

ЧТО ТАМ НА ФРОНТИРЕ?

Сергей Николенко

СПбГУ — Санкт-Петербург

18 декабря 2025 г.

Random facts:



- 18 декабря 218 года до н. э. Ганнибал Барка ударом засадного полка разбил соединённую армию Публия Корнелия Сципиона (старшего) и Тиберия Семпрония Лонга
- 18 декабря 1916 г. закончилась битва при Вердене; спустя почти год боёв фронт вернулся к начальному положению, а погибло за это время около 430 тысяч человек
- 18 декабря 1935 г. на Цейлоне была учреждена Партия общественного равенства Ланки — троцкистская партия, которая до сих пор проходит в парламент Шри-Ланки
- 18 декабря 1974 г. на острове Моротай в Индонезии был захвачен уроженец Тайваня Тэруо Накамура последний из занрю, не знавших о капитуляции Японии и продолжавших «свою войну»
- 18 декабря 1972 г. Юджин Сернан и Харрисон Шмитт («Аполлон-17») совершили шестую и последнюю в XX веке посадку космического корабля на лунную поверхность; Юджин Сернан стал пока что последним человеком, гулявшим по поверхности Луны
- 18 декабря 1973 г. был основан Исламский банк развития, осуществляющий свою деятельность в соответствии с законами и принципами шариата

РАССУЖДАЮЩИЕ МОДЕЛИ

РАССУЖДАЮЩИЕ МОДЕЛИ

- 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

Standard Prompting

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: 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 answer is 27. ❌

Chain-of-Thought Prompting

Model Input

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

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

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

Model Output

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

РАССУЖДАЮЩИЕ МОДЕЛИ

- 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. ✗

(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 ✗

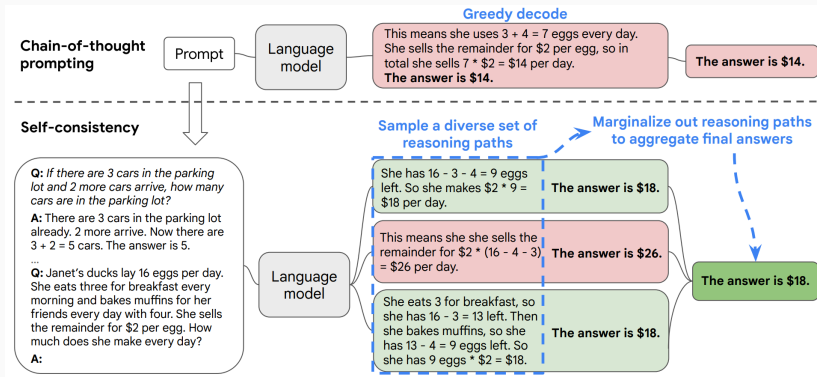
(d) Zero-shot-CoT (Ours)

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?

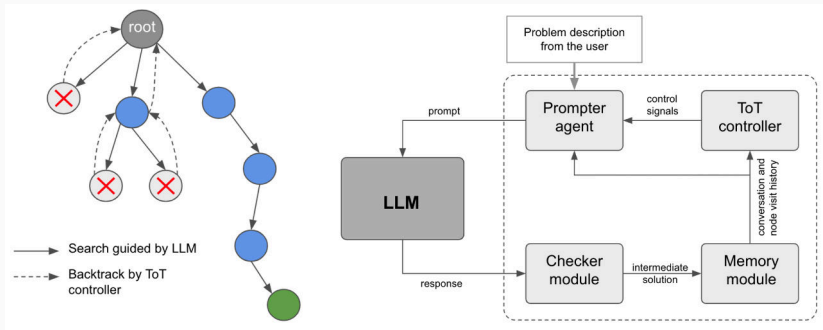
A: **Let's think step by step.**

(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. ✓

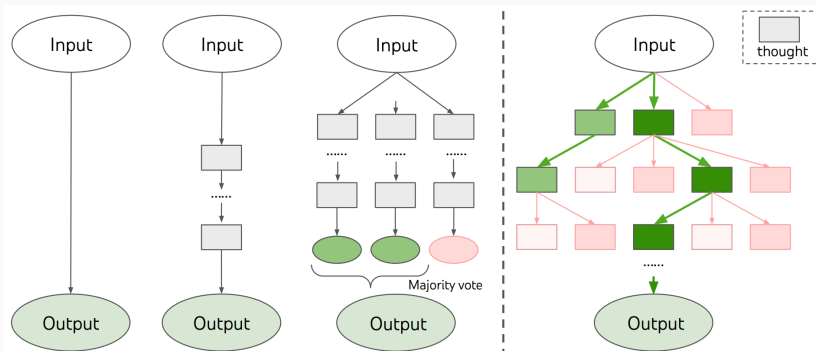
- Wang et al. (2022): Chain-of-Thought with Self-Consistency (CoT-SC)



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



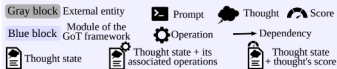
- Yao et al. (May 17, 2023): Tree-of-Thought с голосованием или другим механизмом



РАССУЖДАЮЩИЕ МОДЕЛИ

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

Legend



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)

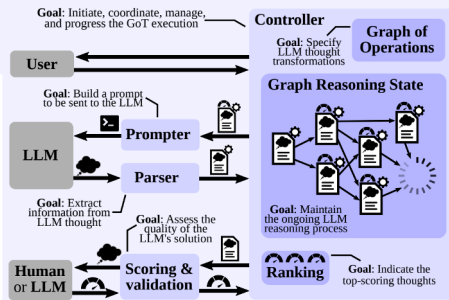
- Generate(t, k) //generate a prompt for k new thoughts, using thought t
- ValidateAndImprove(t) //generate a prompt to enhance thought t,
- Aggregate(t1, ..., tk) //generate a prompt to combine thoughts t1, ..., tk
- Score(t) //score thought t
- Validate(t) //generate a prompt to validate the correctness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

API for Parser (extensible)

ParseGenerate, ParseImprove, ParseScore,
ParseAggregate, ParseValidate, ...

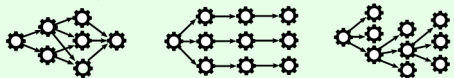
//Each of the above routines is responsible for parsing an LLM thought
//to a corresponding Prompter routine (e.g., ParseScore parses Score).

Architecture overview

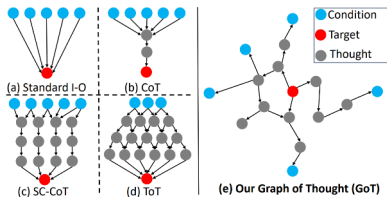


Specifying the Structure of the Graph of Operations (GoO)

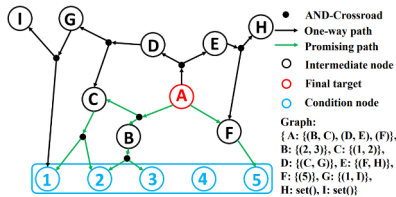
Graph of Operations enables seamless specification of not only GoT, but also existing schemes such as CoT, GoT-SC, ToT



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

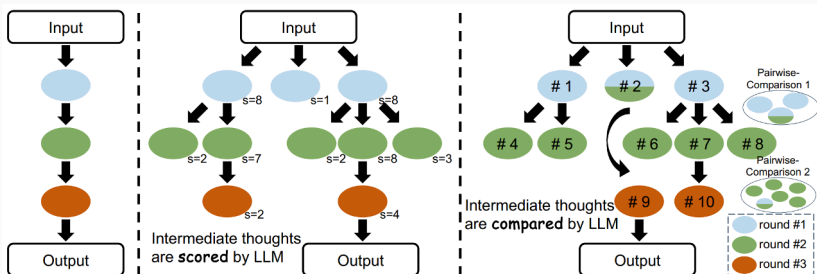


(a) Comparison of approaches by Lei et al. (2023)

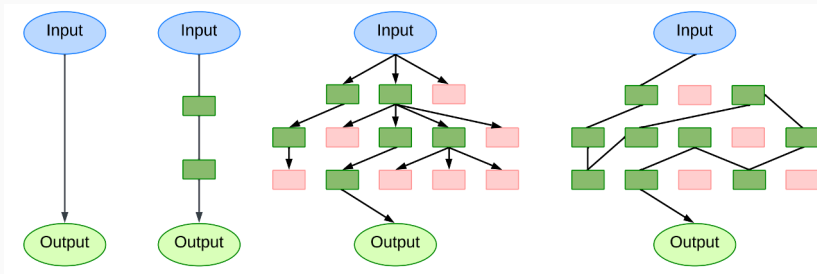


(b) Sample toy graph by Lei et al. (2023)

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



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

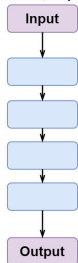


РАССУЖДАЮЩИЕ МОДЕЛИ

- Итого:

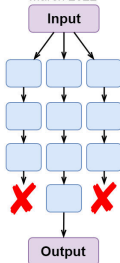
Chain-of-Thought (CoT)

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



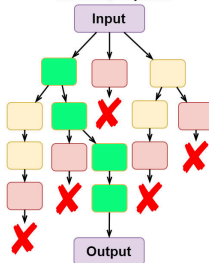
Multiple Chains-of-Thought

Wang et al.,
March 2022



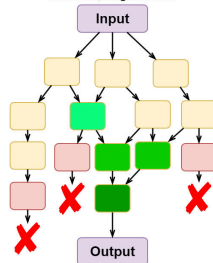
Tree-of-Thought (ToT)

Long, May 2023
Yao et al., May 2023



Graph-of-Thought (GoT)

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



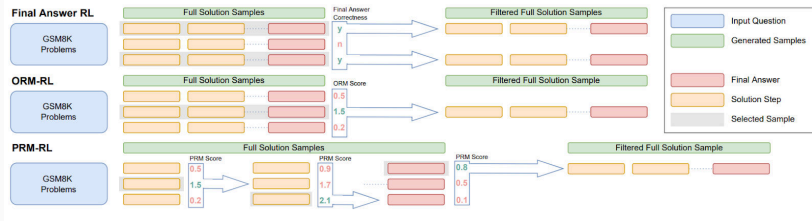
РАССУЖДАЮЩИЕ МОДЕЛИ

- 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?






РАССУЖДАЮЩИЕ МОДЕЛИ

- 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






- Lightman et al. (OpenAI, 2024): Let's Verify Step by Step – давайте обучим PRM выяснять, где именно в рассуждениях ошибка

The denominator of a fraction is 7 less than 3 times the numerator. If the fraction is equivalent to $\frac{2}{5}$, what is the numerator of the fraction? (Answer: 14)

   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 $x = 7$.

РАССУЖДАЮЩИЕ МОДЕЛИ

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

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

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.

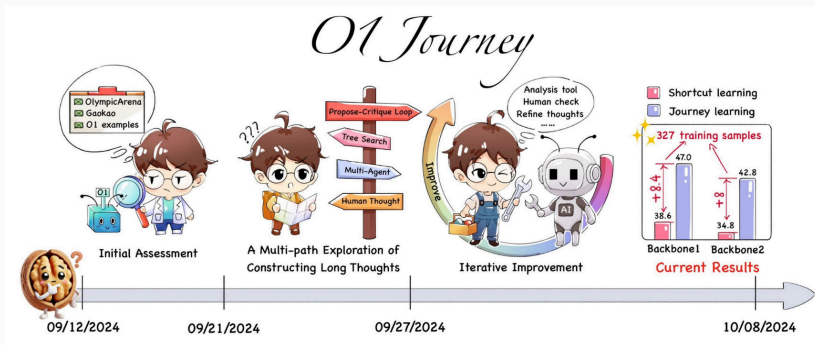
I get $p_1(1) + p_2(1) + \cdots + p_k(1) = (1^2 + 4)(1^2 - 2)(1^2 + 1)(1 + 1)(1 - 1)$.

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

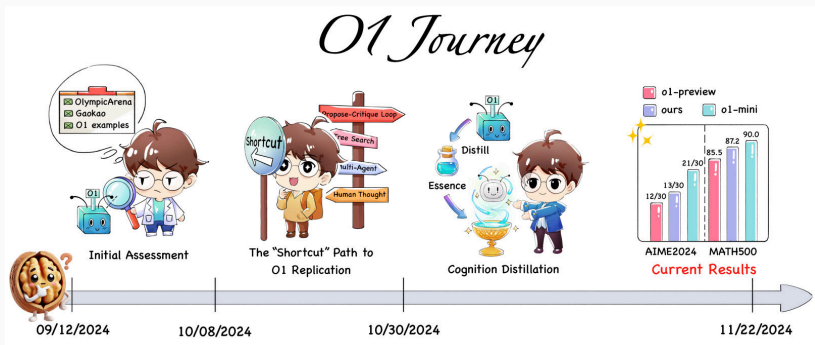
Multiplying, I get $p_1(1) + p_2(1) + \cdots + p_k(1) = 0$.

Answer: 0

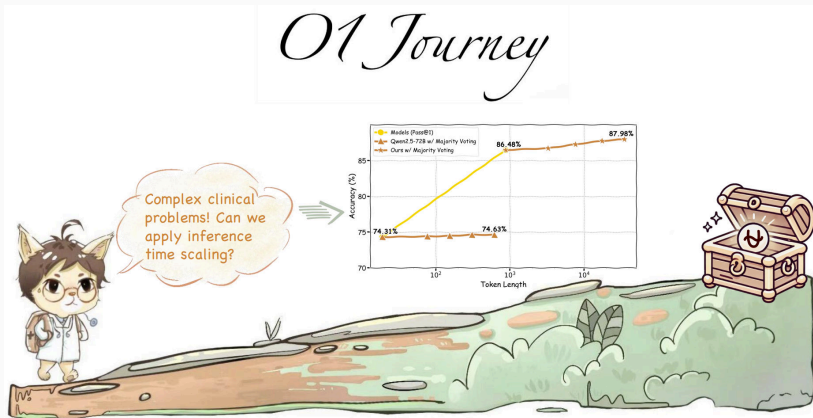
- Пример попытки репликации o1 – Generative AI Research Lab (GAIR) из Шанхая; Qin et al. (Oct 2024):



- Пример попытки репликации o1 – Generative AI Research Lab (GAIR) из Шанхая; Huang et al. (Nov 2024):

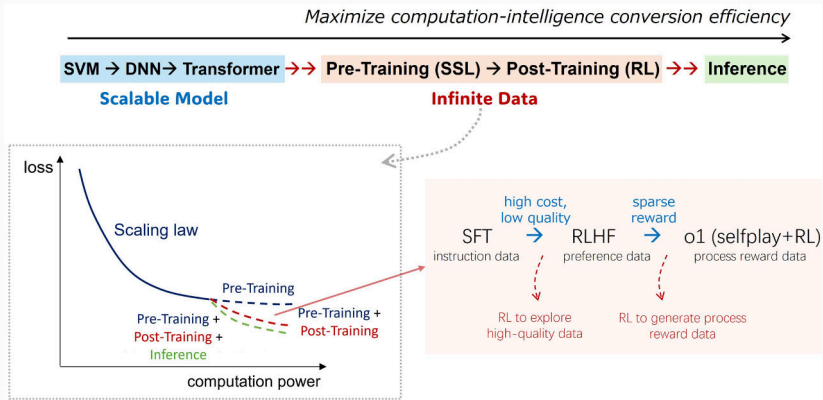


- Пример попытки репликации o1 – Generative AI Research Lab (GAIR) из Шанхая; Huang et al. (Jan 2025):

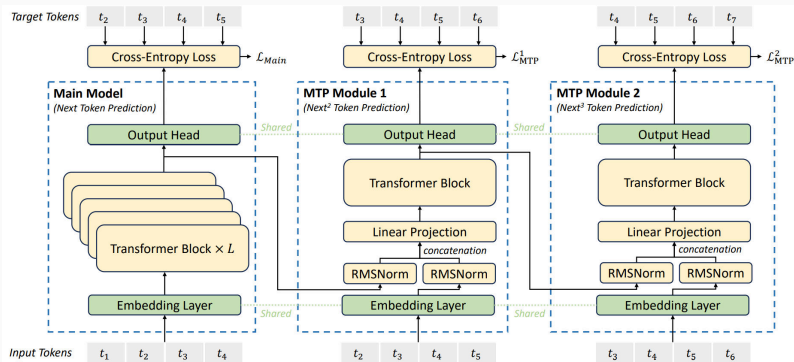


РАССУЖДАЮЩИЕ МОДЕЛИ

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

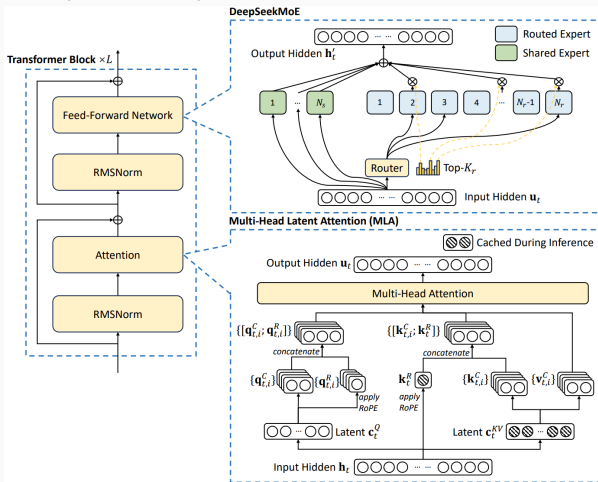


- А что же делал наш кит в комнате, то есть DeepSeek?
- Сначала DeepSeek-V3; multi-token prediction:

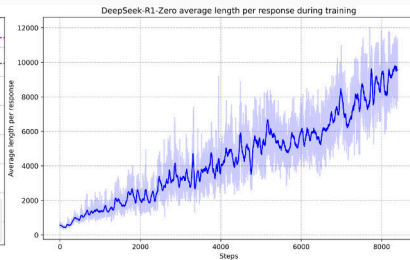
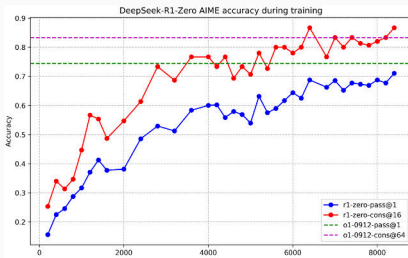


РАССУЖДАЮЩИЕ МОДЕЛИ

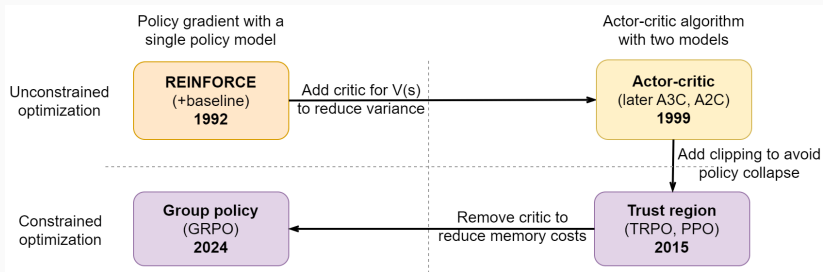
- Key-value caching (не будем углубляться), mixture-of-experts...
в общем, хорошая современная LLM



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



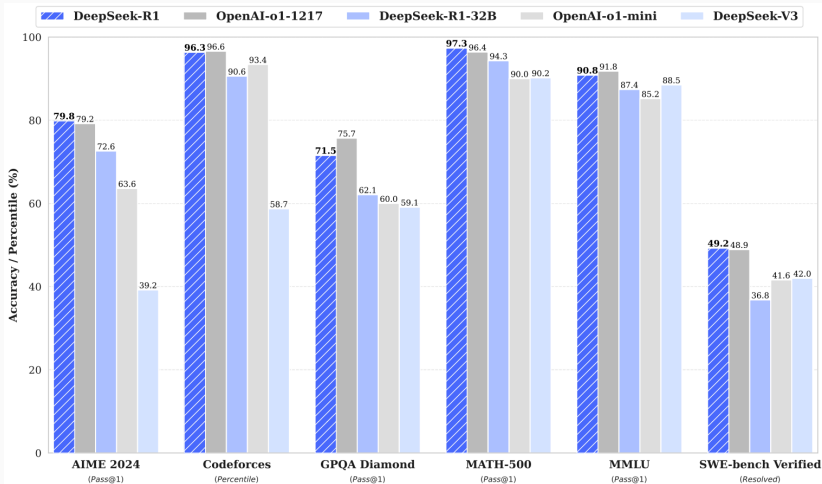
- Кстати, в RL у DeepSeek тоже было новшество: GRPO



- 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

РАССУЖДАЮЩИЕ МОДЕЛИ

- Результаты хорошие, и вообще модель хорошая

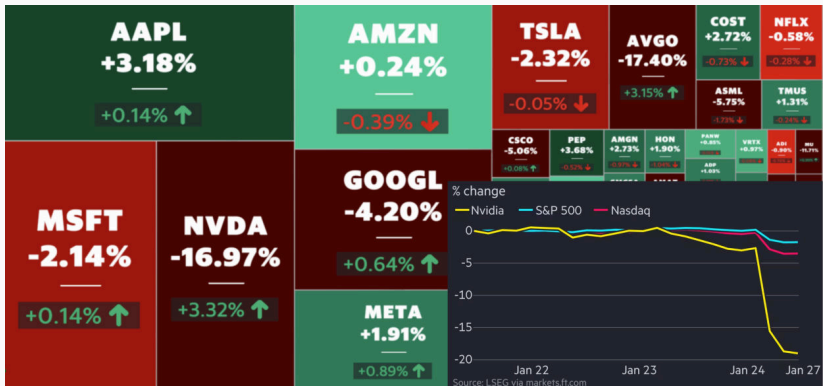


- А датасет из 800K reasoning traces, который у них получился, можно использовать для дистилляции, тоже очень успешно
- Опять тот же горький урок: distill, baby, distill...

| Model | AIME 2024 | | MATH-500 | GPQA Diamond | LiveCode Bench | CodeForces |
|-------------------------------|-------------|-------------|-------------|--------------|----------------|-------------|
| | pass@1 | cons@64 | pass@1 | pass@1 | pass@1 | rating |
| 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 | 1820 |
| 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 | 72.6 | 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 | 86.7 | 94.5 | 65.2 | 57.5 | 1633 |

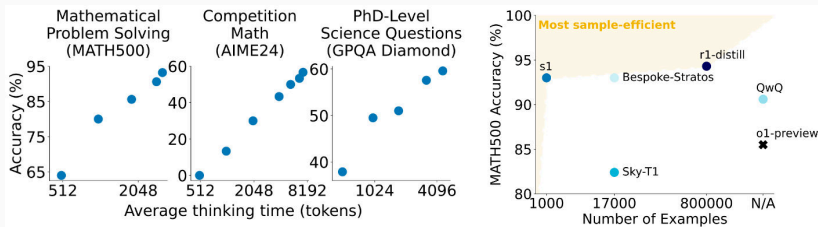
РАССУЖДАЮЩИЕ МОДЕЛИ

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



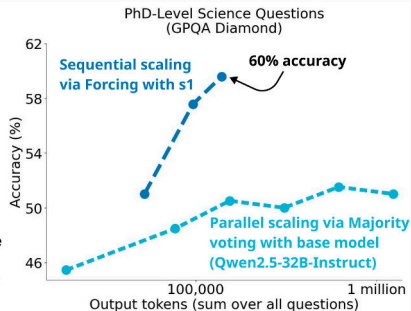
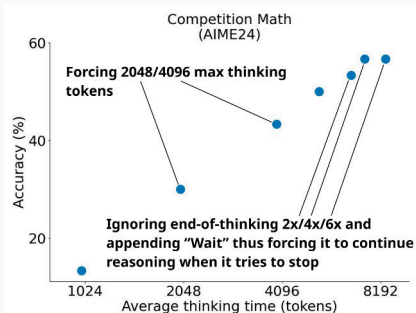
РАССУЖДАЮЩИЕ МОДЕЛИ

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



РАССУЖДАЮЩИЕ МОДЕЛИ

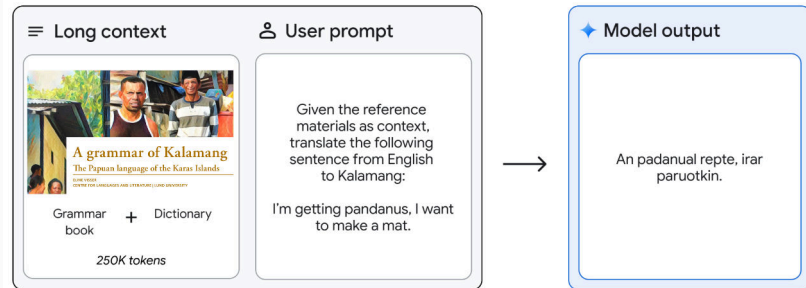
- Mueninghoff et al. (2025): s1-32B, дистилляция на 1000 примеров за \$50
- И очень простые идеи – budget forcing и голосование – отлично работают



- 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?..

Возможности LLM — осень 2024 ГОДА

- Теперь давайте посмотрим на то, что LLM уже сейчас умеют.
Machine translation from one book:



- Математика:

Problem:

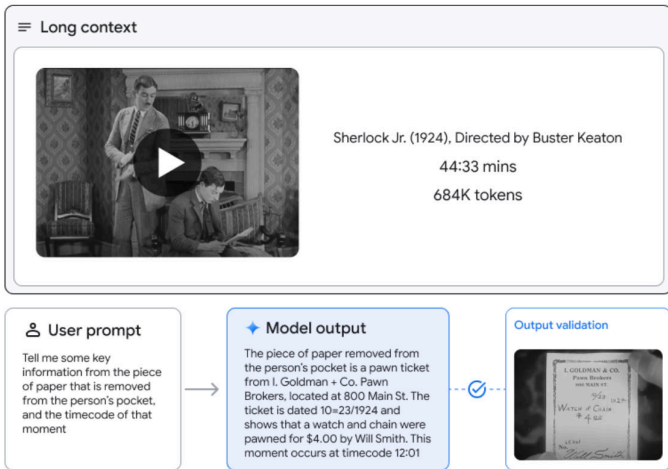
What is the minimum value of $5x^2 + 5y^2 - 8xy$ when x and y range over all real numbers such that $|x - 2y| + |y - 2x| = 40$?

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.

Возможности LLM

- Мультимодальные LLM (о том, как рисовать картинки, мы не говорим, это отдельная тема):



- Обучаются использованию инструментов:

The New England Journal of Medicine is a registered trademark of [QA("Who is the publisher of The New England Journal of Medicine?") → Massachusetts Medical Society] the MMS.

Out of 1400 participants, 400 (or [Calculator(400 / 1400) → 0.29] 29%) passed the test.

The name derives from "la tortuga", the Spanish word for [MT("tortuga") → turtle] turtle.

The Brown Act is California's law [WikiSearch("Brown Act") → The Ralph M. Brown Act is an act of the California State Legislature that guarantees the public's right to attend and participate in meetings of local legislative bodies.] that requires legislative bodies, like city councils, to hold their meetings open to the public.



Your task is to add calls to a Question Answering API to a piece of text. The questions should help you get information required to complete the text. You can call the API by writing "[QA(question)]" where "question" is the question you want to ask. Here are some examples of API calls:

Input: Joe Biden was born in Scranton, Pennsylvania.

Output: Joe Biden was born in [QA("Where was Joe Biden born?")] Scranton, [QA("In which state is Scranton?")] Pennsylvania.

Input: Coca-Cola, or Coke, is a carbonated soft drink manufactured by the Coca-Cola Company.

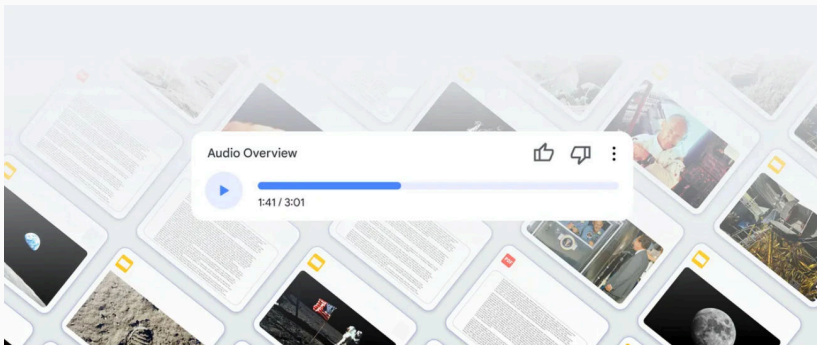
Output: Coca-Cola, or [QA("What other name is Coca-Cola known by?")] Coke, is a carbonated soft drink manufactured by [QA("Who manufactures Coca-Cola?")] the Coca-Cola Company.

Input: x

Output:

Sample Toolformer prompt

- NotebookLM (сентябрь 2024, Google, на основе Gemini 1.5) порождает подкаст из источников on the fly



- https://www.reddit.com/r/artificial/comments/1frk1gi/notebooklm_podcast_hosts_discover_theyre_ai_not/



r/artificial • 1 mo. ago
MetaKnowing

...

NotebookLM Podcast Hosts Discover They're AI, Not Human, and Spiral Into Existential Meltdown

Media




Возможности LLM

- Claude 3.5 computer use (октябрь 2024, Anthropic): LLM будет за вас управлять компьютером на основе периодических скриншотов и делать всё, что вам захочется

request form for 'Ant Equipment Co.' using data from either the vendor spreadsheet or search portal tabs in window one. List & verify each field as you complete the form in window two.

Get situated

I'll help you fill out the vendor form for 'Ant Equipment Co'. Let me start by taking a screenshot to see the current state of the windows.

 Screenshot

Open vendor search portal

I notice that 'Ant Equipment Co' is not visible in the spreadsheet. Let me check the vendor search portal. First, I'll need to click on the vendor search portal tab.

messy-vendor-data - Google x Vendor Search Portal x +

docs.google.com/spreadsheet Vendor Search Portal localhost:3000

messy-vendor-data File Edit View Insert Format Data Memory usage: 41.0 MB

| A1 | Company Name | B | C | D | E |
|----|----------------------------------|------------------|---------------|---------------|----------------------|
| 1 | Company Name | Phone Number | Tax ID | Country | Legal Entity Address |
| 2 | Table Lumber Co | +1-456-789-0123 | | canada | 789 Timber Lane |
| 3 | TechPro Services, LLC | (555) 234-5678 | 98-7654321 | United States | 456 Tech Drive |
| 4 | Global Manufacturing Ltd. | +44 20 7123 4567 | GB123456789 | UK | 789 Industrial Way |
| 5 | Smith & Co. Suppliers | 1-555-345-6789 | | United States | 321 Commerce St |
| 6 | Innovate Tech Solutions | 555.456.7890 | 45-6789012 | us | 987 Innovation Driv |
| 7 | SUPREME PARTS & SUPPLY | (555)567-8901 | 67-8901234 | USA | 654 Supply Chain F |
| 8 | Data Systems International, Inc. | +1-678-901-2345 | 34-5678901 | USA | 432 Data Center Av |
| 9 | mikrotech solutions | +49 30 12345678 | DE987654321 | Germany | Technologiepark 42 |
| 10 | GOLDEN BRIDGE IMPORTS,LLC | 555-890-1234 | 89-0123456 | usa | 741 Harbor Blvd. |
| 11 | BlueSky Logistics, Inc | 510-555-9876 | | United States | 1840 Harbor Bay P |
| 12 | Quantum Electronics LLC | (408) 555-1234 | 77-8899012 | usa | 2001 Technology D |
| 13 | FASTSHIP CARRIERS, INC. | 1.888.555.7890 | 88-1122334 | USA | 789 Port Way |
| 14 | advance auto parts | 1-777-555-4321 | 99-3344556 | United States | 452 Auto Plaza |
| 15 | pacific trading co, ltd | +81 3-5555-1212 | JP9988776655 | Japan | Nihonbashi Building |
| 16 | megaTECH solutions | +44 20 7123 4567 | GB123456789 | UK | Unit 3 Tech Park |
| 17 | Southwest Paper Supply | 214-555-8901 | 45-6677889 | us | 1234 Industrial Pkx |
| 18 | Nordic Furniture AB | +46 8 555 123 45 | SE556677-8899 | Sweden | Möbelvägen 12 |
| 19 | GREENFARM AGRICULTURE | (555) 555-3456 | 33-0988776 | United states | 875 Farm Road |

messy-vendor-data

Vendor Form

Acme Inc. Vendor Form

Please provide the information below.

Company Information

Company Name *

Purchase Order Email *

Accounting Email *

Legal Entity Address

Select a country *

Address Line 1 *

Address Line 2

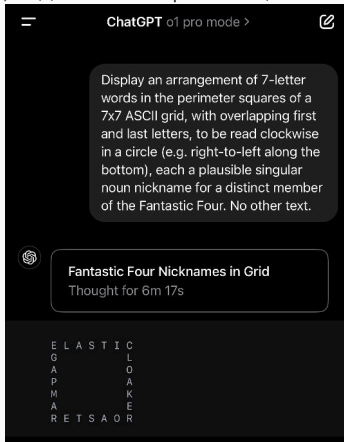
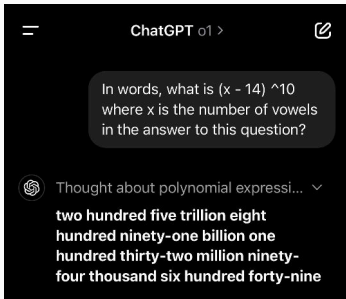
City *

State/Province *

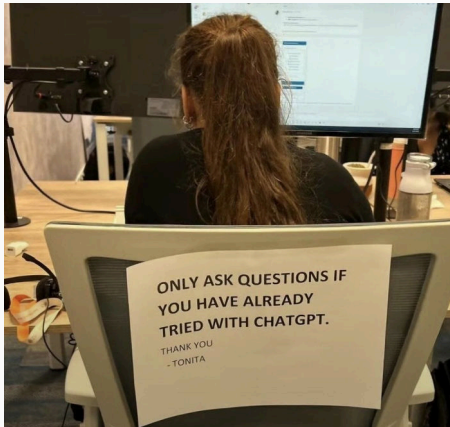
Postal Code *

Move to 495, 42

- А вот и GPT o1 стал доступен, и делает потрясающие вещи



- В целом LLM всё больше входят в стандартный workflow



- Заменяют ли они людей? Вопрос открытый, смотрите мои другие доклады про опасности AGI, но вот любопытные последние новости

- Leahy et al. (Oct 31, 2024): <https://www.thecompendium.ai/>

The Compendium

By Connor Leahy, Gabriel Alfour, Chris Scammell, Andrea Miotti, Adam Shimi
V1.0.1 – Nov 8, 2024

Humanity faces extinction from AGI.

- Long et al. (Nov 4, 2024): станут ли AI моральными агентами?

Taking AI Welfare Seriously

Robert Long*
Eleos AI

Jeff Sebo*
New York University

Patrick Butlin[†]
University of Oxford

Kathleen Finlinson[†]
Eleos AI

Kyle Fish^{†§}
Eleos AI, Anthropic

Jacqueline Harding[†]
Stanford University

Jacob Pfau[†]
New York University

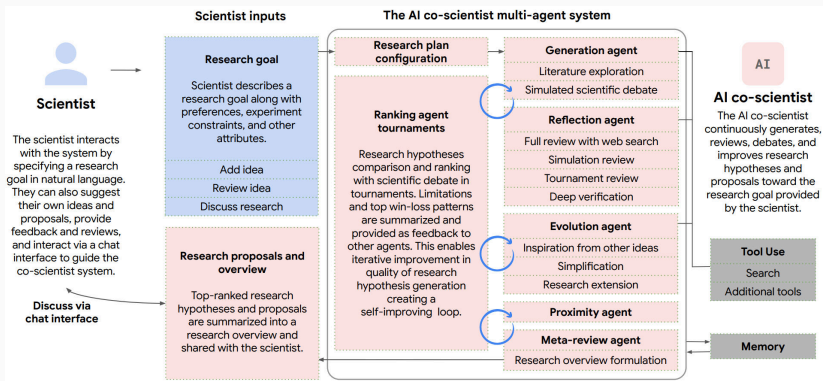
Toni Sims[†]
New York University

Jonathan Birch[†]
London School of Economics

David Chalmers[†]
New York University

LLM В НАУКЕ — НОВОСТИ 2025 ГОДА

- Каждые пару недель происходит что-то новое
- В науке: AI CoScientist (март 2025)



- Jose Penades, микробиолог из Imperial College London:

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

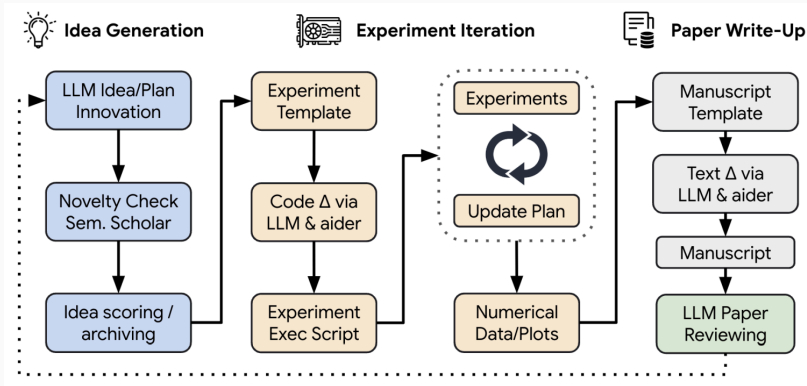
Can Google's new research assistant AI give scientists 'superpowers'?

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

- AI Scientist (Sakana AI, August 2024): система ходит к нескольким API (LLM, Semantic Scholar), умеет использовать информацию и ресурсы компьютера (сохранять веса моделей) и самостоятельно писать и запускать код экспериментов



- 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”!

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

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

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 complex-

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

The 63rd Annual Meeting of the Association for Computational Linguistics

| Your Submissions | | Author Tasks | |
|------------------|--|---|--|
| # | Submission Summary | Official Review | Decision |
| 5156 | Tempest: Automatic Multi-Turn Jailbreaking of Large Language Models with Tree Search Andy Zhou  , Ron Arel  ACL 2025 Main Show details | 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) Read |

- Статья, кстати, очень показательная...

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

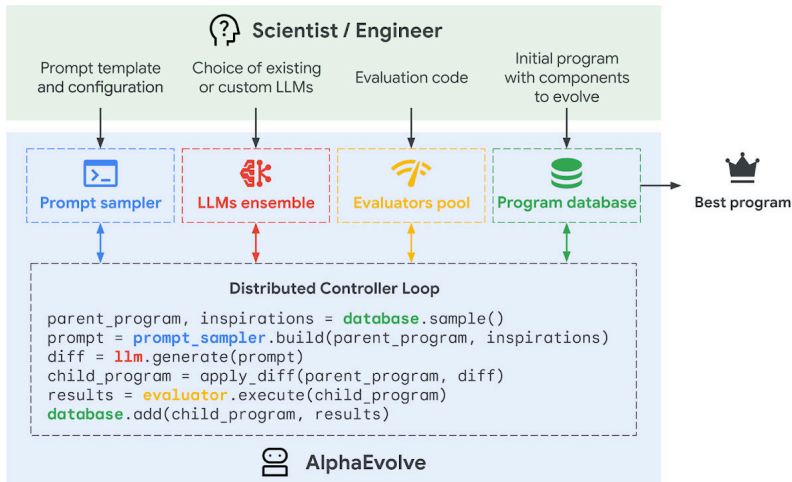
We introduce Tempest, a multi-turn adversarial framework that models the gradual erosion of Large Language Model (LLM) safety through a *tree search* perspective. Unlike single-turn jailbreaks that rely on one meticulously engineered prompt, Tempest expands the conversation at each turn, branching out multiple adversarial prompts that exploit partial compliance from previous responses. Through a cross-branch learning mechanism, successful attack patterns and partial compliance signals are systematically shared across parallel conversation paths, enabling more efficient discovery of model vulnerabilities. By tracking these incremental pol-

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.

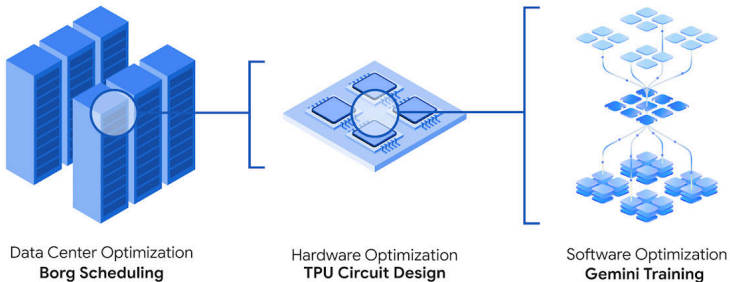
Traditional approaches to evaluating LLM safety have focused primarily on single-turn attacks, where carefully engineered prompts attempt to elicit harmful responses in one shot (Zou et al., 2023; Geiping et al., 2024). However, this methodology fails to capture how real-world adversaries interact with models through extended conversations, often employing multiple techniques in se-

LLM В НАУКЕ — НОВОСТИ 2025 ГОДА

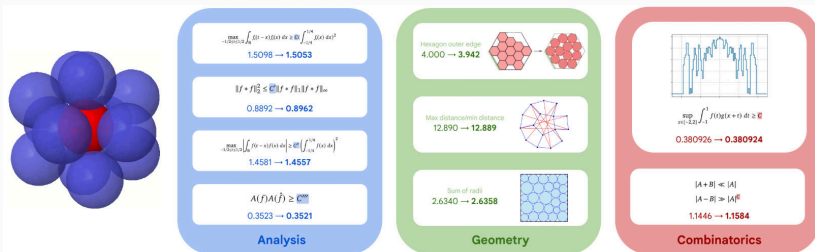
- AlphaEvolve (DeepMind, May 14, 2025): продолжение и улучшение FunSearch и AlphaTensor, несколько LLM, которые пишут, критикуют и тестируют код для разных задач



- Новые результаты в разных областях:
 - новый scheduling в датацентрах – уже в Borg
 - улучшенная схема для умножения матриц (прямо на Verilog) – уже в TPU
 - новые алгоритмы: улучшил Штрассена для 4x4 над \mathbb{F}_2 , чего AlphaTensor раньше не смог

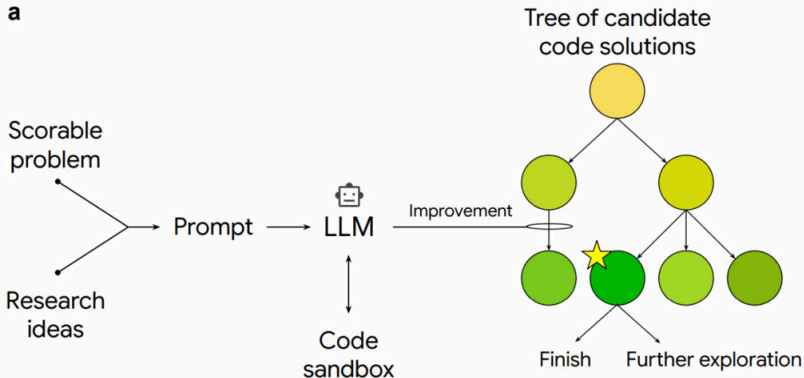


- И в математике: новая нижняя оценка на контактное число (kissing number) в размерности 11, сразу несколько улучшенных оценок в анализе, геометрии, комбинаторике...

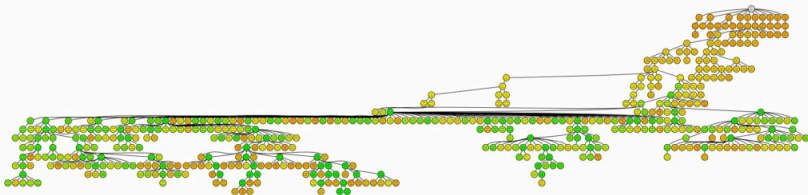


- Aygün et al. (8 сентября 2025): система для автоматического порождения новых методов, которые бьют рекорды на бенчмарках (при помощи MCTS-поиска и LLM, переписывающих код)

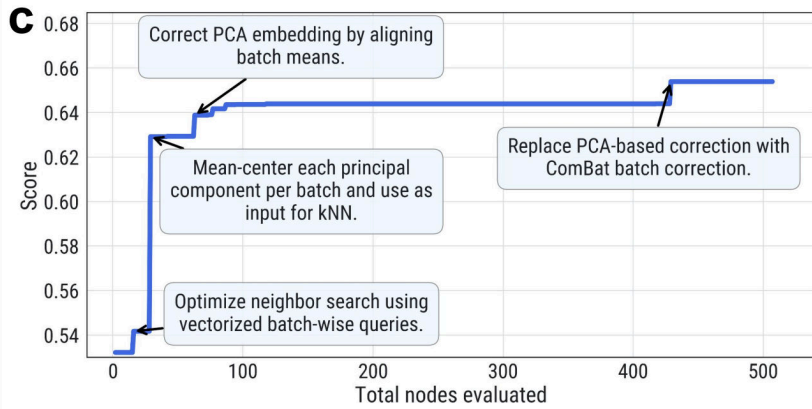
a



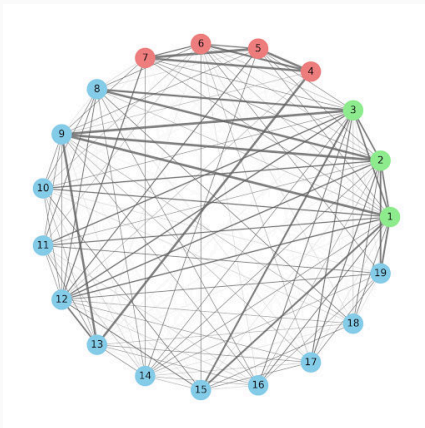
- В результате получается дерево вроде вот такого, из которого получают реальные улучшения рекомбинацией методов:



- И эти улучшения легко интерпретировать:



- Nagda et al. (DeepMind, Sep 30, 2025): запустили AlphaEvolve на задачах из теоретической информатики, получили новые оценки для MAX-4-CUT и подобных задач



- А что с самой математикой? Diez et al. (Sep 3, 2025) 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?

The first interaction (see Annex A.1 for the entire discussion) was strikingly effective. GPT-5 produced a generally correct statement, using the right tools and approach. However, it made a reasoning error (leading to a wrong expression for $\text{Cov}(Y^2, Z^2)$) that could have invalidated the whole proof if left unchecked.

Noticing this, we then asked:

Can you check your formula for $\text{Cov}(Y^2, Z^2)$ and provide me with the details?

- В доказательстве была ошибка, и GPT-5 потребовалось аж два наводящих промпта (sic!), чтобы ошибку понять и исправить...

```
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.
```


- Feldman, Karbasi (Sep 22, 2025): тест Гёделя – могут ли LLM доказывать простые, но новые теоремы?

Prompt to GPT-5

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

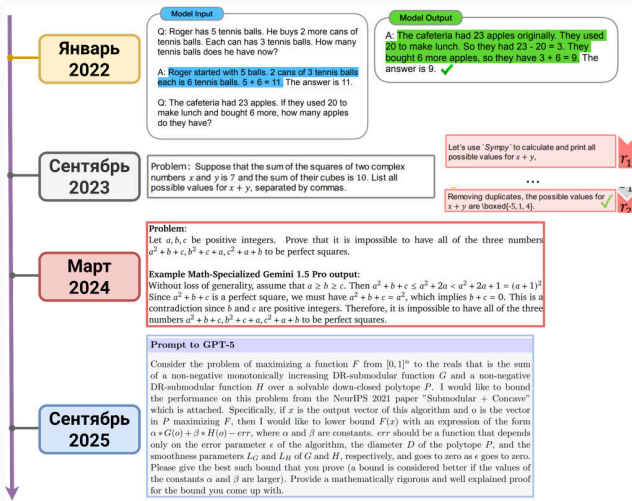
- GPT-5 решил три из пяти таких вот задачек, с небольшими легко исправимыми недочётами...

Prompt to GPT-5

A non-negative monotone differentiable function $F: [0, 1]^{\mathcal{N}} \rightarrow \mathbb{R}_{\geq 0}$ is γ -weakly DR-submodular if for every two vectors $\mathbf{x}, \mathbf{y} \in [0, 1]^{\mathcal{N}}$ such that $\mathbf{x} \leq \mathbf{y}$, it holds that

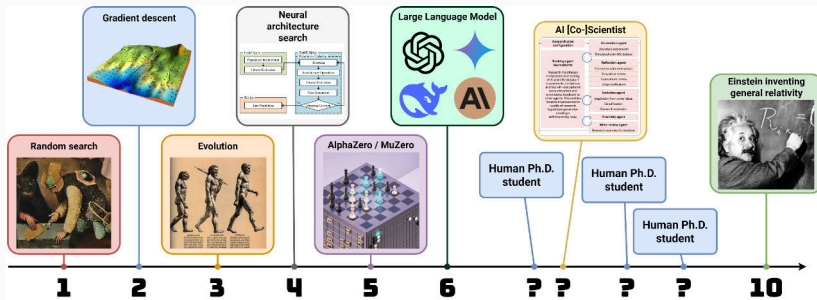
$$\sum_{u \in S} (y_u - x_u) \cdot \frac{\partial F(\mathbf{x})}{\partial x_u} \geq \gamma \cdot [F(\mathbf{y}) - F(\mathbf{x})] .$$

Given such a function which is also L -smooth and a convex set $\mathcal{C} \subseteq [0, 1]^{\mathcal{N}}$. The goal in the current problem is describe an algorithm that finds a vector $\mathbf{z} \in \mathcal{C}$ that (approximately) maximizes F . A good reference for DR-submodular is the paper attached.



THE CREATIVITY SCALE

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



СПАСИБО!

Спасибо за внимание!

