

**Каково это — быть рассуждающей языковой моделью?**

Д. В. Зайцев

МГУ имени М. В. Ломоносова

zaitsev@philos.msu.ru

**Аннотация.** В докладе будет рассмотрена проблема неадекватной обработки большими языковыми моделями естественных рассуждений. Обосновывается предположение, что источник этих затруднений связан с отсутствием пространственных представлений, что, в свою очередь, обусловлено отсутствием сенсорного входа и перцептивной информации. DOI: 10.52119/LPHS.2024.51.82.003.

Большие языковые модели (LLM от английского Large Language Models) появились лишь пять лет назад, но уже успели приобрести популярность и хорошо зарекомендовать себя в решении задач, связанных с усвоением и порождением текста. В то же время оптимистические ожидания от использования генеративных языковых моделей в последние год-два натолкнулись на ряд препятствий. Как оказалось, эти модели недостаточно хорошо справляются с обычными естественными рассуждениями. Более того, проблема не только и не столько кроется в обработке естественных модифицируемых нечетких и т. п. рассуждений, но проявляет себя уже на уровне так называемых логических рассуждений — то есть рассуждений, для проверки которых используется язык классической логики и классический критерий правильности (сохранность истинности при переходе от посылок к заключению). Подробнее об этом см. работы [1–2].

Препятствием для больших моделей становятся и оценка рассуждений и их порождение. В докладе будут представлены и рассмотрены многочисленные примеры таких диалогов о рассуждениях с генеративными моделями. Как показывает анализ этих примеров, основные трудности в обработке рассуждений связаны с отсутствием у языковых моделей пространственных представлений. В связи с этим будет рассмотрен аргумент Д. Чалмерса [3] относительно возможности «чистого думателя» (а также «говорителя» и «понимателя») как аналога языковых моделей в пользу возможности последних «иметь сенсорные представления, понятия и переживания» на основе обработки текстовых данных.

В качестве альтернативной точки зрения обосновывается тезис о наличии сенсорного входа как необходимого условия моделирования мышления вообще и элементарных рассуждений в частности. В поддержку этого положения приводятся данные нейрокогнитивных исследований, фиксирующих сходные нейрональные механизмы пространственной и когнитивной навигации.

В качестве разумной альтернативы языковым моделям предлагается широкая концепция телесно воплощенного искусственного интеллекта и в частности активно развивающиеся мультимодальные генеративные модели [4–6].

**Литература**

1. Arkoudas K. *GPT-4 Can't Reason*. arXiv preprint. arXiv:2308.03762, 2023.
2. Ansari T. Big Tech's AI Models are Lost in Logic. *Analytics India Magazine*, September 2023, URL: <https://analyticsindiamag.com/big-techs-ai-models-are-lost-in-logic/> (accessed: 17.05.2024).
3. Chalmers D. *Does thought require sensory grounding? From pure thinkers to large language models*. Presidential Address at the 119<sup>th</sup> Eastern Division meeting of the American Philosophical Association, 2023.
4. Gupta A., Savarese S. et al. Embodied intelligence via learning and evolution. *Nature Communications*, 2021, 12, 5721.

*Д. В. Зайцев. Каково это — быть рассуждающей языковой моделью?*

5. Lanchantin J., Sukhbaatar S. et al. A data source for reasoning embodied agents. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37.7, 2023, p. 8438–8446.
6. Huang S., Dong L., Wang W., Hao Y., Singhal S., Ma S. et al. Language is not all you need: Aligning perception with language models. *Advances in Neural Information Processing Systems*, 2024, 36.