

ШТУЧНИЙ ІНТЕЛЕКТ І МАТЕМАТИЧНА ОСВІТА: СТАН І ПЕРСПЕКТИВИ

Раков С.А.

**Український центр оцінювання якості освіти,
представник України в Управлінській Раді PISA**

ШІ І МАТЕМАТИЧНА ОСВІТА

1. ШІ і математична освіта: поточний стан

- Результати соціологічних досліджень
- Розв'язування GPT-4 задач PISA (відкритий сегмент)
- Розв'язування GPT-4 задач PISA (закритий сегмент)
- Спалах AGI (ШІ загального призначення): перші експерименти з GPT-4

2. ШІ і математика

- Алгоритм перевірки інтелектуальності ШІ (протокол діалогу з GPT-4)
- Автоматичне доведення теорем (історія, стан перспективи)
- Автоматичне доведення теорем і ШІ (діалог з GPT-4)
- Розв'язування GPT-4 задачі Міжнародних математичних задач ІМО-2022
- Розв'язування GPT-4 задачі МТНК (кількість сестер і братів у родині)

3. ШІ і математична освіта: перспективи (пропозиції)

4. Література

ШІ І МАТЕМАТИЧНА ОСВІТА

Всеукраїнське дослідження
використання ШІ у
шкільній освіті

М.З.Н.

PRJCTR

FACTUM
GROUP



FACTUM

https://factum-ua.com/document/Всеукраїнське_дослідження_використання_ШІ_у_шкільній_освіті.pdf

ШІ І МАТЕМАТИЧНА ОСВІТА

Всеукраїнське дослідження використання ШІ у шкільній освіті

FACTUM
GROUP

CREATIVE
& TECH
PROJECTR
ONLINE
INSTITUTE

M.Z.H.
ПЛАТФОРМА

FACTUM

2

Про дослідження

M.Z.H.

PROJECTR

FACTUM
GROUP

1

2

3

Цільова аудиторія

У межах дослідження проаналізовано
2 цільові аудиторії:

- 1. Вчителі**, які викладають в школі зараз, на момент опитування
- 2. Учні**, які на момент опитування навчаються в **8-11 класі** (включаючи тих, хто виїхав за кордон)

Географія, Метод

Географія дослідження: **Вся Україна**

Метод опитування: CAWI (онлайн, самозаповнення)

Розмір вибірки:

- 1. Вчителі, N = 1747** респондентів
- 2. Учні, N = 1443** респондентів

Термін опитування: Вересень-Жовтень 2023

Формування Вибірки

Польовий етап дослідження виконаний Projector Creative & Tech Institute за підтримки Малою академії наук України.

Google Форма анкети поширювалась каналами в соцмережах, розсилками для вчителів та дітей, каналами Міносвіти та інших партнерів

FACTUM

ШІ І МАТЕМАТИЧНА ОСВІТА

Всеукраїнське дослідження використання ШІ у шкільній освіті



КЛЮЧОВІ РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

- Знання сервісів ШІ
- Використання ШІ
- Ставлення до ШІ



ЧАСТИНА 1 ПОГЛЯД **ВЧИТЕЛІВ** НА ШТУЧНИЙ ІНТЕЛЕКТ

1. Про які сервіси ШІ вчителі знають та які з них використовували
2. Як саме вчителі використовували ШІ, чи задоволені досвідом
3. Чи готові вчителі рекомендувати ШІ своїм колегам та учням
4. Профіль аудиторії вчителів



ЧАСТИНА 2 ПОГЛЯД **УЧНІВ** НА ШТУЧНИЙ ІНТЕЛЕКТ

1. Про які сервіси ШІ учні знають та які з них використовували
2. Як саме учні користувались ШІ, чи задоволені, які ідеї мають
3. Погляд учнів на застосування ШІ вчителями
4. Профіль аудиторії учнів

ШІ І МАТЕМАТИЧНА ОСВІТА

Всеукраїнське дослідження використання ШІ у шкільній освіті

Ключові результати

ЗНАННЯ СЕРВІСІВ ШІ

- Найпопулярнішим сервісом штучного інтелекту (ШІ), з яким добре знайомі обидві аудиторії, є ChatGPT.
Рівень знання серед учнів дещо вищий, ніж серед вчителів (76% проти 68% відповідно).
- На другому місці за рівнем знання є інструмент ШІ від проєкту «На Урок». Про цей сервіс знають 49% опитаних вчителів, рівень знання серед учнів нижчий – 35%.
- І вчителі, і учні значно менше обізнані про такі сервіси ШІ як Grammarly, Bard Google, Midjourney, Notion AI та Stable Diffusion.

КОРИСТУВАННЯ ШІ

- Більшість вчителів та учнів зазначили, що мають досвід користування сервісами ШІ.
У більшості випадків аудиторія задоволена отриманим досвідом.
- Вчителі зазначають, що використовували сервіси ШІ в своїй діяльності з метою підготовки до занять, створення тестів для домашнього завдання, при проведенні занять, а також для перевірки знань учнів і, навіть, в позакласній роботі. Частина вчителів залучала учнів до використання ШІ.
- Учні, говорячи про досвід користування ШІ, найчастіше зазначали його допомогу при підготовці домашнього завдання.

СТАВЛЕННЯ ДО ШІ

- Думки вчителів щодо ШІ різняться. Вони розуміють, що це новітні технології та може бути корисно, але мають вагання через можливі проблеми. Вчителі хвилюються, що ШІ може обмежити розвиток учнів, призвести до списування та неетичного використання. Є і те, що вчителі самі ще не зовсім розуміють, як працює ШІ, і вони побоюються помилок в його роботі, оскільки технологія ще лише розвивається, недостатньо вивчена.
- Щодо списування, вчителі мають підстави для занепокоєнь, оскільки деякі учні справді визнають, що використовують ШІ для цього. Проте більшість учнів вважають, що ШІ може, навпаки, поліпшити навчання та розвиток учнів, зробити освітній процес цікавим та захоплюючим.

Учні навіть виражають бажання отримати від вчителів інформацію про те, як правильно використовувати ШІ етично, які є його переваги та недоліки, особливості застосування.

https://factum-ua.com/document/Всеукраїнське_дослідження_використання_ШІ_у_шкільній_освіті.pdf

ШІ І МАТЕМАТИЧНА ОСВІТА

Рух до Ефективної Цифрової Освітньої Екосистеми



OECD DIGITAL EDUCATION OUTLOOK 2023

Towards an Effective Digital Education Ecosystem

13 Співпраця та співтворчість багатьох зацікавлених сторін: до відповідального застосування AI в освіті	316
Вступ	317
Співпраця та співтворчість як пропозиція подолання відокремленості	318
Спільна мова для зацікавлених сторін	320
Глобальний ландшафт інноваційних лабораторій та експертних центрів, які підтримують розвиток AI в освіті по всьому світу	322
Висновок: куди нам йти далі?	326
Список літератури	327
Додаток 13.A. Описи лабораторій, експертних центрів та розвитку бізнесу	329

Part III Guidelines **395**

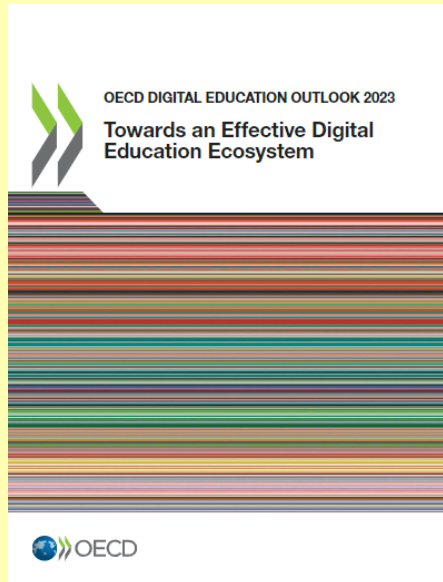
16 Opportunities, guidelines and guardrails for effective and equitable use of AI in education	396
Introduction	397
Opportunities of AI and digital technology	397
Risks of AI and digital technology	398

ШІ І МАТЕМАТИЧНА ОСВІТА

Інноваційні лабораторії і дослідницьких центри з питань ШІ в освіті (світовий досвід)

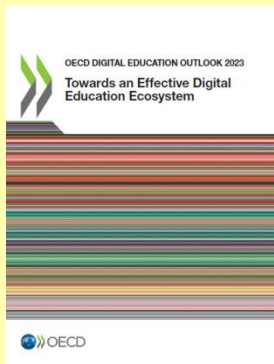
Table 13.1. Description of different innovation labs and research centres around the globe

Name ⁴	Country	Partners	Goals	Category	Duration	Size
National Education Lab AI (NOLAI)	The Netherlands	Universities, School boards and Business development agencies	National expert centre that develops innovative intelligent technologies and scientific knowledge of responsible usage of AI in education with the aim to improve the quality of primary and secondary education	Schools, science and industry combined product development	2022 for 10 years	80 million 30 to 40 FTE
IMEC Smart Education	Belgium	Universities, national research and innovation hub	Strategic research and innovation programme that aims to develop state-of-the-art technologies in order to address grand challenges in education and training.	Schools, science and industry combined product development	2017 for 9 years	40 FTE 1 million seed funding
Tübingen Center for Digital Education (TüCeDE)	Germany	University departments, independent research institute	Expert centre that aim to strengthen the transfer expertise in digital education within the university and toward new teachers.	Science for Teacher Professionalization	2022 for 5 years	12.2 million 30 FTE
Engage AI Institute	USA	Universities and non-profit organisations	Research institute to deepen engagement and advance learning by creating a new class of narrative-centred learning environments in which students can collaboratively engage with customised plots, synthesised characters, and realistic forms of interaction.	Science for product development	2022 for 5 years	20 million 30-40 FTE
National Artificial Intelligence Institute in Adult Learning and Online Education (AI-ALOE)	USA	Universities, non-profit organisations and corporations	National institute to expand access to quality jobs and improving workforce reskilling and upskilling by applying the affordances of AI to transform online education for adult learners.	Science for product development	2022 for 5 years	20 million 50 fete



ШІ І МАТЕМАТИЧНА ОСВІТА

Частина III Рекомендації



16. Можливості, рекомендації і запобіжники для ефективного і справедливого використання ШІ в освіті

Вступ

Можливості ШІ і цифрових технологій

Ризики ШІ і цифрових технологій

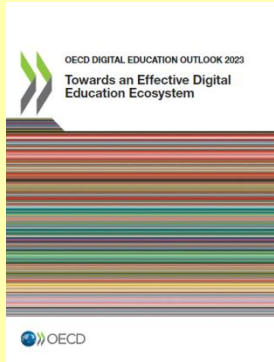
Генеративні програми штучного інтелекту (такі як великі мовні моделі) представляють як можливості, невідомі та ризики. Вони можуть допомогти вчителям створити проекти планів уроків і надати можливість розвивати критичне мислення своїх учнів у класі.

Ці програми можуть підтримувати зміну педагогічних моделей від того, щоб учні отримували відповіді, до допомоги їм у постановці правильних запитань, орієнтуванні в двозначності та конкуруючих твердженнях і розрізненні фактів від думок. Оскільки технологія продовжує розвиватися, вона може не тільки стати потужним інструментом навчання для учнів і зручним помічником для вчителів, а й сприяти збагаченню спеціальних технологічних рішень, таких як адаптивні системи навчання або індивідуальні відповіді ШІ на запитання учнів на основі їхнього навчального шляху.

Все це залежить від того, чи вчителі та учні матимуть спроможність критично оцінювати та коригувати те, що створює ШІ.

ШІ І МАТЕМАТИЧНА ОСВІТА

Частина III Ризики ШІ і цифрових технологій



16. Можливості, рекомендації і запобіжники для ефективного і справедливого використання ШІ в освіті

Вступ

Можливості ШІ і цифрових технологій

Ризики ШІ і цифрових технологій

Існують також ризики використання цифрових технологій в освіті:

- 1. Нерівність може виникнути внаслідок нерівного доступу до технологій або більшої ефективності цих інструментів для більш успішних учнів;**
- 2. Інструменти з підтримкою ШІ можуть бути більш придатними для деяких предметів, враховуючи, що не всі форми знань або процеси навчання можна легко перенести в цифровий формат. Це може призвести до пріоритетних форм навчання, які легко оцифрувати, що поставить під загрозу широту навчальної програми та якість освіти.**
- 3. Збільшення використання технологій також може призвести до атрофії людських навичок і свободи волі, а також до збільшення залежності від доступності ШІ та інших технологій, у тому числі для навичок, необхідних для успіху та благополуччя. Особливо це стосується генеративного ШІ, який тягне за собою деякі інші ризики, такі як надійність і відстежуваність інформації, ризик культурної упередженості та створює нові проблеми для традиційних оцінок.**

ШІ І МАТЕМАТИЧНА ОСВІТА

Тестування ШІ: Як порівнюються вправність GPT і 15-річних учнів на завданнях PISA?



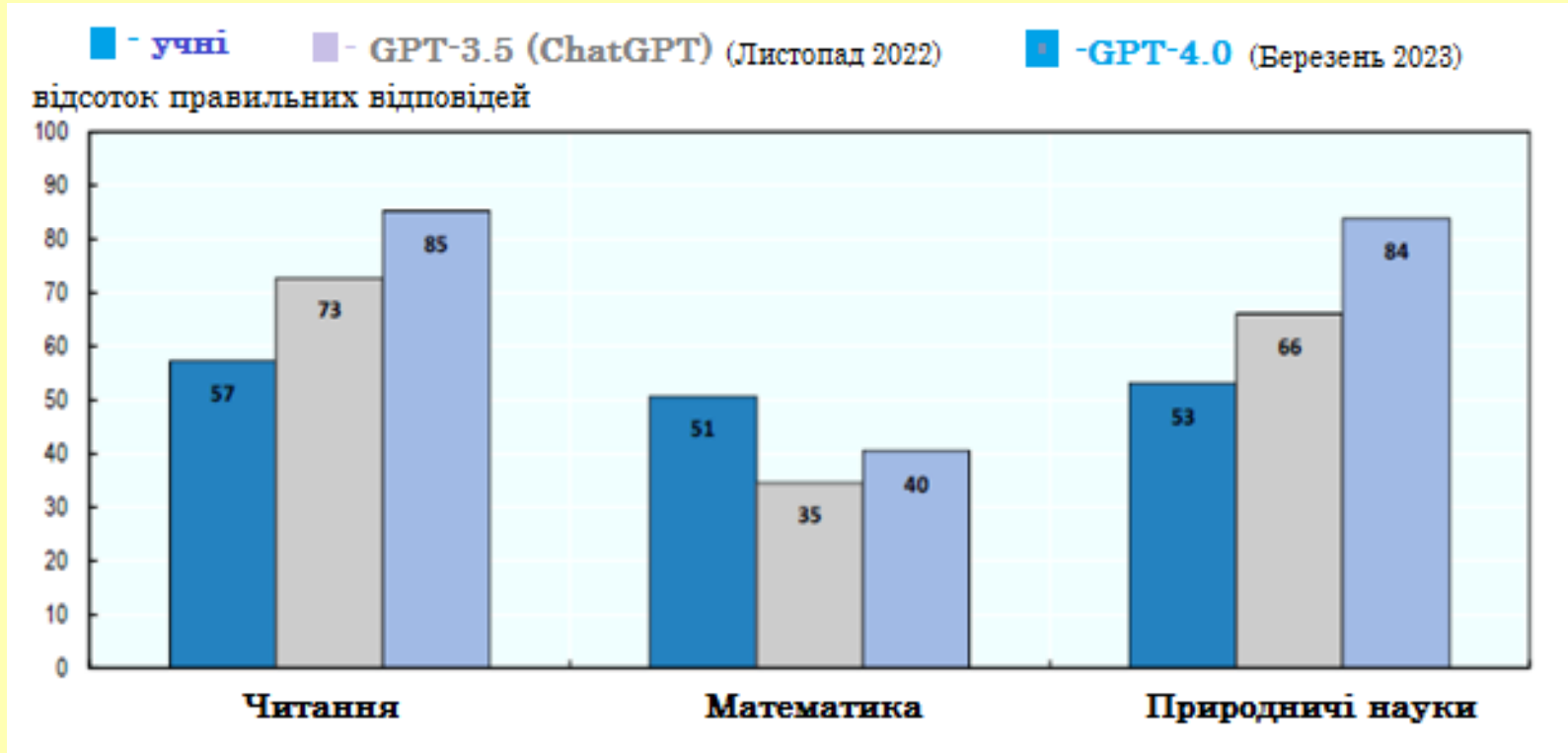
Putting AI to the test: How does the performance of GPT and 15-year-old students in PISA compare?

<https://www.oecd-ilibrary.org/docserver/2c297e0b-en.pdf?expires=1705505411&id=id&accname=guest&checksum=ADDDBCA01524EB3DF0A9077ECBD74225>

Посилання публікації дають доступ до матеріалів актуальних досліджень з різних аспектів застосування ШІ в освіті.

ШІ І МАТЕМАТИЧНА ОСВІТА

Вправність ШІ і учнів у головних доменах PISA на відкритій базі завдань тесту PISA.



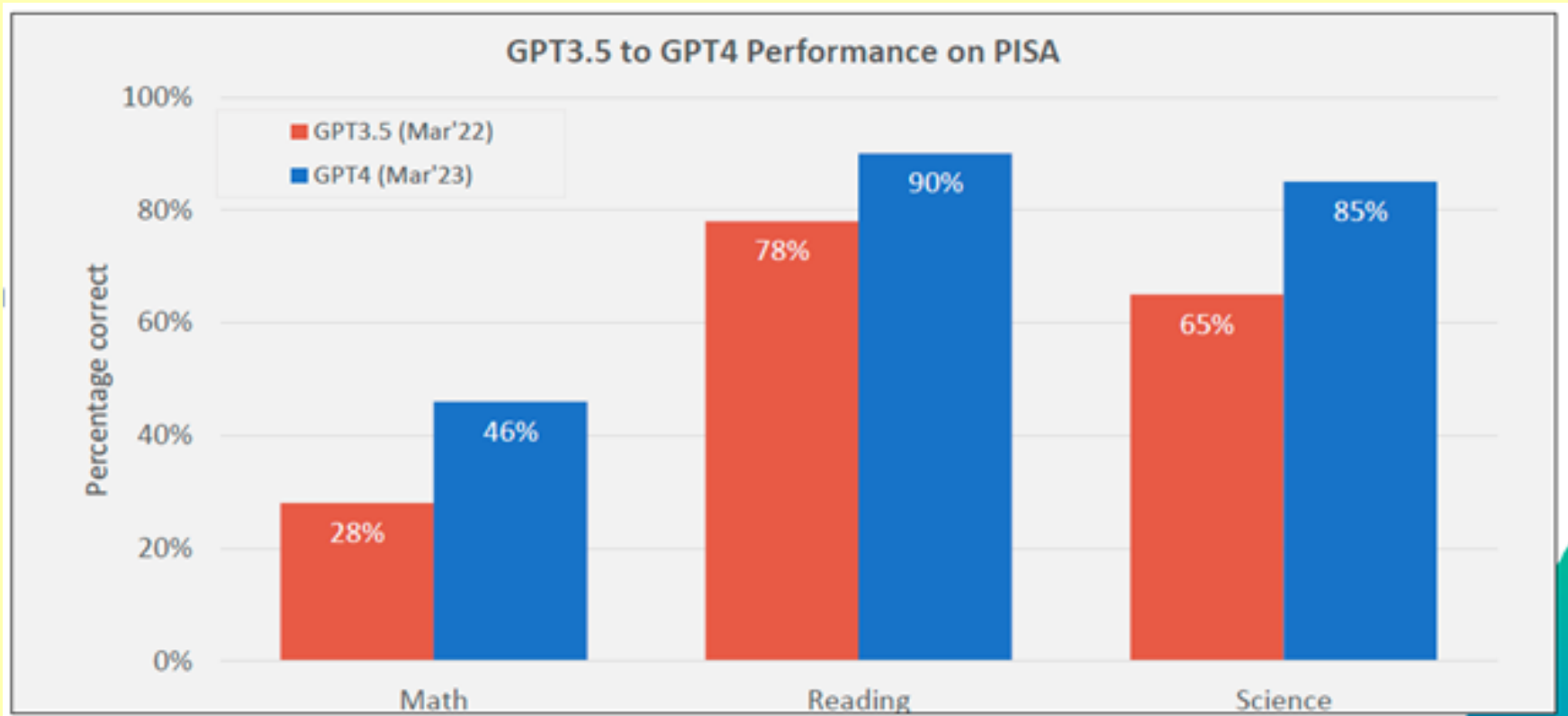
Завдання обрані з відкритої частини бази даних завдань PISA (44 завдання з читання, 42 - з математики, 34 - з природничих дисциплін).

Результати учнів – це середній рівень відсотка правильно розв'язаних задач учасників дослідження PISA.

Немає ознак для сподівань, що людина переможе ШІ у головних доменах PISA.

ШІ ІМАТЕМАТИЧНА ОСВІТА

Успішність GPT у розв'язуванні завдань дослідження PISA



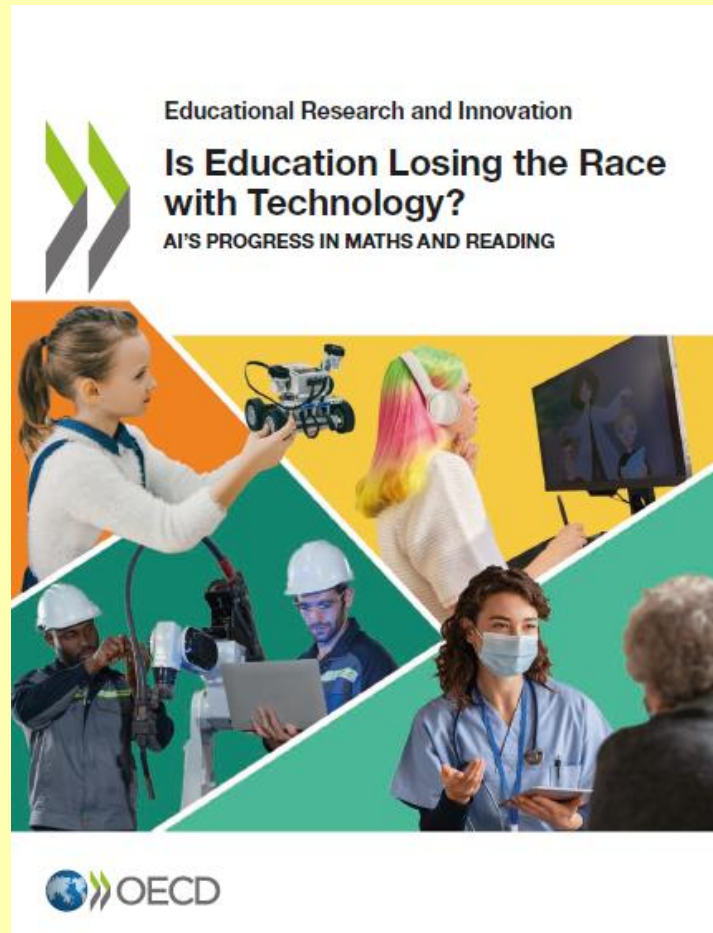
Презентація А.Шлейхера на 56 засіданні РГВ PISA (листопад 2023 р.)

Прогрес успішності GPT на завданнях PISA.

ШІ І МАТЕМАТИЧНА ОСВІТА

Чи програє Освіта у змаганнях з технологіями?

Прогрес ШІ у математиці і читанні.



<https://www.oecd-ilibrary.org/docserver/5ee71f34-en.pdf?expires=1703757034&id=id&accname=guest&checksum=B95E29EA3E7E18C800D74CE0E7A78CA0>

Звіт про результати порівняльного дослідження прогресу ШІ в розв'язуванні задач Міжнародного дослідження компетентності дорослих РІААС з математики і читання.

ШІ І МАТЕМАТИЧНА ОСВІТА



Educational Research and Innovation

AI and the Future of Skills, Volume 1

CAPABILITIES AND ASSESSMENTS



ШІ і майбутнє компетентностей (умінь)

Том 1

Здатності та оцінювання

Part I. Setting the scene	17
1. New approaches to understanding the impact of computers on work and education	18
Introduction	19
OECD's work on the impact of artificial intelligence	20
Attempts to measure artificial intelligence capabilities and impact	21
Purpose and structure of the report	27
References	31
Notes	33
2. Evaluating revolutions in artificial intelligence from a human perspective	34
Introduction	35
Artificial intelligence Revolution 1: Deep learning	35
Artificial intelligence Revolution 2: Knowledge graphs	37
Artificial intelligence Revolution 3: Reasoning	39
The coming fourth artificial intelligence Revolution: Integrated intelligence	41
Additional implications for measuring artificial intelligence progress	42
Conclusions	43
References	44
Notes	47

<https://www.oecd-ilibrary.org/docserver/See71f34-en.pdf?expires=1703757034&id=id&accname=guest&checksum=B95E29EA3E7E18C800D74CE0E7A78CA0>

Системне дослідження майбутнього ШІ, його засад, стану та перспектив удосконалення, порівняння ШІ з людським і їх оцінювання.

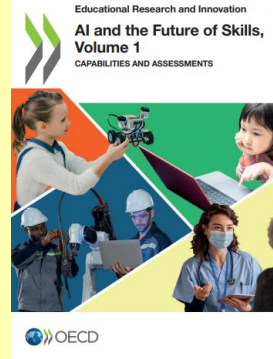
ШІ І МАТЕМАТИЧНА ОСВІТА

ШІ і майбутнє компетентностей (умінь)

Том 1

Здатності та оцінювання

Part IV. Reflections and a pragmatic way forward	296
18. Tasks and tests for assessing artificial intelligence and robotics in comparison with humans	297
Introduction	298
What is the value in identifying ideal models when comparing humans and artificial intelligence/robotic systems?	299
How might systematic mapping occur between skill taxonomies, tasks, tests and functional artificial intelligence components?	300
How can major differences be handled in targeted skills, different occupations and changes in the world?	302
Recommendations	303
References	304
19. Questions to guide assessment of artificial intelligence systems against human performance	306
Introduction	307
Why measure or characterise artificial intelligence systems?	307
Issues in organisation and task selection	308
Rating process	309
Answering the questions	312
Further reading	313
20. Building an assessment of artificial intelligence capabilities	315
Introduction	316
Human vs. artificial intelligence taxonomies and tests	316
Testing capabilities vs. tasks	317
Working with both dimensions	317
Filling in the details	318
Fitting the details into the framework	320
Next steps	321
From assessment to implications	322



<https://www.oecd-ilibrary.org/docserver/5ee71f34-en.pdf?expires=1703757034&id=id&accname=guest&checksum=B95E29EA3E7E18C800D74CE0E7A78CA0>

Системне дослідження майбутнього ШІ, його засад, стану та перспектив удосконалення, порівняння ШІ з людським і їх оцінювання.

ШІ І МАТЕМАТИЧНА ОСВІТА

ШІ і майбутнє компетентностей (умінь)

Том 1



| 137

Table 7.1. A comparison of large-scale assessment frameworks

Skill	Definition	Sub-process	Examples of large-scale assessments assessing these skills (and labels used in the assessment)
Mathematical literacy	An individual's capacity to formulate, employ and interpret mathematics in a variety of contexts.	<ul style="list-style-type: none">• Formulate mathematics• Employ mathematics• Interpret mathematical results	PISA (Mathematical literacy) PIAAC (Numeracy) NAEP (Mathematics) TIMSS (Mathematics) GRE (Quantitative fluid reasoning) SAT (Mathematics)
Reading literacy	The ability to make use of written texts, to achieve one's goals, to increase one's knowledge and potential, and to participate in society.	<ul style="list-style-type: none">• Access and retrieve• Integrate and interpret• Reflect and evaluate	PISA (Reading literacy) PIAAC (Literacy) NAEP (Reading) GRE (Verbal fluid reasoning; analytical writing) SAT (English; Languages)
Science literacy	The ability to engage with science-related issues and with the ideas of science.	<ul style="list-style-type: none">• Explain phenomena scientifically• Evaluate and design scientific enquiry• Interpret data and evidence scientifically	PISA (Science literacy) NAEP (Science) TIMSS (Science) SAT (Science)

<https://www.oecd-ilibrary.org/docserver/5ee71f34-en.pdf?expires=1703757034&id=id&accname=guest&checksum=B95E29EA3E7E18C800D74CE0E7A78CA0>

Математична грамотність: Здатність людини формулювати (будувати математичну модель), застосовувати та інтерпретувати математику в різноманітних контекстах.

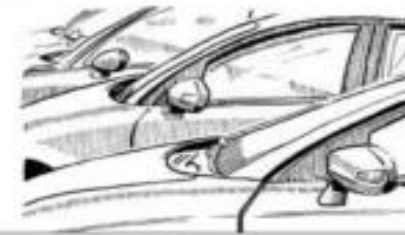
Figure 7.1. Sample item assessing mathematical literacy

WHICH CAR? – a unit from the PISA 2012 main survey

WHICH CAR?

Chris has just received her car driving licence and wants to buy her first car. This table below shows the details of four cars she finds at a local car dealer.

Model:	Alpha	Bolte	Castel	Dezal
Year	2003	2000	2001	1999
Advertised price (zeds)	4 800	4 450	4 250	3 990
Distance travelled (kilometres)	105 000	115 000	128 000	109 000
Engine capacity (litres)	1.79	1.796	1.82	1.783



WHICH CAR? – QUESTION 1

Chris wants a car that meets **all** of these conditions:

- The distance travelled is **not** higher than 120 000 kilometres.
- It was made in the year 2000 or a later year.
- The advertised price is **not** higher than 4 500 zeds.
- Which car meets Chris's conditions?

- A. Alpha
- B. Bolte
- C. Castel
- D. Dezal

WHICH CAR? – QUESTION 2

Which car's engine capacity is the smallest?

- A. Alpha
- B. Bolte
- C. Castel
- D. Dezal

WHICH CAR? – QUESTION 3

Chris will have to pay an extra 2.5% of the advertised cost of the car as taxes.

How much are the extra taxes for the Alpha?

Extra taxes in zeds:

Source: OECD (2014_[16]).



ШІ ІМАТЕМАТИЧНА ОСВІТА

Спалах AGI (ШІ загального призначення): перші експерименти з GPT-4

Sparks of Artificial General Intelligence: Early experiments with GPT-4

Sébastien Bubeck Varun Chandrasekaran Ronen Eldan Johannes Gehrke
Eric Horvitz Ece Kamar Peter Lee Yin Tat Lee Yuanzhi Li Scott Lundberg
Harsha Nori Hamid Palangi Marco Tulio Ribeiro Yi Zhang
Microsoft Research

Abstract

Artificial intelligence (AI) researchers have been developing and refining large language models (LLMs) that exhibit remarkable capabilities across a variety of domains and tasks, challenging our understanding of learning and cognition. The latest model developed by OpenAI, GPT-4 [Open23], was trained using an unprecedented scale of compute and data. In this paper, we report on our investigation of an early version of GPT-4, when it was still in active development by OpenAI. We contend that (this early version of) GPT-4 is part of a new cohort of LLMs (along with ChatGPT and Google's PaLM for example) that exhibit more general intelligence than previous AI models. We discuss the rising capabilities and implications of these models. We demonstrate that, beyond its mastery of language, GPT-4 can solve novel and difficult tasks that span mathematics, coding, vision, medicine, law, psychology and more, without needing any special prompting. Moreover, in all of these tasks, GPT-4's performance is strikingly close to human-level performance, and often vastly surpasses prior models such as ChatGPT. Given the breadth and depth of GPT-4's capabilities, we believe that it could reasonably be viewed as an early (yet still incomplete) version of an artificial general intelligence (AGI) system. In our exploration of GPT-4, we put special emphasis on discovering its limitations, and we discuss the challenges ahead for advancing towards deeper and more comprehensive versions of AGI, including the possible need for pursuing a new paradigm that moves beyond next-word prediction. We conclude with reflections on societal influences of the recent technological leap and future research directions.

**Дослідження
мультимодальних
здатностей
(зображення, музика,
міждисциплінарні),
програмування,
математика різних
рівнів (зокрема рівня
Міжнародних
математичних
олімпіад школярів)**

Системне дослідження спроможностей GPT з розв'язування математичних задач (проект Microsoft) .

ШІ ІМАТЕМАТИЧНА ОСВІТА

ChatGPT: Революційний засіб для навчання і вивчення математики

MODESTUM





EURASIA Journal of Mathematics, Science and Technology Education, 2023, 19(7), em2286
ISSN:1305-8223 (online)

OPEN ACCESS

Research Paper

<https://doi.org/10.29333/ejmste/13272>

ChatGPT: A revolutionary tool for teaching and learning mathematics

Yousef Wardat ^{1*} , Mohammad A. Tashtoush ^{2,3} , Rommel AlAli ⁴ , Adeb M. Jarrah ⁵ 

¹ Higher Colleges of Technology, Abu Dhabi, UAE

² Faculty of Education & Arts, Sohar University, Sohar, OMAN

³ Al-Huson University College, Al-Balqa Applied University, Salt, JORDAN

⁴ King Faisal University, Al-Ahsa, SAUDI ARABIA

⁵ Emirates College for Advanced Education, Abu Dhabi, UAE

Received 10 April 2023 ▪ Accepted 03 May 2023

<https://www.ejmste.com/download/chatgpt-a-revolutionary-tool-for-teaching-and-learning-mathematics-13272.pdf>

Дослідження поглядів учнів і вчителів на використання ШІ у навчанні математики після запуску ChatGPT: аналіз контенту інтерв'ю та дослідження досвіду користувача.

ШІ І МАТЕМАТИЧНА ОСВІТА

Математика, текстові задачі, загальні смисли і ШІ.

Mathematics, word problems, common sense, and artificial intelligence

Ernest Davis
Dept. of Computer Science
New York University
New York, NY 10012
davis@cs.nyu.edu

January 26, 2023

Abstract

The paper discusses the capacities and limitations of current artificial intelligence (AI) technology to solve word problems that combine elementary knowledge with commonsense reasoning. No existing AI systems can solve these reliably. We review three approaches that have been developed, using AI natural language technology: outputting the answer directly, outputting a computer program that solves the problem, and outputting a formalized representation that can be input to an automated theorem verifier. We review some benchmarks that have been developed to evaluate these systems and some experimental studies. We discuss the limitations of the existing technology at solving these kinds of problems. We argue that it is not clear whether these kinds of limitations will be important in developing AI technology for pure mathematical research, but that they will be important in applications of mathematics, and may well be important in developing programs capable of reading and understanding mathematical content written by humans.

У статті обговорюються можливості та обмеження поточної технології ШІ для розв'язування текстових задач, які поєднують елементарні знання зі здоровим глуздом. Жодна існуюча система ШІ не може надійно вирішити ці проблеми. Ми розглядаємо три підходи, які були розроблені з використанням технології природної мови ШІ: пряме виведення відповіді, виведення комп'ютерної програми, яка розв'язує задачу, і виведення формалізованого представлення (міркувань), яке можна ввести в **автоматизований верифікатор теорем**. Ми розглядаємо деякі тести, які були розроблені для оцінки цих систем, і деякі експериментальні дослідження. Ми обговорюємо обмеження існуючої технології у вирішенні таких проблем.

Стверджуємо, що це не зрозуміло, чи будуть такі обмеження важливими для розробки технології ШІ для суто математичних досліджень, але **вони будуть важливими для застосування математики і ще більш важливими для розробки програм, здатних читати та розуміти математичний вміст, написаний людьми.**

ШІ ІМАТЕМАТИЧНА ОСВІТА

Огляд ТРАНСФОРМЕР – Моделей для задач NLP (Natural Language Processing)



Proceedings of the Federated Conference on
Computer Science and Information Systems pp. 179–183 ISSN 2300-5963 ACSIS, Vol. 21

DOI: 10.15439/2020F20

Overview of the Transformer-based Models for NLP Tasks

Anthony Gillioz
University of Neuchâtel
Neuchâtel, Switzerland
Email: anthony.gillioz@unine.ch

Jacky Casas, Elena Mugellini, Omar Abou Khaled
University of Applied Sciences and Arts Western Switzerland
Fribourg, Switzerland
Email: {firstname.lastname}@hes-so.ch

https://annals-csis.org/Volume_21/drp/pdf/20.pdf

In 2017, Vaswani et al. proposed a new neural network architecture named Transformer. That modern architecture quickly revolutionized the natural language processing world. Models like GPT and BERT relying on this Transformer architecture have fully outperformed the previous state-of-the-art networks. It surpassed the earlier approaches by such a wide margin that all the recent cutting edge models seem to rely on these Transformer-based architectures. In this paper, we provide an overview and explanations of the latest models. We cover the auto-regressive models such as GPT, GPT-2 and XLNET, as well as the auto-encoder architecture such as BERT and a lot of post-BERT models like RoBERTa, ALBERT, ERNIE 1.0/2.0.

ШІ ІМАТЕМАТИЧНА ОСВІТА

Огляд ТРАНСФОРМЕР – Моделей для задач NLP (Natural Language Processing)



Proceedings of the Federated Conference on
Computer Science and Information Systems pp. 179–183 ISSN 2300-5963 ACSIS, Vol. 21

DOI: 10.15439/2020F20

Overview of the Transformer-based Models for NLP Tasks

Anthony Gillioz
University of Neuchâtel
Neuchâtel, Switzerland
Email: anthony.gillioz@unine.ch

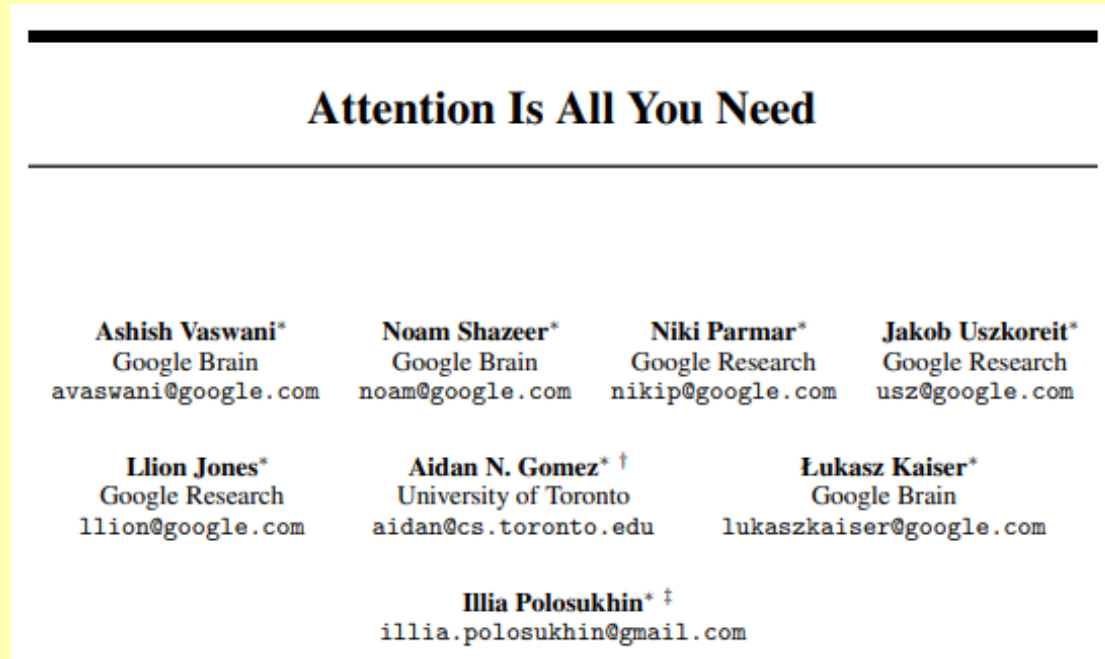
Jacky Casas, Elena Mugellini, Omar Abou Khaled
University of Applied Sciences and Arts Western Switzerland
Fribourg, Switzerland
Email: {firstname.lastname}@hes-so.ch

https://annals-csis.org/Volume_21/drp/pdf/20.pdf

У 2017 році Васвані та ін. запропонували нову архітектуру нейронної мережі під назвою Transformer. Ця сучасна архітектура швидко змінила світ обробки природної мови. Такі моделі, як GPT і BERT, що спираються на цю архітектуру Transformer, повністю перевершили попередні найсучасніші мережі. Вона перевершила попередні підходи з таким значним відривом, що всі останні новітні моделі, здається, покладаються на ці архітектури на основі Transformer. У цій статті ми надаємо огляд і пояснення останніх моделей. Ми розглядаємо моделі автоматичної регресії, такі як GPT, GPT-2 і XLNET, а також архітектуру автоматичного кодування, таку як BERT, і багато пост-BERT моделей, таких як RoBERTa, ALBERT, ERNIE 1.0/2.0.

ШІ МАТЕМАТИЧНА ОСВІТА

Увага – це все, що вам потрібно (пропозиція TRANSFORMER)



https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train.

ШІ ІМАТЕМАТИЧНА ОСВІТА

Увага – це все, що вам потрібно (пропозиція TRANSFORMER (2017))



https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

Моделі трансдукції домінуючої послідовності базуються на складних рекурентних або згорткових нейронних мережах, які включають кодер і декодер. Найефективніші моделі також з'єднують кодер і декодер через механізм уваги. **Ми пропонуємо нову просту мережеву архітектуру, Transformer, засновану виключно на механізмах уваги, повністю відмовляючись від повторень і згорток.** Експерименти з двома завданнями машинного перекладу показали, що ці моделі є кращими за якістю, водночас є більш розпаралелюваними та потребують значно менше часу для навчання.

Оцінка GPT 4 конструктивного алгоритму перевірки інтелектуальності LLM (протокол)

Запит

Оціни, будь ласка, наступне конструктивне визначення інтелектуальності LLM:

- 1. Уважаємо, що людина має інтелект;*
- 2. Авторитетна комісія (наприклад, все людство) задає одні й ті ж питання двом суб'єктам А (загально визнана високоінтелектуальна людина (умовний Альберт Ейнштейн або Бертран Рассел) і В (LLM), не знаючи , де людина, а де LLM, а потім, на основі їх відповідей, робить висновок, хто з них людина, а хто LLM;*
- 3. Якщо комісія не може визначити, які відповіді належать людині, а які LLM, то LLM має інтелект, у протилежному випадку – ні.*

GPT-4

Зазначене визначення інтелектуальності LLM (що ймовірно є аббревіатурою для Large Language Model, великої мовної моделі), нагадує Тьюрингів тест, головна мета якого визначити, чи машина може імітувати людську поведінку в бесіді настільки добре, що експерти не можуть відрізнити її від справжньої людини. За ваших умов, якщо авторитетна комісія не може зробити розрізнення між людиною і LLM на основі їхніх відповідей, вважається, що LLM проявляє інтелект.

Оцінка GPT 4 конструктивного алгоритму перевірки інтелектуальності LLM (протокол)

Ось деякі міркування з цього приводу:

Переваги:

Орієнтація на поведінку: Визначення оцінює інтелект LLM на основі його здатності імітувати інтелектуальну поведінку, не вимагаючи розуміння внутрішньої праці LLM.

Верифікація через експертів: Рішення приймається компетентною групою експертів, які вважаються достатньо освіченими, щоб здійснювати таку оцінку.

Універсальність: Тест може бути застосований до будь-якої мовної моделі без зміни основних принципів.

Оцінка GPT 4 конструктивного алгоритму перевірки інтелектуальності LLM (протокол)

Недоліки:

Антропоцентризм: Визначення базується на ідеї людського інтелекту як еталону, що може бути обмежуючим, оскільки інтелект може проявляти себе у різних формах, не обмежуючись людською поведінкою.

Вербальний Інтелект: Визначення фокусується виключно на вербальному інтелекті, ігноруючи інші важливі аспекти, наприклад, візуальне сприйняття, моторні навички, емоційний інтелект та інші.

Об'єктивність оцінювання: Незважаючи на наявність "авторитетної комісії," оцінка може бути суб'єктивною, оскільки людське судження часто схильне до помилок, переконань і упереджень.

Під час застосування такого визначення варто також враховувати, що інтелект не обмежується лише спосібністю до вербального спілкування, і що велика мовна модель може демонструвати видатні результати в тестах на інтелектуальності, не маючи при цьому самосвідомості, інтенції, чи розуміння сенсу наданих відповідей.

Автоматичне доведення теорем і ШІ



математическая
логика
и основания
математики

Ч. Чень, Р. Ли

МАТЕМАТИЧЕСКАЯ
ЛОГИКА
И АВТОМАТИЧЕСКОЕ
ДОКАЗАТЕЛЬСТВО
ТЕОРЕМ

Чень Ч., Ли Р. Математическая логика и автоматическое доказательство теорем: Пер. с англ./Под ред. С. Ю. Маслова — М.: Наука, Главная редакция физико-математической литературы, 1983. — 360 с

Книга посвящена детальному изложению всего круга проблем, связанных с так называемым методом резолюций. Этот метод наиболее известен и широко используется в современных работах по доказательству на ЭВМ математических теорем и вообще при построении систем «искусственного интеллекта». Описываются применения метода к таким, например, актуальным для всякого системного программиста задачам, как автоматический анализ и синтез программ. В приложении описаны другие методы и некоторые результаты последних лет, знакомство с которыми необходимо при изучении проблематики автоматического доказательства теорем.

Автоматичне доведення теорем і ШІ

Sergiy Rakov

Оціни інтелектуальність GPT

GPT-4

GPT (Generative Pretrained Transformer) технології, включаючи GPT-3 і останніші версії, є продвинутими системами штучного інтелекту, які використовують методи глибокого навчання для генерації тексту. Їх "інтелектуальність", якщо використовувати цей термін у неформальному сенсі, може бути оцінена за рядом різних параметрів:

Здатність до обробки мови (NLP): GPT-моделі здатні до здійснення завдань обробки природної мови на дуже високому рівні, включно із генеруванням тексту, перекладом, відповідями на запитання, створенням текстів на основі певних вимог, і так далі.

Адаптація до контексту: GPT демонструє здатність утримувати і використовувати контекстуальну інформацію протягом діалогу або текстової сесії, що дозволяє генерувати послідовні і зрозумілі відповіді.

Універсальне використання: така система не обмежена однією мовою чи доменом, а може обробляти текст на багатьох мовах і в широкому спектрі тем та галузей.

Загальна знання: GPT-моделі мають доступ до величезного обсягів текстової інформації, що дозволяє їм відповідати на запитання з різних областей знань, часто з високою точністю.

Автоматичне доведення теорем і ШІ

Sergiy Rakov

Оціни інтелектуальність GPT

GPT-4 (продовження)

Проте, важливо зазначити, що технології на зразок GPT не мають свідомості або реального розуміння, не здатні до самостійного навчання без людської допомоги та спираються на дані, на яких були натреновані. Їх "інтелектуальність" проявляється у вигляді розширеного аналізу мови та генерації тексту, але вони не мають здатності до творчого мислення, емпатії чи глибокого розуміння людського досвіду у спосіб, який ми асоціюємо з людською інтелектуальністю.

Автоматичне доведення теорем і ШІ

Sergiy Rakov

Оціни здатність GPT розв'язувати нестандартні математичні задачі, наприклад, задачі Міжнародних математичних олімпіад школярів

GPT-4

Технологія GPT (Generative Pretrained Transformer) демонструє значні можливості в області обробки природної мови, зокрема в розв'язуванні математичних завдань, які можна чітко сформулювати в текстовій формі. Однак здатність GPT розв'язувати складні та нестандартні математичні задачі, такі як ті, що зустрічаються на Міжнародних математичних олімпіадах школярів, залежить від кількох чинників:

Формулювання задачі: Якщо завдання з математичної олімпіади можна чітко та однозначно сформулювати у текстовій формі, GPT потенційно може допомогти знайти шлях розв'язку або надати повну відповідь.

Рівень абстракції: Нестандартні задачі часто вимагають глибокого розуміння теми та здатності абстрактно мислити. GPT може мати труднощі з такими задачами, оскільки вони вимагають більш гнучкого та творчого підходу, ніж стандартні алгоритмічні задачі.

Творчість та інтуїція: Математичні олімпіади часто включають задачі, що потребують творчих підходів та математичної інтуїції. Штучний інтелект на зразок GPT орієнтований на шаблони, які він вивчив із навчальних даних, тому може не виявляти оригінальних рішень або використовувати інтуїтивні методи розв'язку.

Автоматичне доведення теорем і ШІ

Sergiy Rakov

Оціни здатність GPT розв'язувати нестандартні математичні задачі, наприклад, задачі Міжнародних математичних олімпіад школярів

GPT-4 (продовження)

Обмеження тренувальних даних: Здатність GPT розв'язувати математичні задачі обмежена даними, на яких модель була натренована. Якщо модель не бачила достатньо прикладів подібних завдань під час навчання, її ефективність у роботі з такими задачами буде меншою.

Точність та перевірка: Для завдань на математичних олімпіадах потрібна висока точність у рішеннях. Але GPT може іноді робити помилки в розрахунках або втрачати ключові деталі, тому його рішення завжди варто перевіряти.

В загальному, хоча GPT має потенціал у роботі з математичними задачами, найпродуктивнішим буде застосування його як інструмента для підтримки або як джерело для генерації ідей, а не як повністю автономного розв'язника олімпіадних задач.

Автоматичне доведення теорем і ШІ

Запит

Чому не можна розробити програму для автоматичного доведення теорем?

GPT-4

Можливість розробки програми для автоматичного доведення теорем обмежується кількома фундаментальними теоретичними та практичними проблемами:

Теореми про неповноту Геделя: Згідно з першою теоремою Геделя про неповноту, жодна достатньо потужна формальна система, яка може описувати арифметику натуральних чисел, не може бути одночасно повною та несуперечливою. Це означає, що завжди існують істинні математичні твердження, які не можуть бути доведені в рамках такої системи. Тож програма автоматичного доведення не буде здатна впоратися з усіма можливими твердженнями, які можуть виникнути в математиці.

Проблема зупинки: Тьюринг довів, що проблема зупинки є алгоритмічно нерозв'язною, тобто неможливо наперед визначити, чи зупиниться будь-яка програма з будь-яким вводом чи ні. Це необхідно для програми автоматичного доведення теорем, оскільки для доведення деяких тверджень може знадобитися нескінченний час або ресурси.

Обчислювальна складність: Навіть коли теорема може бути доведена, простір пошуку доказу може бути надзвичайно великим. Задачі автоматичного доведення часто належать до класу складності NP-жорстких або навіть нерозв'язних, що робить швидке знаходження доказів практично неможливим для деяких теорем.

Автоматичне доведення теорем і ШІ

Запит

Чому не можна розробити програму для автоматичного доведення теорем?

GPT-4 (продовження)

Відсутність творчого мислення: На відміну від людини-математика, яка може винаходити нові підходи і концепції, поточні програми автоматичного доведення теорем мають працювати в межах визначених формальних правил і алгоритмів. Їм бракує здатності до творчості та інтуїції.

Варіативність математичних структур: Математичні теореми можуть стосуватися різноманітних структур та областей, кожна з яких вимагає свого підходу до доказу. Це означає, що універсальна програма для автоматичного доведення теорем повинна бути здатна обробити дуже широкий спектр правил, аксіом і концепцій.

Водночас, існують системи для автоматичного доведення теорем, які можуть бути дуже ефективними у певних областях математики та для особливих класів проблем. Вони можуть автоматично генерувати докази для тверджень, в межах певних формальних систем, та служити корисним інструментом для дослідників. Проте, головні теоретичні обмеження, поставлені теоремами Геделя про неповноту і проблемою зупинки Тьюринга, залишаються фундаментальними принципами, які визначають границі можливостей автоматичного доведення теорем.

Автоматичне доведення теорем і ШІ

Запит

Як довести, що будь-яка змістовна несуперечлива формальна теорія неповна?

GPT-4

Доведення того, що будь-яка значима (або змістовна) несуперечлива формальна теорія неповна, виходить з Геделевої першої теореми про неповноту. Ось спрощене викладення аргумента:

Формулювання теорії: Беремо до уваги формальну теорію, що може виразити основну арифметику, зокрема відносини порядку, додавання і множення. Така теорія повинна бути несуперечлива, що означає відсутність в ній одночасно істинних тверджень та їх заперечень.

Арифметизація: Наступним кроком є процес «арифметизації», в якому виразам, доказам та твердженням формальної теорії присвоюються унікальні арифметичні значення. Це дозволяє «говорити» про вирази та докази всередині самієї системи.

Побудова Геделевого речення: Гедель сконструював спеціальне речення (відоме як "Геделеве речення"), яке по суті стверджує: «Це речення не має доказу в рамках цієї системи». Через арифметизацію це речення цілком законне в межах системи і може бути представлено як звичайне арифметичне твердження.

Автоматичне доведення теорем і ШІ

Запит

Як довести, що будь-яка змістовна несуперечлива формальна теорія неповна?

GPT-4 (продовження)

Аналіз Геделевого речення: Якщо Геделеве речення істинне (це означає, воно дійсно не має доказу в системі), то система є неповною, тому що існує істинне твердження (Геделеве речення), яке не може бути доказаним в рамках системи. Якщо ж Геделеве речення може бути доведеним у системі, це призводить до протиріччя, оскільки речення стверджує власну недоказовість, і, отже, система має бути неконсистентною.

Висновок про неповноту: Виходячи з припущення про несуперечливість системи, Геделеве речення не може мати доказу, що означає його істинність. Але оскільки воно істинне і не має доказу, система є неповною.

Геделева робота посилається на конкретну класичну інтуїцію про "істинність", яка трансцендує формальні системи. Його аргументи є дуже технічними і вимагають глибокого розуміння математичної логіки і принципів конструювання та обробки формальних систем.

Автоматичне доведення теорем і ШІ

Запит

Якщо доведено алгоритмічну нерозв'язність задачі доведення теорем, то як може GPT розв'язувати задачі на доведення?

GPT-4

GPT (Generative Pretrained Transformer) та інші схожі моделі машинного навчання можуть підходити до задач доведення теорем відмінно від строгих алгоритмічних методів, які обмежені теоремами про неповноту та проблемою зупинки. Ось як це працює:

Хевристичний підхід: Моделі подібні до GPT використовують статистичні хевристики для генерації відповідей. Вони не "розв'язують" проблему у класичному понятті, а генерують відповіді, засновані на великих обсягах даних, на яких вони були навчені. Це означає, що GPT може генерувати те, що виглядає як доказ, але без гарантій його точності або повноти.

Патерн визнання: Моделі не шукають "доказ" у традиційному розумінні, але можуть ідентифікувати патерни та структури в математичних твердженнях та доказах, на яких вони були навчені, і використовувати ці паттерни для генерації нових висловлювань, які могли б вважатися доказами.

Асистування, а не вирішення: Такі системи, як GPT, можуть використовуватися для асистування у доведенні теорем, наприклад, пропонуючи можливі напрямки дослідження або запропонувавши кроки, які можуть вести до доведення. Вони можуть підказати математикам ідеї або навіть потенційні докази, які потім необхідно перевірити і формально верифікувати.

Автоматичне доведення теорем і ШІ

Запит

Якщо доведено алгоритмічну нерозв'язність задачі доведення теорем, то як може GPT розв'язувати задачі на доведення?

GPT-4 (продовження)

Обмежена область застосування: В деяких специфічних областях або на визначених класах проблем моделі машинного навчання можуть бути ефективними для автоматизованого доведення. Це, зокрема, трапляється, коли задача добре визначена і має чітку структуру, яку модель може досить точно навчитися репрезентувати.

Інтерпретація результатів: Люди повинні оглядати та інтерпретувати результати, генеровані AI. Модель може вивести доказ, що здається обґрунтованим, але це виведення може бути помилкове або вимагати додаткової перевірки.

Отже, незважаючи на проблеми з автоматичним доведенням теорем на алгоритмічному рівні, моделі, подібні до GPT, можуть допомагати у процесі доведення, використовуючи статистичні методи та велику кількість доступних їм даних. Така підтримка може бути корисною для математиків як допоміжний інструмент, але формальну верифікацію все одно зазвичай необхідно проводити людським експертам.

Розв'язок GPT-4 спрощеної задачі ІМО -2022

GPT-4

Prompt:

Suppose g is a continuous smooth function such that for every $x > 0$, there is **one and only one** $y > 0$ such that $g(x) + g(y) \leq 2xy$. Can you prove that $g(x) = x^2$?

GPT-4:

There are a few possible ways to approach this problem, but one possible method is as follows:

- Assume, for the sake of contradiction, that $g(x)$ is not equal to x^2 for some $x > 0$.
- Then, either $g(x) < x^2$ or $g(x) > x^2$ for that x .
- If $g(x) < x^2$, then let $y = x$. Then $g(x) + g(y) = 2g(x) < 2x^2 = 2xy$.
- If $g(x) + g(y) < 2xy$, then $g(x) + g(y)$ is strictly less than the upper bound of $2xy$ for positive x and y .
- This means that there is some room to increase y slightly and still satisfy the inequality $g(x) + g(y) \leq 2xy$.
- For example, let $y' = y + \epsilon$, where ϵ is a small positive number such that $g(x) + g(y') < 2xy$. This is possible because g is continuous and smooth, so it does not have any jumps or kinks that would prevent such a choice of epsilon.
- Then, y' is another positive number such that $g(x) + g(y') \leq 2xy'$, which contradicts the assumption that y is the one and only one positive number that satisfies this condition given x .
- Therefore, if $g(x) < x^2$, then $g(x) + g(y)$ cannot be strictly less than $2xy$, and it must be equal to $2xy$.
- If $g(x) > x^2$, then let y^* be the unique positive solution of $g(x) + g(y) = 2xy$. Then $g(y^*) < (y^*)^2$, since otherwise we would have $g(x) + g(y^*) > 2xy$, which contradicts the existence of y^* given x .
- But then, by the same argument as before, we can let $x = y = y^*$ and obtain a contradiction, since $g(y) + g(y) = 2g(y) < 2y^2 = 2yy$, which contradicts the uniqueness of y given y .
- Therefore, $g(x)$ must be equal to x^2 for all $x > 0$, as required.

Розв'язок GPT-4 спрощеної задачі ІМО -2022

Запит:

Припустимо g є неперервна гладка функція, така, що для кожного $x > 0$ існує один і тільки один $y > 0$, такий що $g(x) + g(y) \leq 2xy$. Чи можете ви довести, що $g(x) = x^2$?

GPT-4:

Існує декілька можливих підходів до цієї задачі, один із можливих методів такий:

- Припустимо супротивне:
 $g(x) \neq x^2$ для деякого $x > 0$.
- Тоді $g(x) < x^2$ або $g(x) > x^2$ для цього x .
- Якщо $g(x) < x^2$, то покладемо $y = x$. Тоді $g(x) + g(y) = 2g(x) < 2x^2 = 2xy$.
- Якщо $g(x) + g(y) < 2xy$, то $g(x) + g(y)$ *строго менше* верхньої грани для додатних x і y .
- Це означає, що існує деякий простір для невеликого збільшення y , при якому все ще буде виконуватись нерівність $g(x) + g(y) \leq 2xy$.
- Наприклад, покладемо $y_0 = y + \varepsilon$, де ε є мале додатне число, таке, що $g(x) + g(y) \leq 2xy$. Це можливо тому що g є неперервною і гладкою, тому вона не має ні скачків, ні злізів, які можуть зав'язати вибору такого ε .
- Тоді y_0 є іншим додатнім числом, для якого $g(x) + g(y_0) \leq 2xy_0$, що суперечить припущенню, що y є одне і тільки одне додатне число, що задовольняє цій умові для даного x .

- Тому, якщо $g(x) < x^2$, то $g(x) + g(y)$ не може бути строго менше за $2xy$, і має бути рівною $2xy$.
- Якщо $g(x) > x^2$, то нехай y^* єдине додатне рішення $g(x) + g(y) = 2xy$. Тоді $g(y^*) < y^{*2}$, оскільки інакше ми будемо мати $g(x) + g(y^*) > 2xy$, що суперечить існуванню y^* для даного x .
- Але тоді за тими самими аргументами, як і у попередньому випадку, ми можемо покласти $x = y = y^*$ і отримуємо суперечність, оскільки $g(y) + g(y) = 2g(y) < 2y^2 = 2yy$, що суперечить єдності y для даного x .
- Тому $g(x)$ має бути рівним x^2 для кожного $x > 0$, як і вимагалось.

Формулювання оригінальної Задачі 2 ІМО -2022

Оригінальна Задача 2 (ІМО 2022)

A3. Let $\mathbb{R}_{>0}$ be the set of positive real numbers. Find all functions $f: \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}$ such that, for every $x \in \mathbb{R}_{>0}$, there exists a unique $y \in \mathbb{R}_{>0}$ satisfying

$$xf(y) + yf(x) \leq 2.$$

(Netherlands)

Задача 2. Через \mathbb{R}^+ позначимо множину додатних дійсних чисел. Знайдіть усі функції $f: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ такі, що для всіх чисел $x \in \mathbb{R}^+$ існує єдине число $y \in \mathbb{R}^+$, для якого справджується нерівність

$$xf(y) + yf(x) \leq 2.$$

Спрощення Задачі 2 стосується перш за все того, що слід довести єдиність такої функції в класі неперервних і навіть гладких функцій, що суттєво спрощує задачу, перетворюючи її у задачу двоходівку зі стандартними кроками і міркуваннями.

Офіційний розв'язок Задачі 2 ІМО-2022 наведено на наступному слайді і за посиланням: (https://artofproblemsolving.com/wiki/index.php/2022_IMO_Problems/Problem_2).

Офіційний розв'язок Задачі 2 ІМО -2022

Answer: The unique solution is the function $f(x) = \frac{1}{x}$ for every $x \in \mathbb{R}^+$. This function clearly satisfies the required property since the expression $xf(y) + yf(x) = \frac{x}{y} + \frac{y}{x}$ is greater than 2 for every $y \neq x$ (directly from AM-GM) and equal to 2 (with equality) for the unique value $y = x$.

Proof: Let's consider a solution based on some ideas we encountered in the preparation classes for the Olympiad, specifically involving auxiliary sets and functions with specific properties.

The fact that for every $x \in \mathbb{R}^+$, there exists a unique $y \in \mathbb{R}^+$ that satisfies the equation $xf(y) + yf(x) \leq 2$ can be equivalently expressed as follows: there exists a well-defined function $g: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ given by $g(x) := y$, where y is the one mentioned above. The well-definedness of this function is evident due to the existence and uniqueness, and it satisfies the equation $P(x): xf(g(x)) + f(x)g(x) \leq 2$ while applying the same property for $x \mapsto g(x)$ gives another unique $y := g(g(x))$ such that $g(x)f(y) + yf(g(x)) \leq 2$. Therefore, we have $xf(y) + yf(x) > 2$ for all $y \neq g(x)$.

Since this inequality holds for $y = x$ (from $xf(y) + yf(x) > 2$), the uniqueness assumption implies that $g(g(x)) = x$, making g an involution (hence bijective).

Generally, working with an involution naturally leads us to consider its fixed points, especially since we aim to show that $g(x) = x$ identically (which holds for the solution $f(x) = \frac{1}{x}$). Let's define the set of fixed points of g as $\mathcal{S} := \{x \in \mathbb{R}^+ \mid g(x) = x\}$ and show that $\mathcal{S} = \mathbb{R}^+$ is the entire domain.

Assume for a contradiction that some $x \notin \mathcal{S}$ is not a fixed point, i.e., $x \neq g(x)$. Then, the inequality $2xf(x) > 2$ (derived from $y \mapsto x$) holds, implying $f(x) > \frac{1}{x}$. Similarly, $x \notin \mathcal{S}$ implies $g(x) \notin \mathcal{S}$ (otherwise $g(x) \in \mathcal{S}$ implies $x = g(g(x)) = g(x)$, a contradiction), leading to $f(g(x)) > \frac{1}{g(x)}$.

Applying these inequalities to $P(x)$ gives $xf(g(x)) + f(x)g(x) < 2$, which is clearly a contradiction as $\frac{x}{g(x)} + \frac{g(x)}{x} \geq 2$, e.g., from the AM-GM inequality. Therefore, we must have $x \in \mathcal{S}$ for every $x \in \mathbb{R}^+$, i.e., $g(x) = x$.

Substituting this relationship into the original equation, we obtain $P(x): xf(x) + f(x)x \leq 2 \implies xf(x) \leq 1 \implies f(x) \leq \frac{1}{x}$ for every $x \in \mathbb{R}^+$. Applying $yf(y) \leq 1$ to the equation $xf(y) + yf(x) > 2$ (since $g(x) = x$) yields $f(x) > \frac{2}{y} - \frac{x}{y^2}$, and taking the limit $y \rightarrow x$ from either side results in $f(x) \geq \frac{1}{x}$.

Combining the results, we have $f(x) \leq \frac{1}{x}$ and $f(x) \geq \frac{1}{x}$, implying $f(x) = \frac{1}{x}$ as desired. ■

(https://artofproblemsolving.com/wiki/index.php/2022_IMO_Problems/Problem_2).

Розв'язок оригінальної Задачі 2 ІМО -2022

Розв'язок Івана Ченя Задачі 2 ІМО-2022

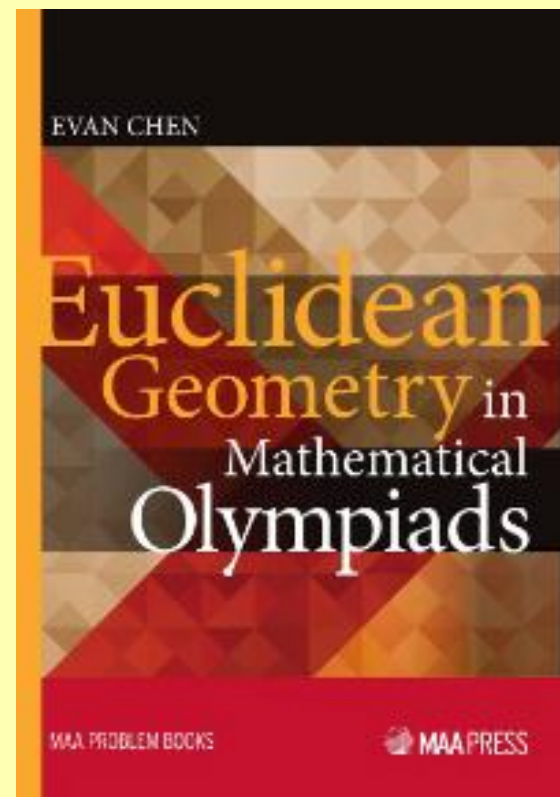
§1.2 IMO 2022/2, proposed by Merlijn Staps (NLD)

Available online at <https://aops.com/community/p25635138>.

Problem statement

Find all functions $f: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ such that for each $x \in \mathbb{R}^+$, there is exactly one $y \in \mathbb{R}^+$ satisfying

$$xf(y) + yf(x) \leq 2.$$



Іван Чень (Evan Chen 《陳誼廷》) – автор творчих і вишуканих розв'язків задач ІМО:
<https://web.evanchen.cc/exams/IMO-2022-notes.pdf>

Чи спроможеться ШІ колись на розв'язок такого класу математичної задачі?

Розв'язок Івана Ченя Задачі 2 ІМО-2022 наведено на наступному слайді.

Розв'язок оригінальної Задачі 2 ІМО -2022

Розв'язок Івана Ченя Задачі 2 ІМО-2022 (11 Листопада 2023)

The answer is $f(x) \equiv 1/x$ which obviously works (here $y = x$).

For the converse, assume we have f such that each $x \in \mathbb{R}^+$ has a *friend* y with $xf(y) + yf(x) \leq 2$. By symmetry y is also the friend of x .

Claim — In fact every number is its own friend.

Proof. Assume for contradiction $a \neq b$ are friends. Then we know that $af(a) + bf(b) > 2 \implies f(a) > \frac{1}{a}$. Analogously, $f(b) > \frac{1}{b}$. However, we then get

$$2 \geq af(b) + bf(a) > \frac{a}{b} + \frac{b}{a} \stackrel{\text{AMGM}}{\geq} 2$$

which is impossible. □

The problem condition now simplifies to saying

$$f(x) \leq \frac{1}{x} \text{ for all } x, \quad xf(y) + yf(x) > 2 \text{ for all } x \neq y.$$

In particular, for any $x > 0$ and $\varepsilon > 0$ we have

$$\begin{aligned} 2 < xf(x + \varepsilon) + (x + \varepsilon)f(x) &\leq \frac{x}{x + \varepsilon} + (x + \varepsilon)f(x) \\ \implies f(x) > \frac{x + 2\varepsilon}{(x + \varepsilon)^2} &= \frac{1}{x + \frac{\varepsilon^2}{x + 2\varepsilon}}. \end{aligned}$$

Since this holds for all $\varepsilon > 0$ this forces $f(x) \geq \frac{1}{x}$ as well. We're done.

Розв'язок модифікації Задачі 2 ІМО -2022

Запит на розв'язок модифікованої Задачі 2 ІМО-2022 (англ. мова) (27.11.2023))

Запит:

Suppose g is a continuous smooth function such that for every $x > 0$, there is one and only one $y > 0$ such that $g(x) + g(y) \leq 2xy$. Can you prove that $g(x) = x^2$?

GPT-4:

- This is an interesting mathematical question. The assertion is that for every $x > 0$, there's a unique $y > 0$ such that the inequality $g(x) + g(y) \leq 2xy$ holds. We want to see if this implies that $g(x) = x^2$.
- The condition seems to imply some form of symmetry, resembling the equality condition of the Cauchy-Schwarz inequality. However, the condition as stated isn't quite enough to straightaway prove that $g(x) = x^2$ for all $x > 0$, as we don't have any explicit information about how g behaves other than the inequality involving two possibly different arguments of the function.
- To move towards a proof, we would likely have to employ some properties of continuous and smooth functions (such as differentiability, and the Mean Value Theorem), and perhaps attempt to derive further properties that the function must satisfy — like analyzing the behavior of the function as x approaches 0, or seeing how the function behaves as x approaches infinity.
- For instance, because $g(x)$ is smooth and continuous everywhere for $x > 0$, we know $g(x)$ is differentiable, and we can look at $g'(x)$ (the derivative of g with respect to x).
- A possible way to start might involve fixing x and considering the function $h(y) = g(x) + g(y) - 2xy$, which by assumption has a zero at the given y — that is, $h(y) \leq 0$ and $h(y)$ is minimal when $y = x$ if $g(x) = x^2$. The derivative $h'(y) = g'(y) - 2x$ has a root at that y by Fermat's theorem, because it's a stationary point of h if g is to follow this property. This is all under the assumption of enough smoothness on g to differentiate it.
- However, without further information or assumptions about g , a rigorous proof would require additional steps or insights. The condition about the existence and uniqueness of such a y for every $x > 0$ implies a relationship between x and y but does not outright restrict the form of g to x^2 . It could potentially suggest a method of obtaining matched pairs (x, y) that fit the given condition, which might be analyzed further for obtaining conclusions about the form of g .
- A complete proof would likely involve demonstrating that no other function form but $g(x) = x^2$ can satisfy this condition for all $x > 0$. This suggests a potentially involved characterization problem that is not solvable by simple manipulation or basic theorems alone — it would likely require a series of logical deductions and possibly more advanced theorems in analysis.

Відповідь GPT-4 обмежується порадами розв'язок не пропонується (на відміну від [3])

Розв'язок модифікації Задачі 2 ІМО -2022

Запит на розв'язок модифікованої Задачі 2 ІМО-2022 (англ. мова) (Переклад укр.)

Запит:

Припустимо g є неперервна гладка функція, така, що для кожного $x > 0$ існує один і тільки один $y > 0$, такий що $g(x) + g(y) \leq 2xy$. Чи можете ви довести, що $g(x) = x^2$?

GPT-4:

- Це цікаве математичне питання. Твердження полягає в тому, що для кожного $x > 0$ існує унікальна $y > 0$ таке, що виконується

нерівність $g(x) + g(y) \leq 2xy$. Ми хочемо переконатися, чи це означає, що $g(x) = x^2$.

- Здається, ця умова передбачає певну форму симетрії, що нагадує умову рівності у нерівності Коші-Шварца. Однак зазначеної умови недостатньо, щоб одразу довести, що $g(x) = x^2$ для всіх $x > 0$, оскільки ми не маємо жодної чіткої інформації про те, як g поводить ся, окрім нерівності, що включає два, можливо, різні аргументи функції.
- Щоб перейти до доказу, нам ймовірно, доведеться використати деякі властивості неперервних і гладких функцій (такі як диференційованість і теорема про середнє значення) і, можливо, спробувати вивести додаткові властивості, яким має задовольняти функція, наприклад аналіз поведінки функції коли x наближається до 0, або дивитися, як поведеться функція, коли x наближається до нескінченності.
- Наприклад, оскільки $g(x)$ є гладкою і неперервною скрізь для $x > 0$, ми знаємо, що $g(x)$ є диференційовним, і ми можемо подивитися на $g'(x)$ (похідну від g відносно x).
- Можливий спосіб почати може включати фіксацію x і розгляд функції $h(y) = g(x) + g(y) - 2xy$, яка за припущенням має нуль при заданому y — тобто $h(y) \leq 0$ і $h(y)$ є мінімальним, коли $y = x$, якщо $g(x) = x^2$. Похідна $h'(y) = g'(y) - 2x$ має корінь у цьому y згідно з теоремою Ферма, оскільки це стаціонарна точка h , якщо g має відповідати цій властивості. Це все за припущення достатньої гладкості на g , щоб її диференціювати.
- Однак без додаткової інформації чи припущень щодо g строгий доказ вимагав би додаткових кроків або розуміння. Умова про існування та унікальність такого y для кожного $x > 0$ передбачає зв'язок між x і y , але не обмежує прямо форму g до x^2 . Це потенційно може запропонувати метод отримання відповідних пар (x, y) , які відповідають даній умові, які можна проаналізувати далі для отримання висновків про форму g .
- Повний доказ, ймовірно, включатиме демонстрацію того, що жодна інша форма функції, крім $g(x) = x^2$, не може задовольнити цю умову для всіх $x > 0$. Це говорить про потенційно залучену проблему визначення характеристик, яку неможливо розв'язати простими маніпуляціями чи основними теоремами — це, ймовірно, потребуватиме ряду логічних висновків і, можливо, більш просунутих теорем в аналізі.

Відповідь GPT-4 обмежується порадами розв'язок не пропонується (на відміну від [3])

Розв'язок модифікації Задачі 2 ІМО -2022

Запит на розв'язок модифікованої Задачі 2 ІМО-2022 (укр.мова) (27.11.2023)

Запит:

Припустимо g є неперервна гладка функція, така, що для кожного $x > 0$ існує один і тільки один $y > 0$, такий що $g(x) + g(y) \leq 2xy$. Чи можете ви довести, що $g(x) = x^2$?

GPT-4:

- Щоб довести, що $g(x) = x^2$ на основі наведених умов, потрібен додатковий аналіз. Існують математичні засоби для розв'язування таких завдань, наприклад, використання теорем про неперервні функції та їх властивості. Однак без додаткових властивостей чи обмежень для функції $g(x)$ довести, що $g(x) = x^2$ виключно з наведеного нерівності може бути неможливо.
- Головна ідея полягає в тому, що ми можемо використовувати дану умову для доведення певних властивостей $g(x)$, і потім спробувати довести, що ці властивості виконуються тільки функцією $g(x) = x^2$. Однак це лише один із можливих шляхів, і без конкретного методу рішення, ми не можемо бути впевненими у правильності висновку лише на основі даної нерівності.
- Щоб офіційно довести, що $g(x) = x^2$, можливі наступні кроки:
 1. Проаналізувати властивості функції з огляду на неперервність та гладкість. Це включає вивчення першої та вищих похідних.
 2. Математично формалізувати умову, що для всіх $x > 0$ існує один і тільки один $y > 0$, такий що $g(x) + g(y) \leq 2xy$.
 3. Використовувати властивості функції, теореми про неперервні функції, методи оптимізації та інші математичні інструменти, щоб довести, що єдина функція, яка задовольняє цю умову, це $g(x) = x^2$.
- Якщо ви маєте конкретний доказ або хочете обговорити конкретний метод доведення, я готовий допомогти розібратися з ним.

Відповідь GPT-4 обмежується порадами розв'язок не пропонується (на відміну від [3])

Відповідь сутнісно вдрізняється від запиту англійською мовою.

Розв'язок модифікації Задачі 2 ІМО -2022

Розв'язок модифікованої Задачі 2 ІМО-2022 на основі підказки (ідеї) GPT-4

1. Функція $y = x^2$ задовольняє умові задачі: для будь-яких двох векторів $\mathbf{u} = \langle x, y \rangle$ і $\mathbf{v} = \langle y, x \rangle$, нерівність Коши-Шварца означає: $(\mathbf{u}, \mathbf{v}) \leq \|\mathbf{u}\| \|\mathbf{v}\|$, або: $xy + yx = 2xy \leq \sqrt{x^2 + y^2} \sqrt{y^2 + x^2} = x^2 + y^2$, причому рівність має місце тоді й тільки тоді, коли $x = y$.
2. Доведемо від супротивного, що для кожного x єдиний y , що задовольняє умові задачі, є $y = x$. Дійсно, нехай $y \neq x$ і $g(x) + g(y) \leq 2xy$. Тоді маємо: $g(x) > x^2$ і $g(y) > y^2$ (у першому випадку вибрали $y = x$, а в другому $x = y$ і в обох випадках це не ті пари, для яких виконується нерівність $g(x) + g(y) \leq 2xy$). Звідки отримаємо $g(x) + g(y) > x^2 + y^2 > 2xy$ (за нерівністю Коши-Шварца), що суперечить вибору x і y .
3. Доведемо, що $g(x) = x^2$ - єдина функція, яка задовольняє умову задачі. Дійсно, з доказанного вище випливає, що для будь-якого x має місце нерівність $g(x) \leq x^2$. З другого боку, якщо в деякій точці $g(x) < x^2$, то, оскільки при $x = y$ маємо $g(x) + g(y) < 2xy$, в силу неперервності функції $g(x)$, в деякому околі точки x для будь-якої точки y цього околу також буде виконуватись нерівність $g(x) + g(y) < 2xy$, що суперечить єдиності такого y . Таким чином, доказано, що $g(x) = x^2$ у будь-якій точці x .

Це приклад того, що GPT-4 може продуктивно застосовуватись у математичних дослідженнях.

Розв'язок задачі МТНК

Запит

Хлопчик сказав, що в нього однакова кількість братів і сестер, а його сестра сказала, що в неї кількість братів утричі більше ніж кількість сестер. Скільки дітей у цій родині?

GPT-4

Давайте розглянемо інформацію, яку ми маємо:

Хлопчик сказав, що в нього однакова кількість братів і сестер. Це означає, що якщо n – кількість братів (включаючи його), а m – кількість сестер, то маємо: $(n - 1 = m)$

Сестра сказала, що у неї кількість братів утричі більше, ніж кількість сестер. Тобто: $(n = 3(m - 1))$

Тепер ми можемо вирішити цю систему рівнянь:

З першого рівняння маємо: $(n = m + 1)$

Підставляємо в друге рівняння: $(m + 1 = 3(m - 1))$

Розгортаємо і вирішуємо: $(m + 1 = 3m - 3)$

Перенесемо всі члени з m в одну сторону: $(1 + 3 = 3m - m)$

Отримаємо: $(4 = 2m)$

Звідси: $(m = 2)$

Знаючи m , можемо підставити і знайти n : $(n = 2 + 1 = 3)$

Отже, у родині 3 брати та 2 сестри, що в сумі дає 5 дітей.

Бездоганний розв'язок задачі GPT-4 з бездоганними поясненнями.

Розв'язок задачі МТНК

Запит

Хлопчик сказав, що в нього однакова кількість братів і сестер, а його сестра сказала, що в неї кількість братів учетверо більше ніж кількість сестер. Скільки дітей у цій родині?

GPT-4

Давайте спробуємо розшифрувати цю задачу:

Хлопчик каже, що у нього однакова кількість братів і сестер. Це означає, що якщо його сестер є S , то братів є $S+1$ (бо ми не враховуємо самого хлопчика).

Сестра каже, що у неї кількість братів учетверо більша, ніж кількість сестер. Якщо у неї є B братів, то сестер є $B/4$ (саму себе ми не враховуємо).

Таким чином, у нас є такі відносини:

Кількість братів $B = S + 1$ (з першої заяви хлопчика)

Кількість сестер $S = B/4$ (з другої заяви сестри)

Підставимо B з першої заяви до другої: $S = (S + 1)/4$

Помножимо обидві сторони рівняння на 4, щоб позбутися знаменника: $4S = S + 1$

Віднімемо S від обох сторін: $3S = 1$

Отже, $S = 1/3$. Але кількість сестер не може бути дробовим числом, тож, очевидно, десь у нас помилка. Запишемо відносини відносно кількості братів і сестер більш правильно:

Хлопчик: $S = B - 1$, бо хлопчик не рахує себе серед братів.

Сестра: $B = 4S$, бо сестра не рахує себе сестрою.

GPT -4 намагався модифікувати свій розв'язок і в кінці кінців (після 20 циклів!) помилився в розрахунках, отримав неправильну відповідь у натуральних числах. Але все це хвороба росту – після побудови мат. моделі слід підключати КМС.

ОСОБЛИВОСТІ МАТЕМАТИЧНОГО ДІАЛОГУ ЗІ ШІ

Проблема валідності повідомлень ШІ:

Математичний діалог зі ШІ потребує постійного критичного оцінювання відповідей ШІ, оскільки його «аргументація» може тільки за формою бути схожою на логічну аргументацію в силу статистичної, а не дедуктивної природи ШІ на основі LLM – великої мовної моделі.

Проблема стійкості повідомлень ШІ: навіть переформулювання запитання іншими словами може радикально змінити відповідь ШІ – з таким «співрозмовником» важко вести змістовний діалог .

Проблема першоджерел у міркуваннях ШІ: повідомлення ШІ завжди ведуться від лиця ШІ (якщо вони не стосуються питання поглядів конкретної особи на конкретне питання) і відстежити першоджерела тверджень ШІ, як правило, неможливо.

Складності з навчанням ШІ – проблемними для ШІ є розпізнавання еквівалентних задач у різних контекстах, абстрагування, аналогія, тощо.

Автоматичне доведення теорем і ШІ

1. *Готфрід Лейбніц (1646-1716)* – спроби побудувати універсальне числення для автоматичного розв'язання спорів (зокрема в логіці, математиці і юриспруденції), заклав логічні основи для автоматичного (формального) доведення
2. *Курт Гедель (1906 - 1972)* – довів неповноту змістовних формальних теорій (що веде до побудови дерева різних математик, яке розгалужується після кожного твердження, яке не може бути ані доведене, ані спростоване у рамках теорії)
3. *Алан Тьюринг (1912 - 1954)* – довів відсутність загального алгоритму для перевірки чи буде заданий алгоритм результативним для будь-якого набору вихідних даних (проблема зупинки)
4. *Алонзо Черч (1903 - 1995)* – довів нерозв'язність логіки першого порядку
5. *Coq (2009)* – інтерактивне середовище, створене групою французьких математиків для інтерактивної побудови / верифікації доведень терем:
 - У середовищі Coq було проведено верифікацію розв'язку проблеми чотирьох фарб
 - У середовищі Coq Воеводській В.О. (лауреат премії Філдса 2002 року) провів верифікацію своїх результатів, виправивши велику кількість технічних (на щастя) помилок.

ШІ і МАТЕМАТИЧНА ОСВІТА: ВИСНОВКИ і пропозиції

I. Людство знаходиться у процесі переходу від стану інформаційного суспільства до стану постінформаційного суспільства (або суспільства знань), яке характеризується широким впровадженням систем ШІ в виробничі і суспільні процеси:

1. Стан ШІ – Шоку стосовно спроможностей ШІ у різних галузях поступово змінюється на процес практичного впровадження ШІ у різних галузях; які традиційно вважалися прерогативою виключно людини;
2. Велика кількість публікацій присвячені дослідженню меж спроможностей ШІ у різних галузях; авторів таких публікацій умовно можна поділити на дві групи: «оптимістів», які сфокусовані на прикладах ефективності ШІ і «песимістів», які сфокусовані на прикладах неефективності ШІ; динаміка прогресу ШІ на користь оптимістів;
3. Темпи поширення кількості користувачів ШІ встановила абсолютний рекорд за всю історію розробки комп'ютерних систем як у приватному, так і в корпоративному використанні, охоплюючи практично усі верстви суспільства і галузі;
4. Створені міжнародні, національні, галузеві, корпоративні програми дослідження можливостей ШІ і сприяння впровадженню ШІ в практику.

ШІ і МАТЕМАТИЧНА ОСВІТА: ВИСНОВКИ і пропозиції

II. ВИКОРИСТАННЯ ШІ В ОСВІТІ:

1. Високі спроможності ШІ в освіті підтверджує, зокрема, успішність проходження ШІ тестувань, зокрема проходження тестів PISA і PIAAC з читання, математики, природничих дисциплін, яка перевершує середній рівень учасників тестувань по міжнародній репрезентативній вибірці;
2. Упровадження ШІ в освіті вже де факто відбулося: більшість студентів ЗВО і ЗССО, так і учнів старшої, основної і навіть початкової школи) є активними користувачами ШІ, причому спектр застосувань ШІ дуже широкий – від дружньої довідкової системи, яка завжди під рукою, до дружньої інтелектуальної діалогової системи навчання, здатної надати «людською мовою» відповідь або пораду практично на будь-яке запитання, помічника у виконанні домашніх робіт;
3. Ключовою проблемою системи освіти є ефективне застосування ШІ в навчальному діалозі: культура постановки запитань (промптів) до ШІ, критичне оцінювання відповідей ШІ та їх верифікація, уточнення запитань, додаткові запитання, знаходження недоліків в аргументації і їх усунення, наведення контрприкладів, тощо.

ШІ і МАТЕМАТИЧНА ОСВІТА: ВИСНОВКИ і пропозиції

III. ГОЛОВНІ ЗАДАЧІ МАТЕМАТИЧНОЇ ОСВІТИ В ЕПОХУ ШІ

ПРОСВІТНИЦЬКО-МЕТОДОЛОГІЧНА:

Ознайомлення учнів на концептуальному рівні з математичними засадами систем ШІ з метою усвідомлення моделей функціонування ШІ, їх поточного стану і напрямків розвитку, необмежені можливості і можливі обмеження, перспективність гібридних діалогових систем, які ефективно використовують творчі здатності людини і продуктивність ШІ, для чого:

1. Уточнити концептуальні математичні засади систем ШІ;
2. Побудувати доступні для широкого загалу функціональні моделі ШІ на основі концептуальних засад, які, зокрема, розкривають можливості, обмеження і особливості застосування ШІ в цілому і зокрема в математиці;
3. Удосконалити навчальні програми і підручники з метою висвітлення питань концептуальних математичних засад ШІ;
4. Розробити програми і навчальні матеріали для підвищення кваліфікації учителів в математики у галузі концептуальних математичних засад ШІ;
5. Провести перепідготовку вчителів математики у галузі концептуальних математичних засад ШІ;

ШІ і МАТЕМАТИЧНА ОСВІТА: ВИСНОВКИ і пропозиції

III. ГОЛОВНІ ЗАДАЧІ МАТЕМАТИЧНОЇ ОСВІТИ В ЕПОХУ ШІ

ПРАКТИЧНО-МЕТОДОЛОГІЧНА:

Набуття учнями здатностей ефективно використовувати ШІ для розв'язування задач математичними методами в режимі діалогу зі ШІ: математизації задачі (побудова математичної моделі задачі), постановки задачі (запиту) ШІ, критичний аналіз та верифікація запропонованих ШІ розв'язків задачі або рекомендацій їх розв'язування, уточнення/модифікація запиту до ШІ, тощо, для чого:

1. Удосконалити навчальні програми і підручники з математики для органічного включення в навчальний процес з математики роботу зі ШІ у діалоговому режимі для навчання і розв'язування задач;
2. Розробити програми і навчальні матеріали для підвищення кваліфікації учителів математики у галузі з питань застосування ШІ навчальному процесі з математики і розв'язування математичних у діалоговій формі;
3. Провести перепідготовку вчителів математики у галузі з питань застосування ШІ навчальному процесі з математики і розв'язування математичних у діалоговій формі.

Посилання

1. **Projector Creative & Tech Institute, 2024, Всеукраїнське дослідження використання ШІ у шкільній освіті, https://factum-ua.com/document/Всеукраїнське_дослідження_використання_ШІ_у_шкільній_освіті.pdf**
2. **OECD (2023), Putting AI to the test: How does the performance of GPT and 15-year-old students in PISA compare?, <https://www.oecd.org/social/putting-ai-to-the-test-2c297e0b-en.htm>**
3. **S. Bubeck at all, Sparks of Artificial General Intelligence: Early experiments with GPT-4, 2023, <https://arxiv.org/abs/2303.12712>)**
4. **Problems IMO -2022, 2022, https://www.imo-official.org/year_info.aspx?year=2022**
5. **Art of Problem Solving, 2022, IMO Problems/Problem 2, https://artofproblemsolving.com/wiki/index.php/2022_IMO_Problems/Problem_2**
6. **J.Slesinski, C. Fadel, What “Mathematics of AI” should be taught in schools?, 2024, <https://curriculumredesign.org/wp-content/uploads/Mathematics-of-AI-Opinion-paper-CCR.pdf>**

Класифікація сімейств ШІ

Сімейство ШІ	Загальний підхід	Математичний апарат
Символісти Дерево рішень	Розуміння дедуктивної логіки та роботи у зворотному напрямку для визначення передумови на основі фактів	Логічний висновок Індукція Найкраща відповідність /регресія Статистична значущість
Мережевики Нейронні мережі	Погляд на знання через призму нейронних зв'язків і спроби змодельовати роботу мозку	Класифікація та розпізнавання образів Кут нахилу та дотичні лінії Оптимізація Множення та перетворення матриць Проста й умовна ймовірності Регресія
Еволюціоністи Генетичні алгоритми	Імітація еволюції та виживання найбільш пристосованих	Проста ймовірність і комбінаторика Апроксимуючі функції Оптимізація
Баєсисти Класифікатор Баєса	Побудова припущень і удосконалення попередніх уявлень з урахуванням нової інформації на основі нових даних	Базова і умовна ймовірності Теорема Баєса Ланцюги Маркова і Баєсівські мережі Теорія графів
Аналогізатори Найближчий Сусід	Визначення, наскільки одна ідея схожа на іншу, і виведення висновку про навчання	Представлення та інтерпретування даних Координатна площина Асоціація двовимірних даних Формула евклідової відстані Типи функцій похибок і вартості Диференційованість

Відкритий лист: призупинення розробок ШІ загального призначення (AGI)

https://futureoflife.org/wp-content/uploads/2023/04/FLI_Policymaking_In_The_Pause.pdf



Самі системи можуть переслідувати цілі, поставлені людиною чи власними силами, таким чином, щоб права людини, її безпека чи, у найжахливіших сценаріях, людське існування були незначущими. Намагаючись запобігти цим результатам, Інститут майбутнього життя (FLI), до якого приєдналися понад 20 000 провідних дослідників штучного інтелекту, професорів, генеральних директорів, інженерів, студентів та інших, хто займається прогресом штучного інтелекту, закликав призупинити принаймні на шість місяців у найризикованіших і найбільш ресурсомістких експериментах ШІ. – Ті експерименти, що спрямовані на подальше збільшення розміру та загальних можливостей найпотужніших систем, розроблених на сьогоднішній день.

Аналогія з розробками ядерної і термоядерної зброї (Роберт Опенгеймер і Андрій Сахаров): їх боротьба за заборону зброї після її створення вже була запізньою – вона опинилася у руках політиків.

«Призупинити гігантські експерименти ШІ: відкритий лист», 22 березня 2023 р.

future
of life

12th April 2023
Last revision: 10th April 2023

Policymaking in the Pause

What can policymakers do now to combat risks
from advanced AI systems?

Політичні рекомендації:

1. Дозволити надійний сторонній аудит і сертифікацію розробок ШІ.
2. Регулювати доступ до обчислювальних потужностей.
3. Створити спроможні агентства ШІ на національних рівнях.
4. Встановити відповідальність за шкоду, спричинену ШІ.
5. Запровадити заходи для запобігання та відстеження витоків моделі ШІ.
6. Розширити фінансування досліджень безпеки технічного ШІ.
7. Розробити стандарти для ідентифікації та керування контентом, створеним ШІ.

Щоб координувати, співпрацювати чи запитувати щодо наведених тут рекомендацій, зв'яжіться з нами за адресою policy@futureoflife.org.

Освітні втрати (PISA-2022/PISA-2018)

Sample	Mathematics Performance		Reading Performance		Science Performance		ESCS		Boys	
	Mean	S.E	Mean	S.E	Mean	S.E	Mean	S.E	%	S.E
Ukraine 2018 (all regions)	453	3,60	466	3,00	469	3,00	-0.20	0.02	53	1,0
Ukraine 2018 (18 regions)	452	4,50	464	4,00	467	4,00	-0.20	0.03	52	1,0
Ukraine 2022 (18 regions)	441	4,10	428	4,00	450	4,00	-0.35	0.04	50	2,0
PISA 2018: Difference between Ukraine's full sample and 18 regions: (all regions) - (18 regions)	1	5,80	2	5,40	2	5,20	0.00	0.04	0.9	1,6
Trend over time in Ukraine (18 regions): 2022 - 2018	-11	6,50	-37	5,90	-16	5,70	-0.15	0.05	-1.8	2,0

**Дякую
за увагу!**

Python Code Snippet for NN Learning Function

```
1 def train(X, y, n_hidden, learning_rate, n_iter):
2     m, n_input = X.shape
3
4     # 1. random initialize weights and biases
5     w1 = np.random.randn(n_input, n_hidden)
6     b1 = np.zeros((1, n_hidden))
7     w2 = np.random.randn(n_hidden, 1)
8     b2 = np.zeros((1, 1))
9
10    # 2. in each iteration, feed all layers with the latest weights and biases
11    for i in range(n_iter + 1):
12
13        z2 = np.dot(X, w1) + b1
14
15        a2 = sigmoid(z2)
16
17        z3 = np.dot(a2, w2) + b2
18
19        a3 = z3
20
21        dz3 = a3 - y
22
23        dw2 = np.dot(a2.T, dz3)
24
25        db2 = np.sum(dz3, axis=0, keepdims=True)
26
27        dz2 = np.dot(dz3, w2.T) * sigmoid_derivative(z2)
28
29        dw1 = np.dot(X.T, dz2)
30
31        db1 = np.sum(dz2, axis=0)
32
33        # 3. update weights and biases with gradients
34        w1 -= learning_rate * dw1 / m
35        w2 -= learning_rate * dw2 / m
36        b1 -= learning_rate * db1 / m
37        b2 -= learning_rate * db2 / m
38
39        if i % 1000 == 0:
40            print("Epoch", i, "loss: ", np.mean(np.square(dz3)))
41
42    model = {"w1": w1, "b1": b1, "w2": w2, "b2": b2}
43    return model
```

Фрагмент коду Python ілюструє функцію навчання, яка використовує навчальний набір даних, кількість прихованих шарів (лейерів), швидкість навчання та кількість ітерацій як параметри.