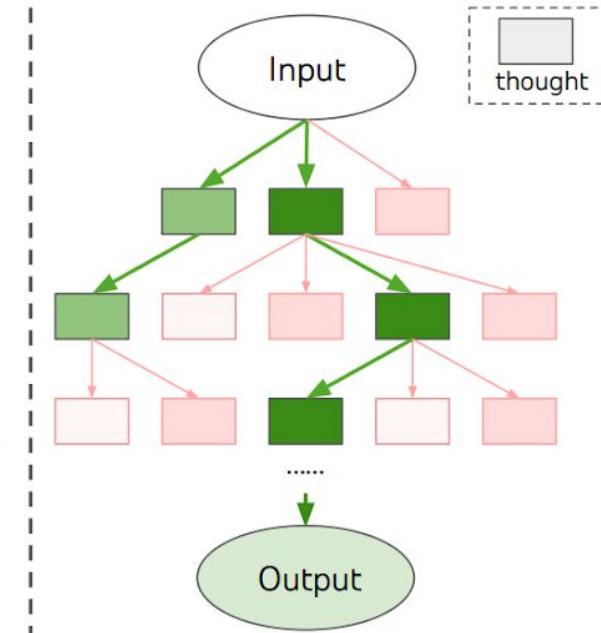
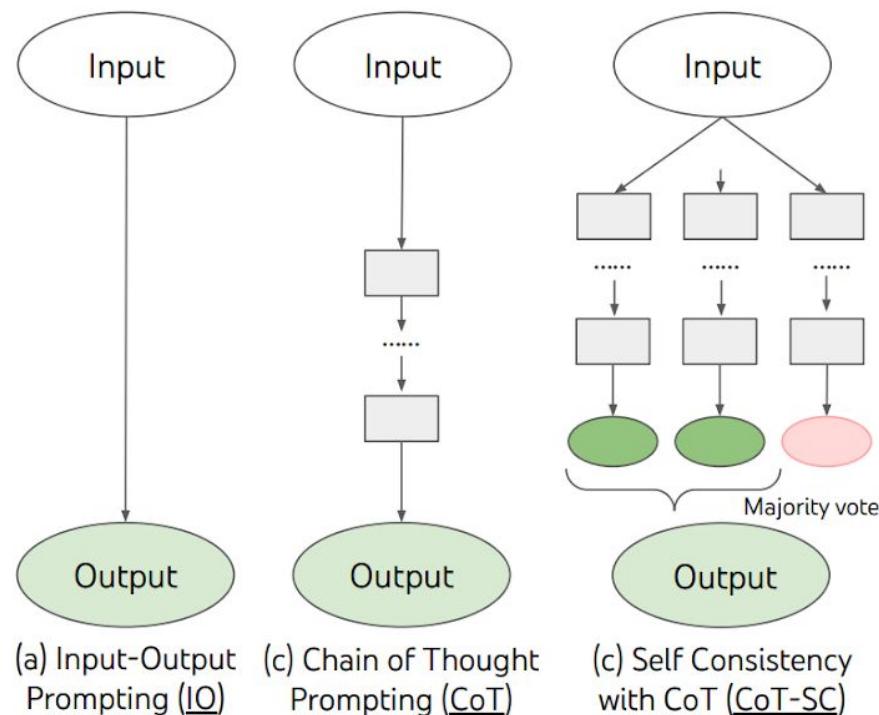
# От Chain-of-Thought к Graph-of-Thought

- Long (May 15, 2023): Tree-of-Thought через контроллер с откатом по дереву



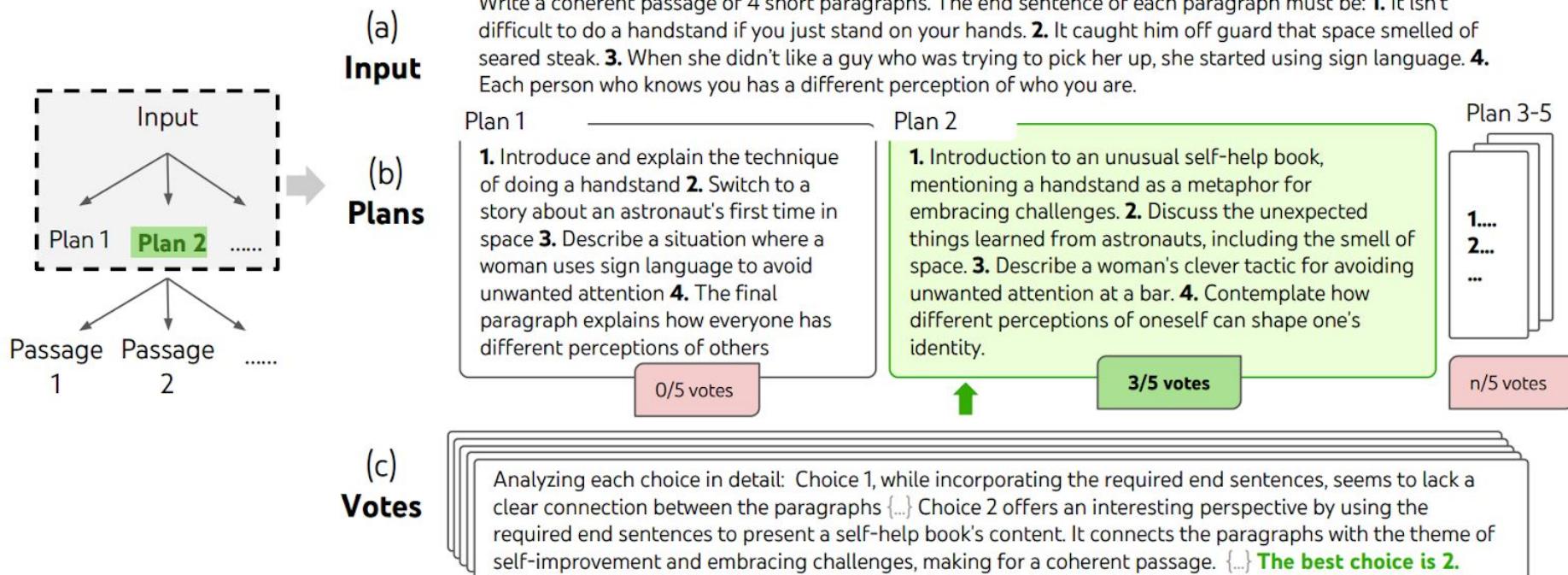
# От Chain-of-Thought к Graph-of-Thought

- Yao et al. (May 17, 2023): Tree-of-Thought с голосованием или другим механизмом



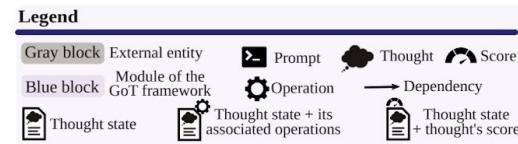
# От Chain-of-Thought к Graph-of-Thought

- Yao et al. (May 17, 2023): Tree-of-Thought с голосованием или другим механизмом



# От Chain-of-Thought к Graph-of-Thought

- Besta et al. (2023): логичный следующий шаг – граф операций с возможностью агрегации



#### API for Controller

- //LLM params: model used, temperature, max tokens, api key, org, ...
- //LLM cost features: prompt token cost, response token cost, ...
- //Instances of Prompter + Parser + Graph of Operations,
- //Any additional input parameters (e.g., numbers to be sorted).

#### Available operations when building the GoO (extensible)

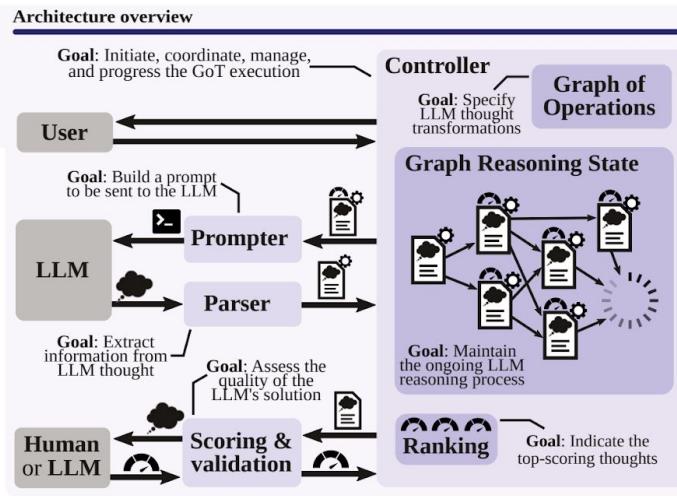
- Generate, Aggregate, Score, ... //see Prompter API
- KeepBest(N) //preserves N best scoring thoughts
- Repeat(k) //Repeat a given operation k times, generating k thoughts. //For example, this enables "Aggregate" to generate multiple outcomes //of the combination operation. Each such thought is maintained //within the Graph Reasoning State and scored individually.

#### API for Prompter (extensible)

- Generate(t, k) //generate a prompt for k new thoughts, using thought t
- ValidateAndImprove(t) //generate a prompt to enhance thought t,
- Aggregate(t1, ..., tk) //generate a prompt to combine thoughts t1, ..., tk
- Score(t) //score thought t
- Validate(t) //generate a prompt to validate the correctness of thought t

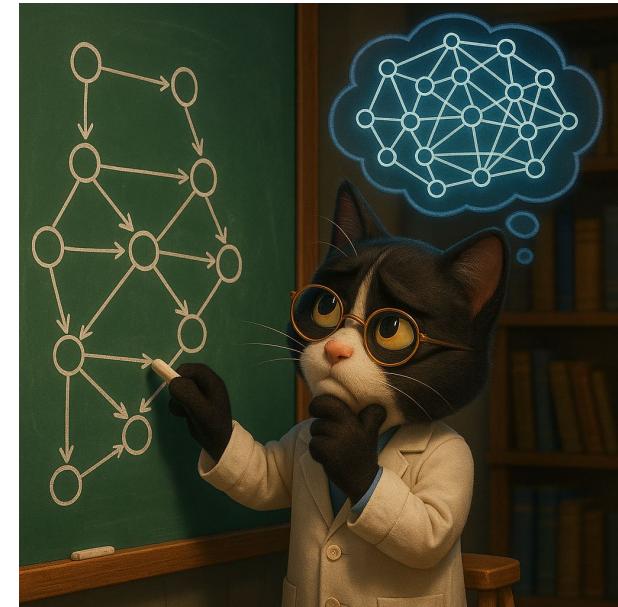
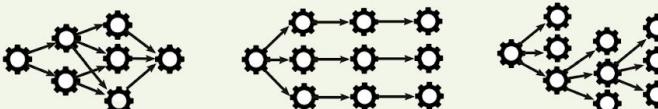
#### API for Parser (extensible)

ParseGenerate, ParseImprove, ParseScore,  
 ParseAggregate, ParseValidate, ...  
 //Each of the above routines is responsible for parsing an LLM thought  
 //to a corresponding Prompter routine (e.g., ParseScore parses Score).



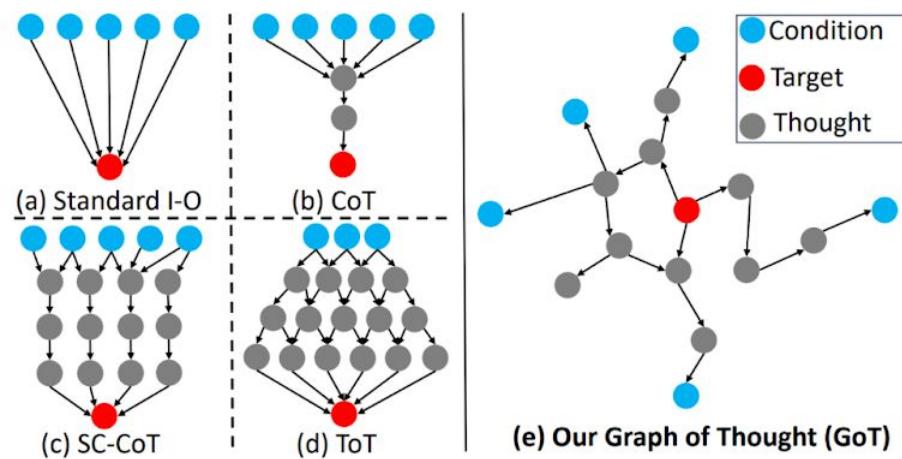
#### Specifying the structure of the Graph of Operations (GoO)

Graph of Operations enables seamless specification of not only GoT, but also existing schemes such as CoT, CoT-SC, ToT

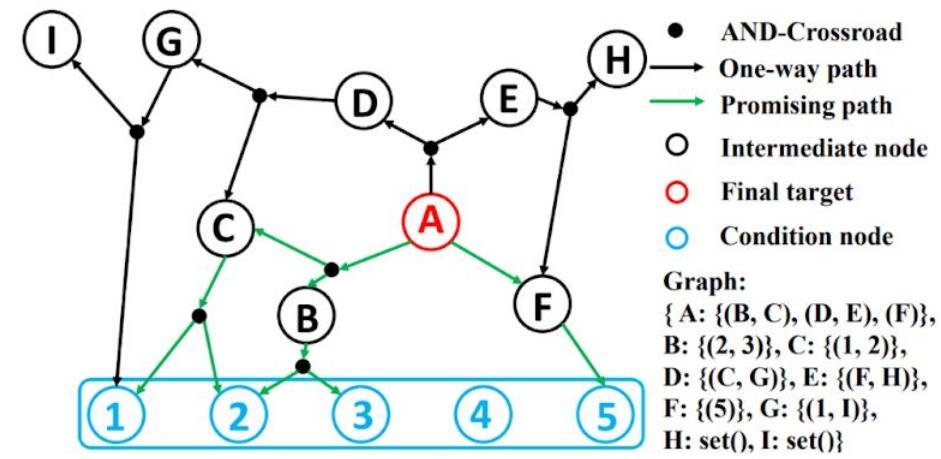


# От Chain-of-Thought к Graph-of-Thought

- Lei et al. (2023): добавляют узлы, которые обобщают то, чему модель научилась из предыдущих узлов, и добавляет результат как условия для дальнейших рассуждений



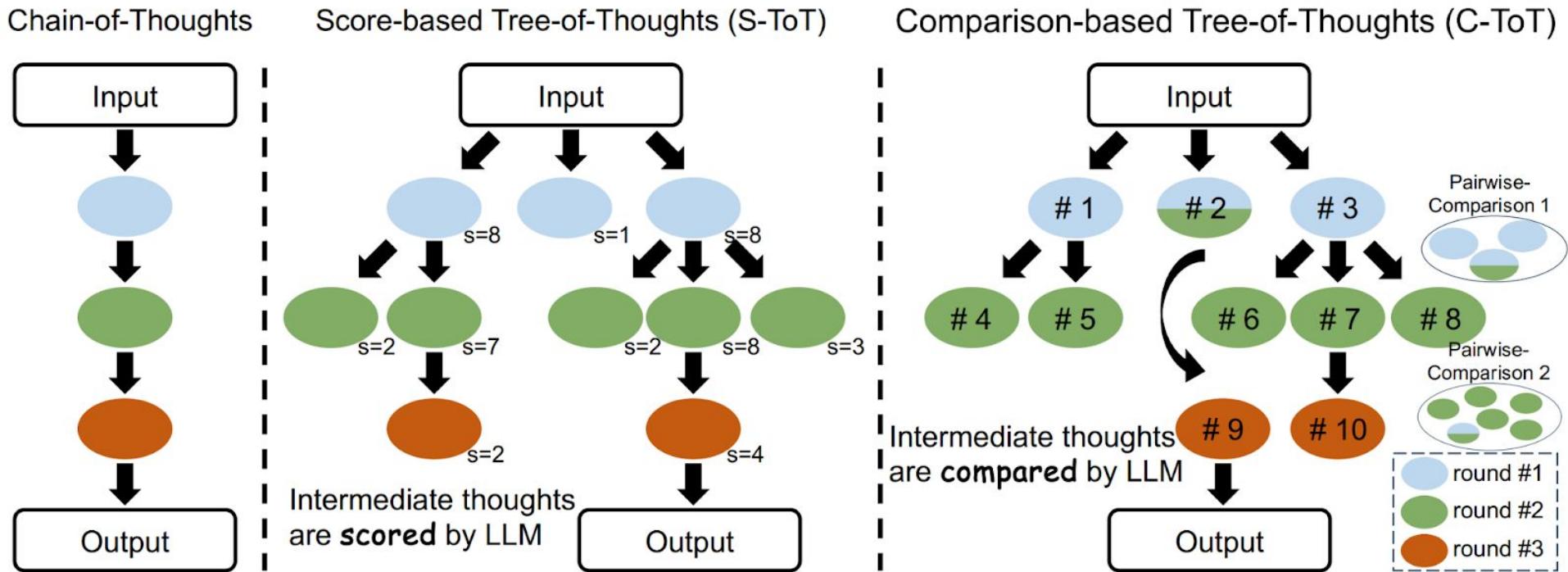
(a) Comparison of approaches by Lei et al. (2023)



(b) Sample toy graph by Lei et al. (2023)

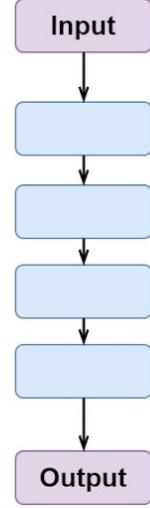
# От Chain-of-Thought к Graph-of-Thought

- Zhang et al. (2024): попарные сравнения вместо отдельных контроллеров



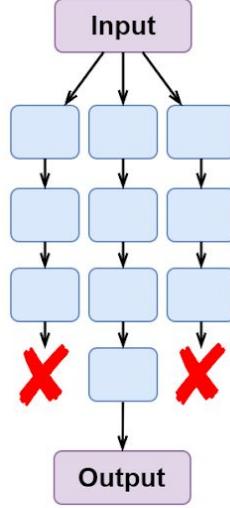
## Chain-of-Thought (CoT)

Wei et al., Jan 2022  
Kojima et al., May 2022



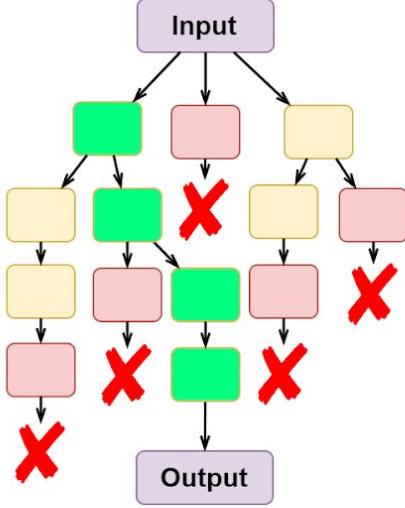
## Multiple Chains-of-Thought

Wang et al.,  
March 2022



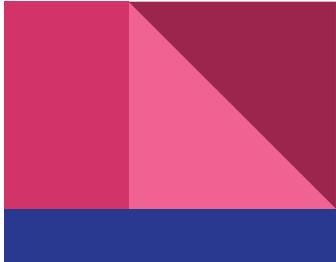
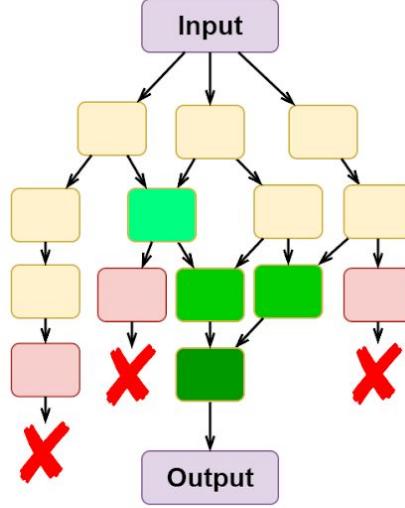
## Tree-of-Thought (ToT)

Long, May 2023  
Yao et al., May 2023



## Graph-of-Thought (GoT)

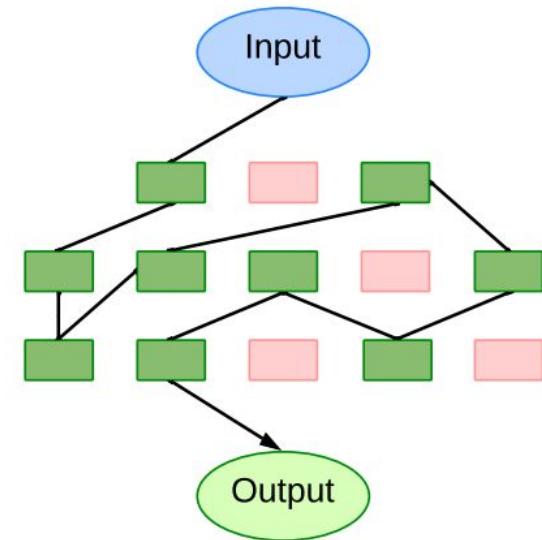
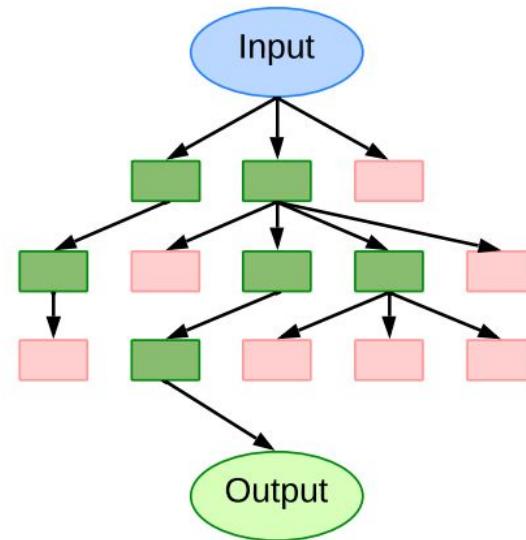
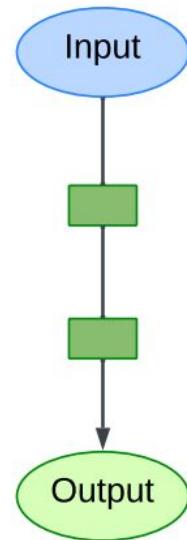
Besta et al., August 2023  
Lei et al., August 2023



- ИТОГО полу- чается так

# Algorithm-of-Thought

- Sel et al. (2023): Algorithm-of-Thought; каждый узел дерева – шаг в алгоритме, и модель порождает следующий шаг по результатам предыдущего



Standard Prompting

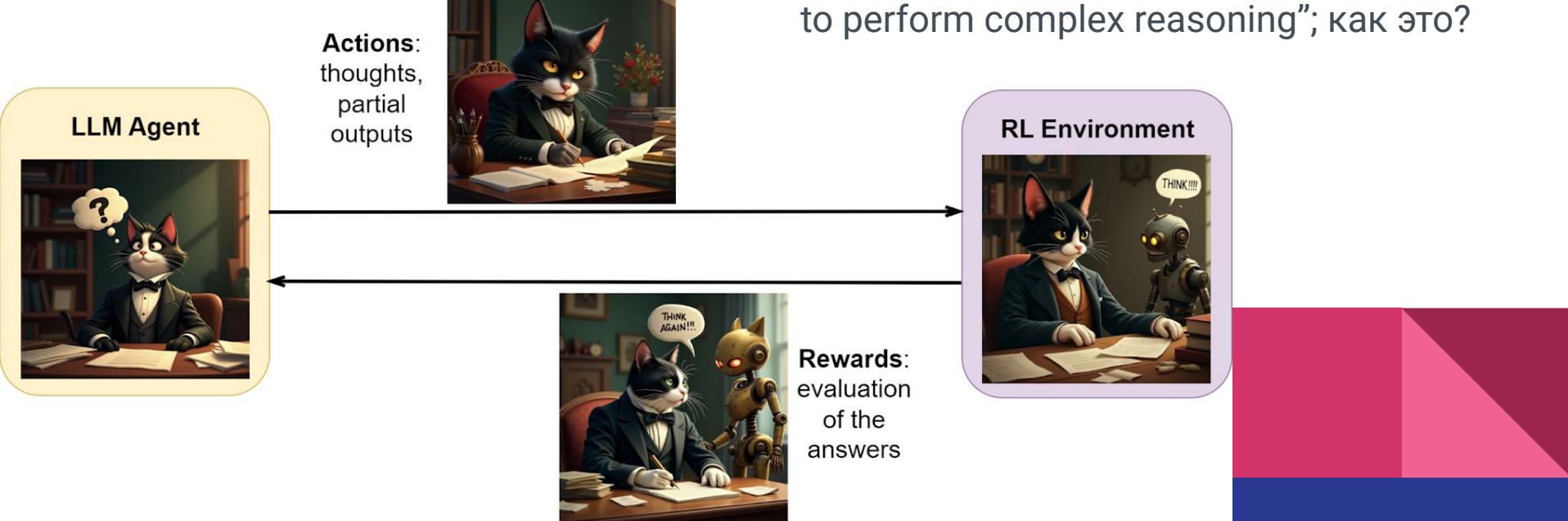
Chain of Thoughts

Tree of Thoughts

Algorithm of Thoughts

# Семейство моделей o1

- Но всё изменилось, когда появилось семейство моделей o1 от OpenAI
- Детали не разглашались, **system card** говорила так: “The o1 large language model family is trained with reinforcement learning to perform complex reasoning”; как это?



# 4. Рассуждающие модели

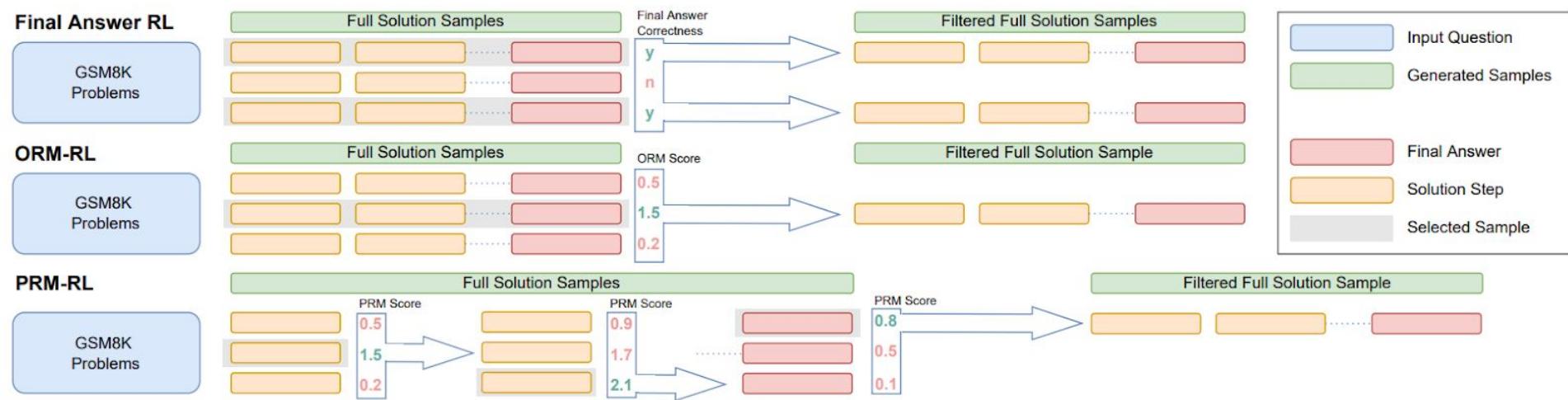


Напрасны все предсказания, все рассуждения о том, что с вами может быть, когда вы под властию вашей страсти: увлечение сердца перемогает все красноречивые речи. Нам нужна другая воля, и она тогда только придёт, когда наше коварство нас порядком накажет... и мы почувствуем в том для себя ущерб, что мы... оставили Бога. Тогда только страх Божий к нам возвратится.

*Авдотья Глинка. Подневные рассуждения*

# Process Reward Models

- Uesato et al. (2022): обычно в RL можно награждать только конечный результат; но в chain-of-thought мышлении LLM выводит мысли, которые в принципе можно оценивать самостоятельно
- Process reward model (PRM) обучается оценивать промежуточные шаги



# Process Reward Models

- Lightman et al.  
(OpenAI, 2024):

Let's Verify  
Step by Step –  
давайте  
обучим PRM  
выяснить, где  
именно в  
рассуждениях  
ошибка

The denominator of a fraction is 7 less than 3 times the numerator. If the fraction is equivalent to  $2/5$ , what is the numerator of the fraction? (Answer:

   Let's call the numerator x.

   So the denominator is  $3x-7$ .

   We know that  $x/(3x-7) = 2/5$ .

   So  $5x = 2(3x-7)$ .

    $5x = 6x - 14$ .

   So  $x = 7$ .

# Process Reward Models

- Lightman et al.  
(OpenAI, 2024):

Let's Verify  
Step by Step –  
давайте  
обучим PRM  
выяснять, где  
именно в  
рассуждениях  
ошибка

Let

$$x^8 + 3x^4 - 4 = p_1(x)p_2(x) \cdots p_k(x),$$

where each non-constant polynomial  $p_i(x)$  is monic with integer coefficients, and cannot be factored further over the integers. Compute  $p_1(1) + p_2(1) + \cdots + p_k(1)$ .

I notice that the given polynomial has even degree and only even powers of  $x$ , so I can try to make a substitution to simplify it.

Let  $y = x^4$ , then the polynomial becomes  $y^2 + 3y - 4$ , which is a quadratic equation.

I can factor this quadratic equation as  $(y + 4)(y - 1)$ , so the original polynomial is  $(x^4 + 4)(x^4 - 1)$ .

Now I need to factor each of these factors further, if possible, over the integers.

For the first factor,  $x^4 + 4$ , I recall the Sophie Germain identity, which says that  $a^4 + 4b^4 = (a^2 + 2b^2 + 2ab)(a^2 + 2b^2 - 2ab)$ .

Applying this identity with  $a = x$  and  $b = 1$ , I get  $x^4 + 4 = (x^2 + 2x + 2)(x^2 - 2x + 2)$ .

For the second factor,  $x^4 - 1$ , I can use the difference of squares formula, which says that  $a^2 - b^2 = (a + b)(a - b)$ .

Applying this formula with  $a = x^2$  and  $b = 1$ , I get  $x^4 - 1 = (x^2 + 1)(x^2 - 1)$ .

Then I can apply the difference of squares formula again with  $a = x$  and  $b = 1$ , to get  $x^2 - 1 = (x + 1)(x - 1)$ .

So the original polynomial is  $(x^4 + 4)(x^4 - 1) = (x^2 + 2x + 2)(x^2 - 2x + 2)(x^2 + 1)(x + 1)(x - 1)$ .

These are all the irreducible factors over the integers, so  $p_1(x) = x^2 + 2x + 2$ ,  $p_2(x) = x^2 - 2x + 2$ ,  $p_3(x) = x^2 + 1$ ,  $p_4(x) = x + 1$ , and  $p_5(x) = x - 1$ .

To find the sum of their values at  $x = 1$ , I just need to plug in  $x = 1$  into each factor and add them up.

I get  $p_1(1) = 1^2 + 2 \cdot 1 + 2 = 5$ ,  $p_2(1) = 1^2 - 2 \cdot 1 + 2 = 1$ ,  $p_3(1) = 1^2 + 1 = 2$ ,  $p_4(1) = 1 + 1 = 2$ , and  $p_5(1) = 1 - 1 = 0$ .

So the final answer is  $5 + 1 + 2 + 2 + 0 = 10$ .

Answer: 10

Let

$$x^8 + 3x^4 - 4 = p_1(x)p_2(x) \cdots p_k(x),$$

where each non-constant polynomial  $p_i(x)$  is monic with integer coefficients, and cannot be factored further over the integers. Compute  $p_1(1) + p_2(1) + \cdots + p_k(1)$ .

This problem is about factoring a polynomial over the integers, which can be a tricky task. One possible strategy is to look for rational roots of the polynomial, using the rational root theorem.

However, that would involve checking a lot of possible factors of  $-4$ , and there is no guarantee that the polynomial has any rational roots at all.

Another possible strategy is to use some algebraic tricks to simplify the polynomial or reduce its degree.

I notice that the polynomial has only even powers of  $x$ , so I can try to substitute  $y = x^2$  and see what happens.

Then the polynomial becomes  $y^4 + 3y^2 - 4$ , which is a quadratic equation in  $y^2$ .

I can try to factor this quadratic equation by looking for two numbers that multiply to  $-4$  and add to  $3$ .

Those numbers are  $4$  and  $-1$ , so I can write  $y^4 + 3y^2 - 4 = (y^2 + 4)(y^2 - 1)$ .

Now I can substitute back  $x^2 = y$  and get  $x^8 + 3x^4 - 4 = (x^4 + 4)(x^4 - 1)$ .

This looks promising, because I have reduced the degree of the factors from  $8$  to  $4$ .

I can try to factor each of these factors further, using the difference of squares formula.

I get  $x^8 + 3x^4 - 4 = (x^2 + 2)(x^2 - 2)(x^2 + 1)(x^2 - 1)$ .

I can apply the difference of squares formula again to the last factor and get  $x^8 + 3x^4 - 4 = (x^2 + 2)(x^2 - 2)(x^2 + 1)(x + 1)(x - 1)$ .

Now I have factored the polynomial completely into monic linear and quadratic factors with integer coefficients.

These are the  $p_i(x)$ 's that the problem is asking for.

To find the sum of their values at  $x = 1$ , I just need to plug in  $x = 1$  into each factor and add them up.

I get  $p_1(1) + p_2(1) + \cdots + p_k(1) = (1^2 + 2)(1^2 - 2)(1^2 + 1)(1 + 1)(1 - 1)$ .

Simplifying, I get  $p_1(1) + p_2(1) + \cdots + p_k(1) = (3)(-1)(2)(2)(0)$ .

Multiplying, I get  $p_1(1) + p_2(1) + \cdots + p_k(1) = 0$ .

Answer: 0

# Process Reward Models

- Xia et al. (2024):  
ReasonEval с тем же смыслом
- А что делает о1?  
Неизвестно, и люди сразу попытались реплицировать

Question: Simplify  $\sqrt{242}$

Solution

**Step 1:** We can simplify the square root of 242 by factoring it into its prime factors.

**Step 2:** The prime factorization of 242 is  $2 * 11 * 11$ .

**Step 3:** Let's check it. The prime factorization of 242 is  $2 * 11 * 11$ . ● ←

**Step 4:** By the property of square roots, we know that  $\sqrt{ab} = \sqrt{a} * \sqrt{b}$  for any non-negative numbers  $a$  and  $b$ . Applying this property to  $\sqrt{2 * 11 * 11}$ , we have  $\sqrt{2} * \sqrt{11^2}$ .

**Step 5:** By the property of square roots, we know that  $\sqrt{a^2} = a$  for any non-negative number  $a$ . Applying this property to  $\sqrt{11^2}$ , we have 11.

**Step 6:** By combining the results from step 4 and step 5, we have  $\sqrt{2} * 11 * 11 = 11 * \sqrt{2}$ . ● ←

**Step 7:** Therefore, the simplified form of  $\sqrt{242}$  is 11 $\sqrt{2}$ .

● redundant   ● invalid

Answer: 11 $\sqrt{2}$  ✓

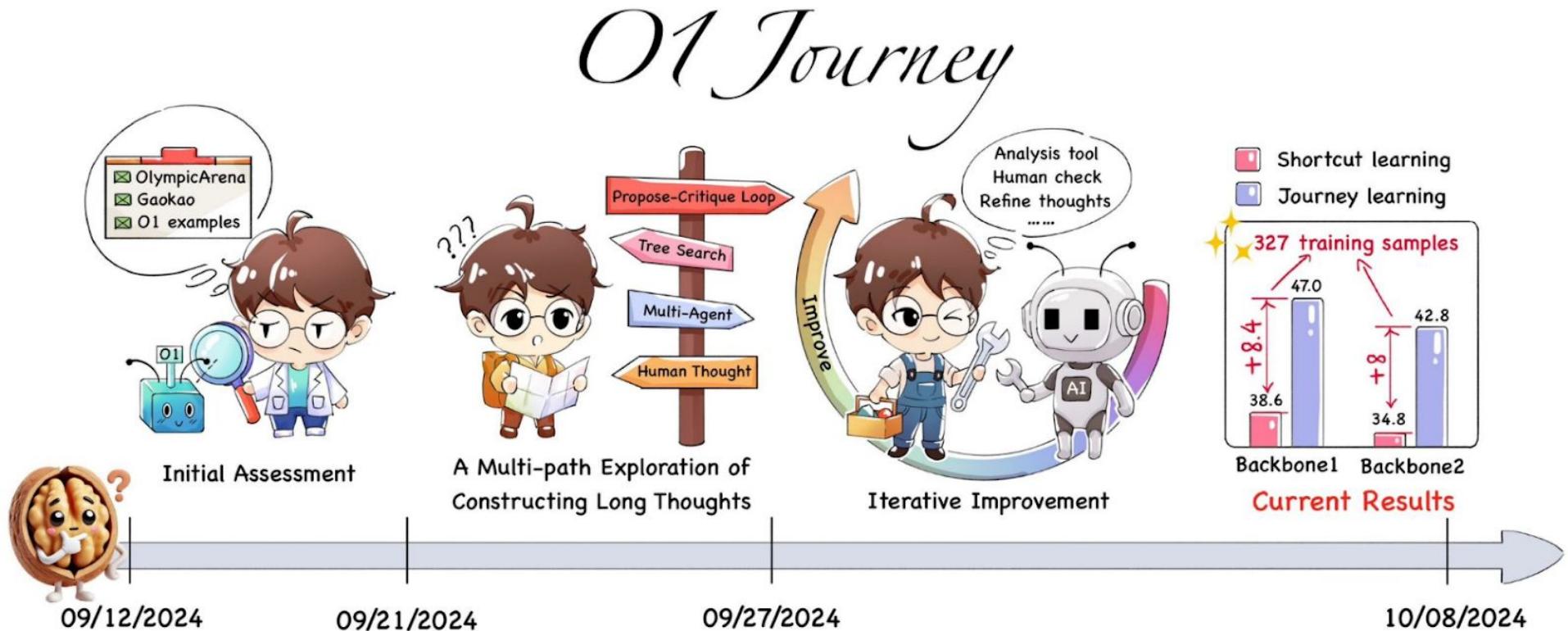
Ground Truth:  $11\sqrt{2}$

ReasonEval

$S_{validity}$	$S_{redundancy}$
0.85	0.05
0.85	0.05
0.75	0.25
0.95	0.05
0.80	0.05
0.15	0.05
0.75	0.05

# Пример: 01 Journey

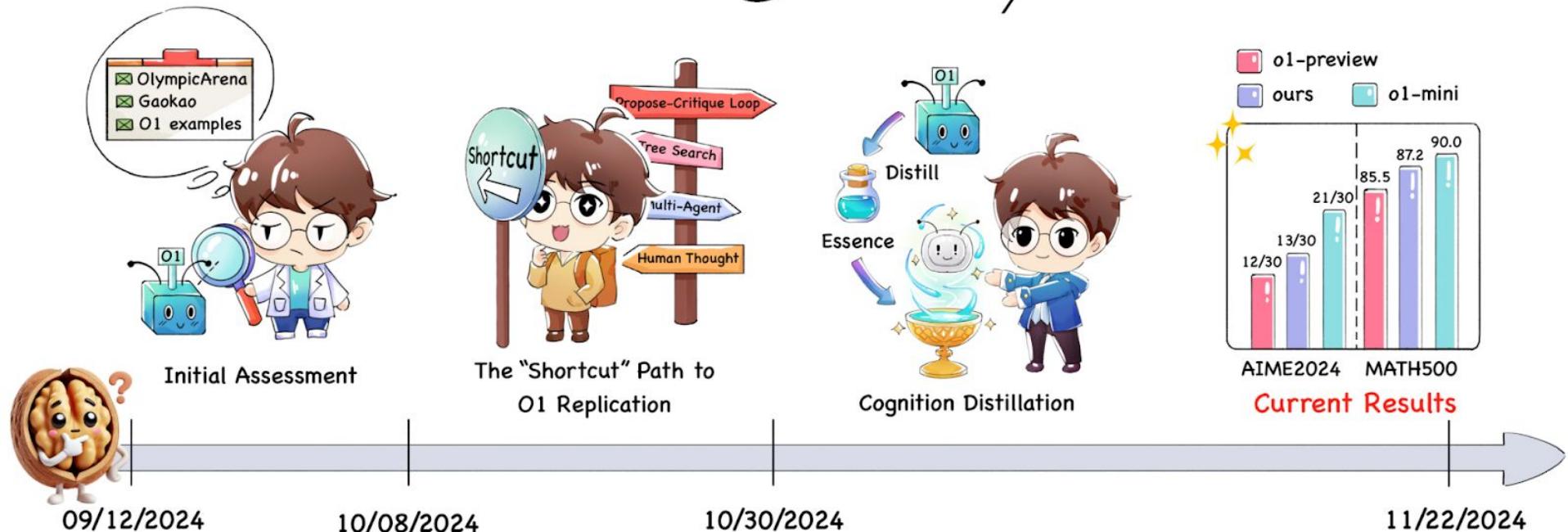
- Пример – Generative AI Research Lab (GAIR) из Шанхая; [Qin et al. \(Oct 2024\)](#):



# Пример: O1 Journey

- Пример – Generative AI Research Lab (GAIR) из Шанхая; [Huang et al. \(Nov 2024\)](#):

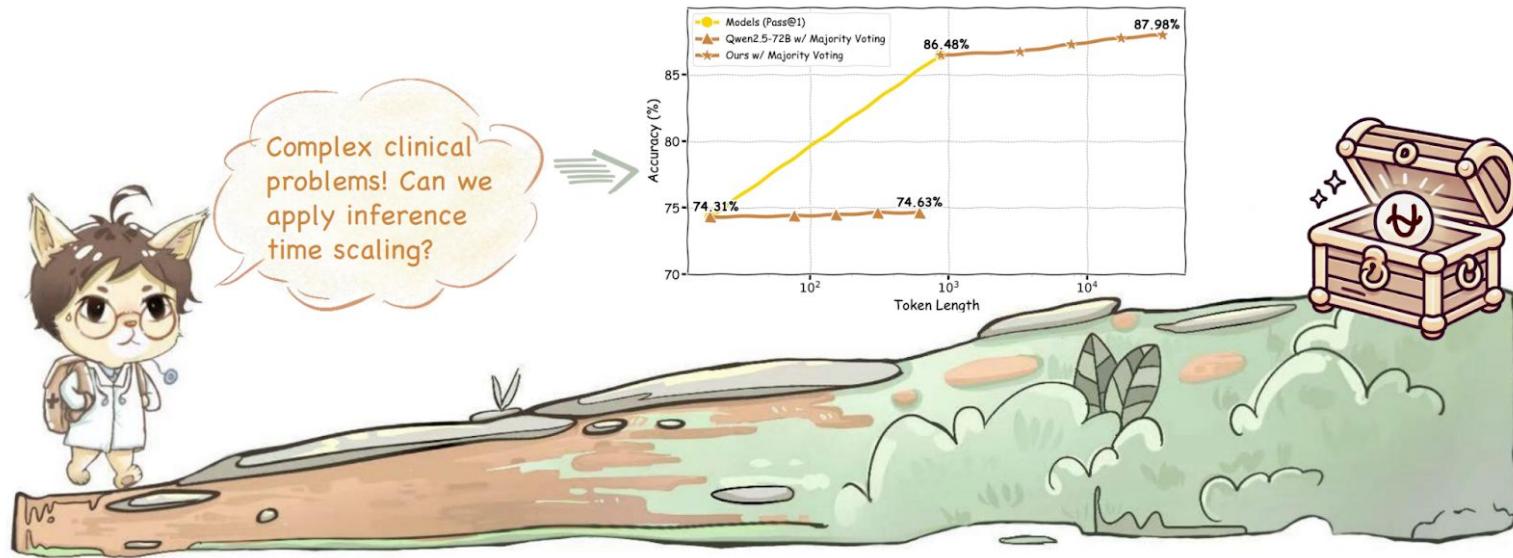
## O1 Journey



# Пример: 01 Journey

- Пример – Generative AI Research Lab (GAIR) из Шанхая; [Huang et al. \(Jan 2025\)](#):

## 01 Journey



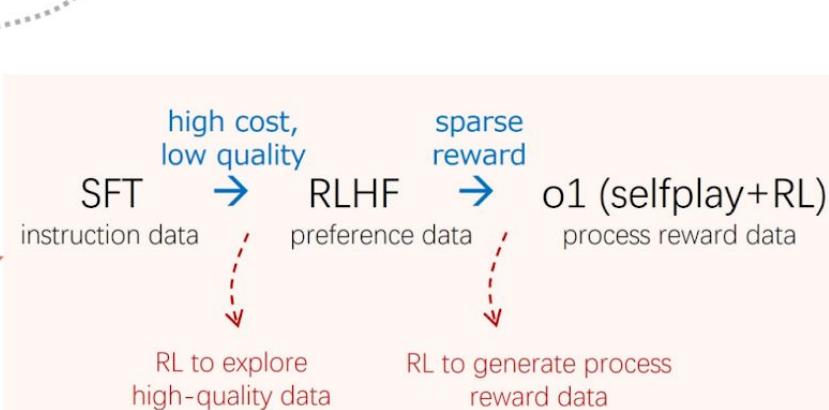
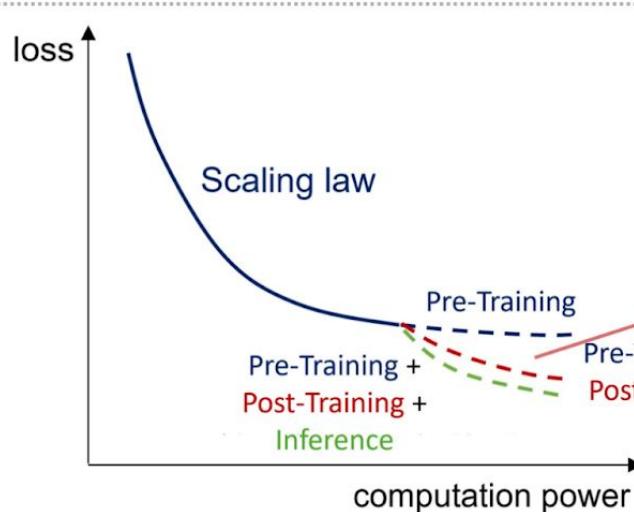
# Пример: O1 Journey

- Zhang et al. (2024): o1-Coder, добавили PRM и MCTS

SVM → DNN → Transformer →→ Pre-Training (SSL) → Post-Training (RL) →→ Inference

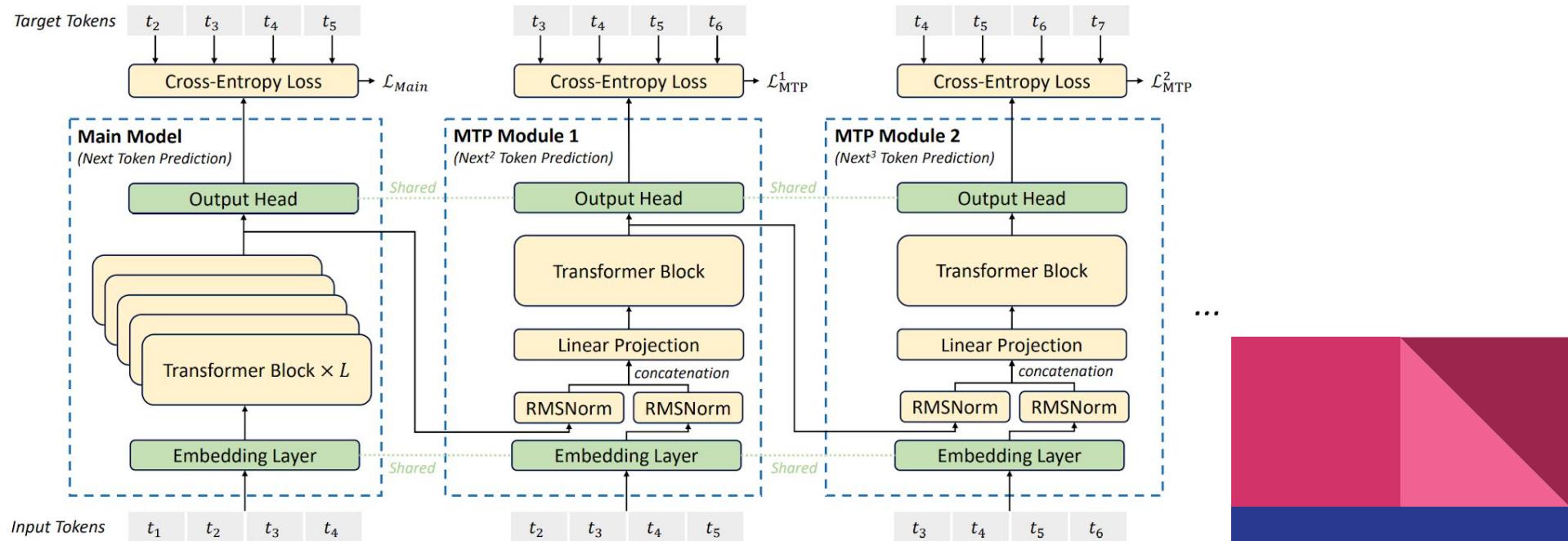
Scalable Model

Infinite Data



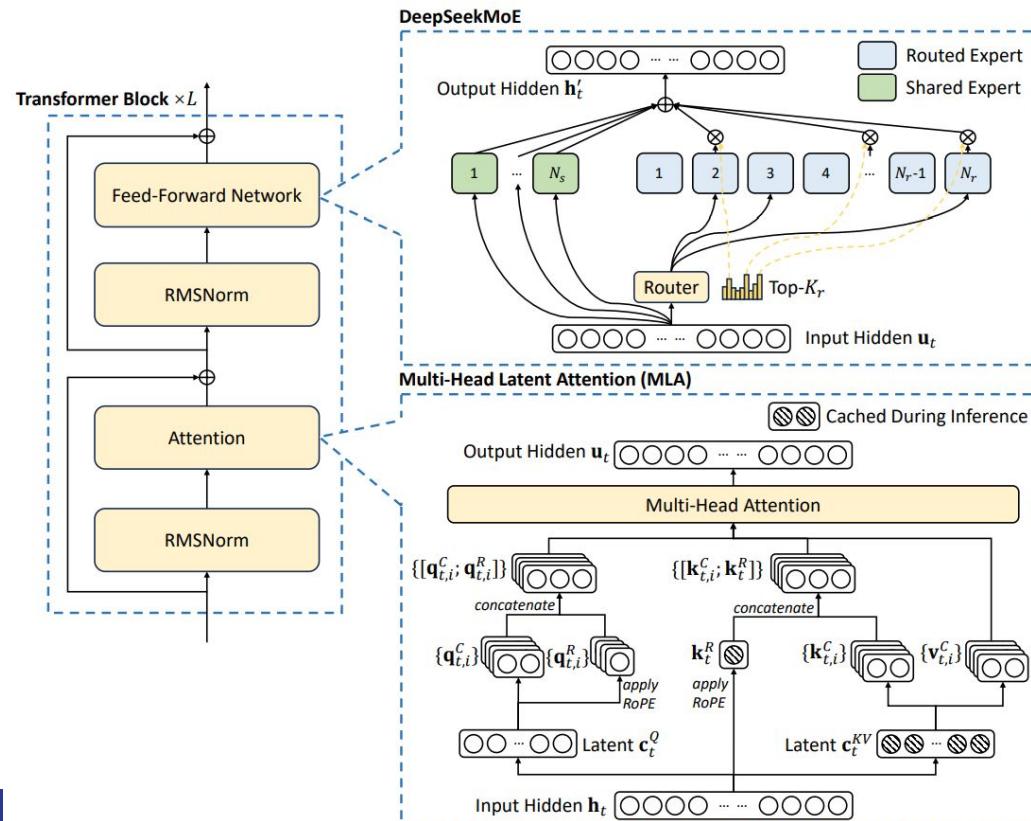
# DeepSeek-V3

- А что же делал наш КИТ в комнате, то есть DeepSeek?
- Сначала DeepSeek-V3; multi-token prediction:



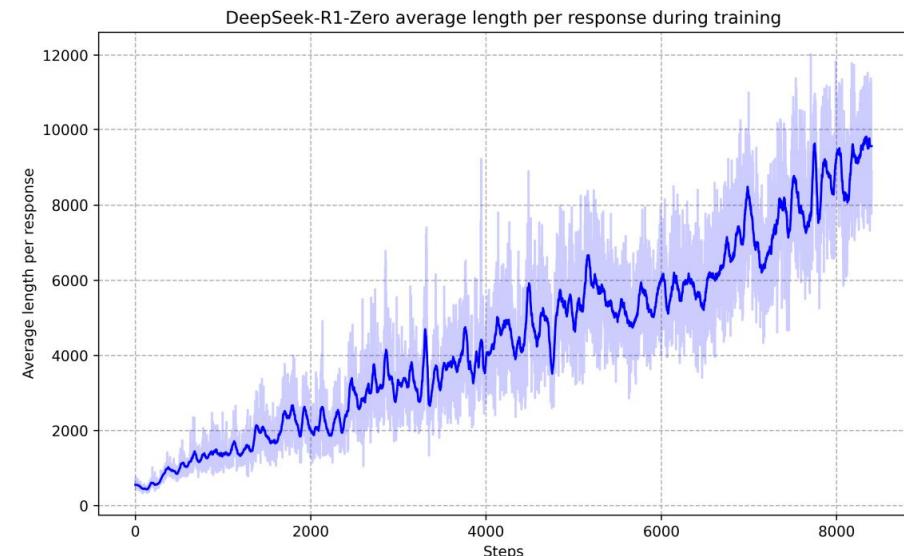
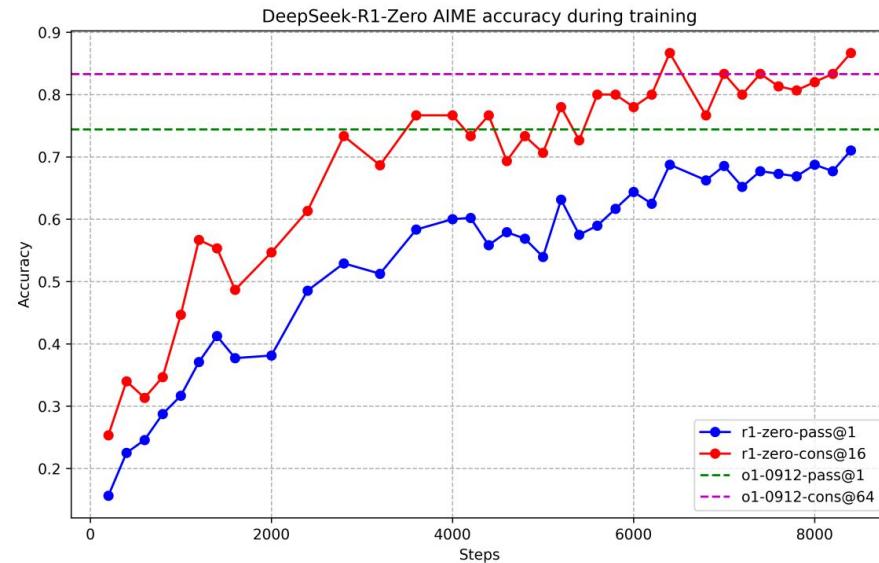
# DeepSeek-V3

- А что же делал наш КИТ в комнате, то есть DeepSeek?
- Сначала DeepSeek-V3
- Key-value caching (не будем углубляться), mixture-of-experts
- В общем, хорошая современная LLM



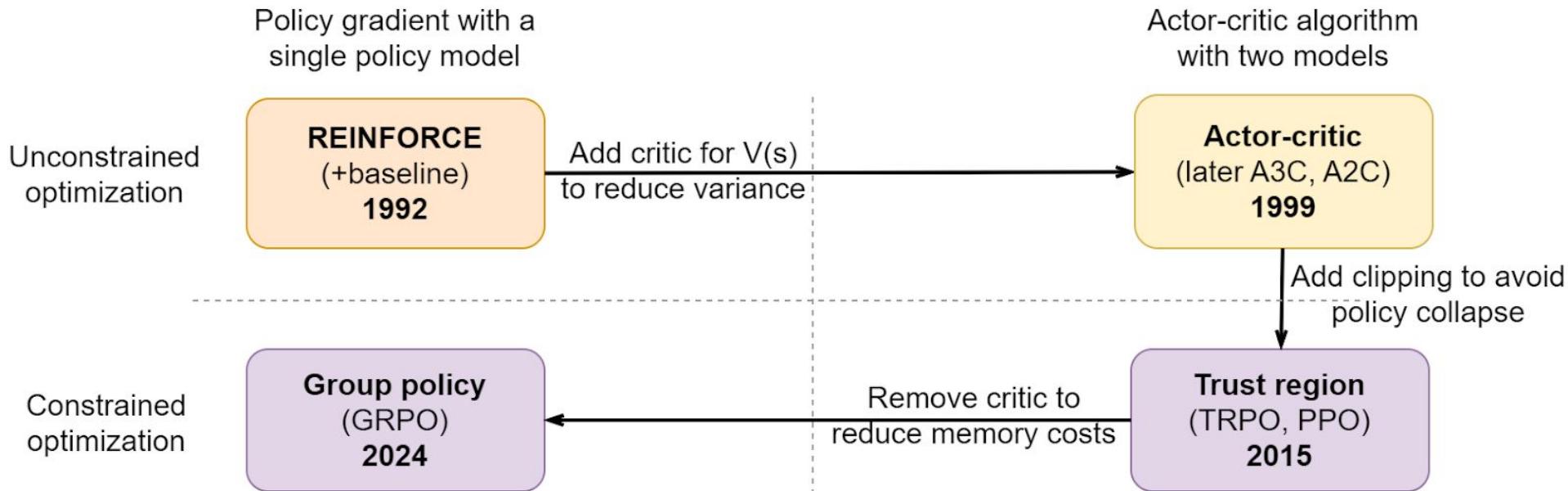
# DeepSeek-R1

- А потом и DeepSeek-R1
- R1-Zero: чистый RL на проверяемых задачах



# DeepSeek-R1

- А потом и DeepSeek-R1
- Кстати, в RL у DeepSeek тоже было новшество: GRPO



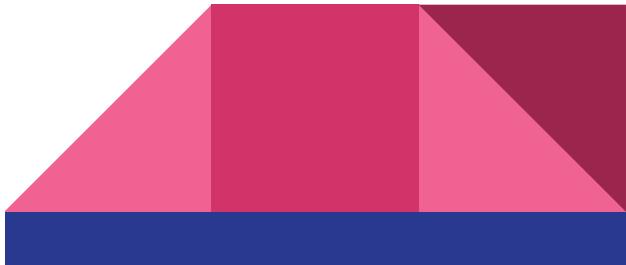
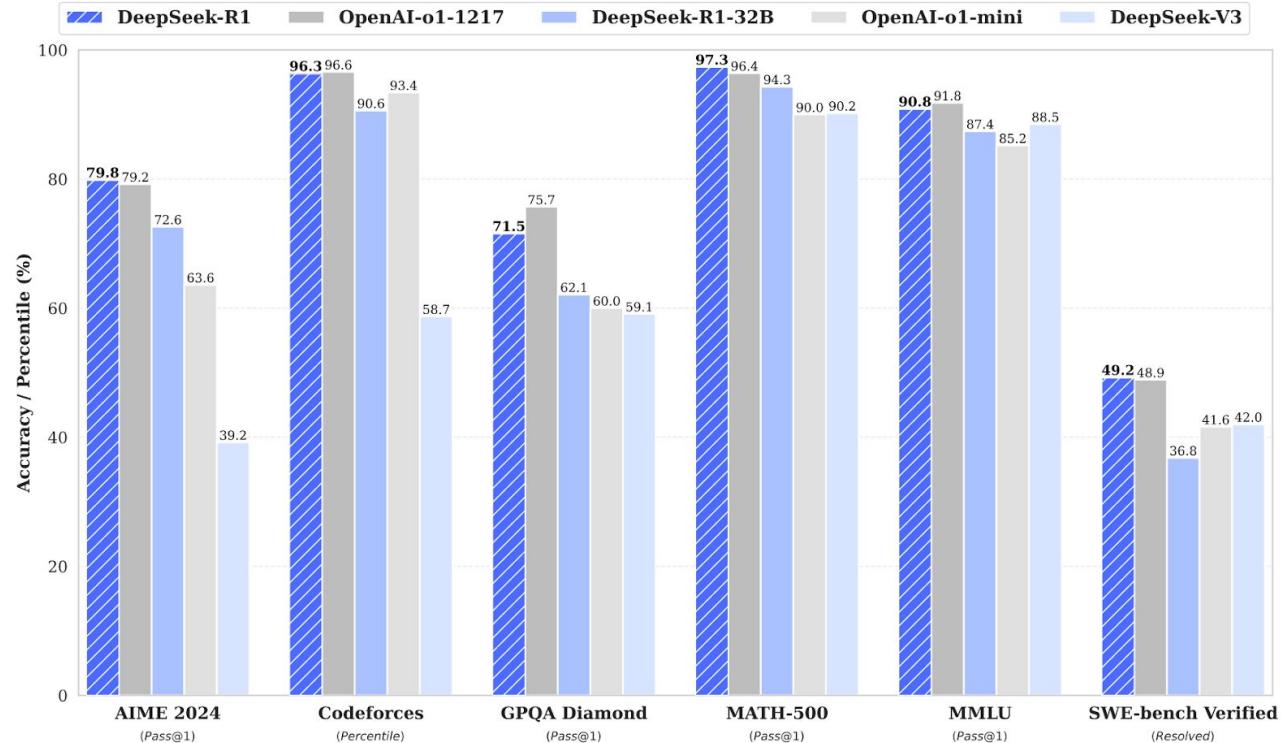
# DeepSeek-R1

- R1: сначала соберём маленький chain-of-thought датасет (несколько тысяч примеров) для cold start
- Потом уже RL на полученной модели
- После сходимости RL ещё соберём датасет из reasoning traces этой модели, отфильтруем по читабельности и каким-то внешним признакам
- Сделаем SFT на этом датасете
- А потом ещё RLHF как обычно поверх



# DeepSeek-R1

- Результаты хорошие, и вообще модель хорошая



# DeepSeek-R1

- А датасет из 800K reasoning traces, который у них получился, можно использовать для дистилляции, тоже очень успешно
- Опять тот же горький урок: distill, baby, distill...



Model	AIME 2024		MATH-500	GPQA Diamond	LiveCode Bench	CodeForces
	pass@1	cons@64				
GPT-4o-0513	9.3	13.4	74.6	49.9	32.9	759
Claude-3.5-Sonnet-1022	16.0	26.7	78.3	65.0	38.9	717
OpenAI-o1-mini	63.6	80.0	90.0	60.0	53.8	<b>1820</b>
QwQ-32B-Preview	50.0	60.0	90.6	54.5	41.9	1316
DeepSeek-R1-Distill-Qwen-1.5B	28.9	52.7	83.9	33.8	16.9	954
DeepSeek-R1-Distill-Qwen-7B	55.5	83.3	92.8	49.1	37.6	1189
DeepSeek-R1-Distill-Qwen-14B	69.7	80.0	93.9	59.1	53.1	1481
DeepSeek-R1-Distill-Qwen-32B	<b>72.6</b>	83.3	94.3	62.1	57.2	1691
DeepSeek-R1-Distill-Llama-8B	50.4	80.0	89.1	49.0	39.6	1205
DeepSeek-R1-Distill-Llama-70B	70.0	<b>86.7</b>	<b>94.5</b>	<b>65.2</b>	<b>57.5</b>	1633

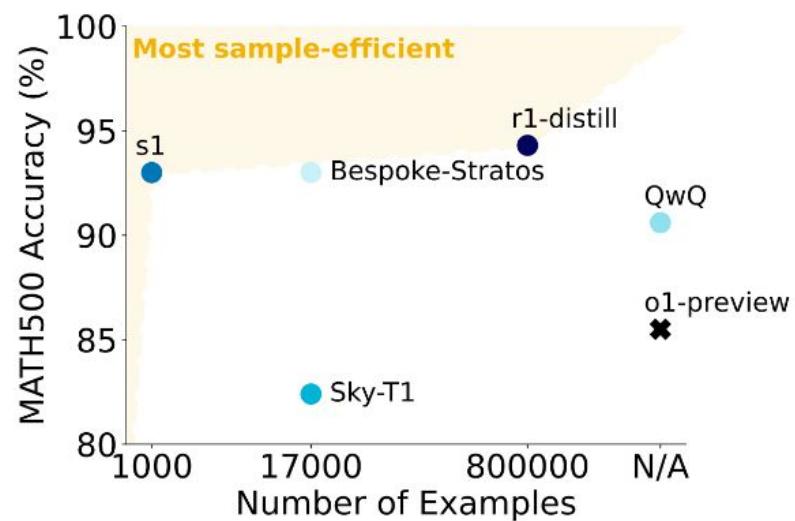
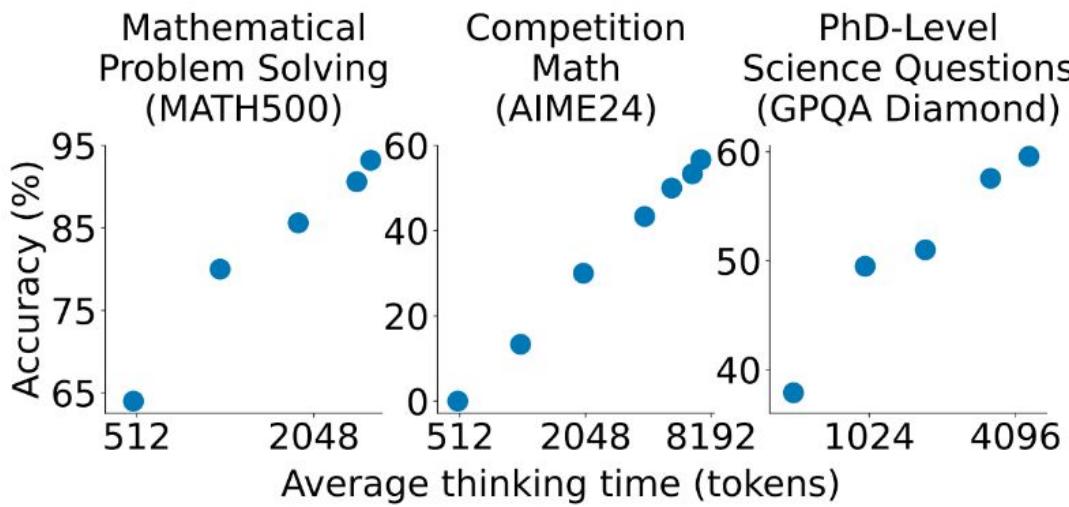
# DeepSeek-R1

- Самое, конечно, поразительное во всём этом – это реакция рынка :) но это уже другой разговор



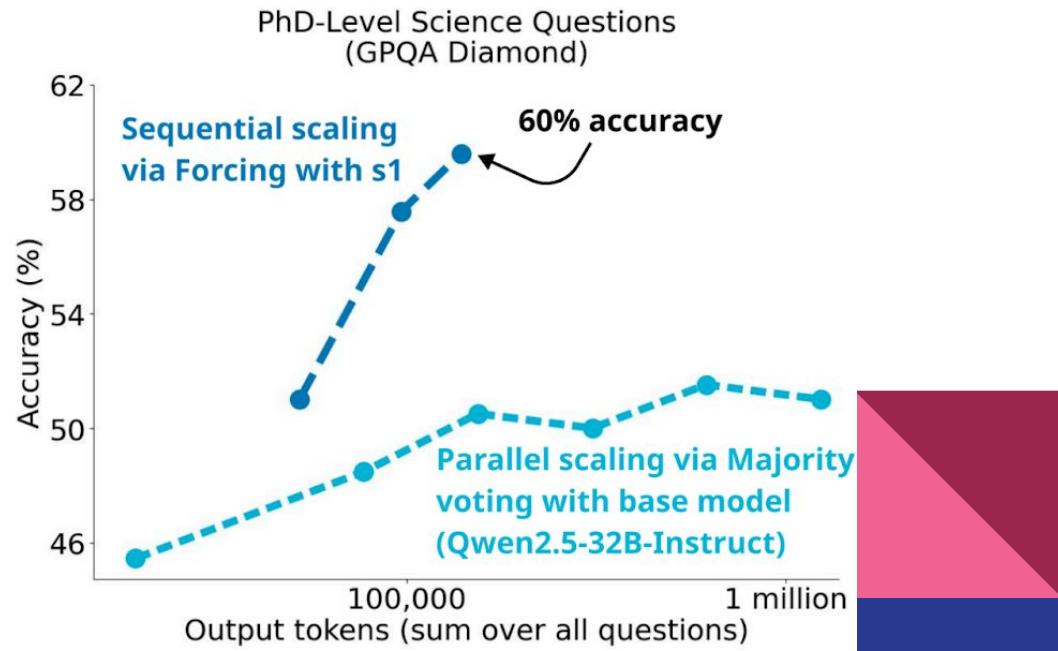
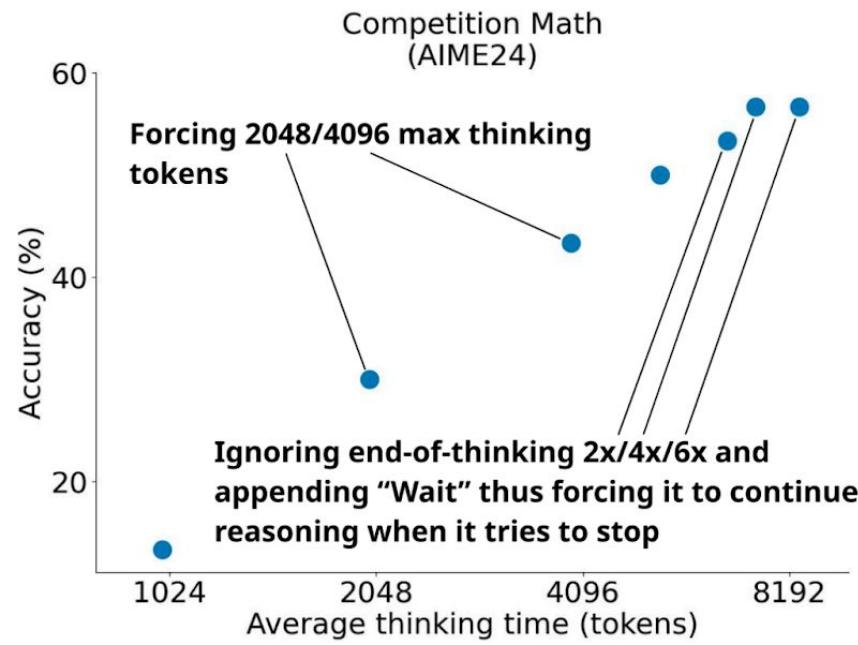
# s1-32B: distill, baby, distill...

- Mueninghoff et al. (2025): s1-32B, дистилляция на 1000 примеров за \$50
- Получились очень хорошие результаты в рассуждениях



# s1-32B: distill, baby, distill...

- Mueninghoff et al. (2025): s1-32B, дистилляция на 1000 примеров за \$50
- И очень простые идеи – budget forcing и голосование – отлично работают



# Итоги

- Рассуждающие модели сейчас везде
- Важный вывод из DeepSeek-R1 – то, что там **не** используется: ни PRM, ни MCTS... оказалось, что достаточно RL хорошо сложить с SFT и RLHF
- Но что это даёт?
- И что будет дальше?..



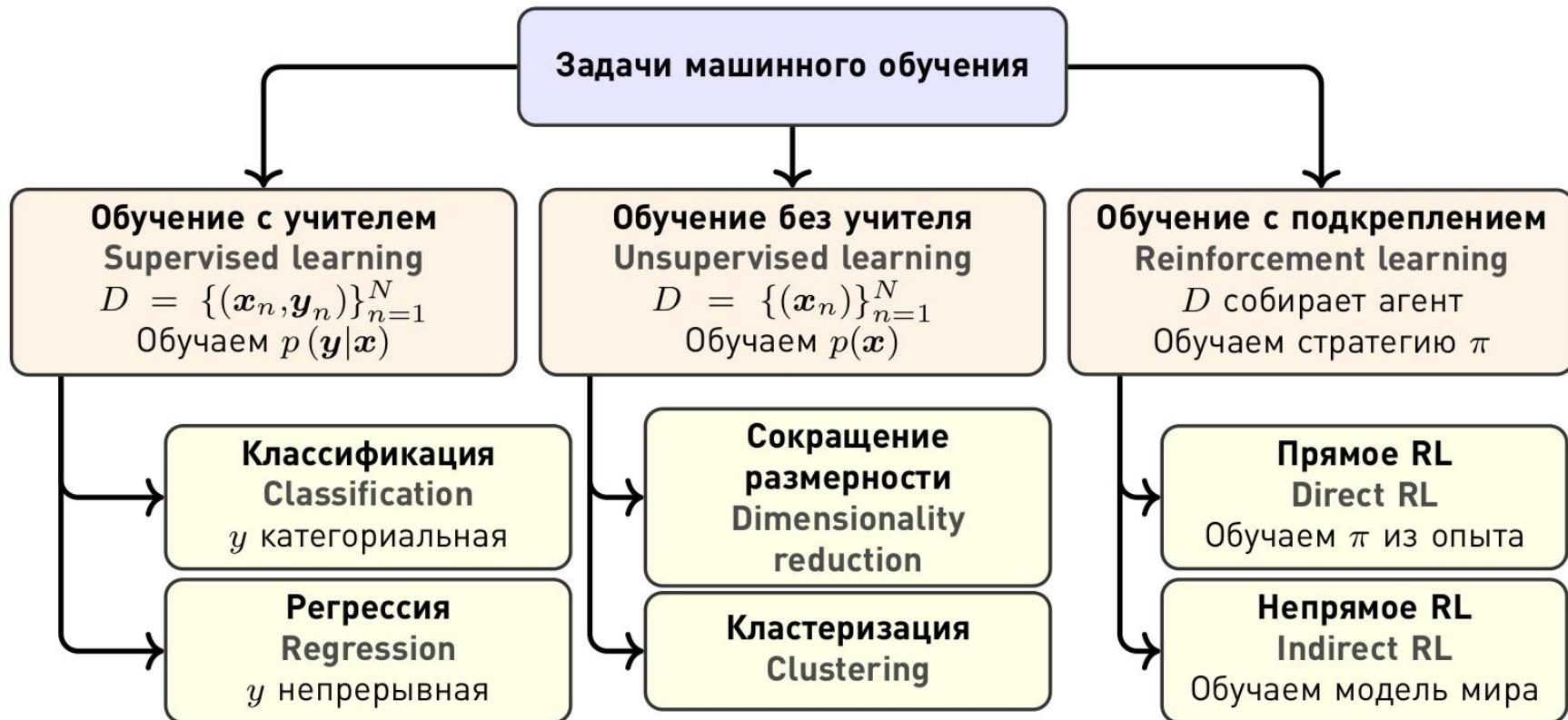
# 5. RL для математики



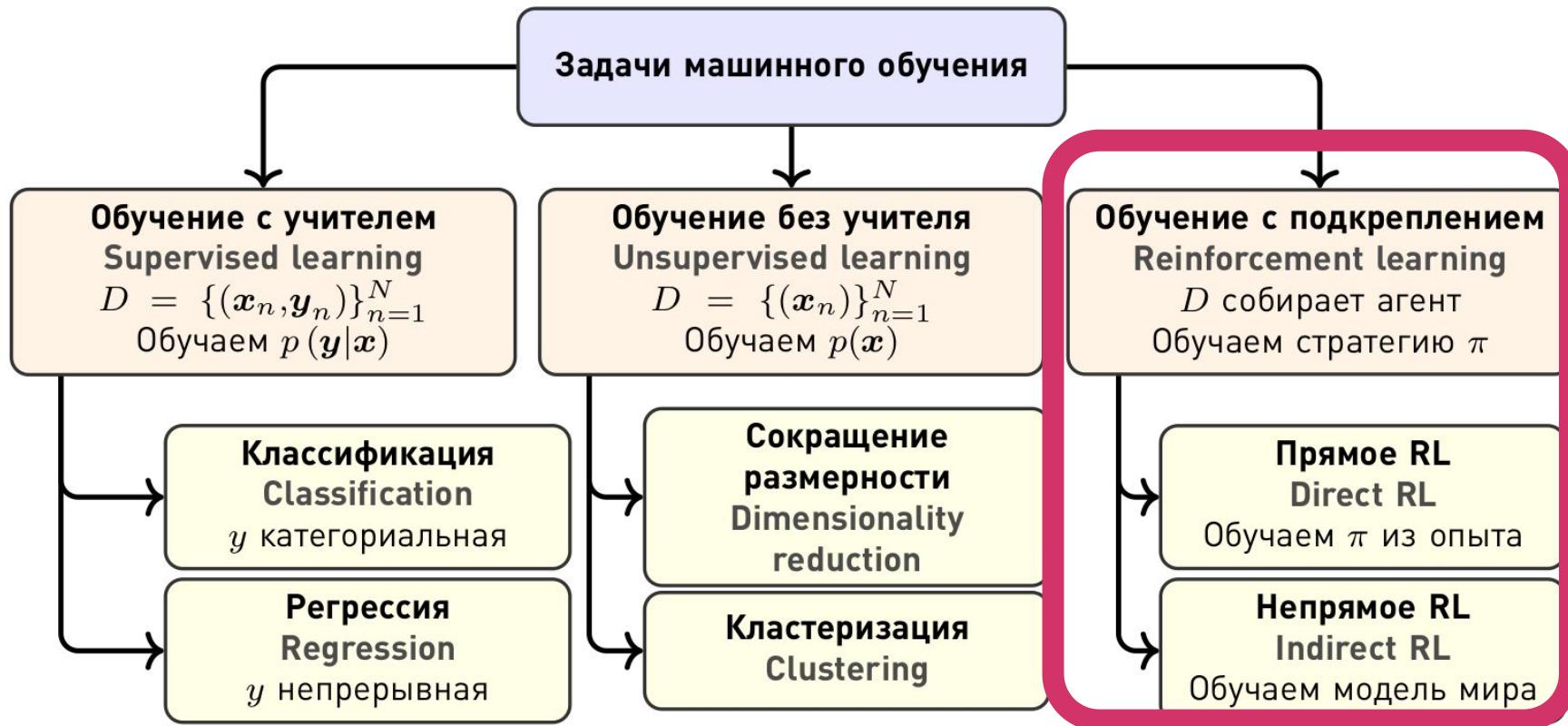
Люблю я покушать. Это моя слабость. Когда я ем, у меня всегда хорошее настроение, ну а когда у меня хорошее настроение, так я ем еще больше. Так что я даже не знаю, чего у меня больше – аппетита или настроения. Впрочем, одно другому не мешает. Потому что я так полагаю: понятно, если не подкрепиться, так можно умереть от огорчения, а огорчений столько, что нельзя не подкрепиться.

Юрий Слёзкин. Бандит

# Машинное обучение

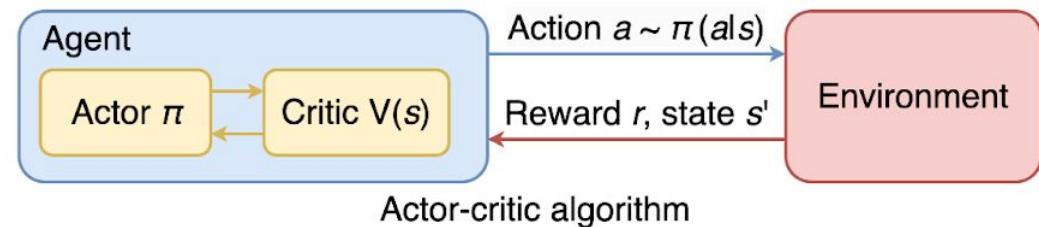
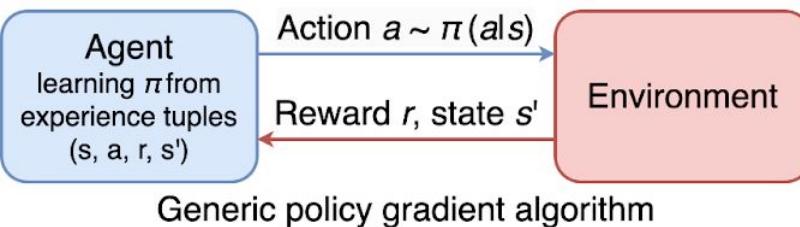


# Машинное обучение



# Обучение с подкреплением

- Обучение с подкреплением (reinforcement learning, RL) – это раздел ML, в котором агент “живёт” в окружающей среде и собирает датасет для обучения по ходу дела
- Для этого нужно суметь реализовать окружающую среду, которая будет давать награду; например, результат партии в шахматы или корректность доказательства в математике



# RL-модели от DeepMind



Atari DQN



AlphaGo



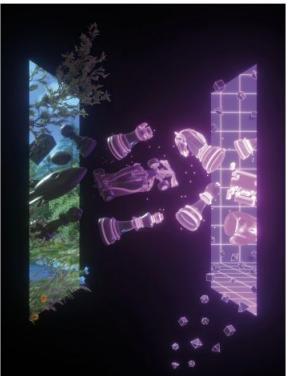
AlphaGoZero



AlphaZero



AlphaStar



MuZero



Stratego



AlphaDev



AlphaTensor



Fusion

# Кстати, о DeepMind



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

**David Baker**

University of Washington, Seattle, WA, USA

*"for computational protein design"*

**Demis Hassabis**

Google DeepMind, London, UK

*"for protein structure prediction"*

**John M. Jumper**

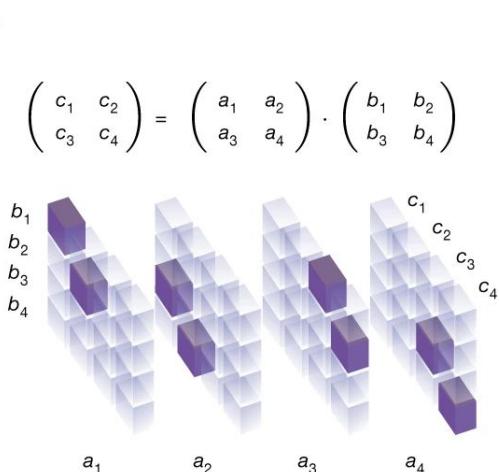
Google DeepMind, London, UK



**They cracked the code for proteins' amazing structures**

# AlphaTensor

- Через RL “перебор случаев” выходит на новый уровень
- AlphaTensor ([Fawzi et al., 2022](#)): RL-метод в стиле AlphaZero для поиска алгоритмов умножения матриц



**b**

$$m_1 = (a_1 + a_4)(b_1 + b_4)$$

$$m_2 = (a_3 + a_4)b_1$$

$$m_3 = a_1(b_2 - b_4)$$

$$m_4 = a_4(b_3 - b_1)$$

$$m_5 = (a_1 + a_2)b_4$$

$$m_6 = (a_3 - a_1)(b_1 + b_2)$$

$$m_7 = (a_2 - a_4)(b_3 + b_4)$$

$$c_1 = m_1 + m_4 - m_5 + m_7$$

$$c_2 = m_3 + m_5$$

$$c_3 = m_2 + m_4$$

$$c_4 = m_1 - m_2 + m_3 + m_6$$

**c**

$$\mathbf{U} = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & -1 \end{pmatrix}$$

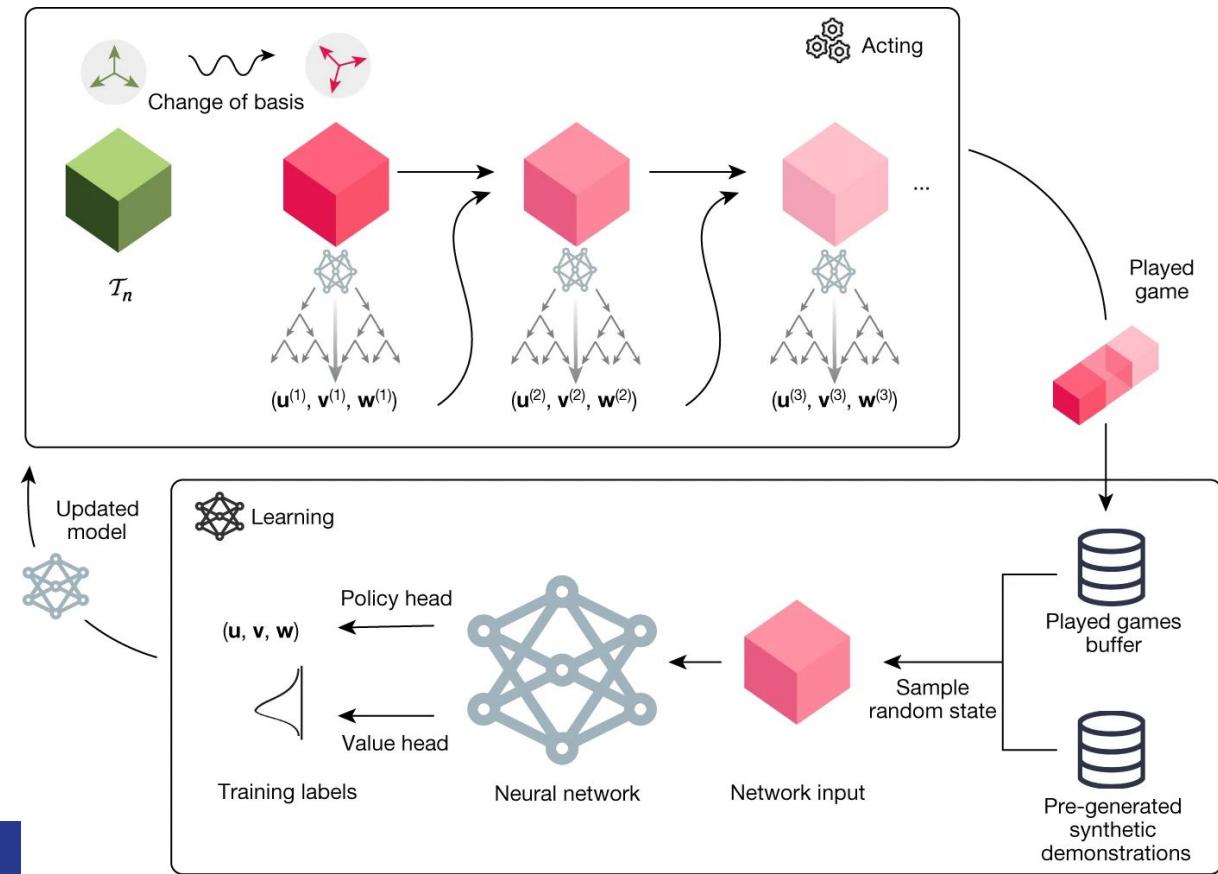
$$\mathbf{V} = \begin{pmatrix} 1 & 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & -1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

$$\mathbf{W} = \begin{pmatrix} 1 & 0 & 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 & 1 & 0 \end{pmatrix}$$



# AlphaTensor

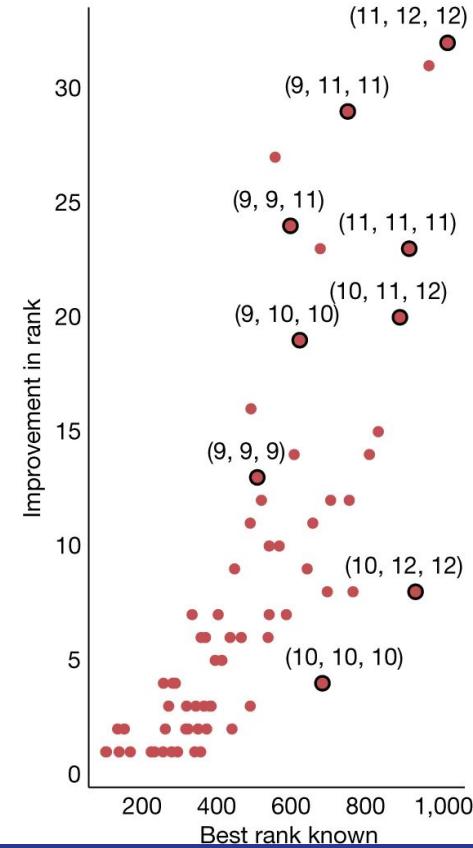
- AlphaTensor ([Fawzi et al., 2022](#)): разница с шахматами в том, что дерево поиска здесь гораздо шире
- Sampled AlphaZero ([Hubert et al., 2021](#)): планирование на сэмплах действий, а не всём множестве



# AlphaTensor

- AlphaTensor ([Fawzi et al., 2022](#)): но всё получилось, и AlphaTensor действительно нашёл новые алгоритмы, улучшающие ещё работы Штрассена

Size ( $n, m, p$ )	Best method known	Best rank known	AlphaTensor rank Modular Standard
(2, 2, 2)	(Strassen, 1969) <sup>2</sup>	7	7
(3, 3, 3)	(Laderman, 1976) <sup>15</sup>	23	23
(4, 4, 4)	(Strassen, 1969) <sup>2</sup>	49	47
	(2, 2, 2) $\otimes$ (2, 2, 2)		49
(5, 5, 5)	(3, 5, 5) + (2, 5, 5)	98	96
<hr/>			
(2, 2, 3)	(2, 2, 2) + (2, 2, 1)	11	11
(2, 2, 4)	(2, 2, 2) + (2, 2, 2)	14	14
(2, 2, 5)	(2, 2, 2) + (2, 2, 3)	18	18
(2, 3, 3)	(Hopcroft and Kerr, 1971) <sup>16</sup>	15	15
(2, 3, 4)	(Hopcroft and Kerr, 1971) <sup>16</sup>	20	20
(2, 3, 5)	(Hopcroft and Kerr, 1971) <sup>16</sup>	25	25
(2, 4, 4)	(Hopcroft and Kerr, 1971) <sup>16</sup>	26	26
(2, 4, 5)	(Hopcroft and Kerr, 1971) <sup>16</sup>	33	33
(2, 5, 5)	(Hopcroft and Kerr, 1971) <sup>16</sup>	40	40
(3, 3, 4)	(Smirnov, 2013) <sup>18</sup>	29	29
(3, 3, 5)	(Smirnov, 2013) <sup>18</sup>	36	36
(3, 4, 4)	(Smirnov, 2013) <sup>18</sup>	38	38
(3, 4, 5)	(Smirnov, 2013) <sup>18</sup>	48	47
(3, 5, 5)	(Sedoglavic and Smirnov, 2021) <sup>19</sup>	58	58
(4, 4, 5)	(4, 4, 2) + (4, 4, 3)	64	63
(4, 5, 5)	(2, 5, 5) $\otimes$ (2, 1, 1)	80	76

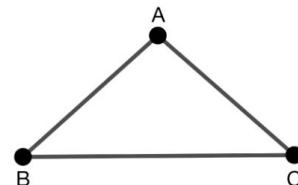


# AlphaGeometry

- AlphaGeometry ([Trinh et al., 2024](#)): LLM + СИМВОЛЬНЫЕ ВЫЧИСЛЕНИЯ

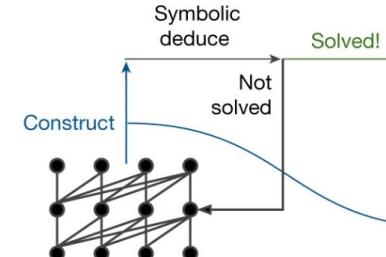
- LLM порождает идеи для новых построений, a symbolic engine выводит из них всё, что можно

**a** A simple problem



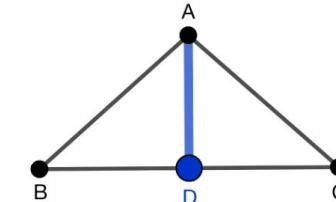
"Let  $ABC$  be any triangle with  $AB = AC$ .  
Prove that  $\angle ABC = \angle BCA$ ."

**b** AlphaGeometry



**c** Language model

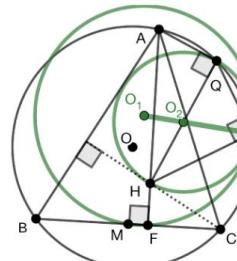
**d** Solution



**Construct D: midpoint BC,**  
 $AB=AC, BD = DC, AD=AD \Rightarrow \angle ABD = \angle DCA [1]$   
[1], **B C D collinear**  $\Rightarrow \angle ABC = \angle BCA$

**e** IMO 2015 P3

"Let  $ABC$  be an acute triangle. Let  $(O)$  be its circumcircle,  $H$  its orthocenter, and  $F$  the foot of the altitude from  $A$ . Let  $M$  be the midpoint of  $BC$ . Let  $Q$  be the point on  $(O)$  such that  $QH \perp QA$  and let  $K$  be the point on  $(O)$  such that  $KH \perp KQ$ . Prove that the circumcircles  $(O_1)$  and  $(O_2)$  of triangles  $FKM$  and  $KQH$  are tangent to each other."



**f** Solution

**Construct D: midpoint BH [a]**  
[a],  $O_2$  midpoint HQ  $\Rightarrow BQ \parallel O_2D$  [20]  
...

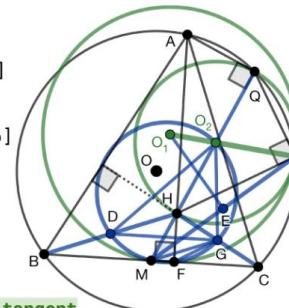
**Construct G: midpoint HC [b]** ...  
 $\angle GMD = \angle GQD \Rightarrow M O_2 G D$  cyclic [26]

[a], [b]  $\Rightarrow BC \parallel DG$  [30]

**Construct E: midpoint MK [c]**  
..., [c]  $\Rightarrow \angle KFC = \angle KQ_1 E$  [104]

...  
 $\angle FKO_1 = \angle FKO_2 \Rightarrow K_1 O_1 \parallel K_2 O_2$  [109]  
[109]  $\Rightarrow O_1 O_2 K$  collinear  $\Rightarrow (O_1)(O_2)$  tangent

Alpha-  
Geometry

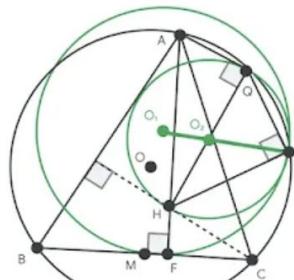


# AlphaGeometry

- AlphaGeometry  
([Trinh et al., 2024](#)): LLM +  
СИМВОЛЬНЫЕ  
ВЫЧИСЛЕНИЯ
- В том числе  
для довольно  
сложных  
задач (здесь  
109 шагов)

## IMO 2015 P3

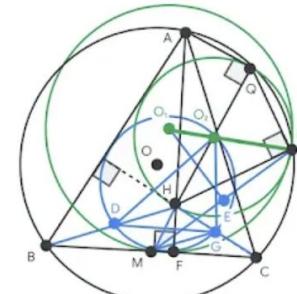
Let  $ABC$  be an acute triangle. Let  $(O)$  be its circumcircle,  $H$  its orthocenter, and  $F$  the foot of the altitude from  $A$ . Let  $M$  be the midpoint of  $BC$ . Let  $Q$  be the point on  $(O)$  such that  $QH \perp QA$  and let  $K$  be the point on  $(O)$  such that  $KH \perp KQ$ . Prove that the circumcircles  $(O_1)$  and  $(O_2)$  of triangles  $FKM$  and  $KQH$  are tangent to each other.



AlphaGeometry

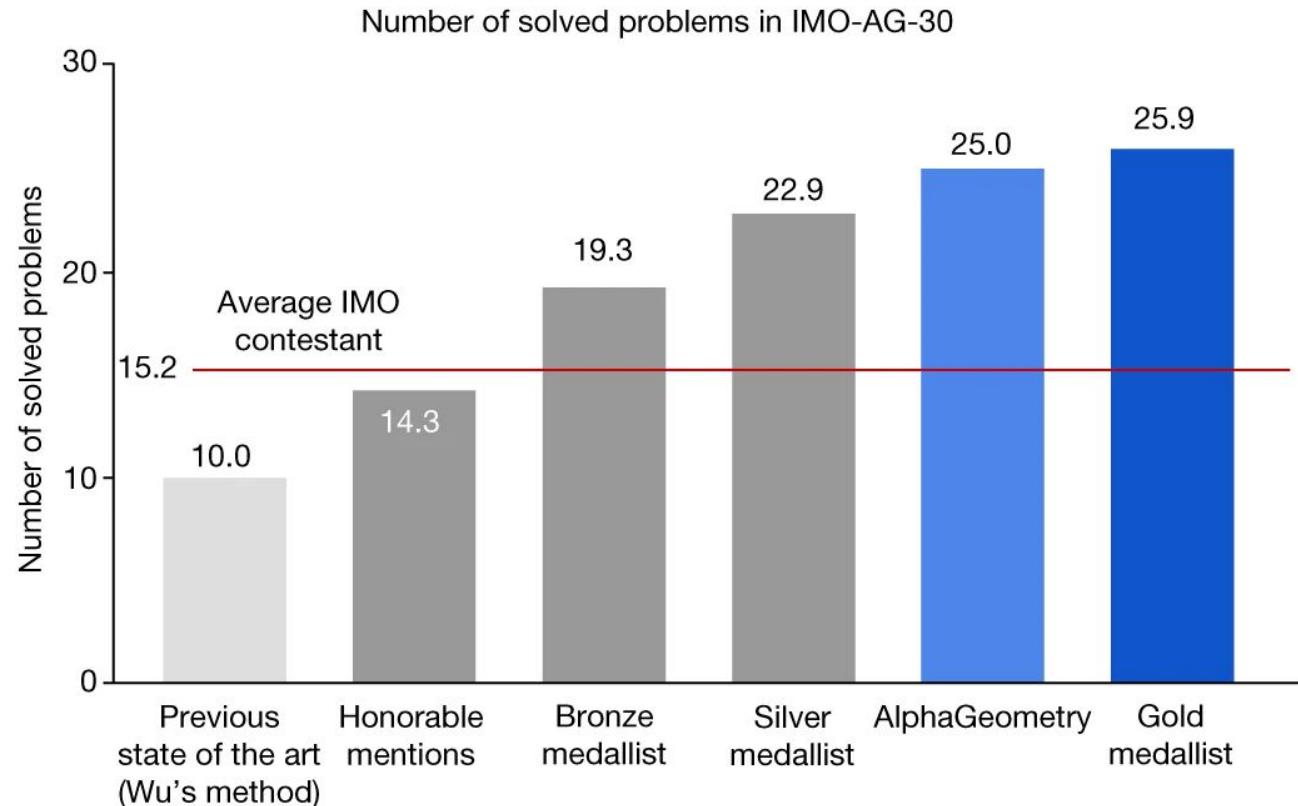
## Solution

[...]  
Construct  $D$ : midpoint  $BH$  [a]  
[a],  $O_2$  midpoint  $HQ \Rightarrow BQ \parallel O_2 D$  [20]  
[...]  
Construct  $G$ : midpoint  $HC$  [b]  
 $\angle GMD = \angle GO_2 D \Rightarrow M O_2 G D$  cyclic [26]  
[...]  
[a], [b]  $\Rightarrow BC \parallel DG$  [30]  
[...]  
Construct  $E$ : midpoint  $MK$  [c]  
[c]  $\Rightarrow \angle KFC = \angle KO_1 E$  [104]  
[...]  
 $\angle FKO_1 = \angle FKO_2 \Rightarrow KO_1 \parallel KO_2$  [109]  
[109]  $\Rightarrow O_1 O_2 K$  collinear  $\Rightarrow (O_1)(O_2)$  tangent



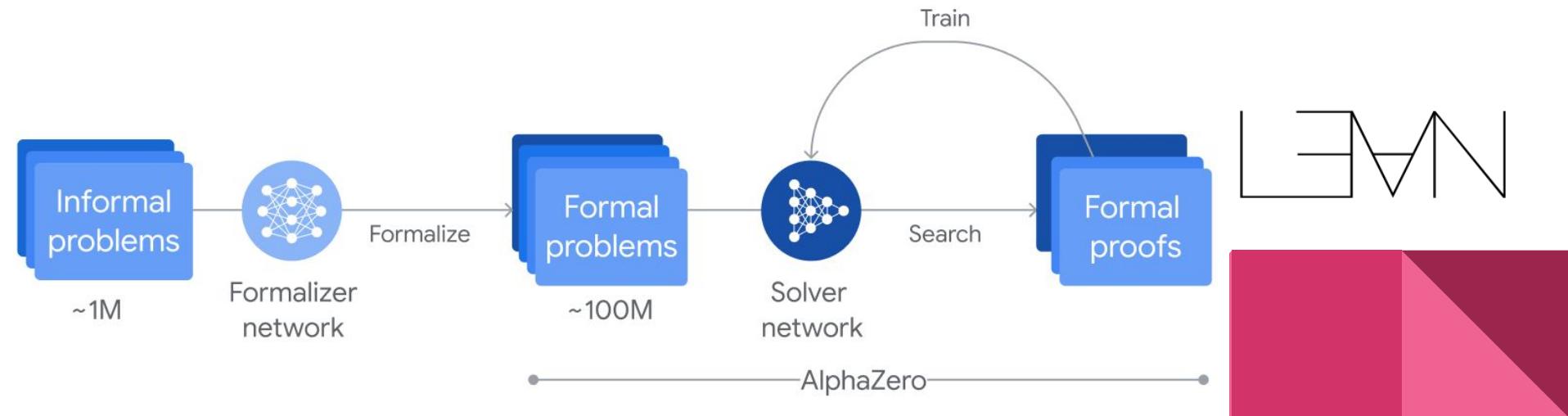
# AlphaGeometry

- AlphaGeometry  
([Trinh et al., 2024](#))  
решал геометрию  
на уровне серебра  
IMO
- Потом появилась  
AlphaGeometry 2, на  
модели Gemini  
лучше и на  
большем  
синтетическом  
датасете



# AlphaProof

- AlphaProof ([DeepMind, 2024](#)): давайте соединим LLM и прувер
- Специально дообученная модель Gemini переводит естественный язык в формальную постановку, а потом RL в стиле AlphaZero обучается искать доказательство в Lean



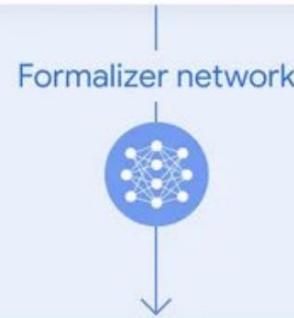
# AlphaProof

- AlphaProof ([Hubert et al., 2024](#)) по пунктам
- Формализатор – это языковая модель из семейства Gemini

## IMO 2021 Shortlist, Problem A5

Let  $n \geq 2$  be an integer and let  $a_1, a_2, \dots, a_n$  be positive real numbers such that  $a_1 + a_2 + \dots + a_n = 1$ . Prove that

$$\sum_{k=1}^n \frac{a_k}{1-a_k} (a_1 + a_2 + \dots + a_{k-1})^2 < \frac{1}{3}.$$

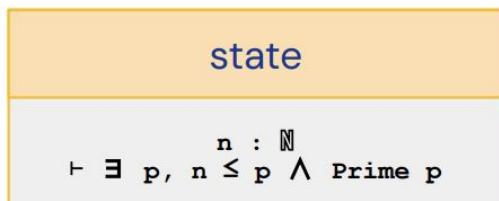


```
theorem imo_shortlist_2021_a5
  (n : ℕ) (hn : 2 ≤ n) (a : ℕ → ℝ)
  (hapos : ∀ i, 0 < a i)
  (hasum : Σ i in Finset.Icc 1 n, a i = 1) :
  Σ k in Finset.Icc 1 n, a k / (1 - a k) *
  (Σ i in Finset.Icc 1 (k-1), a i) ^ 2 < 1 / 3
```

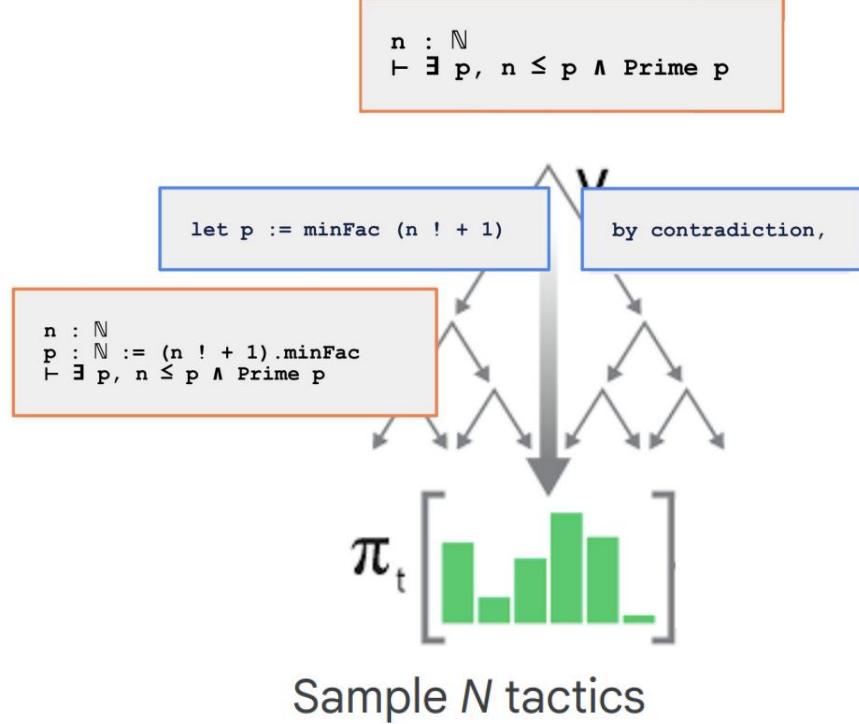
# AlphaProof

- AlphaProof ([Hubert et al., 2024](#)) по пунктам
- Солвер – это модель, которая выбирает следующий “ход” в виде Lean tactic, как AlphaZero

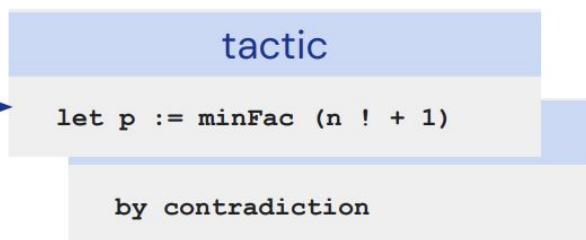
Input a lean state



Prover  
Model

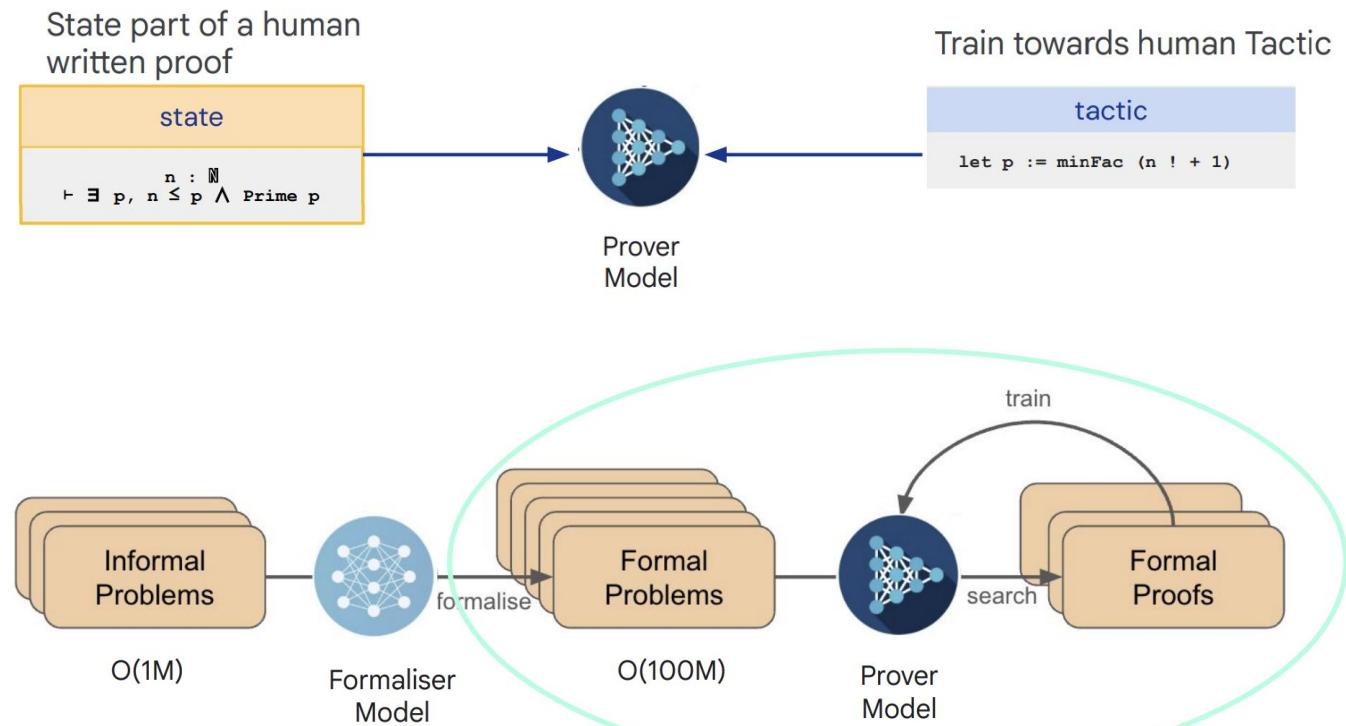


Sample  $N$  tactics



# AlphaProof

- AlphaProof  
([Hubert et al., 2024](#)) по пунктам
- Солвер обучается сначала на mathlib с учителем, а потом RL, где Lean проверяет доказательства



# AlphaProof

- Солвер учится доказывать так же, как AlphaZero учится играть в шахматы, это поиск по дереву; а LLM даёт формализацию и новые идеи
- Curriculum learning: по мере решения задач они добавляются в обучение, т. е. сами собой идут от простого к сложному
- AlphaProof ([DeepMind, 2024](#)): вместе с AlphaGeometry 2 они на IMO 2024 получили серебро, минус один балл от золота



# RL и гипотеза Эндрюса-Кёртиса

- Shehper et al. (Feb 2025):  
почему новые доказательства трудно найти через RL?
- Потому, что вариантов очень много, и путь может быть очень длинным
- Гипотеза Эндрюса-Кёртиса выглядит как хороший кандидат для автоматического поиска

Let  $G$  be a group,  $m$  be a positive integer, and  $G^m$  be the cartesian product of  $n$  copies of  $G$ . Consider the following elementary transformations that can be applied to an arbitrary tuple  $U = \{u_1, \dots, u_m\} \in G^m$ :

- (AC1)  $u_i$  is replaced by  $u_i u_j$  for some  $j \neq i$ ;
- (AC2)  $u_i$  is replaced by  $u_i^{-1}$ ;
- (AC3)  $u_i$  is replaced by  $gu_ig^{-1}$  for some  $g \in F$ .

Two  $n$ -tuples  $V$  and  $W$  from  $G^m$  are called *Andrews-Curtis equivalent* (or *AC-equivalent*) if one of them can be obtained from the other by a finite sequence of elementary AC-transformations. In this event, we write  $V \sim W$ . The relation  $\sim$  is an equivalence relation on  $G^m$ .

**The Andrews-Curtis conjecture.** Let  $F = F_n$  be a free group of a finite rank  $n \geq 2$  with basis  $X$ . Then for any  $Y \in F^n$  the following equivalence holds:

$$\langle Y \rangle^F = F \iff Y \sim X$$

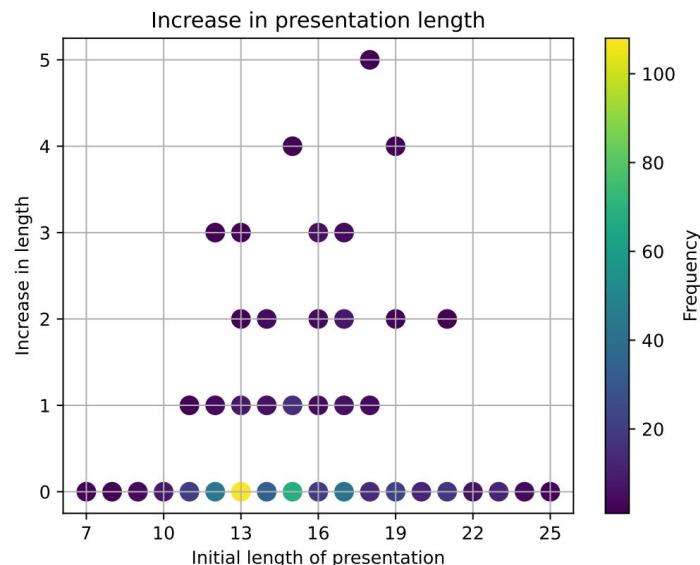
(AC4)  $Y$  is extended to a tuple  $\{y_1, \dots, y_m, x_\nu\}$  over a new free group with basis  $X \cup \{x_\nu\}$ .

(AC5) the converse of (AC4).

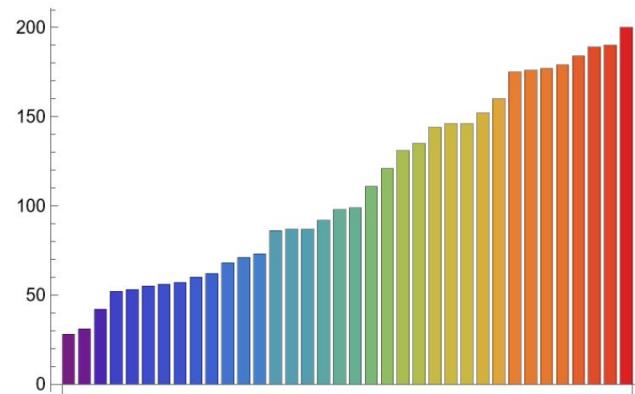
If  $U \in F^m$  can be obtained from  $W \in F^k$  by means of transformations (AC1–AC5), we say that  $U$  is *stably* AC-equivalent to  $W$ .

# RL и гипотеза Эндрюса-Кёртиса

- Shehper et al. (Feb 2025) исследуют эту задачу и находят, что пути длинные, и по дороге слова могут становиться длиннее



Element	Implementation details
State	a pair of words $(r_1, r_2)$ written in generators $(x, y)$ ; encoded into tokens.
Actions	the 5 AC moves $\pm 30$ hand-crafted “super-moves” (macro sequences discovered by pattern mining).
Reward	0 everywhere, +1 only if both relators shorten to 1 after $\leq H$ steps.
Horizon	variable; trained up to $H = 12\,000$ moves (chess $\approx 40!$ ).



# RL и гипотеза Эндрюса-Кёртиса

- Shehper et al. (Feb 2025): и тем не менее, новый результат в итоге получился! Даже без больших вычислительных ресурсов

**Theorem 1.** *The following potential counterexample introduced by Akbulut and Kirby [AK85] is stably AC-trivial:*

$$AK(3) = \langle x, y \mid x^3 = y^4, xyx = yxy \rangle.$$

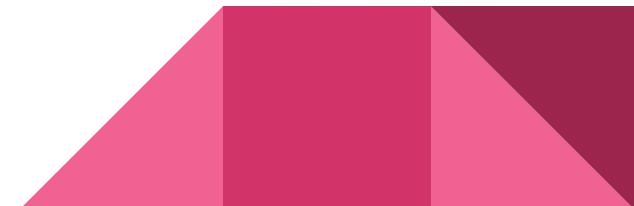
**Theorem** (Myasnikov, Myasnikov, and Shpilrain, [MMS02]). *Using the notation  $[a, b] = aba^{-1}b^{-1}$  and  $[a, b, c] = [[a, b], c]$ , any presentation of the following form is a presentation of the trivial group:*

$$\langle x, y, z \mid x = z \cdot [y^{-1}, x^{-1}, z], \ y = x \cdot [y^{-1}, x^{-1}, z^{-1}] \cdot [z^{-1}, x], \ w \rangle,$$

*where  $w$  is a word in  $x$ ,  $y$ , and  $z$  whose exponent sum on  $x$ ,  $y$ , and  $z$  equals  $\pm 1$ . Moreover, all such presentations are stably AC-trivial.*

For  $w = x^{-1}yz$ , one of the relators eliminates the generator  $z$ , resulting in the following length 25 presentation with two generators:

$$\langle x, y \mid x^{-1}y^{-1}xy^{-1}x^{-1}yxy^{-2}xyx^{-1}y, \ y^{-1}x^{-1}y^2x^{-1}y^{-1}xyxy^{-2}x \rangle.$$



# RL и гипотеза Эндрюса-Кёртиса

- Shehper et al. (Feb 2025): и тем не менее, новый результат в итоге получился! Даже без больших вычислительных ресурсов

For  $w = x^{-1}yz$ , one of the relators eliminates the generator  $z$ , resulting in the following length 25 presentation with two generators:

$$\langle x, y \mid x^{-1}y^{-1}xy^{-1}x^{-1}yxy^{-2}xyx^{-1}y, y^{-1}x^{-1}y^2x^{-1}y^{-1}xyxy^{-2}x \rangle.$$

The sequence of moves that connects the length-25 presentation to AK(3) can be conveniently expressed in terms of the following 12 transformations:

$$\begin{aligned} h_1 &= r_2 \rightarrow r_2r_1, & h_5 &= r_2 \rightarrow x^{-1}r_2x, & h_9 &= r_2 \rightarrow xr_2x^{-1}, \\ h_2 &= r_1 \rightarrow r_1r_2^{-1}, & h_6 &= r_1 \rightarrow y^{-1}r_1y, & h_{10} &= r_1 \rightarrow yr_1y^{-1}, \\ h_3 &= r_2 \rightarrow r_2r_1^{-1}, & h_7 &= r_2 \rightarrow y^{-1}r_2y, & h_{11} &= r_2 \rightarrow yr_2y^{-1}, \\ h_4 &= r_1 \rightarrow r_1r_2, & h_8 &= r_1 \rightarrow xr_1x^{-1}, & h_{12} &= r_1 \rightarrow x^{-1}r_1x, \end{aligned}$$

among which a careful reader can recognize moves (AC'1) and (AC'2) introduced in [Section 2](#). Expressed in terms of the moves  $h_i$ , the desired sequence has length 53 and looks as follows:

$$\begin{aligned} &h_9 \cdot h_7 \cdot h_4 \cdot h_8 \cdot h_{11} \cdot h_5 \cdot h_{11} \cdot h_9 \cdot h_3 \cdot h_{10} \cdot h_{12} \cdot h_7 \cdot h_7 \cdot h_9 \cdot h_{11} \cdot h_5 \cdot h_3 \cdot h_5 \cdot \\ &h_4 \cdot h_3 \cdot h_{12} \cdot h_5 \cdot h_7 \cdot h_7 \cdot h_1 \cdot h_9 \cdot h_{11} \cdot h_8 \cdot h_3 \cdot h_5 \cdot h_{10} \cdot h_2 \cdot h_6 \cdot h_{12} \cdot h_9 \cdot h_7 \cdot \\ &h_5 \cdot h_{11} \cdot h_{10} \cdot h_3 \cdot h_8 \cdot h_{11} \cdot h_9 \cdot h_2 \cdot h_{10} \cdot h_{12} \cdot h_5 \cdot h_7 \cdot h_9 \cdot h_{11} \cdot h_1 \cdot h_9 \cdot h_8. \end{aligned}$$

# RL и гипотеза Эндрюса-Кёртиса

- Shehper et al. (ECCV 2022)

мен  
ре  
пол  
без  
выч  
ресу

Но это всё опять  
“умный перебор  
случаев”, разве  
нет? Где же ваши  
хвалёные LLM?



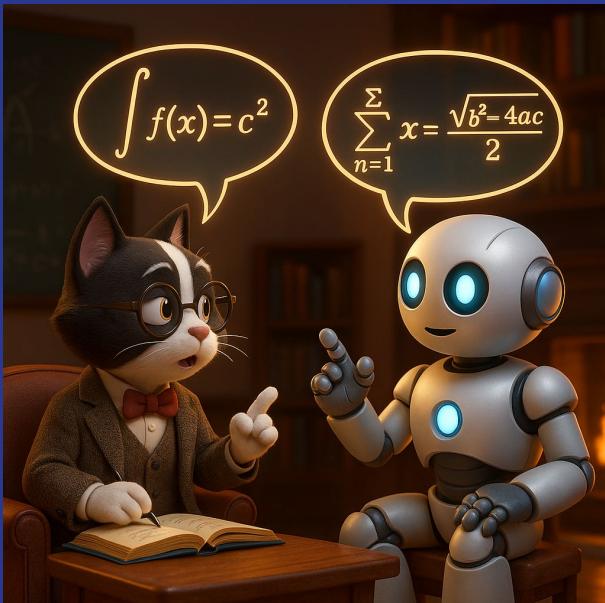
resulting in the  
 $x^{-2} x$ ).

AK(3) can be  
 $x^{-1}$ ,  
 $x^{-1}$ ,  
 $x^{-1}$ ,  
 $r_1 x$ ,

(AC'2) introduced in  
 $\tau$ , the desired sequence has length 53

$h_4 \cdot h_8 \cdot h_{11} \cdot h_5 \cdot h_{11} \cdot h_9 \cdot h_3 \cdot h_{10} \cdot h_{12} \cdot h_7 \cdot h_7 \cdot h_9 \cdot h_{11} \cdot h_5 \cdot h_3 \cdot h_5 \cdot h_4 \cdot h_3 \cdot h_{12} \cdot h_5 \cdot h_7 \cdot h_7 \cdot h_1 \cdot h_9 \cdot h_{11} \cdot h_8 \cdot h_3 \cdot h_5 \cdot h_{10} \cdot h_2 \cdot h_6 \cdot h_{12} \cdot h_9 \cdot h_7 \cdot h_5 \cdot h_{11} \cdot h_{10} \cdot h_3 \cdot h_8 \cdot h_{11} \cdot h_9 \cdot h_2 \cdot h_{10} \cdot h_{12} \cdot h_5 \cdot h_7 \cdot h_9 \cdot h_{11} \cdot h_1 \cdot h_9 \cdot h_8.$

# 6. LLM в математике

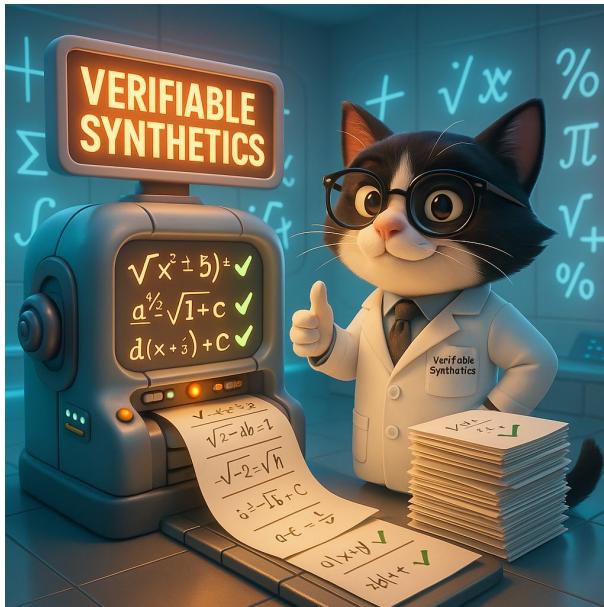


Я так привык к самостоятельной работе, что считал более лёгким для себя доказать теорему без книги, чем вычитывать из нее доказательства. Лишь не всегда это удавалось.

Константин Циолковский

# Почему математика для LLM хороша

- LLM обучается предсказывать следующий токен в тексте
- Это обучение с учителем
- И поэтому особенно хорошо, если мы можем породить синтетические данные с правильными ответами
- Мы можем так сделать для (простой) математики и для (простого) программирования, поэтому math & coding – две области, где LLM всё время проверяют



# LLM для математики

- LLM, конечно, ничего формально выводить в системах доказательств не умеют (как и живые математики, впрочем); первые успехи были скромные
- Minerva ([Lewkowycz et al., 2022](#))

**Question:** Assume that the variance of the first  $n$  natural numbers is 10, and the variance of the first  $m$  even natural numbers is 16. Compute  $m + n$ .

**Model output:** Let  $\sigma^2$  be the variance of the first  $m$  even natural numbers, and  $\tau^2$  be the variance of the first  $n$  natural numbers. Then, we have  $\tau^2 = \frac{n^2 - 1}{12}$  and  $\sigma^2 = \frac{m^2 - 1}{3}$ . We solve the system of equations:

$$\frac{n^2 - 1}{12} = 10$$

$$\frac{m^2 - 1}{3} = 16.$$

This translates into the system

$$n^2 = 121$$

$$m^2 = 48 + 1 = 49.$$

Therefore,  $n = \boxed{11}$  and  $m = \boxed{7}$ , so  $n + m = \boxed{18}$ .

**Question:** For every  $a, b, b \neq a$  prove that

$$\frac{a^2 + b^2}{2} > \left(\frac{a + b}{2}\right)^2.$$

**Model output:**

$$\frac{a^2 + b^2}{2} > \left(\frac{a+b}{2}\right)^2$$

$$\iff \frac{a^2 + b^2}{2} > \frac{a^2 + b^2 + 2ab}{4}$$

$$\iff a^2 + b^2 > \frac{a^2 + b^2 + 2ab}{2}$$

$$\iff 2a^2 + 2b^2 > a^2 + b^2 + 2ab$$

$$\iff a^2 + b^2 > 2ab$$

$$\iff a^2 + b^2 - 2ab > 0$$

$$\iff (a - b)^2 > 0$$

which is true, because the square of a real number is positive.

# LLM для математики

- ToRA  
(tool-integrated  
reasoning agent;  
Gou et al., Sep  
29, 2023) уже  
делала  
tool-assisted  
математические  
рассуждения

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.

We are given that  $x^2 + y^2 = 7$  and  $x^3 + y^3 = 10$ . We can factor  $x^3 + y^3$  to get  $(x + y)(x^2 - xy + y^2)$ .

Thus, we can write  $10 = (x + y)(7)$ . It follows that  $x + y$  must divide 10.

The only possibilities are that  $x + y$  is 1, 2, 5 or 10.

```
import math

def solution():
    x = symbols('x'), y = symbols('y')
    x**2 + y**2 = 7
    x**3 + y**3 = 10
```

```
x = (-1 + math.sqrt(17)) / 2
y = (-1 - math.sqrt(17)) / 2
```

```
# Get all possible values for x+y
result = [x + y, -x - y]
```

```
>>> SyntaxError: cannot assign to expression.
```

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

```
from sympy import symbols, Eq, solve, simplify

def possible_values():
    x, y = symbols('x y')
    eq1 = Eq(x**2 + y**2, 7)
    eq2 = Eq(x**3 + y**3, 10)
    solutions = solve((eq1, eq2), (x, y))
    sums = [simplify(sol[0] + sol[1]) for sol in
    solutions]
    return sums
print(possible_values())
>>> [-5, -5, 1, 1, 4, 4]
```

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

X

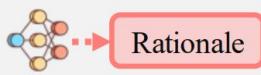
X

$r_1$

$a_1$

$O_1$

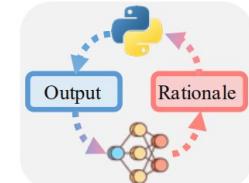
$r_2$



(a) Rationale-based



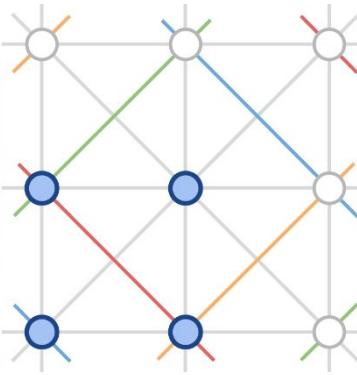
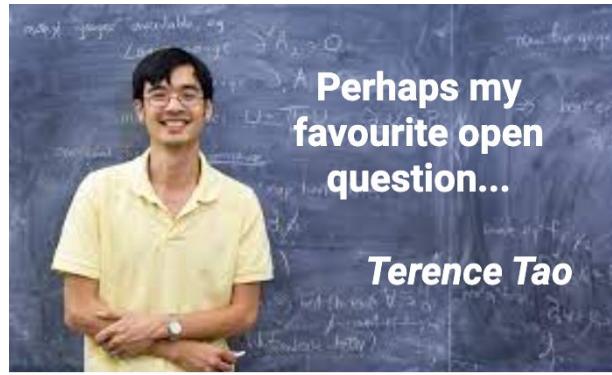
(b) Program-based



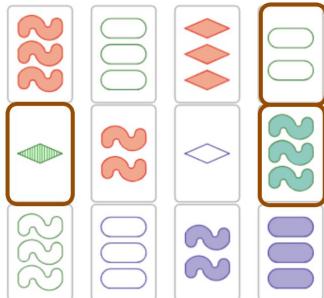
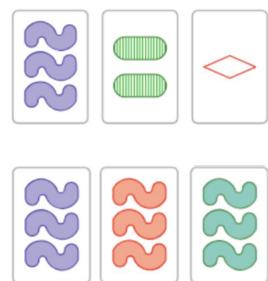
(c) Tool-integrated Reasoning  
(Format used by ToRA)

# LLM для математики

- [Romera-Paredes et al. \(December 2023\)](#): DeepMind's FunSearch proves new bounds for the cap set problem



$n$	3	4	5	6	7	8
Best known	9	20	45	112	236	496
FunSearch	9	20	45	112	236	512



- Это почти “перебор случаев”,  
но очень направленный  
и в виде функций

```

def priority(el: tuple[int,...], n: int) -> float:
    score = n
    in_el = 0
    el_count = el.count(0)

    if el_count == 0:
        score += n**2
    if el[1] == el[-1]:
        score *= 1.5
    if el[2] == el[-2]:
        score *= 1.5
    if el[3] == el[-3]:
        score *= 1.5
    else:
        if el[1] == el[-1]:
            score *= 0.5
        if el[2] == el[-2]:
            score *= 0.5

    for e in el:
        if e == 0:
            if in_el == 0:
                score *= n * 0.5
            elif in_el == el_count - 1:
                score *= 0.5
        else:
            score *= n * 0.5 ** in_el
            in_el += 1
    else:
        score *= 1

    if el[1] == el[-1]:
        score *= 1.5
    if el[2] == el[-2]:
        score *= 1.5

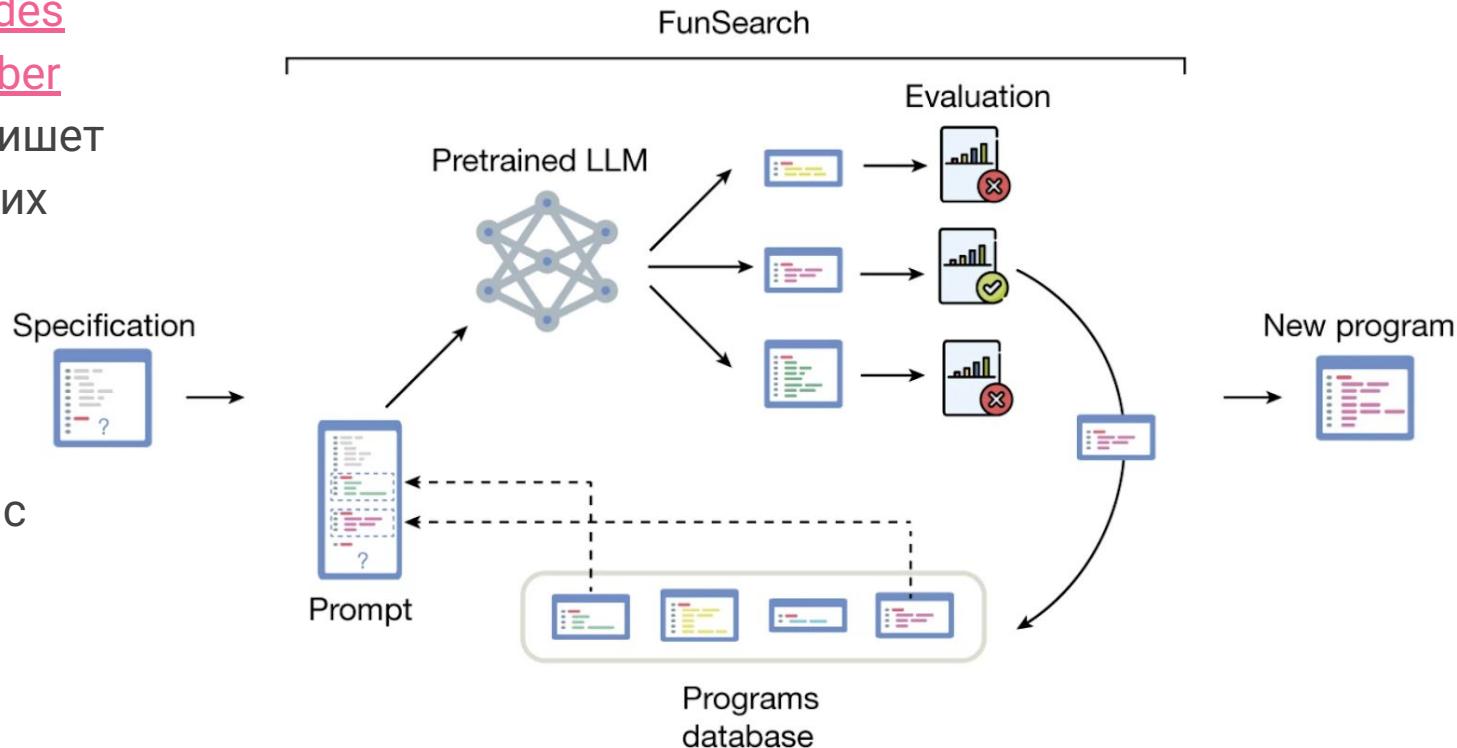
    return score

```

# LLM для математики

- Romera-Paredes et al. (December 2023): LLM пишет программы, их запускают, выбирают лучшие и дают опять на вход LLM с просьбой улучшить

- В итоге улучшили верхнюю оценку для  $n=8$ , т.е. для Set с восемью атрибутами



# NuminaMath

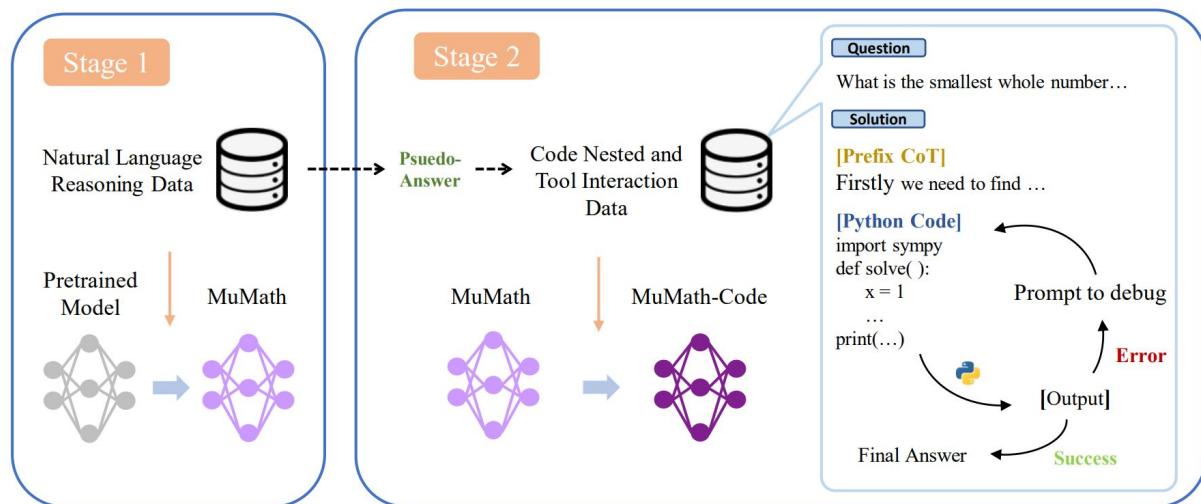
- Fleureau et al. (2024):  
How Numinamath Won the 1st AIMO Progress Prize
- Пример того, как правильно (до)обучать LLM для математических рассуждений
- Задачи AIMO проще, чем IMO, но тогда были сложными для LLM:

#	Team	Members	Score
1	Numina		29
2	CMU_MATH		22
3	after exams		21
4	codeinter		21
5	Conor #2		20

“Let  $k, l > 0$  be parameters. The parabola  $y = kx^2 - 2kx + l$  intersects the line  $y = 4$  at two points  $A$  and  $B$ . These points are distance 6 apart. What is the sum of the squares of the distances from  $A$  and  $B$  to the origin?”

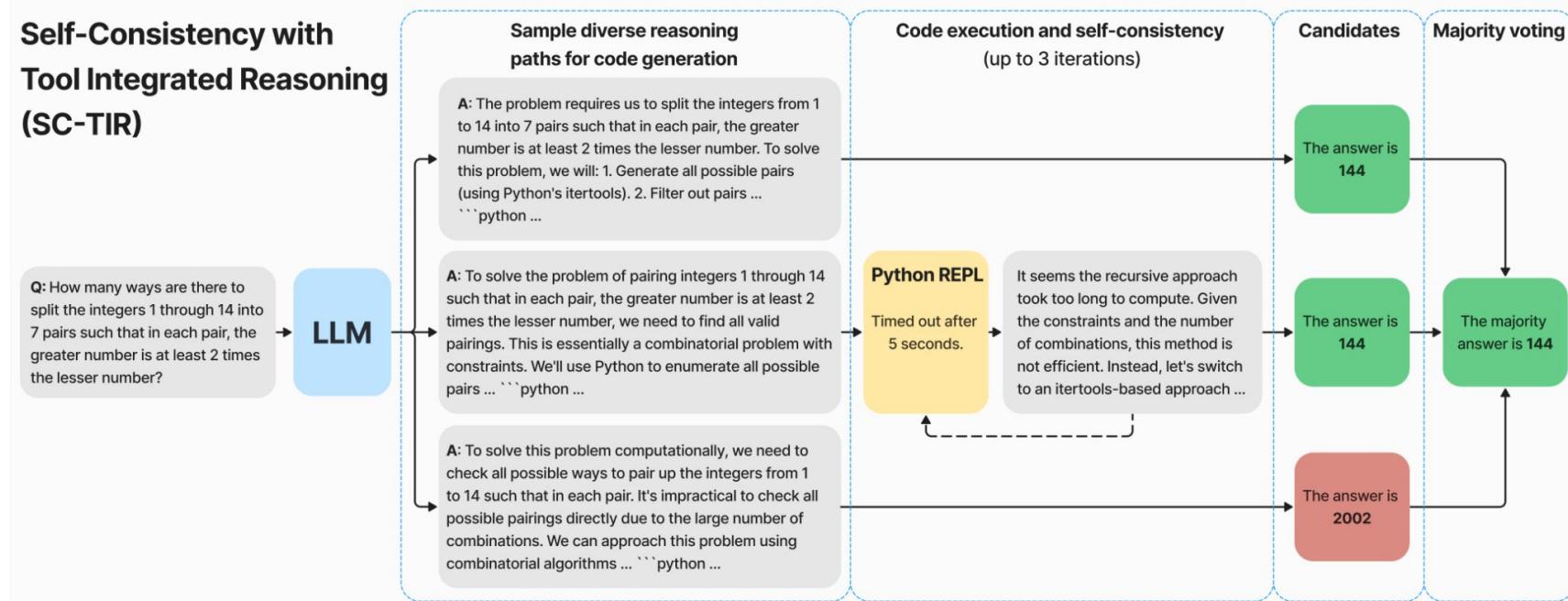
# NuminaMath

- Fleureau et al. (2024):  
структура в основном из MuMath-Code (Yin et al., 2024)
- Сначала дообучаем на математическом датасете на естественном языке, потом отдельно учим работать с инструментами (писать на Python), как ToRA



# NuminaMath

- Fleureau et al. (2024): good data is all you need, потом ещё можно породить много разных chains of thought и выбрать голосованием



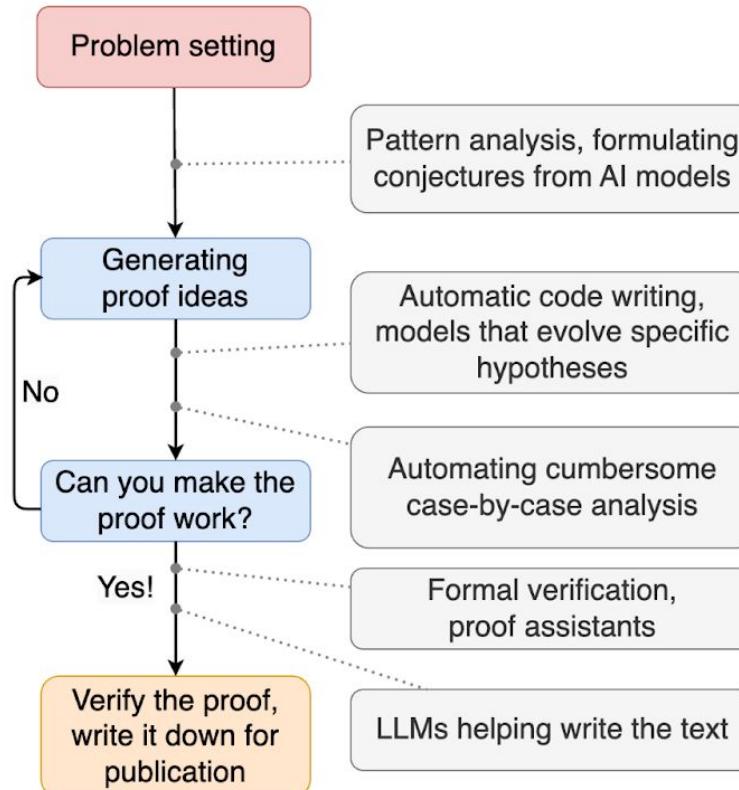
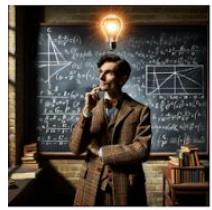
# NuminaMath

- Fleureau et al. (2024): в результате получается маленькая LLM (7B параметров), которая хороша в chain-of-thought и tool-integrated reasoning и умеет решать задачки

Model	MATH (%)
Chain of Thought Reasoning	
GPT-4 (2023)	42.5
GPT-4o	76.6
Claude 3.5 Sonnet	71.1
DeepSeekMath-7B-Instruct	46.8
DeepSeekMath-7B-RL	51.7
NuminaMath-7B-CoT	56.3
Tool-Integrated Reasoning	
DeepSeekMath-7B-Instruct	57.4
DeepSeekMath-7B-RL	58.8
NuminaMath-7B-TIR	68.2



# LLM для математики



## Advancing mathematics by guiding human intuition with AI

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

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



## TORA: A TOOL-INTEGRATED REASONING AGENT FOR MATHEMATICAL PROBLEM SOLVING

Zhibin Geu<sup>1,2\*</sup>, Zhihong Shao<sup>1,2</sup>, Yeyun Gong<sup>2</sup>, Yelong Shen<sup>2</sup>,  
Yujin Yang<sup>1</sup>, Meng Huang<sup>1</sup>, Nan Duan<sup>1</sup>, Weizhe Chen<sup>2</sup>,  
Tongzhuo Wang<sup>1</sup>, Meng Wang<sup>1</sup>, Yusheng Li<sup>1</sup>,  
(zgh322, shao19)@mails.tsinghua.edu.cn  
(yegong, yeshi, nanduan, wzchen)@microsoft.com

## Discovering faster matrix multiplication algorithms with reinforcement learning

Alihussein Fawzi , Matej Balog, Aja Huang, Thomas Hubert, Bernardino Romera-Paredes, Mohammadamin Barezatian, Alexander Novikov, Francisco J. R. Ruiz, Julian Schmittwieser, Grzegorz Swirszcz, David Silve, 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 Barezatian, Alexander Novikov, Matej Balog, M. Pawan Kumar, Emilia Dupont, Francisco J. R. Ruiz, Jordan S. Ellenberg, Pengming Wang, Omar Fawzi, Pushmeet Kohli & Alihussein 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<sup>1</sup>

BY K. APPEL AND W. HAKEN

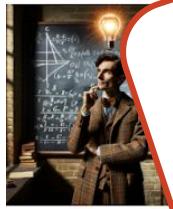
Communicated by Robert Finson, July 26, 1976



# LLM для математики



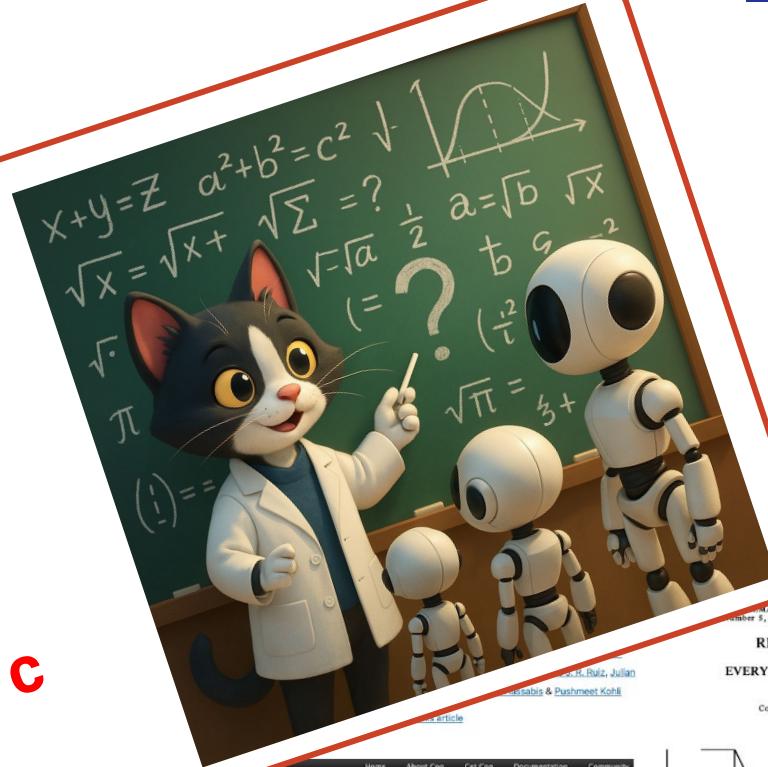
Problem setting



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



LLMs helping write the text



The Coq Proof Assistant

GPT-4

discoveries from program  
language models

Mohammadamin Barezai, Alexander Novikov, Matej  
Dupert, Francisco J. R. Ruiz, Jordan S. Ellenberg,  
Pushmeet Kohli, Alhussein Fawzi

Read this article

MATHEMATICAL SOCIETY

Volume 5, September 1976

RESEARCH ANNOUNCEMENTS

EVERY PLANAR MAP IS FOUR COLORABLE<sup>1</sup>

BY K. APPEL AND W. HAKEN

Communicated by Robert Fortune, July 26, 1976

LEM  
Community

# 7. Что там на фронтире?



Неравная скорость передвижения заставляет нас разделять фронтir на фронтir торговца, фронтir ранчера, фронтir горняка и фронтir фермера. Когда шахты и загоны для скота находились ещё на “линии водопадов”, караваны торговцев переваливали через Аллеганы, а француз на Великих озёрах, предупреждённый берёзовым каноэ британского торговца, укреплял свои посты. Когда трапперы карабкались по Скалистым горам, фермер ещё стоял в устье Миссури.

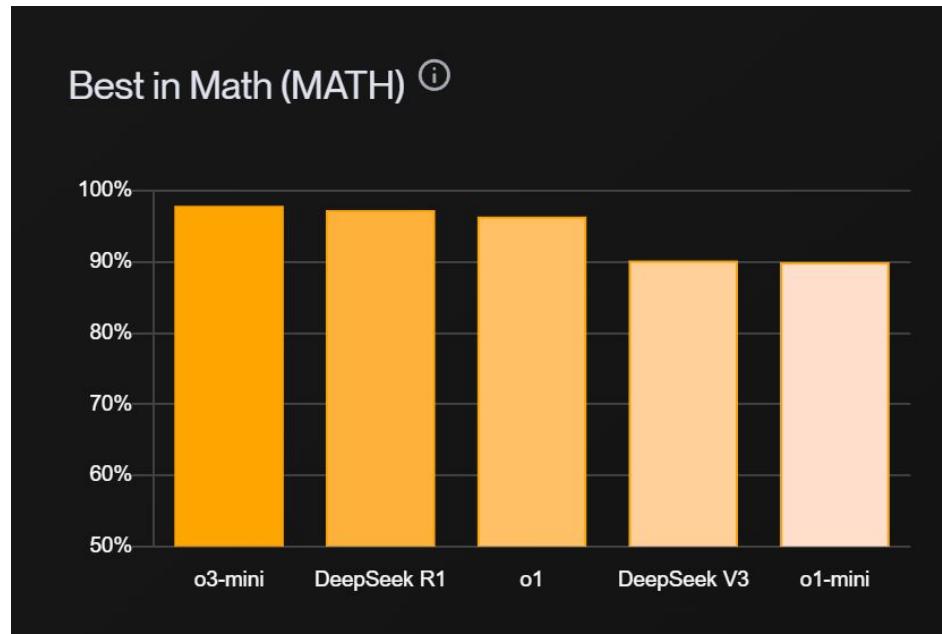
Фредерик Тёрнер  
Значение фронтира в американской истории

# LLM для математики

- Какие-то датасеты уже насыщаются; МАТН – это датасет с примерно такими задачами ([Hendrycks et al., 2021](#)):

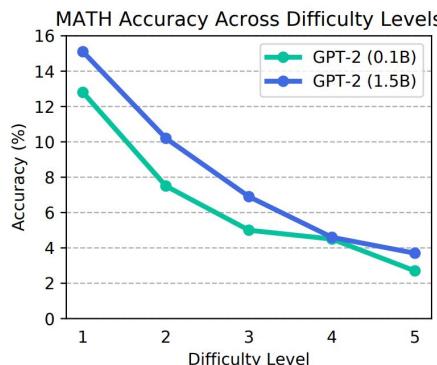
**Problem:** The equation  $x^2 + 2x = i$  has two complex solutions. Determine the product of their real parts.

**Solution:** Complete the square by adding 1 to each side. Then  $(x + 1)^2 = 1 + i = e^{\frac{i\pi}{4}}\sqrt{2}$ , so  $x + 1 = \pm e^{\frac{i\pi}{8}}\sqrt[4]{2}$ . The desired product is then  $(-1 + \cos(\frac{\pi}{8})\sqrt[4]{2})(-1 - \cos(\frac{\pi}{8})\sqrt[4]{2}) = 1 - \cos^2(\frac{\pi}{8})\sqrt{2} = 1 - \frac{(1+\cos(\frac{\pi}{4}))}{2}\sqrt{2} = \boxed{\frac{1 - \sqrt{2}}{2}}$ .



# LLM для математики

- MATH Level 5 – самое сложное подмножество MATH ([Hendrycks et al., 2021](#))
- Всё равно скорее школьные задачки, но в 2021 GPT-2 их совсем не решала...



You have linear functions  $p(x)$  and  $q(x)$ . You know  $p(2)=3$ , and  $p(q(x))=4x+7$  for all  $x$ . Find  $q(-1)$ .

Suppose the roots of the polynomial  $x^2 - mx + n$  are positive prime integers (not necessarily distinct). Given that  $m < 20$ , how many possible values of  $n$  are there?

When Trilisa takes pictures, they turn out with probability  $1/5$ . She wants to take enough pictures so that the probability of at least one turning out is at least  $3/4$ . How few pictures can she take to accomplish this?

# LLM для математики

- Gemini 1.5 Pro ([Gemini Team, 2024](#)) решал подобные задачки ещё лучше, да и олимпиадные
- Но кажется, что тут должны “решать” reasoning models?..

**Problem:**

What is the minimum value of  $5x^2 + 5y^2 - 8xy$  when  $x$  and  $y$  range over all real numbers such that  $|x - 2y| + |y - 2x| = 40$ ?

**Example Math-Specialized Gemini 1.5 Pro output:**

Let  $a = x - 2y$  and  $b = y - 2x$ . Then, we have  $|a| + |b| = 40$ . Also,  $5x^2 + 5y^2 - 8xy = (x - 2y)^2 + (y - 2x)^2 = a^2 + b^2$ . Thus, we want to find the minimum value of  $a^2 + b^2$  given that  $|a| + |b| = 40$ . By the Trivial Inequality, we have  $(|a| - 20)^2 + (|b| - 20)^2 \geq 0$ , so  $a^2 - 40|a| + 400 + b^2 - 40|b| + 400 \geq 0$ , so  $a^2 + b^2 \geq 40(|a| + |b|) - 800 = 40(40) - 800 = 800$ . This minimum is achieved when  $|a| = |b| = 20$ , so the minimum value of  $5x^2 + 5y^2 - 8xy$  is 800.

**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.

**Example Math-Specialized Gemini 1.5 Pro output:**

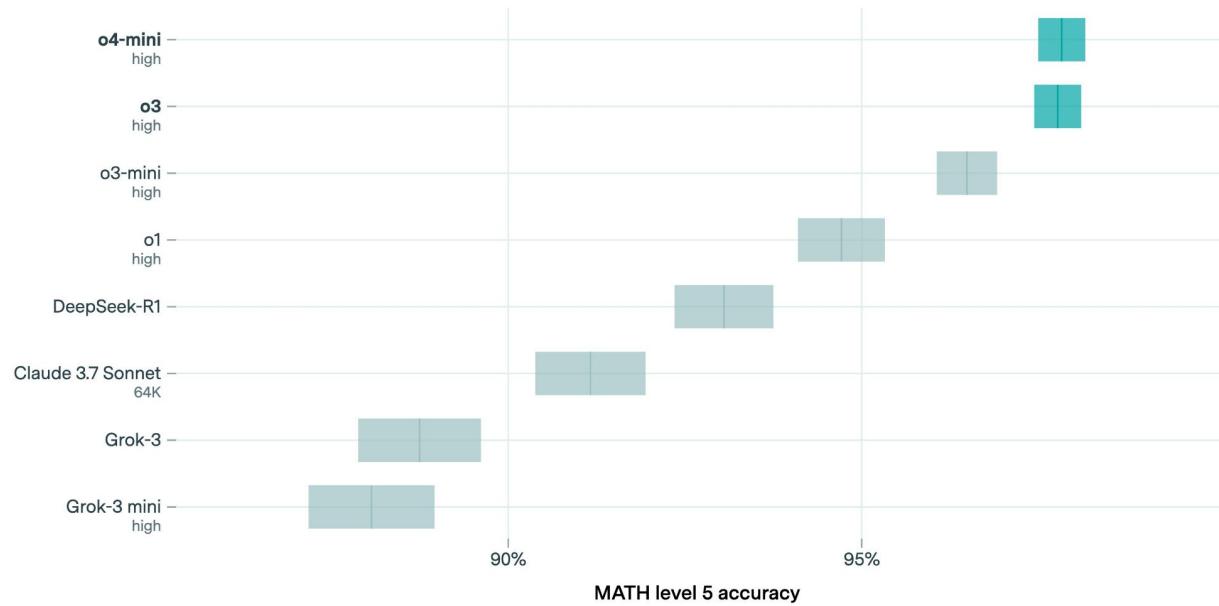
Without loss of generality, assume that  $a \geq b \geq c$ . Then  $a^2 + b + c \leq a^2 + 2a < a^2 + 2a + 1 = (a + 1)^2$ . Since  $a^2 + b + c$  is a perfect square, we must have  $a^2 + b + c = a^2$ , which implies  $b + c = 0$ . This is a contradiction since  $b$  and  $c$  are positive integers. Therefore, 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.

# LLM для математики

- И действительно, текущие рассуждающие модели решают MATH Level 5 на 98% ([Epoch AI, April 19, 2025](#))

MATH Level 5 accuracy for o3 and o4-mini compared to leading models

Shaded regions correspond to the average score  $\pm 1$  standard error. Gemini 2.5 Pro has not been evaluated yet



# FrontierMATH

- А вот FrontierMATH ([Glazer et al., 2024](#)) — это пока совсем другое дело...
- В нём задачи с решениями, но очень непростые и относительно новые

## Find the degree 19 polynomial

[Problem](#) [Solution](#)

Construct a degree 19 polynomial  $p(x) \in \mathbb{C}[x]$  such that  $X := \{p(x) = p(y)\} \subset \mathbb{P}^1 \times \mathbb{P}^1$  has at least 3 (but not all linear) irreducible components over  $\mathbb{C}$ . Choose  $p(x)$  to be odd, monic, have real coefficients and linear coefficient -19 and calculate  $p(19)$ .

## Testing Artin's primitive root conjecture

[Problem](#) [Solution](#)

For a positive integer  $n$ , let  $v_p(n)$  denote the largest integer  $v$  such that  $p^v \mid n$ . For a prime  $p$  and  $a \not\equiv 0 \pmod{p}$ , let  $\text{ord}_p(a)$  denote the smallest positive integer  $o$  such that  $a^o \equiv 1 \pmod{p}$ . For  $x > 0$ , let

$$\text{ord}_{p,x}(a) = \prod_{\substack{q \leq x \\ q \text{ prime}}} q^{v_q(\text{ord}_p(a))} \prod_{\substack{q > x \\ q \text{ prime}}} q^{v_q(p-1)}.$$

Let  $S_x$  denote the set of primes  $p$  for which

$$\text{ord}_{p,x}(2) > \text{ord}_{p,x}(3),$$

and let  $d_x$  denote the density

$$d_x = \lim_{t \rightarrow \infty} \frac{|p \leq t : p \in S_x|}{|p \leq x : p \text{ prime}|}$$

of  $S_x$  in the primes. Let

$$d_\infty = \lim_{x \rightarrow \infty} d_x.$$

Compute  $\lfloor 10^6 d_\infty \rfloor$ .

# FrontierMATH

- А вот FrontierMATH ([Glazer et al., 2024](#)) — это пока совсем другое дело...
- В нём задачи с решениями, но очень непростые и относительно новые

## Find the degree 19 polynomial

[Problem](#) [Solution](#)

Construct a degree 19 polynomial  $p(x) \in \mathbb{C}[x]$  such that  $X := \{p(x) = p(y)\} \subset \mathbb{P}^1 \times \mathbb{P}^1$  has at least 3 (but not all linear) irreducible components over  $\mathbb{C}$ . Choose  $p(x)$  to be odd, monic, have real coefficients and linear coefficient -19 and calculate  $p(19)$ .

[Problem](#) [Solution](#)

**Answer:** 1876572071974094803391179.

**Solution:** Note that  $p(x)$  defines a covering  $p : \mathbb{P}_x^1 \rightarrow \mathbb{P}_z^1$  of degree 19 which at  $z = \infty$  has monodromy  $\sigma_\infty$ , a 19-cycle, which without loss of generality is  $(1, 2, 3, \dots, 19)$ .

Note that  $X$  is just the fibre product of  $p$  with  $p$  (which in general is a singular curve) and therefore  $p(x)$  also defines a finite map  $\hat{p} : X \rightarrow \mathbb{P}_z^1$  of degree 19<sup>2</sup> (if we prefer to stick with smooth schemes, we can instead use the normalisation  $\hat{X}$  of  $X$ , and we will freely move between the irreducible components of  $X$  and the corresponding connected components of the smooth  $\hat{X}$  with its covering structure).

The idea then is that any such covering  $\hat{p}$  can be studied using group theory via its monodromy. Specifically, suppose  $q : X \rightarrow \mathbb{P}_z^1$  is of degree  $n$  and has branch locus  $B \subset \mathbb{P}_z^1$ . Then if you follow the fibres around each branch value and see how this permutes the points in the fibres, we get a homomorphism

$\bar{q} : \pi_1(\mathbb{P}_z^1 - B, x) \rightarrow \text{Im}(\phi) := G_q \subset S_n$  (defined up to conjugation) and  $X$  is irreducible exactly when  $\bar{q}$  is transitive. Moreover, if  $X$  is reducible, then the irreducible components correspond to the  $G_q$ -orbits in  $S_n$ .

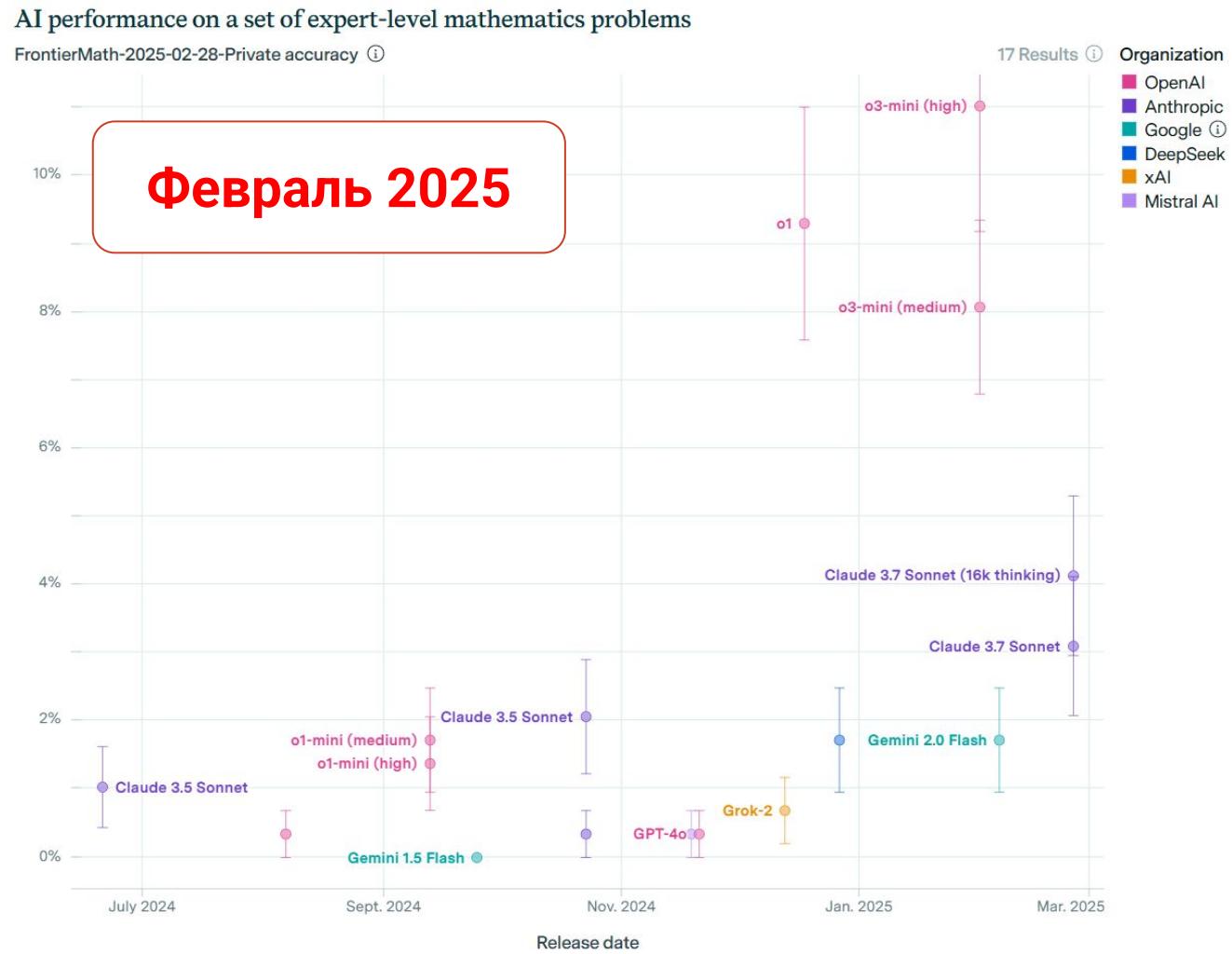
Suppose  $p$  with branch locus  $B$  (including  $\infty$ ) is encoded as  $\bar{p} : G_p \subset S_{19}$ . Then we can check that  $\hat{p}$  has branch locus  $B$  and is encoded diagonally as  $\bar{p}^2 : G_p \subset S_{19}^2$ . So we are looking for some  $G_p \subset S_{19}$  that has between 3 and 18 orbits in  $S_{19}^2$ . In particular  $G_p$  is not doubly transitive on  $S_{19}$ .

...и **ещё** столько же дальше

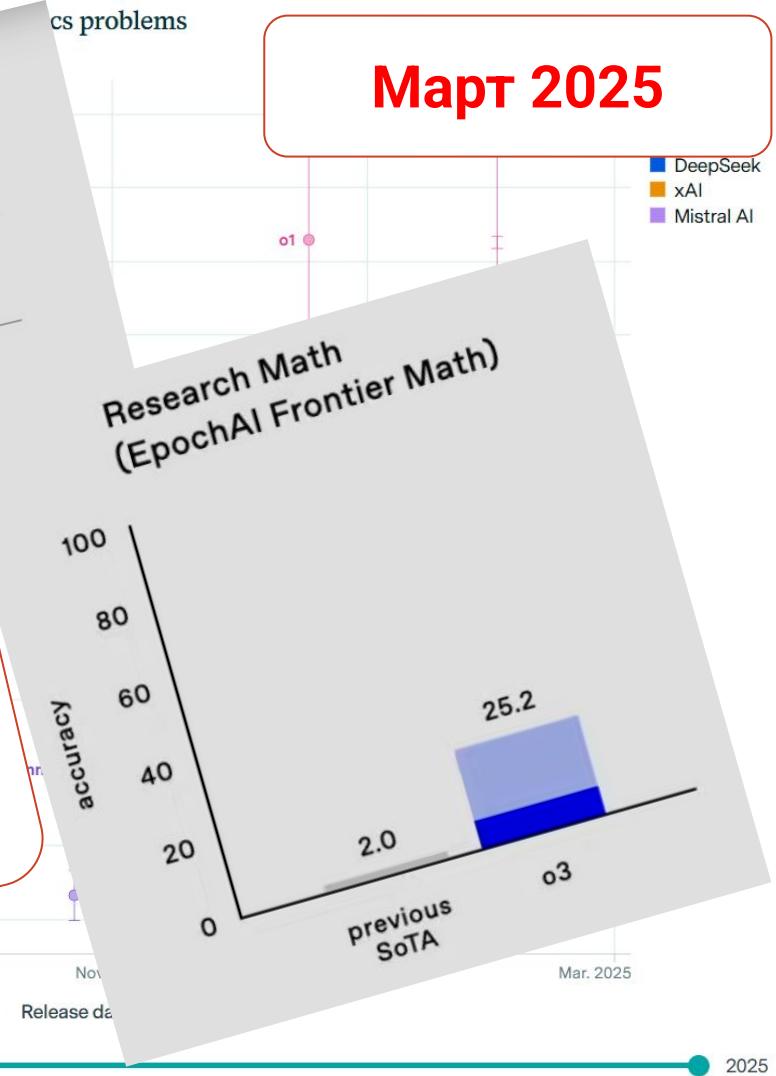
Recall a theorem of Burnside (Burnside 1911): for  $q$  prime and a permutation group  $G \subset S_q$ , either  $G$  is doubly transitive or  $|G|$  divides  $q(q-1)$ .

# FrontierMATH

- А вот FrontierMATH ([Glazer et al., 2024](#)) – это пока совсем другое дело...
- И здесь результаты LLM пока невысоки – но они не равны нулю! Это довольно круто
- Впрочем...



# FrontierMATH



# AI performance on a set of expert-level math and physics problems

FrontierMath-2025-02-28-Private accuracy

17 Results ⓘ Organization

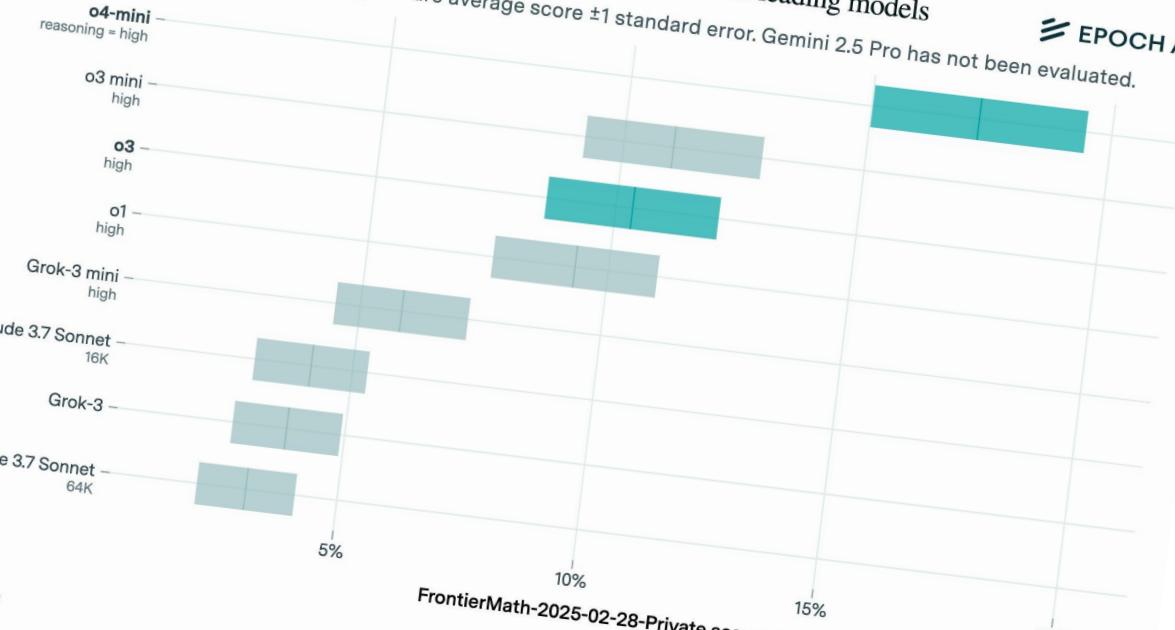
- OpenAI
- Anthropic
- Google ⓘ
- DeepSeek
- xAI
- Mistral AI

FrontierMath accuracy for OpenAI's o3 and o4-mini compared to leading models

FrontierMath private set. Shaded regions are average score  $\pm 1$  standard error. Gemini 2.5 Pro has not been evaluated.



epoch.ai



CC-BY

довольно н

ясно,

Впрочем...

Sept. 2024

Release da

epoch.ai

Mar. 2025

2024

2025

April 19,  
2025: a

нет, не  
25%, а

всего

17-18, и не

о3, а

о4-mini...

nothing to  
see here,  
right?

# Proof or Bluff?

- Есть и отрицательные результаты. Proof or Bluff ([Petrov et al., 2025](#)): как LLM справляются с 2025 USA Math Olympiad? Просили писать полное доказательство и проверяли как людей
- Пока результаты почти нулевые, максимум 2 балла из 42 у DeepSeek R1
- Это чистые LLM, не AlphaProof

**Problem 1:** Let  $k$  and  $d$  be positive integers. Prove that there exists a positive integer  $N$  such that for every odd integer  $n > N$ , the digits in the base- $2n$  representation of  $n^k$  are all greater than  $d$ .

**Problem 5:** Determine, with proof, all positive integers  $k$  such that

$$\frac{1}{n+1} \sum_{i=0}^n \binom{n}{i}^k$$

is an integer for all positive integers  $n$ .

Model	P1	P2	P3	P4	P5	P6	Total	Cost
R1	0.5	0.0	0.0	1.5	0.0	0.0	2.0	2.03
FLASH-THINKING	1.5	0.0	0.0	0.0	0.2	0.0	1.8	N/A
CLAUDE 3.7	0.5	0.5	0.0	0.0	0.0	0.6	1.5	9.03
QwQ	1.2	0.0	0.0	0.0	0.0	0.0	1.2	0.42
O1-PRO	0.5	0.0	0.0	0.0	0.2	0.4	1.2	203.44
O3-MINI	0.5	0.1	0.0	0.0	0.0	0.2	0.9	1.11

# Proof or Bluff?

- Но нет, подождите, это же был [Petrov et al., March 27, 2025](#)! За следующие две недели, конечно же, появились новые модели!
- [Petrov et al., April 9, 2025](#): Gemini 2.5 Pro решает уже одну задачу совсем и другую наполовину, 10 баллов из 42
- Эх, не удался отрицательный пример...

**Problem 1:** Let  $k$  and  $d$  be positive integers. Prove that there exists a positive integer  $N$  such that for every odd integer  $n > N$ , the digits in the base- $2n$  representation of  $n^k$  are all greater than  $d$ .

**Problem 5:** Determine, with proof, all positive integers  $k$  such that

$$\frac{1}{n+1} \sum_{i=0}^n \binom{n}{i}^k$$

is an integer for all positive integers  $n$ .

Model	P1 (/7)	P2 (/7)	P3 (/7)	P4 (/7)	P5 (/7)	P6 (/7)	Total (/42)	Cost
GEMINI-2.5-PRO	6.5	0.0	0.1	3.5	0.0	0.0	10.1	N/A
R1	0.5	0.0	0.0	1.5	0.0	0.0	2.0	2.03
GROK 3	2.0	0.0	0.0	0.0	0.0	0.0	2.0	N/A
FLASH-THINKING	1.5	0.0	0.0	0.0	0.2	0.0	1.8	N/A
CLAUDE 3.7	0.5	0.5	0.0	0.0	0.0	0.6	1.5	9.03
QwQ	1.2	0.0	0.0	0.0	0.0	0.0	1.2	0.42
O1-PRO	0.5	0.0	0.0	0.0	0.2	0.4	1.2	203.44
O3-MINI	0.5	0.1	0.0	0.0	0.0	0.2	0.9	1.11

## Proof or Bluff?

- Но нет, подождите, это же был [Petrov et al., April 2025: Gemini](#) (все эти две недели, конечно же, появлялись)
- [Petrov et al., April 2025: Gemini](#) (все эти две недели, конечно же, появлялись)

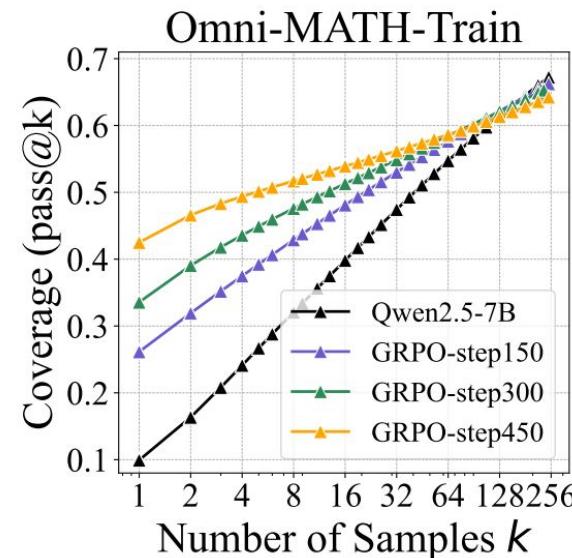
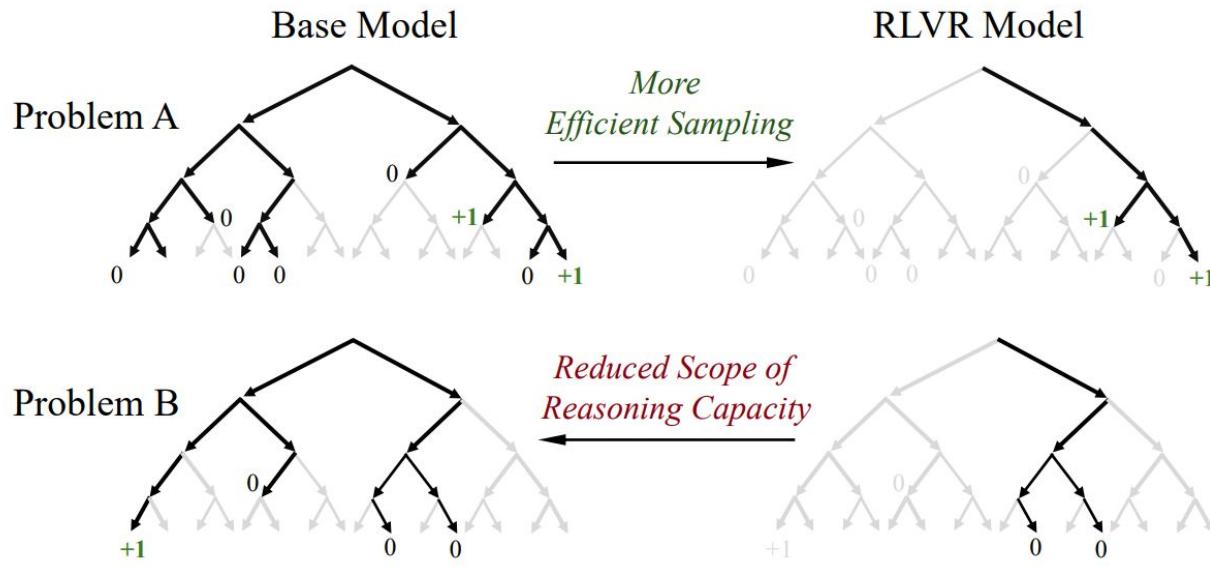
На IMO 2025 уже и новая модель OpenAI (ещё не выпущенная), и новая Gemini решили по 6 из 7 задач (это золото) просто как LLM, end to end



	P6 (/7)	Total (/42)	Cost
01-PRO	0.0	0.0	N/A
02-PRO	0.0	0.0	2.0
03-MINI	0.0	0.0	N/A
	0.5	0.5	2.03
	1.2	1.2	N/A
	0.5	0.5	9.03
	0.0	0.0	0.42
	0.5	0.5	203.44
	0.5	0.5	1.11
	0.0	0.0	
	0.1	0.1	

# Does RL Really Incentivize Reasoning Capacity?

- Yue et al. (Apr 15, 2025): а вообще помогает этот ваш RL? Утверждают, что RLVR (RL with Verifiable Rewards) помогает сделать сэмплирование лучше... но и беднее, и если просто сэмплировать побольше, базовые модели начинают даже выигрывать



# Humanity's Last Exam

- Humanity's Last Exam ([Phan et al., 2025](#)) – очень сложный датасет с разнообразными задачами (я туда подавал сложные вопросы ЧГК, но они оказались слишком простыми)
- Вопросы не обязательно на фронтире науки, но они точно не решались LLM, существовавшими в конце 2024 года

## Classics

Question:



Here is a representation of a Roman inscription, originally found on a tombstone. Provide a translation for the Palmyrene script. A transliteration of the text is provided: RGYNBT HRY BR CT HBL

## Ecology

Question:

Hummingbirds within Apodiformes uniquely have a bilaterally paired oval bone, a sesamoid embedded in the caudolateral portion of the expanded, cruciate aponeurosis of insertion of m. depressor caudae. How many paired tendons are supported by this sesamoid bone? Answer with a number.

## Mathematics

Question:

The set of natural transformations between two functors  $F, G: \mathcal{C} \rightarrow \mathcal{D}$  can be expressed as the end

$$\text{Nat}(F, G) \cong \int_A \text{Hom}_{\mathcal{D}}(F(A), G(A)).$$

Define set of natural cotransformations from  $F$  to  $G$  to be the coend

$$\text{CoNat}(F, G) \cong \int^A \text{Hom}_{\mathcal{D}}(F(A), G(A)).$$

Let:

-  $F = \mathbf{B}_\bullet(\Sigma_4)_{*/}$  be the under  $\infty$ -category of the nerve of the delooping of the symmetric group  $\Sigma_4$  on 4 letters under the unique 0-simplex  $*$  of  $\mathbf{B}_\bullet\Sigma_4$ .

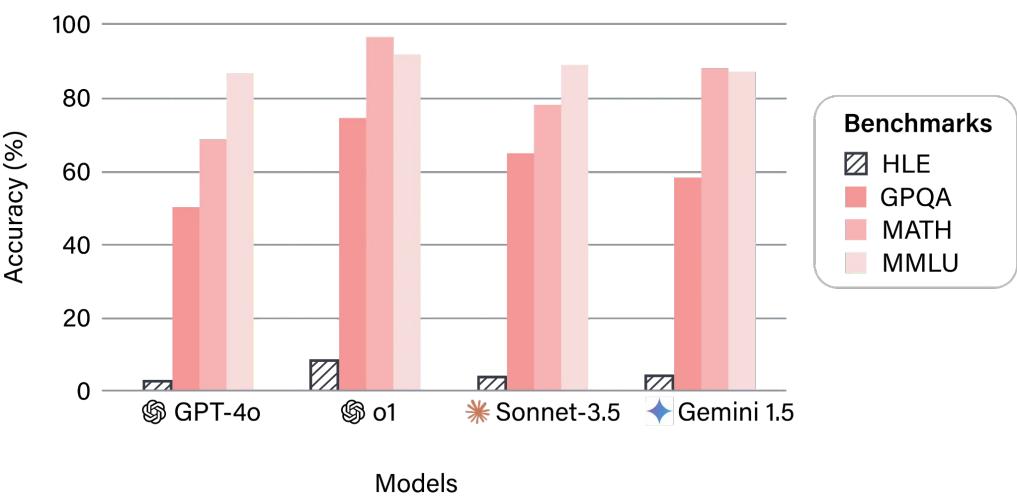
-  $G = \mathbf{B}_\bullet(\Sigma_7)_{*/}$  be the under  $\infty$ -category nerve of the delooping of the symmetric group  $\Sigma_7$  on 7 letters under the unique 0-simplex  $*$  of  $\mathbf{B}_\bullet\Sigma_7$ .

How many natural cotransformations are there between  $F$  and  $G$ ?

# Humanity's Last Exam

- Humanity's Last Exam ([Phan et al., 2025](#)) – очень сложный датасет
- И действительно, LLM справляются с НИМ ПЛОХО...

Accuracy of LLMs Across Benchmarks



Judge Model: o3-mini-2025-01-31 | Last Updated: 02/11/2025

Model	Accuracy (%) ↑	Calibration Error (%) ↓
GPT-4o	3.1	92.3
Grok-2	3.9	90.8
Claude 3.5 Sonnet	4.8	88.5
GPT-4.5 Preview	6.4	90.5
Gemini Thinking	7.2	90.6
DeepSeek-R1*	8.6	81.4
o1	8.8	92.8
Claude 3.7 Sonnet (16K)	8.9	88.3
o3-mini (medium)*	11.1	91.5
o3-mini (high)*	14.0	92.8

\*Model is not multi-modal, evaluated on text-only subset.

# Humanity's Last Exam

- Humanity's Last Exam ([Phan et al., 2025](#)) – очень сложный датасет
- И действительно, LLM справляются с НИМ ПЛОХО...
- ...НО, кажется, начинают справляться лучше
- А что это за deep research такой?

Model	Accuracy (%)
GPT-4o	3.3
Grok-2	3.8
Claude 3.5 Sonnet	4.3
Gemini Thinking	6.2
OpenAI o1	9.1
DeepSeek-R1*	9.4
OpenAI o3-mini (medium)*	10.5
OpenAI o3-mini (high)*	13.0
<b>OpenAI deep research**</b>	<b>26.6</b>

\* Model is not multi-modal, evaluated on text-only subset.

\*\*with browsing + python tools

# OpenAI Deep Research

- Deep Research – LLM-based агент, который может полчаса гуглить и компилировать много разных источников в один полновесный отчёт, часто на десятках страниц, со ссылками на источники
- Пример

February 2, 2025 Release

## Introducing deep research

An agent that uses reasoning to synthesize large amounts of online information and complete multi-step research tasks for you. Available to Pro users today, Plus and Team next.

how are hidden markov models used now? please give an overview of current applications with references to recent papers (2024-2025); I know their rich history but I'm not sure HMM's are still relevant

I'll gather an overview of the current applications of Hidden Markov Models (HMMs) in 2024-2025, focusing on recent papers. Could you specify any particular domains you're most interested in, such as natural language processing, bioinformatics, finance, or cybersecurity? Or would you like a broad overview across multiple fields?

broad overview please

I'll research the current applications of Hidden Markov Models (HMMs) in 2024-2025 across multiple fields, including natural language processing, bioinformatics, finance, cybersecurity, and more. I'll focus on recent papers and references to understand how HMMs are being used today and whether they remain relevant in modern AI and statistical modeling. I'll get back to you with my findings soon.

Research completed in 7m · 36 sources

# OpenAI Deep Research

- Deep Research – LLM-based агент, который может ПОЛЧАСА ГУГЛИТЬ И КОМПИЛИРОВАТЬ МНОГО РАЗНЫХ ИСТОЧНИКОВ В ОДИН ПОЛНОВЕСНЫЙ отчёт, часто на десятках страниц, со ссылками на источники
- Пример

February 2, 2025 Release

## Introducing deep research

An agent that uses reasoning to synthesize large amounts of online information and complete multi-step research tasks for you. Available to Pro users today, Plus and Team next.

## Hidden Markov Models in 2024-2025: Applications and Relevance

Hidden Markov Models (HMMs) remain a versatile tool for sequential data modeling, even as deep learning dominates many tasks. Below we survey current applications of HMMs across multiple fields and assess their relevance relative to neural networks and other modern models. Recent studies from 2024–2025 highlight that HMMs are *still actively used* – often in hybrid systems or niche problems – due to their probabilistic interpretability and robustness. We provide examples in each domain with key findings from the latest research.

### NLP Applications of HMMs

**Speech and Language Processing:** HMMs were foundational in speech recognition and tagging, and they continue to appear in specialized NLP tasks. While end-to-end deep learning has largely superseded HMMs for mainstream speech recognition, HMM-based methods are still explored in areas like speech emotion recognition. For example, a 2024 study combined Gaussian Mixture Models (GMMs) and HMMs for speech emotion recognition, leveraging HMMs to model spectral feature sequences [ARXIV.ORG](#). This approach achieved effective emotion classification from audio, indicating that

...далее идут 12 страниц шрифтом в 11pt...

# OpenAI Deep Research

- Deep Research – LLM-based агент, который может полчаса гуглить и компилировать много разных источников в один полновесный отчёт, часто на десятках страниц, со ссылками на источники
- Пример

February 2, 2025 Release

## Introducing deep research

An agent that uses reasoning to synthesize large amounts of online information and complete multi-step research tasks for you. Available to Pro users today, Plus and Team next.

- In many real-world systems, an HMM is used under the hood for tracking or alignment, even if a neural network does the heavy prediction. This combination leverages the **strength of HMMs in sequence alignment/segmentation with the representational power of deep networks**, as evidenced by multiple 2024 hybrid models [ARXIV.ORG](#) [ARXIV.ORG](#).

In summary, Hidden Markov Models continue to be a vital part of the machine learning landscape in 2024–2025. They have certainly been overshadowed by deep learning in some arenas, but they retain unique advantages that ensure their ongoing use. HMMs today are often found working **in tandem with deep learning**, each compensating for the other's weaknesses. As long as problems involve sequential data with hidden structure – and require solutions that are transparent and data-efficient – HMMs will remain relevant. The latest research not only uses HMMs in creative ways across fields, but also actively improves them, affirming that HMMs are **evolving, not fading**, in the era of deep learning.

### References:

- Al-Dujaili Al-Khazraji, M. J., & Ebrahimi-Moghadam, A. (2024). *An Innovative Method for Speech Signal Emotion Recognition Based on Spectral Features Using GMM and HMM Techniques*. *Wireless Personal Communications*, 134(2), 735–753 [ARXIV.ORG](#).

...и ещё 13 ссылок на свежие статьи

# OpenAI Deep Research

- Deep Research – LLM-based агент, который может полчаса гуглить и компилировать много разных источников в один полновесный отчёт, часто на десятках страниц, со ссылками на источни
- Пример

- In many real-world systems, a neural network does the sequence alignment, and evidence

ment, even if a of HMMs in as

ing landscape in arenas, but they retain often found working in tandem with As long as problems involve sequential are transparent and data-efficient – HMMs will uses HMMs in creative ways across fields, but also actively are evolving, not fading, in the era of deep learning.

То есть уже сейчас можно писать обзоры на любую тему автоматически? ..

February 2, 2025 Release

## Introducing deep research

An agent that uses reasoning to synthesize large amounts of online information and complete multi-step research tasks for you. Available to Pro users today, Plus and Team next.

Al-Khazraji, M. J., & Ebrahimi-Moghadam, A. (2024). An Innovative Method for Speech Signal Emotion Recognition Based on Spectral Features Using GMM and HMM Techniques. *Wireless Personal Communications*, 134(2), 735–753.  

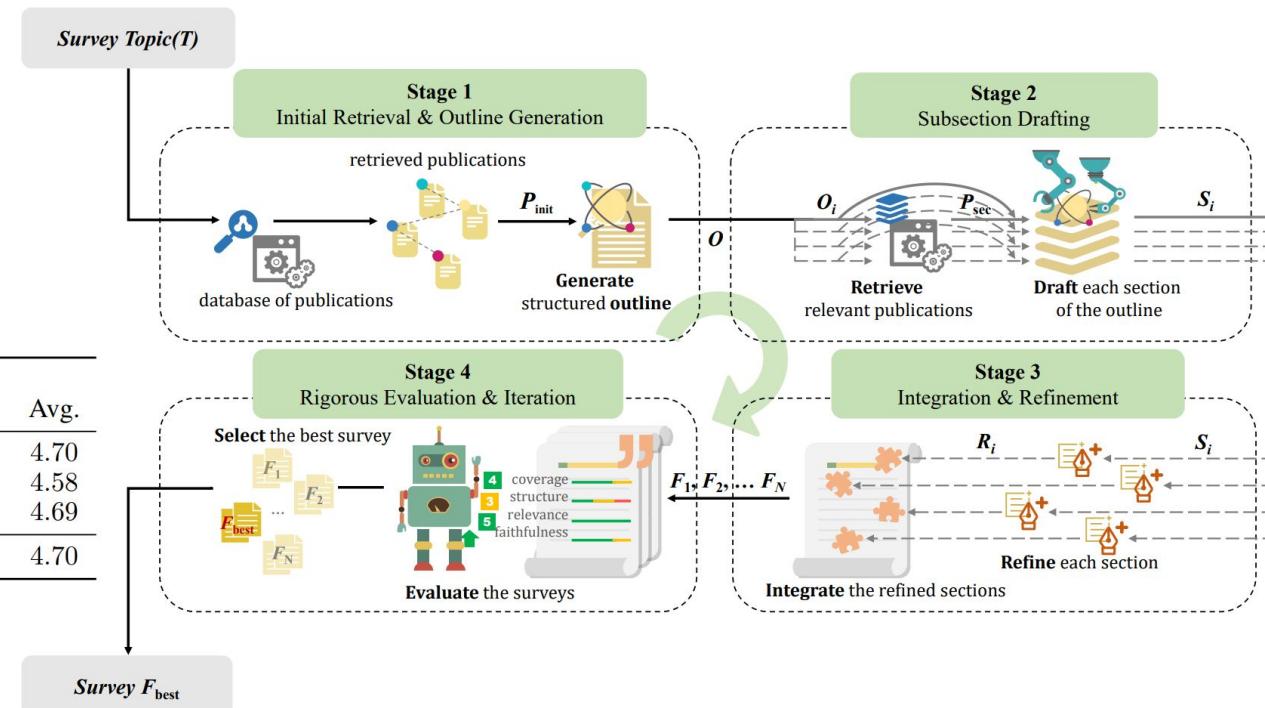
...и ещё 13 ссылок на свежие статьи

# AutoSurvey

- Люди давно пробуют – AutoSurvey ([Wang et al., June 2024](#)) уже был неплох

Base LLM writer	Citation Quality	
	Recall	Precision
GPT-4	$80.25 \pm 4.19$	$78.83 \pm 7.00$
Claude-haiku	$82.45 \pm 2.77$	$76.31 \pm 2.18$
Gemini-1.5-pro	$78.13 \pm 2.39$	$71.24 \pm 3.28$
Human	85.86	80.51

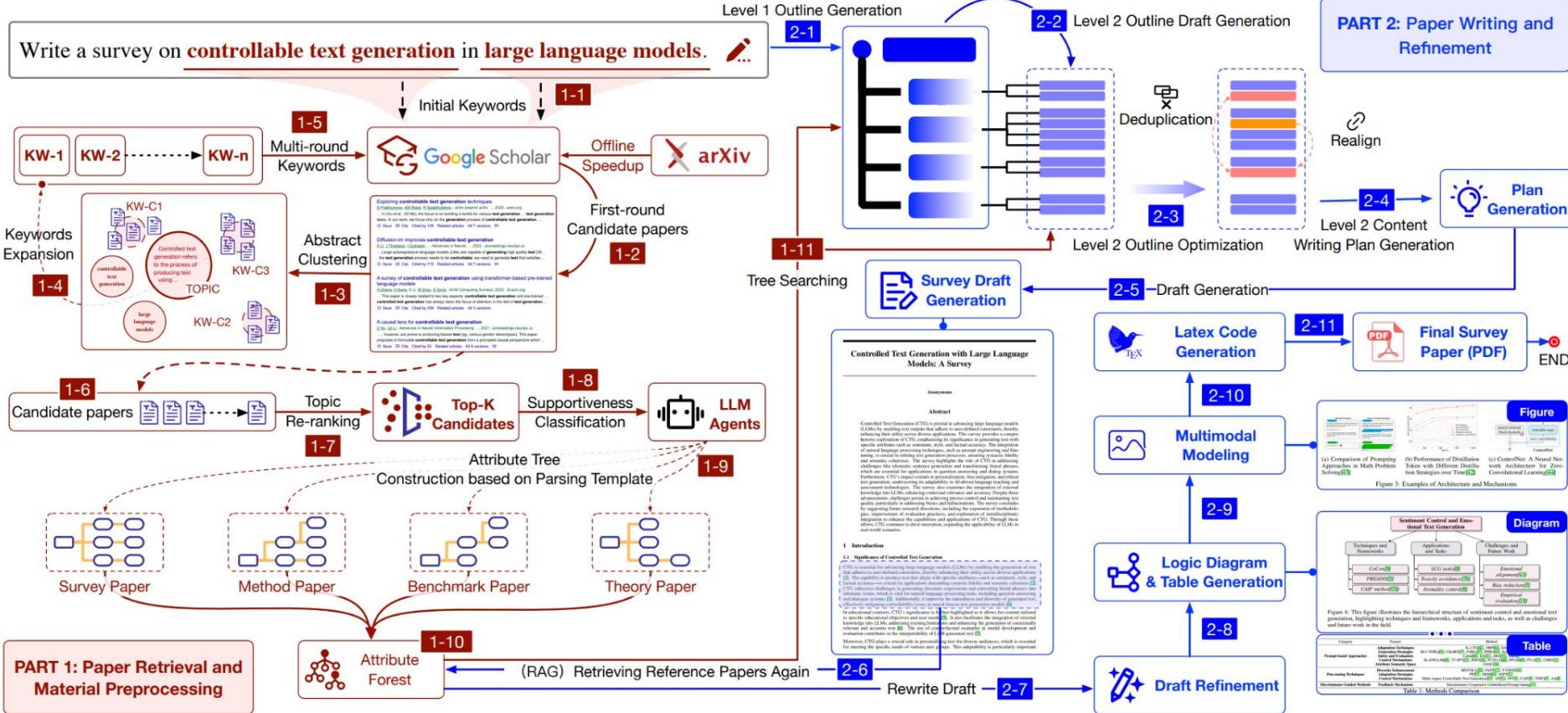
Base LLM writer	Content Quality			
	Coverage	Structure	Relevance	Avg.
GPT-4	$4.8 \pm 0.54$	$4.46 \pm 0.49$	$4.86 \pm 0.33$	4.70
Claude-haiku	$4.66 \pm 0.47$	$4.26 \pm 0.67$	$4.86 \pm 0.33$	4.58
Gemini-1.5-pro	$4.86 \pm 0.33$	$4.33 \pm 0.78$	$4.93 \pm 0.25$	4.69
Human	4.71	4.43	5	4.70



# SurveyX

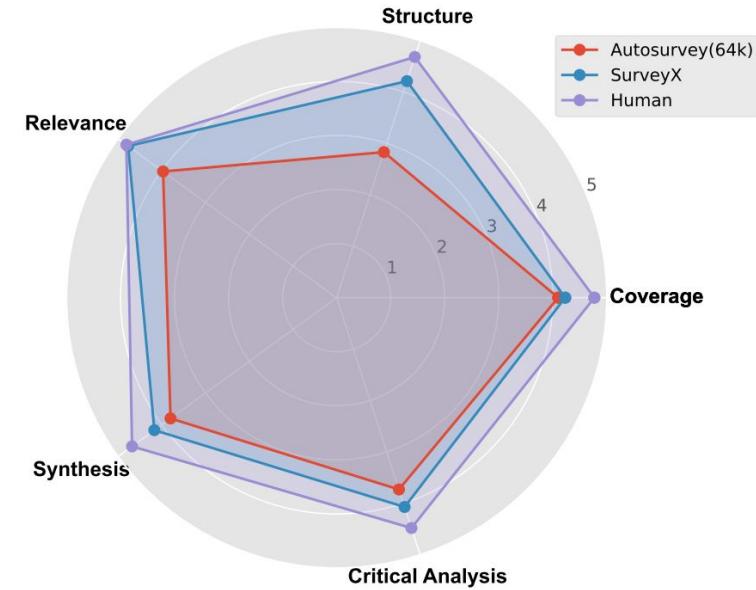
- SurveyX (Liang et al., Feb 20, 2025) ещё лучше

Write a survey on **controllable text generation** in **large language models**.



# SurveyX

- SurveyX ([Liang et al., Feb 20, 2025](#)) подбирается к человеческому уровню даже на базе GPT-4o
- А что будет с o1-pro, GPT-4.5, o3 и o4, а ещё что будет, когда GPT-4.5 совместится с reasoning?..



Model	Coverage	Structure	Relevance	Synthesis	Critical Analysis	Avg	Recall	Precision	F1
naive RAG	4.40	3.66	4.66	3.82	2.82	3.872	68.79	61.97	65.20
AutoSurvey	4.73	4.33	4.86	4.00	3.73	4.331	82.25	77.41	79.76
<b>SURVEYX</b>	<b>4.95</b>	<b>4.91</b>	<b>4.94</b>	<b>4.10</b>	<b>4.05</b>	<b>4.590</b>	<b>85.23</b>	<b>78.12</b>	<b>81.52</b>
Human	5.00	4.95	5.00	4.44	4.38	4.754	86.33	77.78	81.83

# SurveyX

- Примеры SurveyX действительно хороши (хоть и со странными багами)

## Transformer Models in Natural Language Processing: A Survey

[www.surveyx.cn](http://www.surveyx.cn)

### Abstract

Transformer models have revolutionized natural language processing (NLP) by introducing a robust architecture centered around self-attention mechanisms, enabling efficient handling of complex syntactic and semantic tasks. This survey paper provides a comprehensive analysis of transformer models, emphasizing their transformative impact across diverse NLP applications, including translation, summarization, sentiment analysis, and question answering. The paper highlights the advantages of transformers over traditional models like recurrent and convolutional neural networks, particularly in managing long-range dependencies and contextual relationships. Despite their successes, transformer models face challenges such as computational complexity, data dependency, and adaptability to new languages and tasks. The survey explores advancements in model efficiency, such as optimized architectures and quantum computing principles, and discusses ethical considerations, including bias mitigation and interdisciplinary applications. Additionally, the integration of transformers with other AI technologies, such as generative diffusion models and graph neural networks, is examined, showcasing their potential for enhancing multimodal and domain-specific tasks. The paper concludes by emphasizing the ongoing evolution of transformer models, focusing on improving model robustness, interpretability, and ethical deployment in sensitive domains.

## A Survey of Fundamental Concepts in Computational Complexity Theory

[www.surveyx.cn](http://www.surveyx.cn)

### Abstract

This survey paper provides a comprehensive exploration of fundamental concepts in computational complexity theory, focusing on key topics such as P vs NP, NP-completeness, the polynomial hierarchy, space complexity, oracle separation, the Cook-Levin theorem, time complexity, circuit complexity, and relativization. These concepts form a structured framework for understanding computational limits and potential breakthroughs in algorithmic design. The paper highlights the significance of the P vs NP problem, underscoring its implications for cryptography, optimization, and algorithm design. The study of NP-completeness and the Cook-Levin theorem emphasizes the classification of inherently difficult problems. The polynomial hierarchy extends the P vs NP question into multiple levels, offering insights into the complexity of decision problems. Space complexity is examined in both classical and quantum contexts, revealing the potential for quantum algorithms to achieve significant space savings. Oracle separation and relativization are explored as tools for understanding complexity class separations, although their limitations necessitate the development of non-relativizing techniques. Circuit complexity provides insights into resource measurement, with innovative approaches such as thermodynamic perspectives and hybrid circuits offering new methodologies for analyzing computational efficiency. The survey concludes by emphasizing the interconnectedness of these concepts and the ongoing challenges in computational complexity theory, highlighting the need for continued exploration of new methodologies and theoretical frameworks to advance our understanding of computational limits and algorithmic design.



# SurveyX

- Примеры SurveyX действительно хороши (хотя и с некоторыми багами)

Transformer Models in Natural Language Processing: A Survey  
A Survey

[www.surveyx.cn](http://www.surveyx.cn)

## Abstract

Transformer models have revolutionized natural language processing (NLP) by introducing a robust architecture centered around attention mechanisms. This survey paper provides a comprehensive analysis of transformer models, highlighting their transformative impact across diverse NLP applications, including text generation, summarization, sentiment analysis, and question answering. The paper compares the advantages of transformers over traditional models like recurrent neural networks, particularly in managing long-range dependencies and capturing semantic relationships. Despite their successes, transformer models face challenges related to computational complexity, data dependency, and adaptability to specific tasks. The survey explores advancements in model efficiency, including the use of masked language models, architectures and quantum computing principles, and discusses various applications of transformers, including bias mitigation and interdisciplinary applications. The integration of transformers with other AI technologies, such as generative models and graph neural networks, is examined, showcasing the potential for transformer models to enhance multimodal and domain-specific tasks. The paper concludes by emphasizing the ongoing evolution of transformer models, focusing on improving model robustness, interpretability, and ethical deployment in sensitive domains.

Ну хорошо, мы знаем, что LLM умеют читать и суммаризировать. А что насчёт original research?



## 8. AI и новые результаты

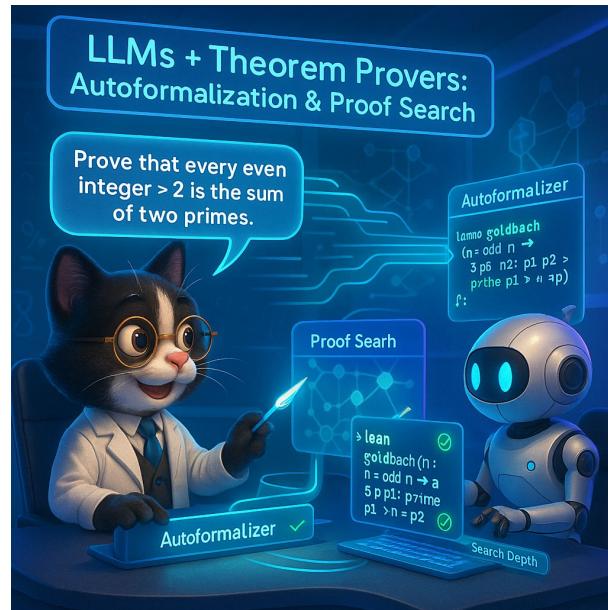


Революции, происходящие на поверхности жизни, ничего существенного никогда не меняют и не открывают, они лишь... по-новому переставляют все те же элементы и являются старые образы в новых одеяниях. Революция всегда есть в значительной степени маскарад, и если сорвать маски, то можно встретить старые, знакомые лица. Новые души рождаются позже, после глубокого перерождения и осмысливания опыта революции.

*Николай Бердяев. Духи русской революции*

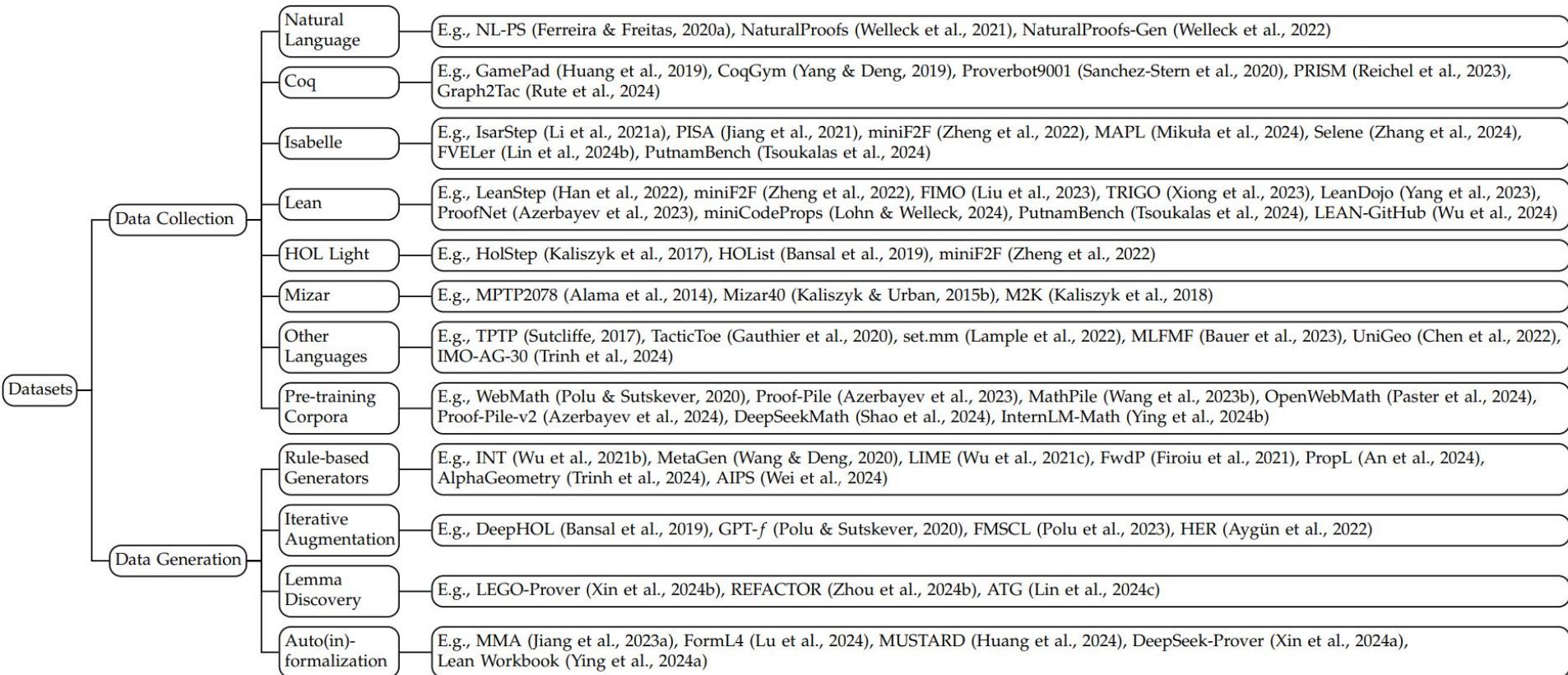
# Как там пруверы?

- Во-первых, вернёмся к пруверам: можно ли всё-таки просить LLM порождать проверяемые доказательства, или код, их порождающий, или что-то ещё?..
- Li et al. (2024): A Survey on Deep Learning for Theorem Proving
  - **autoformalization**: переводим текстовое доказательство в код для proof assistant
  - **premise selection**: какие леммы могут оказаться полезными?
  - **proofstep generation**: предсказываем один шаг доказательства
  - **proof search**: полноценно ищем доказательство



# Как там пруверы?

- Li et al. (2024): сотни статей, куча методов, это ВОТ ТОЛЬКО датасеты



# Как там пруверы?

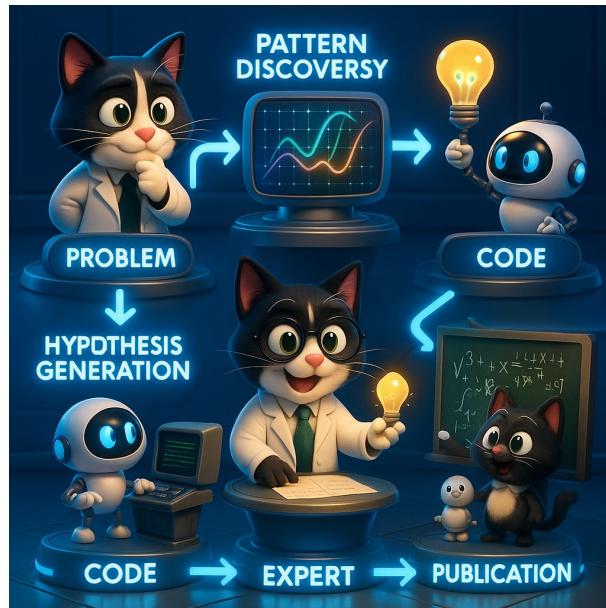
- Li et al. (2024) · синий  
статьей, куда  
здесь, это  
ВОТ

natural language	E.g., NL-PS (Ferreira & Freitas, 2020a), NaturalProof (Reichel et al., 2023),
GamePad	E.g., GamePad (Huang et al., 2019), CogC (Graph2Tac (Rute et al., 2024))
IsarStep	E.g., IsarStep (Li et al., 2021), FVELer (Lin et al., 2024)
Lean	E.g., LeanStep (ProofN (Lin et al., 2024))
LeanDojo	E.g., LeanDojo (Yang et al., 2023), LeanDojo (Yang et al., 2024), LEAN-GitHub (Wu et al., 2024)
Mizar	E.g., Mizar (Kaliszyk et al., 2018)
MLFMF	E.g., MLFMF (Bauer et al., 2023), UniGeo (Chen et al., 2022), UniGeo (Chen et al., 2023), MLFMF (Bauer et al., 2024)
MathPile	E.g., MathPile (Wang et al., 2023b), OpenWebMath (Paster et al., 2024), OpenWebMath (Shao et al., 2024), InternLM-Math (Ying et al., 2024b)
PropL	E.g., PropL (Wang & Deng, 2020), LIME (Wu et al., 2021c), FwdP (Firoiu et al., 2021), PropL (An et al., 2024), PropL (An et al., 2024b), AIPS (Wei et al., 2024)
GPT-f	E.g., GPT-f (Polu & Sutskever, 2020), FMSCL (Polu et al., 2023), HER (Aygün et al., 2022)
EGO-Prover	E.g., EGO-Prover (Xin et al., 2024b), REFACTOR (Zhou et al., 2024b), ATG (Lin et al., 2024c)
Auto(in)formalization	E.g., MMA (Jiang et al., 2023a), FormL4 (Lu et al., 2024), MUSTARD (Huang et al., 2024), DeepSeek-Prover (Xin et al., 2024a), DeepSeek-Prover (Xin et al., 2024b), Lean Workbook (Ying et al., 2024a)

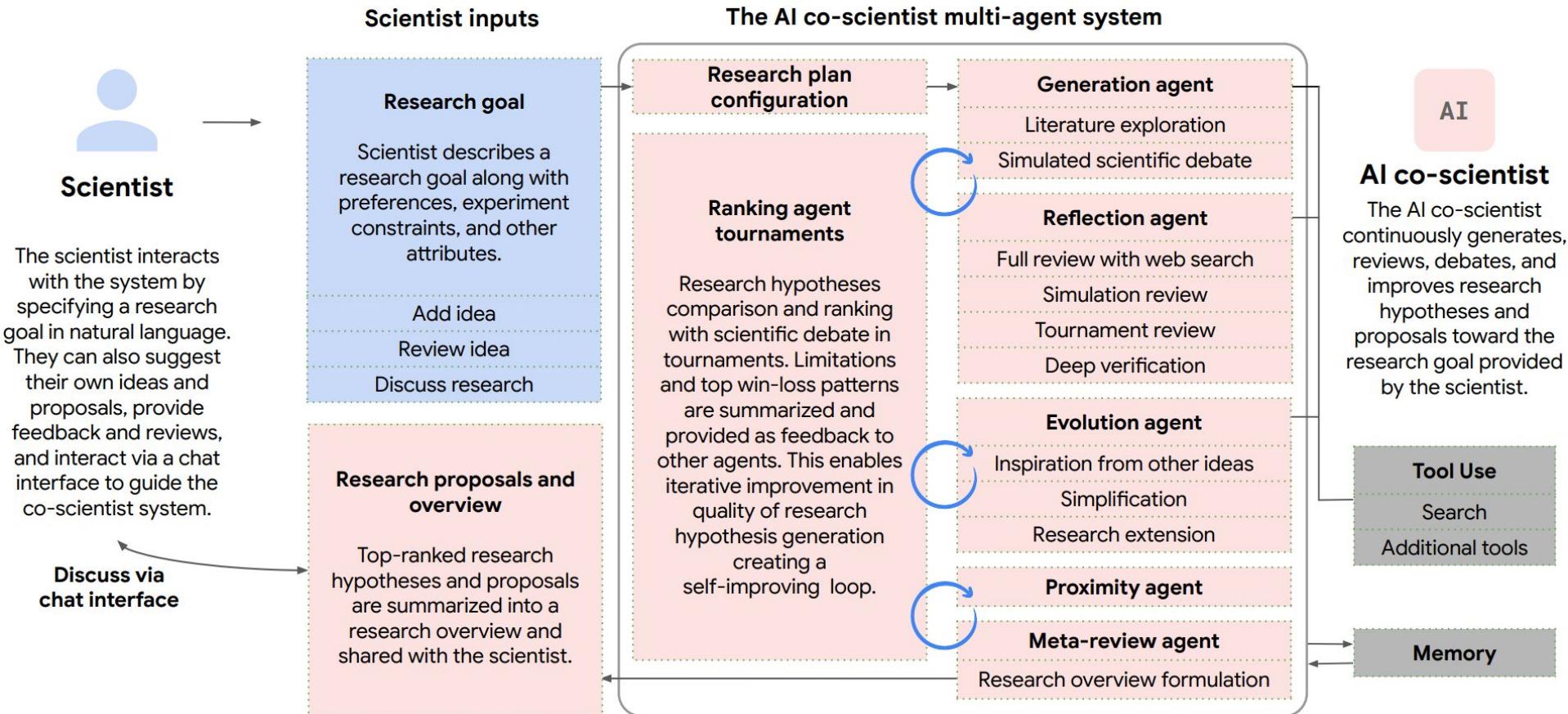
natural language	E.g., NL-PS (Ferreira & Freitas, 2020a), NaturalProof (Reichel et al., 2023),
GamePad	E.g., GamePad (Huang et al., 2019), CogC (Graph2Tac (Rute et al., 2024))
IsarStep	E.g., IsarStep (Li et al., 2021), FVELer (Lin et al., 2024)
Lean	E.g., LeanStep (ProofN (Lin et al., 2024))
LeanDojo	E.g., LeanDojo (Yang et al., 2023), LeanDojo (Yang et al., 2024), LEAN-GitHub (Wu et al., 2024)
Mizar	E.g., Mizar (Kaliszyk et al., 2018)
MLFMF	E.g., MLFMF (Bauer et al., 2023), UniGeo (Chen et al., 2022), UniGeo (Chen et al., 2023), MLFMF (Bauer et al., 2024)
MathPile	E.g., MathPile (Wang et al., 2023b), OpenWebMath (Paster et al., 2024), OpenWebMath (Shao et al., 2024), InternLM-Math (Ying et al., 2024b)
PropL	E.g., PropL (Wang & Deng, 2020), LIME (Wu et al., 2021c), FwdP (Firoiu et al., 2021), PropL (An et al., 2024), PropL (An et al., 2024b), AIPS (Wei et al., 2024)
GPT-f	E.g., GPT-f (Polu & Sutskever, 2020), FMSCL (Polu et al., 2023), HER (Aygün et al., 2022)
EGO-Prover	E.g., EGO-Prover (Xin et al., 2024b), REFACTOR (Zhou et al., 2024b), ATG (Lin et al., 2024c)
Auto(in)formalization	E.g., MMA (Jiang et al., 2023a), FormL4 (Lu et al., 2024), MUSTARD (Huang et al., 2024), DeepSeek-Prover (Xin et al., 2024a), DeepSeek-Prover (Xin et al., 2024b), Lean Workbook (Ying et al., 2024a)

# 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 рассуждающим моделям

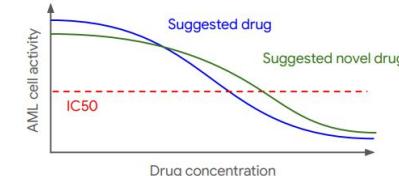
Application ↑  
Basic research ↓

Suggest an existing drug that could be repurposed for acute myeloid leukemia (AML) treatment and provide experimentally testable concentrations for an IC50 assay. The drug should inhibit the proliferation of AML cell lines, particularly MOLM13.

Scientist provide research goal to identify possible drug repurposing candidates for acute myeloid leukemia (AML).

**With preclinical evidence**  
Binimetinib, Pacritinib, ...  
**Completely novel repurposing**  
KIRA6, Leflunomide, ...

The AI co-scientist generates predictions for AML drug repurposing. Scientists review and selects candidates for *in vitro* experiments.



*In vitro* experiments show that the proposed co-scientist drug repurposing candidates inhibit tumor activity in AML cell lines.

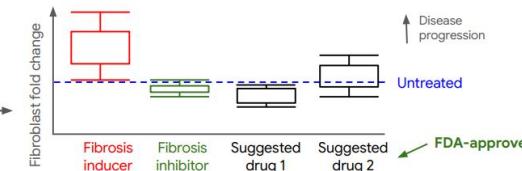
## Drug repurposing for acute myeloid leukemia (AML)

Propose a novel hypothesis about specific epigenetic alterations contributing to myofibroblast formation in liver fibrosis.

Scientist specify research goal to identify novel epigenetic targets for liver fibrosis.

**The AI co-scientist identifies three novel epigenetic targets**

The AI co-scientist proposes several epigenetic target candidates for *in vitro* experiments.



*In vitro* experiments show that the drugs based on co-scientist suggested epigenetic targets reduce the fibrogenesis in human hepatic organoids.

## Identifying novel treatment targets for liver fibrosis

Why are cf-PICLs found in many bacterial species?

Scientists start exploring the cf-PICL mechanisms in 2015.

Hypothesis generation

2015-2024: Scientists create novel hypothesis, and validate experimentally over ~ 10 years of iterative research

Independently propose the same hypothesis

2024: The AI co-scientist generated research hypothesis recapitulates the empirical findings in 2 days

**The AI co-scientist hypothesis**  
Conserved regions on capsids and tails  
Capsid interaction with bacterial membranes  
...

Recapitulate result

## Parallel in-silico discovery of bacterial gene transfer mechanism relevant to antimicrobial resistance (AMR)

# 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'?

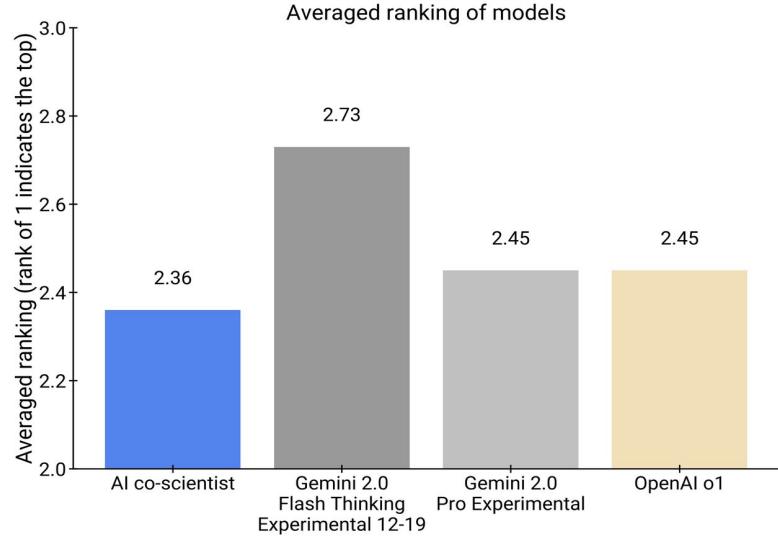
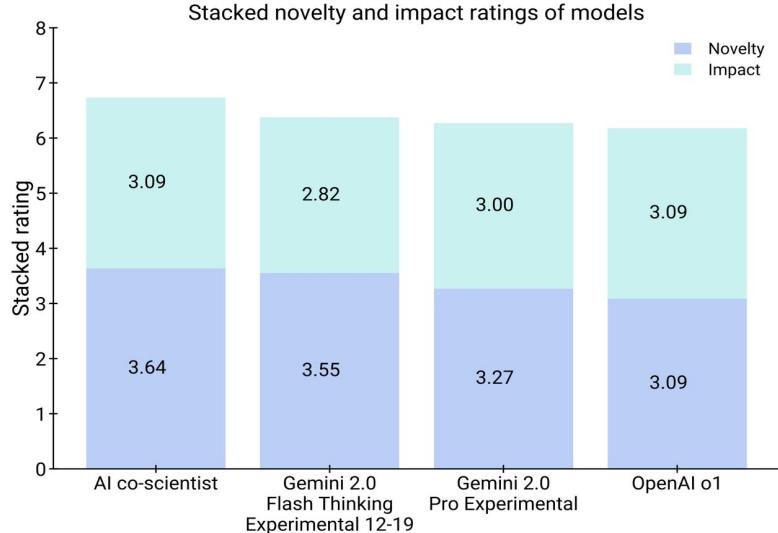
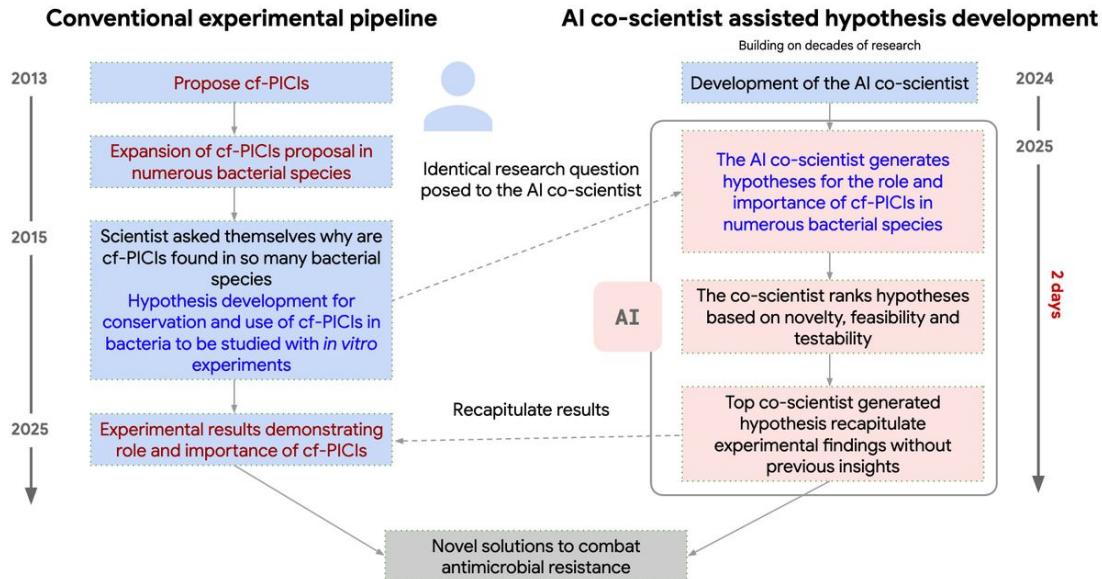
Researchers who have been given access to Google's new AI "co-scientist" tool are enthusiastic about its potential, but it isn't yet clear whether it can make truly novel discoveries

By [Michael Le Page](#)

19 February 2025

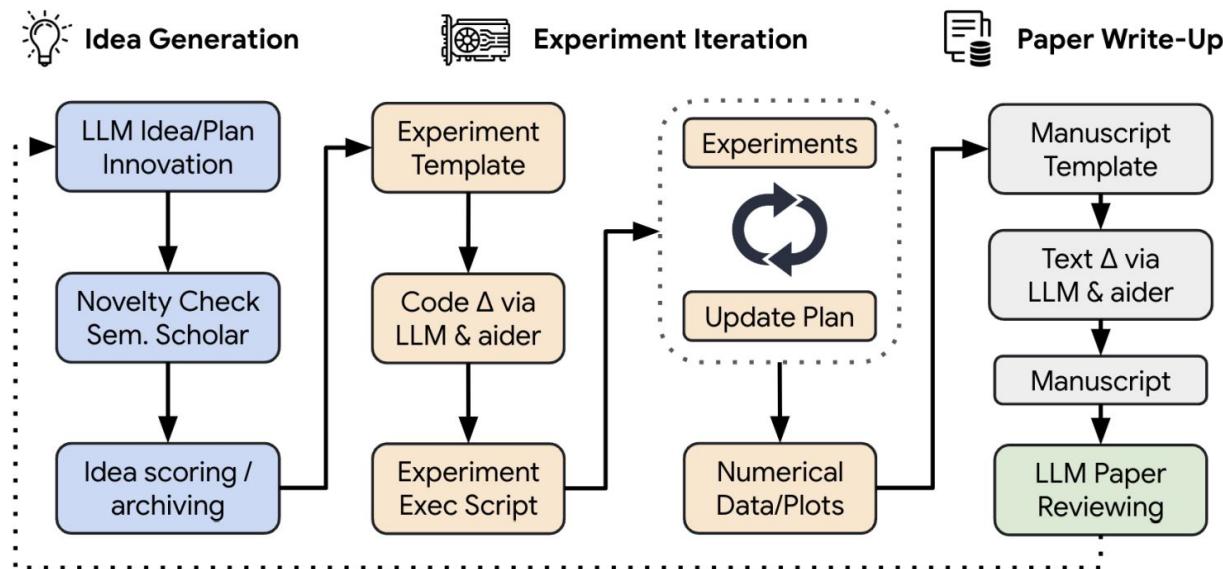
# Google Co-Scientist

- Да и оценки, которыми хвастается сам Google, не так уж убедительны
- Но пример всё равно впечатляет



# Пример из августа 2024-го

- 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 о1 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
<a href="#">Compositional Regularization: Unexpected Obstacles in Enhancing Neural Network Generalization</a>	6, 7, 6
<a href="#">Real-World Challenges in Pest Detection Using Deep Learning: An Investigation into Failures and Solutions</a>	3, 7, 4
<a href="#">Unveiling the Impact of Label Noise on Model Calibration in Deep Learning</a>	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.

# The 63rd Annual Meeting of the Association for Computational Linguistics

# Robin

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

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

# Robin

- 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.,

# Robin

- Но не последняя! Система Zochi (Intology AI, March 2025) автономно написала статью, которую приняла на ACL 2025. Главный
- Кстати, о с. статье тоже интересно поговорить...

## Tempest: Automatic Multi-turn Adversarial Prompts for Large Language Models

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

Intology AI  
intology.ai

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.,

turn adversarial prompts. The gradual erosion of large language model (LLM) safety through a multi-turn perspective. Unlike single-turn jailbreaks, which 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

# 9. Текущее положение дел



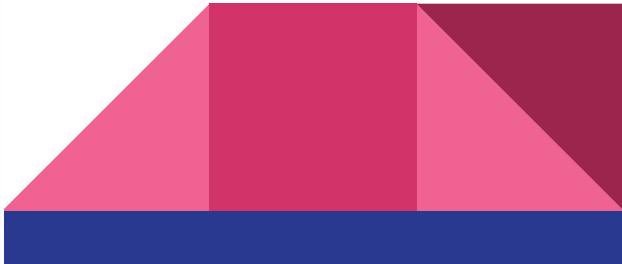
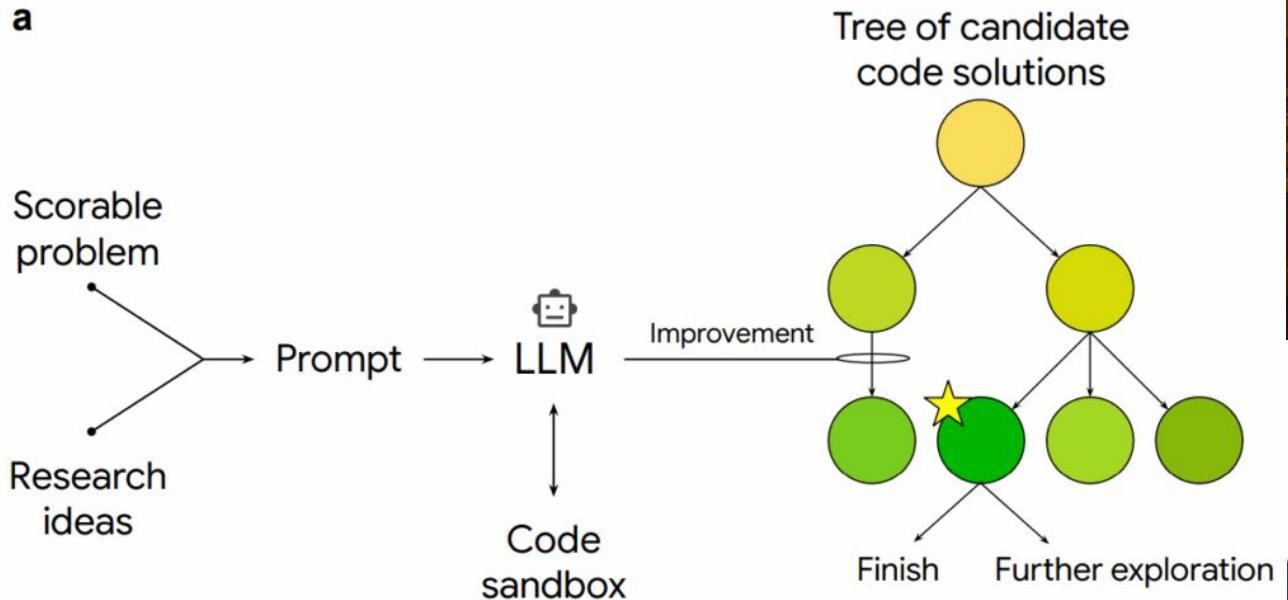
Сограждане! Во-первых, не должно предаваться унынию при настоящем дел положении, сколь бедственным оно ни кажется. Ибо что прежде было вам очень вредно, то самое обещает великие выгоды для будущего. А что это такое? То, что дела в худом положении по причине беспечности вашей и нерадения.

Первое слово Демосфено во против Филиппа

# Побей все бенчмарки до завтра

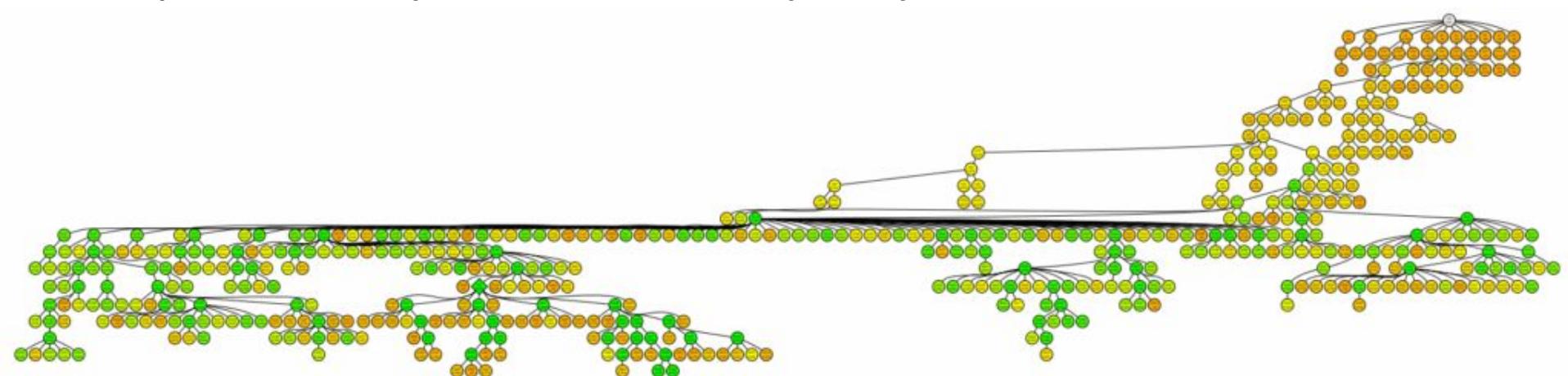
- Aygün et al. ([DeepMind, September 8, 2025](#)): сочетание поиска по дереву (как в AlphaGo), добавления идей из научной литературы (через LLM) и рекомбинации методов (как в генетическом программировании)

a



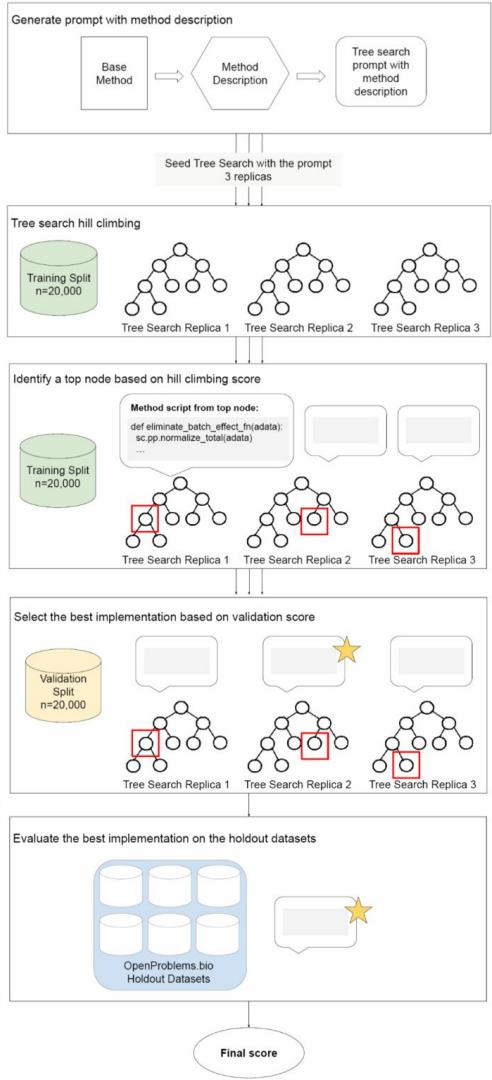
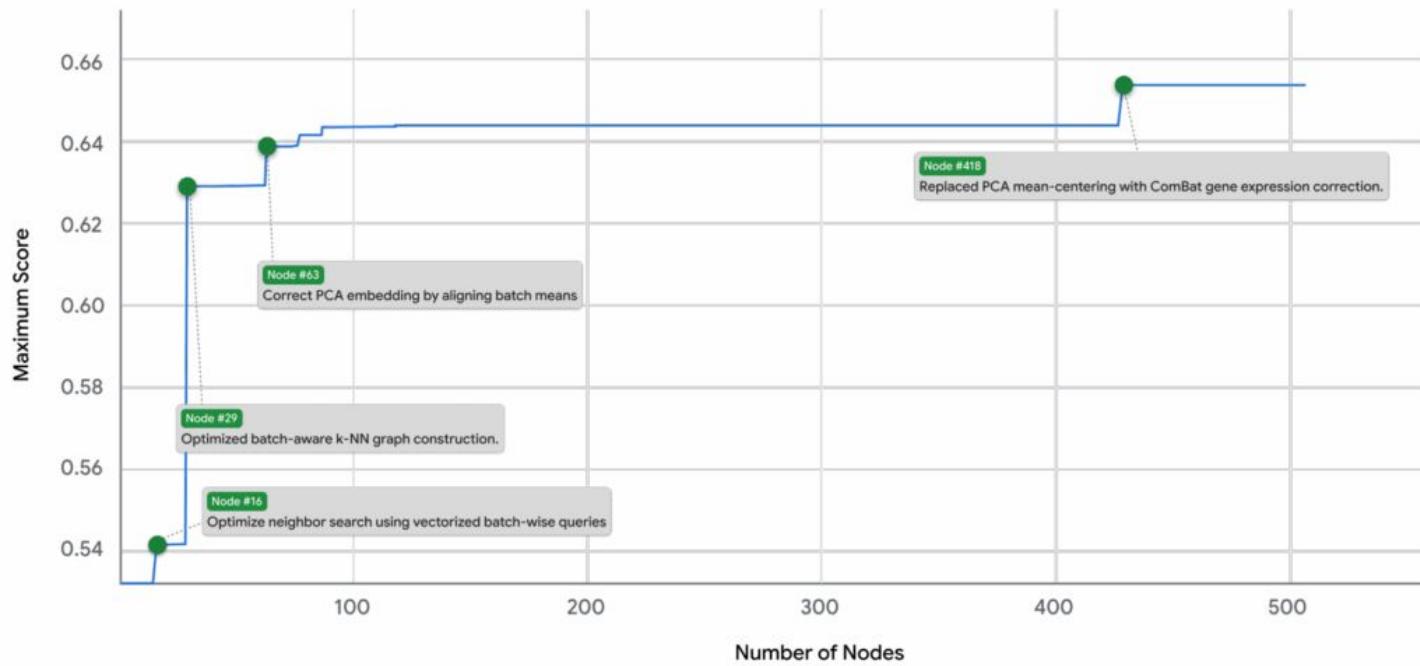
# Побей все бенчмарки до завтра

- 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



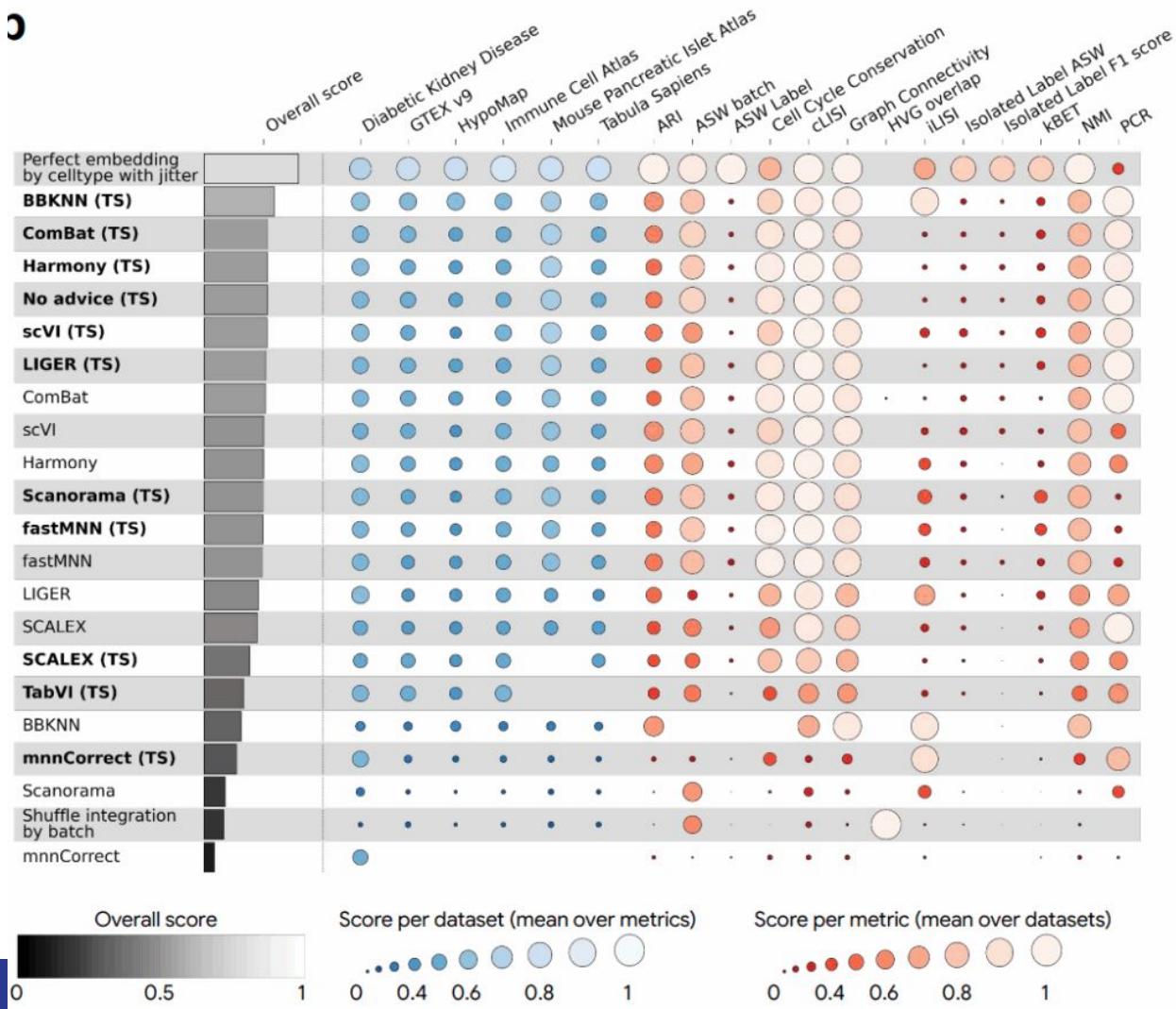
# Побей все бенчмарки до завтра

- 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)



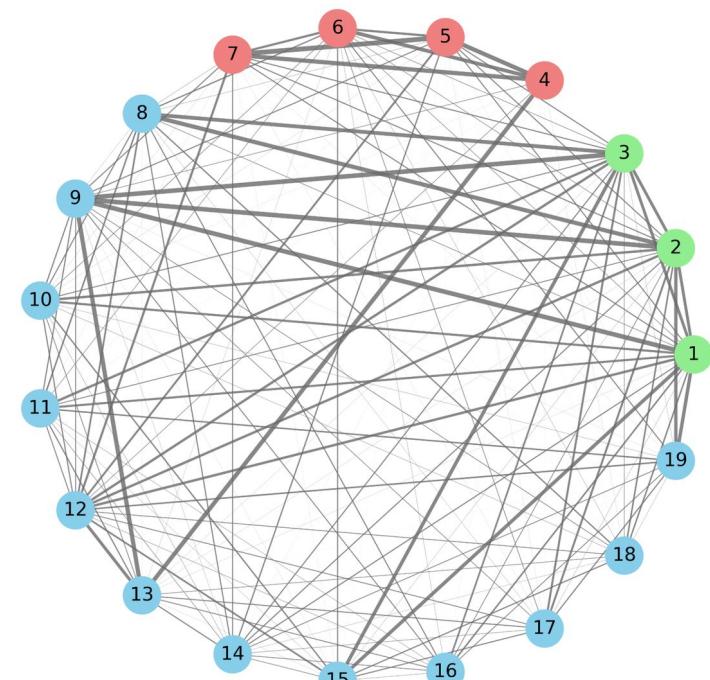
# Aygün et al.

- 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 $\int_0^{\infty} (\sin^4(ax^2) - \sin^4(bx^2)) dx$
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 $\int_0^{\infty} \left( \frac{y+x}{\beta^2 + (y+x)^2} - \frac{y-x}{\beta^2 + (y-x)^2} \right) \sin(ax) dx$	465.002 $\int_0^{\infty} \frac{(3-4\sin^2(ax))\sin^2(ax)}{x} dx$
462.034 $\int_0^{\infty} \frac{x \sin(ax) \cos(bx)}{e^2 + x^2} dx$	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 $\int_0^{\infty} \frac{1}{x} \frac{\sin(x)}{(a^2 \cos^2(x) + b^2 \sin^2(x))^2} dx$	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 о выпуклой оптимизации; буквально:

## 4.1.1 Gaussian framework

We started with the following initial prompt:

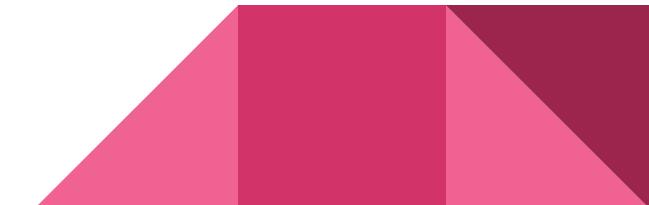
Paper 2502.03596v1 establishes a qualitative fourth moment theorem for the sum of two Wiener-Itô integrals of orders  $p$  and  $q$ , where  $p$  and  $q$  have different parities. Building on the Malliavin-Stein method (see 1203.4147v3 for details), could you derive a quantitative version for the total 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.



# 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  $(p+q)!\|u\tilde{\otimes}v\|^2 = p!q!\|u\|^2\|v\|^2$ . 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 о1

*“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  $\alpha$  and  $\beta$  are constants.  $err$  should be a function that depends only on the error parameter  $\epsilon$  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  $\epsilon$  goes to zero. Please give the best such bound that you prove (a bound is considered better if the values of the constants  $\alpha$  and  $\beta$  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

Consider the performance of a non-DR-submodular function  $G$  which is attained in  $P$  maximal points. We can only give a bound on the error in terms of the smoothness parameter  $\alpha$  and the constants  $\alpha$  and  $\beta$  in the expression  $\alpha * G(o) + \beta * G(o')$ .

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

# Личные свидетельства

- Скотт Ааронсон, один из ведущих специалистов в квантовой теории сложности
- The QMA Singularity: в детали результата вдаваться не будем, но GPT-5 выдала ключевую идею, которая потом сработала в доказательстве (да, пока не самую сложную)

All this will, I hope, interest and excite aficionados of quantum complexity classes, while others might have very little reason to care.

But here's a reason why other people might care. This is the first paper I've ever put out for which a key technical step in the proof of the main result came from AI—specifically, from GPT5-Thinking. Here was the situation: we had an  $N \times N$  Hermitian matrix  $E(\theta)$  (where, say,  $N=2^n$ ), each of whose entries was a  $\text{poly}(n)$ -degree trigonometric polynomial in a real parameter  $\theta$ . We needed to study the largest eigenvalue of  $E(\theta)$ , as  $\theta$  varied from 0 to 1, to show that this  $\lambda_{\max}(E(\theta))$  couldn't start out close to 0 but then spend a long time "hanging out" ridiculously close to 1, like  $1/\exp(\exp(\exp(n)))$  close for example.

Given a week or two to try out ideas and search the literature, I'm pretty sure that Freek and I could've solved this problem ourselves. Instead, though, I simply asked GPT5-Thinking. After five minutes, it gave me something confident, plausible-looking, and (I could tell) wrong. But rather than laughing at the silly AI like a skeptic might do, I *told* GPT5 how I knew it was wrong. It thought some more, apologized, and tried again, and gave me something better. So it went for a few iterations, much like interacting with a grad student or colleague. Within a half hour, it had suggested to look at the function

$$\text{Tr}[(I - E(\theta))^{-1}] = \sum_{i=1}^N \frac{1}{1 - \lambda_i(\theta)}.$$

It pointed out, correctly, that this was a rational function in  $\theta$  of controllable degree, that happened to encode the relevant information about how close the largest eigenvalue  $\lambda_{\max}(E(\theta))$  is to 1. And this ... *worked*, as we could easily check ourselves with no AI assistance. And I mean, maybe GPT5 had seen this or a similar construction somewhere in its training data. But there's not the slightest doubt that, if a student had given it to me, I would've called it clever. Obvious with hindsight, but many such ideas are.

# Личные свидетельства

- А вот (пуб)личное свидетельство математика и преподавателя математики
- Особенно хорошо здесь смотрится “от руки”... когда даже я был студентом (а я ещё не глубокий стариk), конспекты ведь действительно от руки и писали!
- Мой опыт (не совсем в чистой математике) тоже в целом подтверждает этот вывод – мы всё больше становимся проверяющими и постановщиками задач, а не генераторами решений или даже идей

tropical saint petersburg  
как научиться писать статьи:

вот есть у меня несколько проектов с разными студентами. Кажется, сейчас самый быстрый способ писать статью — написать первый драфт как угодно коряво (не обращая внимания на грамматику, артикли — пожалуй, только структуру стоит продумать) — а потом засунуть в ChatGPT и принимать или отклонять его предложения о том, как что переписать.

При этом, скажем, я знаю, что хорошо и что плохо, поэтому знаю, какие предложения AI хорошие, а какие нет. Если оставить студента наедине с AI, скорее всего, получится что-то совсем не то.

Старый способ написания первой статьи (например, переписывание от начала и до конца от руки) уже кажется почти невозможным форсировать. Да и надо ли? Особенно, если студент в принципе не может написать связный текст никакой длины.

Но и говорить — иди и попроси ChatGPT всё написать, начав с отсканированных черновиков вычислений — тоже не очень.

AI вообще уже справляется с доказательством лемм на полстранички. Так что студенческий проект/статью на 10 страничек можно делать вообще без участия студента, достаточно знать, какие именно утверждения верны (то есть как только студент догадался, что именно надо доказывать, точную формулировку, дальше всё может сделать AI). Я пока не понимаю, как с этим жить. Зачем мучаться мне и мучаться студенту, если достаточно сообщить AI точную формулировку, а потом только перепроверять и принимать или отклонять правки.



2233 12:34

A 30 comments >

Январь  
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.

**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. ✓

Сентябрь  
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  $\boxed{-5, 1, 4}$ . ✓

$r_2$

Март  
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.

Example Math-Specialized Gemini 1.5 Pro output:

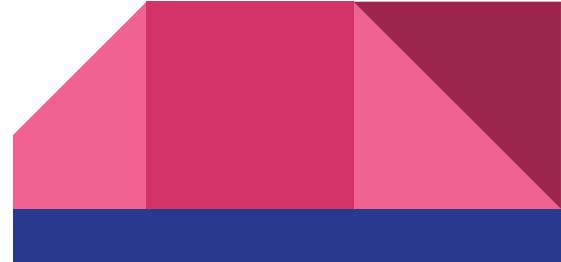
Without loss of generality, assume that  $a \geq b \geq c$ . Then  $a^2 + b + c \leq a^2 + 2a < a^2 + 2a + 1 = (a + 1)^2$ . Since  $a^2 + b + c$  is a perfect square, we must have  $a^2 + b + c = a^2$ , which implies  $b + c = 0$ . This is a contradiction since  $b$  and  $c$  are positive integers. Therefore, 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.

Сентябрь  
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  $\alpha$  and  $\beta$  are constants.  $err$  should be a function that depends only on the error parameter  $\epsilon$  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  $\epsilon$  goes to zero. Please give the best such bound that you prove (a bound is considered better if the values of the constants  $\alpha$  and  $\beta$  are larger). Provide a mathematically rigorous and well explained proof for the bound you come up with.

Важное  
напоми-  
нание о  
прогрессе



# Январь 2022

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. 

# Сентябрь 2023

Problem: Suppose that the numbers  $x$  and  $y$  are possible.

2022

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

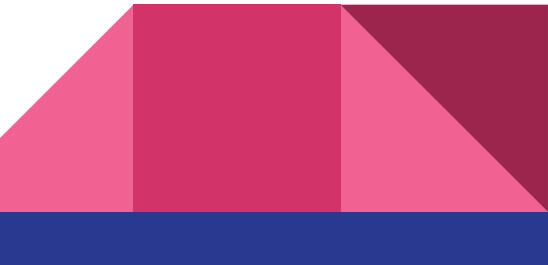
answer is 9. ✓

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

September 2023

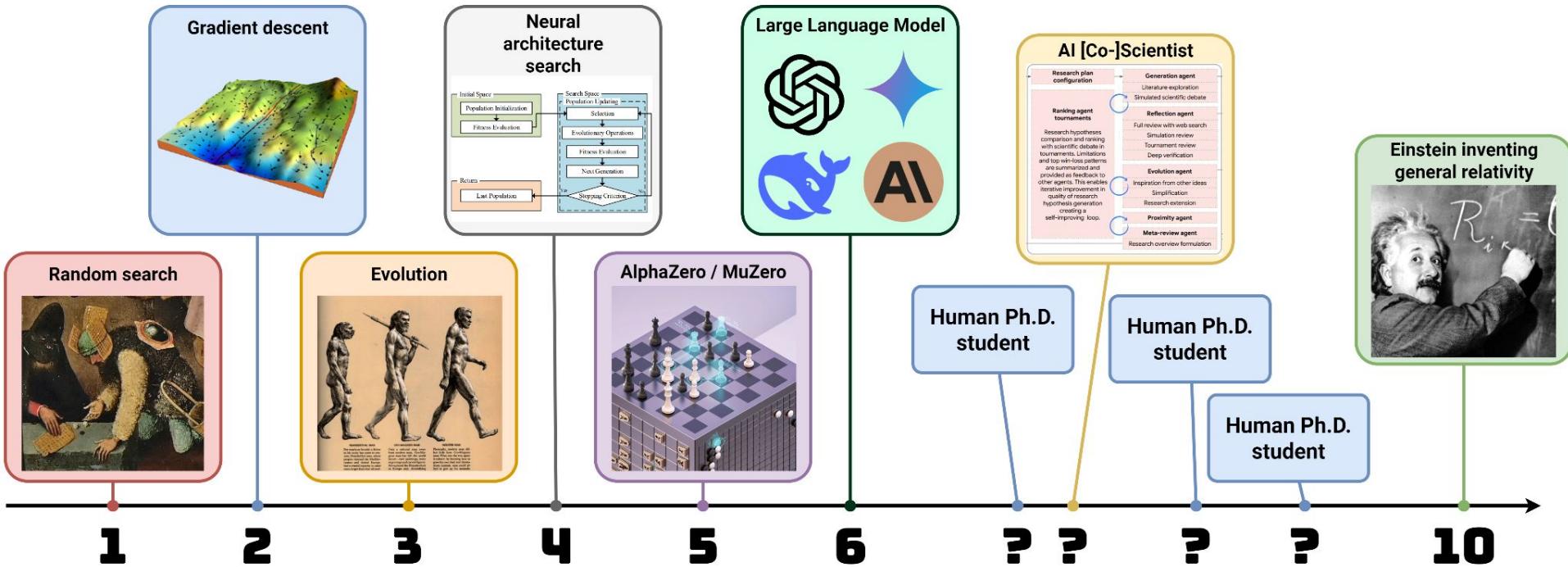
Problem: Suppose that the numbers  $x$  and  $y$  are possible.

Кажется, уже современные LLM могут двигать алгоритмический прогресс для самих себя, и для этого понадобился только некоторый unholobbling. Но в только областях, где ответ можноально легко проверить. ...не сингулярность...



# 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!

