

УДК 004.83

doi: 10.15622/rcai.2025.019

## ДОГАДКА, ЭМПИРИЧЕСКАЯ ПРОВЕРКА И ОБЪЯСНЕНИЕ ПРИ АВТОМАТИЗИРОВАННОМ РЕШЕНИИ ЗАДАЧ

С.С. Курбатов (*curbatow.serg@yandex.ru*)

Научно-исследовательский центр электронной  
вычислительной техники, Москва

Анализируется различие в подходах к автоматизированному решению задач с использованием LLM и базирующемся на моделировании рассуждений. Анализ проводится на примере решения олимпиадной геометрической задачи. Рассматриваются такие аспекты как переход от текста задачи к компьютерному представлению, взаимодействие с нейросетью в процессе решения, качество объяснения доказательства, визуализация и её роль для эмпирических догадок.

**Ключевые слова:** LLM, Моделирование рассуждений, Олимпиадная задача.

### Введение

Интерес к компьютерному моделированию рассуждений, заложенный ещё в пионерских работах Д.А. Поспелова, в последнее время возрос в связи с трудностями объяснения выводов и рекомендаций в нейронных сетях.

Использование больших лингвистических моделей (LLM) для решения сложных проблем инициировало комплексный подход, сочетающий поиск в нейронных сетях с логической (символьной) обработкой. В [Chen et. al., 2025] обсуждаются современные эффективные модели рассуждений (DeepSeek-R1-Zero и DeepSeek-R1), использующие обучение с подкреплением (RL). Отмечаются проблемы моделей и их преодоление путём многоступенчатого обучения и специального старта перед RL. В [Ma et. al., 2025] подчеркивается, что современные масштабные модели рассуждений (DeepSeek-AI, OpenAI o1 и ряд других) при подходе к сложным задачам в процессе поиска генерируют длинные цепочки рассуждений («размышлений»), комбинируя их с возвратом назад и «самооценкой». Тем не менее обход промежуточной логической обработки с

помощью простой подсказки («NoThinking», по терминологии авторов результата) показал эффективность в ряде сложных рассуждений, включая решение математических задач, формальное доказательство теорем и кодирование.

OpenAI o1 обучена для выполнения сложных рассуждений, прежде чем ответить пользователю система может создать длинную внутреннюю цепочку ассоциативных связей [Chen et. al., 2024]. По утверждению авторов модель ИИ, использующая такую цепочку, может рассуждать над сложными задачами и решать более трудные проблемы, чем предыдущие модели в области науки, кодирования и математики. Фреймворк ReSearch рассматривает поисковые операции в LLM как неотъемлемые компоненты цепочки рассуждений. Вопрос о специальном контроле над LLM, чтобы улучшить моделирование рассуждений и сделать их более антропоморфными, исследуется в [Højer et. al., 2025]. Такой контроль позволяет улучшить производительность в конкретных задачах, избегая трудоёмкого дообучения LLM. Обращение ученых к моделям ИИ для формирования новых гипотез, а частности, в таких областях как химия, биология, медицина обсуждается в [Bajorath, 2025]. Автор предупреждает о возможных заблуждениях при интерпретации результатов предсказательных алгоритмов. Подчеркивается необходимость объяснения предсказаний на уровне причинно-следственных связей.

Философские аспекты моделей для научных рассуждений и их роль в образовательных контекстах с эпистемологических позиций обсуждаются в [Rost, et. al., 2022]. Отмечается, что с этих позиций к числу необходимых характеристик знания должно относиться не только соответствие информации реальному положению дел, но и её обоснованность. Такая трактовка способствует более последовательному теоретическому пониманию моделирования и интерпретации результатов эмпирических исследований.

Вопросы естественно-научного образования в аспекте понимания того, как ученые мыслят и рассуждают обсуждаются в [Krell et. al., 2022]. Авторы рассматривают феномен научного мышления как компетенции, а не только как обладание способностями и знаниями. Подчёркивается, что такая интерпретация важна не только для ученого, но и для любого человека при принятии взвешенных решений в обыденной жизни. Абдуктивный подход к моделированию рассуждений исследуется в [Urmeier zu Belzen et. al., 2021]. Подход предполагает объяснение с помощью причины (отличается от индуктивных и дедуктивных рассуждений). Указывается роль подхода для компетентности в области моделирования рассуждений.

Подчёркивая ограничения на математические рассуждения моделей в [Mirzadeh et. al.] отмечается, что существующие LLM не выполняют полные логические рассуждения – они просто копируют шаги рассуждения,

извлекаемые из данных после обучения. Различные перефразировки одного и того же запроса (даже изменение только числовых характеристик) могут существенно повлиять на качество ответа. В относительно узкой предметной области взаимодействие LLM и символического логического вывода («движки») демонстрируется в [Trinh et. al., 2024], [Zhang et. al., 2024]. Тем не менее эти работы будут нам интересны именно в аспекте такого взаимодействия и моделирования рассуждений.

Основная цель предлагаемого исследования – сопоставить возможности моделирования рассуждений в нейросетевой модели (LLM) и в системе, использующей традиционные методы ИИ. В нашей работе подчеркивается, что потенциал традиционных методов недооценён в связи с преобладающей тенденцией использования нейросетей. Не менее важна и прикладная сторона – ознакомление учащихся с элементами ИИ, непосредственно относящихся к учебному процессу, а также возможность прямого участия в компьютерной доработке. Выбранная предметная область – автоматическое решение сложных геометрических задач. Именно в геометрии наглядность и интуитивная очевидность сочетается с логической глубиной. В общей постановке проблема моделирования рассуждений весьма сложна. Ограничения предметной области позволяют сосредоточиться на рассуждениях в «чистом» виде, не выходящая тем не менее суть дела. Далее обсуждаются детали сопоставления.

## 1. Методология

Концепции моделирования рассуждений пока не развиты в направлении упрощения алгоритмизации и масштабных прикладных возможностей. Во введении выделены работы, ориентированные на моделирование рассуждений в бурно развивающейся области нейронных сетей и больших лингвистических моделей (LLM). После пионерских работ Д.А. Поспелова велась многолетняя атака на эту проблему. В частности, в одной из работ в соавторстве с Д.А. Поспеловым анализировалась проблема представления знаний о времени и пространстве, предлагалась каузальная логика, формализующая причинно-следственные отношения.

В работах Э.В. Попова исследовалась модель участника общения с учетом его знаний о языке и предметной области, Р. Шенк выдвинул и экспериментально обосновал идею концептуальной обработки информации, использующей механизм умозаключений. Создавались системы, использующие логический вывод, экспертные системы, базы знаний (онтологии), условно формальные системы (когнитивные, индуктивные, по аналогии, правдоподобные умозаключения и т.п.), в той или иной степени затрагивающие проблему моделирования рассуждений. Однако, несмотря на отдельные и довольно серьёзные успехи, необходимость полноценного решения проблемы компьютерного моделирования рассуждений сохраняет свою актуальность.

Моделирование рассуждений затрагивает и проблему объяснимого ИИ (ХАИ), поскольку сформулированные в [Phillips et al., 2020] общие принципы (Объяснение, Понимание, Точность объяснения, Границы знания) безусловно значимы для рассуждений. Для конкретных областей ХАИ даже разработаны фреймворки на базе Python [Hu et al., 2021], однако они ориентированы скорее на интеграцию алгоритмов и обобщение API для данных. Разумеется, эти технологические аспекты важны, особенно при модификации системы, однако они лежат несколько в стороне от собственно моделирования рассуждений.

В медицинских приложениях используются вероятностные модели, основанные на байесовских сетях [Pradeepta, 2022], отражающих причинно-следственные связи и неопределённости. Более антропоморфные подходы обычно используют интерпретируемые визуализации (графы, таблицы, изображения) или тексты, призванные «объяснить», как система получила свои результаты. Используемый нами подход базируется на интерактивной визуализации, позволяющий кликом мыши по объекту чертежа выдать обоснование его появления.

В развиваемом автором подходе методологической основой является онтология, концентрирующая в когнитивных схемах лингвистические, логические (теоремы, их связи с ЕЯ и выводом) и визуальные знания, а также возможности их взаимодействия для эффективного вывода, ориентированного на человека (объяснительные возможности). Подход реализуется в компьютерной системе, базирующейся на идеях известного ученого и педагога Пойа [Polya, 1981]. Возникающие при этом проблемы требуют комплексного привлечения различных методов, некоторые аспекты рассмотрены в [Kurbatov et. al., 2024], [Kurbatov, 2023]. Пойа подчеркивал, что важно не только собственно решение, но и то, какими рассуждениями обоснованы шаги решения (выбор теорем, дополнительные построения и т.д.), какие догадки привлекаются, какие правдоподобные умозаключения используются. Все эти соображения сохраняют свою значимость и для ХАИ, но требуют детальной и весьма сложной проработки в конкретных приложениях.

Рассуждения на уровне «модель мира» практически не рассматриваются в рамках данной работы. Но в определённом смысле чертёж является «моделью мира» для семантического представления, определяя как «физически» можно менять «конструкцию» чертежа, сохраняя заложенные в семантике условия. Инструментальные средства реализации, используемые в проводимом исследовании: VBA EXCEL, JavaScript, библиотека JSXGRAPH. Средства программирования выбраны с учётом ориентации прикладной ориентации – создание образовательного ресурса для средней школы.

## 2. Эксперимент

С учетом выбранной предметной области сопоставление проведем на основе различного решения олимпиадной геометрической задачи [Trinh et. al., 2024]. Задача в течении ряда лет не поддавалась автоматическому решению, но недавно была решена системой AlphaGeometry с использованием LLM. Предварительное обучение языковой модели проводилось на 100 миллионах синтетически сгенерированных доказательств. Модель взаимодействует с решателем для генерации дополнительных построений и дедуктивных шагов.

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

Фрагмент решения приведён на рис. 1.

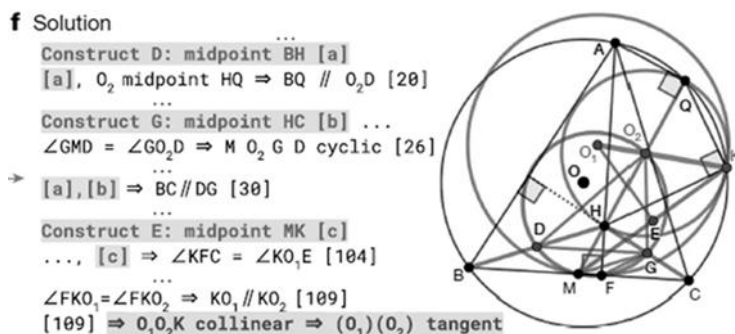


Рис. 1. Фрагмент решения (AlphaGeometry)

AlphaGeometry использует специальные средства («разности зависимостей») при генерации синтетических доказательств и выполняет около 10 миллионов синтетических шагов доказательства. При этом строятся вспомогательные точки, которые оцениваются как фактор очень сильного ветвления при чисто символьном выводе. Решение содержит 109 шагов, причём шаг 26 – спорный (рис. 1). Подробнее некоторые детали и сравнительные характеристики подходов рассматриваются в разделе 3.

Развиваемый нами подход предполагает ввод текста задачи на естественном языке (ЕЯ), лингвистическую обработку для получения семантического представления, автоматическое решение и его интерактивная визуализация. Подход базируется на компьютерном воплощении концепции Пойа, который подчеркивал, что в работе ученого догадка почти всегда

предшествует доказательству. Далее акцентируется идейная сторона нашего решения – догадки и их обоснования, а не теоремы – это техническая сторона. В силу некоторые моменты решения даются в упрощенном варианте.

В результате лингвистической обработки текста задачи (в русском переводе) формируется семантическое представление и визуализируется интерактивный чертёж, аналогичный приведённому на рис. 1. На первом шаге в нашем решении выполняется эмпирическая проверка корректности. Точки  $K$ ,  $O_1$  и  $O_2$  находятся на одной прямой (с точностью возможностей графики), но главное, что при изменении параметров треугольника это расположение сохраняется. AlphaGeometry доказывает касание окружностей коллинеарностью  $KO_1O_2$ , наша система рассматривает равенство касательных (следует, если допустить доказываемое). Догадка и её эмпирическое подтверждение отображены на рис. 2.

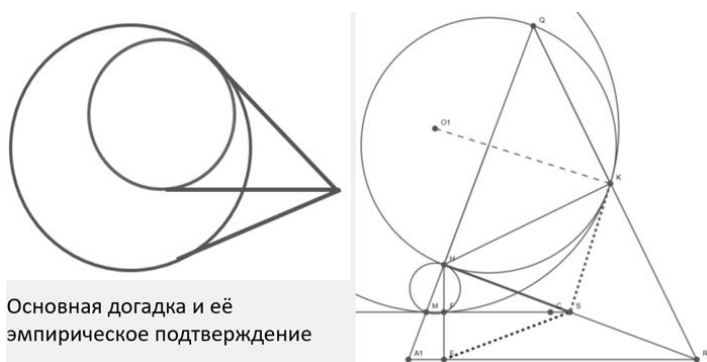


Рис. 2. Догадка и эмпирика

Элементы семантического описания и теоремы базы знаний снабжены ЕЯ-описаниями, что позволяет организовать их взаимодействие с помощью ключевых слов, перифраз или ассоциаций. Например, «три равных касательных из точки» имеет ассоциацию «окружность с центром в точке и радиусом, равным касательной». ЕЯ-описание семантической структуры задачи содержит фрагменты «окружности и равные касательные», «секущая окружности по точкам  $M$  и  $F$ » и т.п. По этим фрагментам из базы знаний извлекаются теоремы, которые могут оказаться полезными. Пример такой теоремы – на рис. 3.

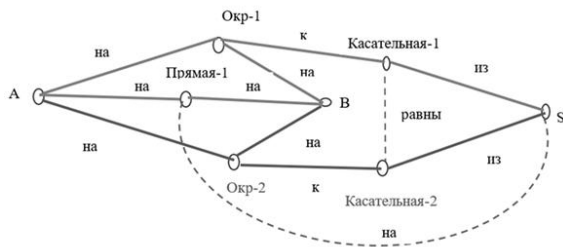


Рис. 3. Граф теоремы и формализация ассоциации «часть подсказывает целое»

Важный шаг – нахождение центра эмпирически подтвержденной окружности отражён на рис. 4.

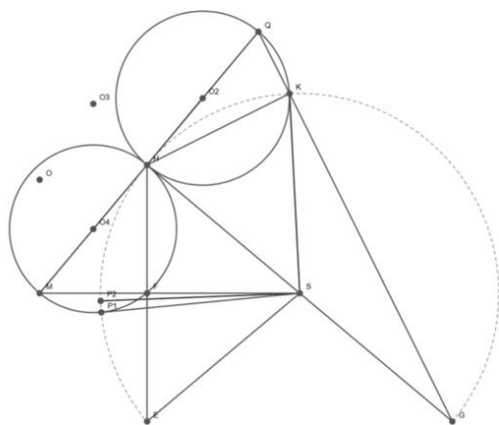


Рис. 4. Ключевая окружность с равными касательными

Текст теоремы: «общая секущая двух окружностей включает геометрическое место точек с равными касательными к этим окружностям». Разумеется, таких теорем может быть несколько. Важно, что если теорема не может быть непосредственно применена к текущей структуре задачи, то она может дать рекомендацию о включении в семантическую структуру недостающего фрагмента, с обоснованием ассоциации. Поясним конкретнее использование ассоциации. Пусть для структуры на рис. 3. в текущем семантическом представлении найдены элементы, которые обозначают подграф, выделенный из общего графа теоремы. Тогда оставшаяся часть структуры является кандидатом на дополнительное построение (окружность и касательная).

После ряда неудачных попыток решатель, используя общее утверждение «центр окружности лежит на пересечении перпендикуляра из середины хорды и диаметра», строит центр  $S$ . Далее использование ассоциации (рис. 3.) позволяет построить окружность  $MFN$ , имеющую общую хорду с окружностью  $MFK$ . Это позволяет применить перспективную теорему и построить точки  $P1$  и  $P2$  (равные касательные на рис. 4). И, наконец, применяя свойства равных касательных, решатель доказывает их совпадение в точке  $K$ . Это завершает решение задачи (рис. 5).

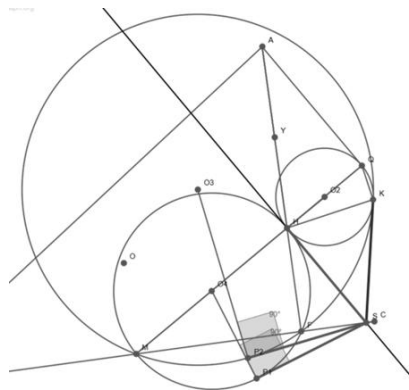


Рис. 5. Решение получено

Процесс решения выполнялся в диалоговом режиме. Однако практически все обращения к системе не являются подсказками решателю, шаги формируются с использованием эвристических, эмпирических и логических обоснований.

### 3. Обсуждение

Сравнение подходов проводится на весьма общем уровне, поскольку из публикаций неясны вопросы семантического представления, возможности интерактивной визуализации и эмпирической верификации, методы обоснования шагов решения, перспективы для образовательных целей. Можно согласиться с разработчиками AlphaGeometry в том, что современные подходы к геометрии (в области автоматизации решения задач) опираются в основном на символьные методы и разработанные человеком жестко закодированные эвристики поиска. Тем не менее развиваемый нами подход ориентирован на существенное уменьшение жесткости эвристик за счёт ассоциаций, эмпирических характеристик (чертёж, а в пределе рецепторы), онтология, снабжённая ЕЯ-описаниями объектов. Последний пункт важен, поскольку для пользователя соответствующие применениям эвристик рассуждения должны быть предъявлены *вербально*.



## Заключение

Комплексное использование нейронных сетей и традиционных методов ИИ (базы знаний, когнитивные подходы, прикладные онтологии и т.п.) обладает значительным потенциалом. Автор согласен с доктором Барнеттом [Matthew Barnett, 2025], в том, что в ближайшие три года будут разработаны системы ИИ, способные превзойти лучших математиков-людей в доказательстве произвольных математических теорем. Согласен автор и в том, что экономически ценные возможности ИИ будут отставать, а надежные агенты компьютерного управления крупными объектами появятся значительно позже, чем высококачественные модели математических рассуждений.

Прикладная значимость развиваемого исследования для образовательного процесса состоит в том, что постановка цели, средства её достижения (теоремы, доп. построения), учёт наводящих и индуктивных соображений (догадок), четкое разграничение догадки и доказательства, структуризация доказательства, вычленение ключевых моментов для устранения когнитивного диссонанса важны для любого специалиста. Создание образовательного ресурса, базирующегося на результатах данной работы, позволит продвинуть элементы ИИ непосредственно в среднее школьное образование.

## Список литературы

- [Chen et. al., 2025] Chen M., Tworek J., Jun H., Yuan Q., and others. DeepSeek-R1. Incentivizing Reasoning Capability in LLMs via Reinforcement Learning. – <https://arxiv.org/abs/2501.12948>, Submitted on 22 Jan 2025.
- [Ma et. al., 2025] Wenjie Ma, Jingxuan He, Charlie Snell, Tyler Griggs, Sewon Min, Matei Zaharia. Reasoning Models Can Be Effective Without Thinking. – <https://doi.org/10.48550/arXiv.2504.09858>, Submitted on 14 Apr 2025.
- [Chen et. al., 2024] Mingyang Chen, Tianpeng Li, Haoze Sun, Yijie Zhou, Chenzheng Zhu and others: Learning to reason with LLMs, September 12, 2024, Release. – <https://openai.com/index/learning-to-reason-with-llms/>.
- [Huang et. al., 2006] Huang W., Hong S.H., Eades P. Predicting Graph Reading Performance: A Cognitive Approach // In: Proc. Asia Pacific Symposium on Information Visualization (APVIS2006), Tokyo, Japan, 2006. – P. 207-216. – doi: 10.1145/1151903.1151933 (статья в сборнике трудов конференции на англ. языке)
- [Højer et. al., 2025] Bertram Højer, Oliver Jarvis, Stefan Heinrich. Improving Reasoning Performance in Large Language Models via Representation Engineering. – <https://doi.org/10.48550/arXiv.2504.19483>, Submitted on 28 Apr 2025.
- [Bajorath, 2025] Jürgen Bajorath. From scientific theory to duality of predictive artificial intelligence models, April 2025 Cell Reports Physical Science 6(4):102516, [Iman Mirzadeh, et. al.] Iman Mirzadeh, Keivan Alizadeh, Hooman Shahrokhi, et. al. GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language Models, Submitted on 7 Oct 2024, preprint. – <https://arxiv.org/abs/2410.05229>.

- [Rost et. al., 2022] Marvin Rost, Tarja Knuuttila: Models as Epistemic Artifacts for Scientific Reasoning in Science Education Research, *Educ. Sci.* – 2022. – 12(4), 276; [Iman Mirzadeh, et. al.] Iman Mirzadeh, Keivan Alizadeh, Hooman Shahrokhi, et. al. GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language Models, Submitted on 7 Oct 2024, preprint. – <https://arxiv.org/abs/2410.05229>.
- [Krell et. al., 2022] Moritz Krell, Andreas Vorholzer, Andreas Nehring. Scientific Reasoning in Science Education: From Global Measures to Fine-Grained Descriptions of Students' Competencies // *Education Sciences.* – January 2022 – 12(97). – P. 1-8. – DOI: 10.3390/educsci12020097.
- [Upmeier zu Belzen, et. al., 2021] Annette Upmeier zu Belzen, Paul Engelschalt, Dirk Krüger: Modeling as Scientific Reasoning – The Role of Abductive Reasoning for Modeling Competence // *Education Sciences.* – September 2021. – 11(9):495. – DOI: 10.3390/educsci11090495.
- [Mirzadeh et. al.] Iman Mirzadeh, Keivan Alizadeh, Hooman Shahrokhi, et. al. GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language Models, Submitted on 7 Oct 2024, preprint, [Iman Mirzadeh, et. al.] Iman Mirzadeh, Keivan Alizadeh, Hooman Shahrokhi, et. al. GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language Models, Submitted on 7 Oct 2024, preprint. – <https://arxiv.org/abs/2410.05229>.
- [Trinh et. al., 2024] Trinh T.H., Wu Y., Le Q.V., et al. Solving Olympiad geometry without human demonstrations // *Nature.* – 2024. – 625. – P. 476-482 – <https://doi.org/10.1038/s41586-023-06747-5>.
- [Zhang et. al., 2024] Chi Zhang, Jiajun Song, Siyu Li, et. al. Proposing and solving olympiad geometry with guided tree search. – <https://arxiv.org/abs/2412.10673>, Submitted on 14 Dec 2024.
- [Polya, 1981] Polya G. *Mathematical Discovery: On Understanding, Learning and Teaching Problem, Solving (Combined Edition).* – Wiley, New York, 1981. – 432 p.
- [Kurbatov et. al., 2024] Sergey S. Kurbatov, Mikhail A. Gilmendinov. Cognitive Dissonance in Solving Planimetric Problems // In: *Proceedings of 8th Computational Methods in Systems and Software.* – 2024. – Vol. 2. – P. 163-172. – [https://doi.org/10.1007/978-3-031-77411-9\\_15](https://doi.org/10.1007/978-3-031-77411-9_15).
- [Kurbatov, 2023] Kurbatov S. Paraphrasing in the system of automatic solution of planimetric problems: data analytics in system engineering // In: *Proceedings of 7th Computational Methods in Systems and Software.* – 2023. – Vol. 3. – P. 217-225. – [https://doi.org/10.1007/978-3-031-53552-9\\_20](https://doi.org/10.1007/978-3-031-53552-9_20).
- [Phillips et al., 2020] Jonathon Phillips P., et al. Four Principles of Explainable Artificial Intelligence, NIST. – 2020. – <https://doi.org/10.6028/NIST.IR.8312-draft>.
- [Hu et al., 2021] Brian Hu, et al. XAITK: the explainable AI toolkit // *Applied AI Letters.* – October 2021. – Vol. 2(4). – <https://doi.org/10.1002/ail2.40>.
- [Pradeepta, 2022] Pradeepta M. *Practical Explainable AI Using Python: Artificial Intelligence Model Explanations Using Python-based Libraries.* – Apress Media LLC, 2022. – 356 p. – ISBN13: 978-1-4842-7157-5.
- [Barnett, 2025] Matthew Barnett: The promise of reasoning models, Gradient Updates – Epoch AI, Feb 28, 2025. – <https://epoch.ai/gradient-updates/the-promise-of-reasoning-models>.