УДК 167

НАКЛАДЫВАЕТ ЛИ АЛГОРИТМОЦЕНТРИЗМ ОГРАНИЧЕНИЯ НА ФУНКЦИОНАЛ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА?

В.А. Бажанов

ФГБОУ ВО «Ульяновский государственный университет»,
г. Ульяновск, Россия

Поступила в редакцию: 11.08.24 В окончательном варианте: 12.09.24

Аннотация. В статье предпринимается попытка проанализировать функционал искусственного интеллекта (ИИ) с точки зрения особенностей работы алгоритмов, лежащих в его основании. Обсуждается концепция алгоритмоцентризма, показано, что границы функционала согласно этой концепции существуют и с ними необходимо считаться, чтобы реальнее представлять возможности ИИ. Возможности последнего также сравниваются с возможностями естественного интеллекта. Выражается мнение, что ИИ может стать эффективным инструментом, дополняющим естественный интеллект, но вряд ли сможет претендовать на его универсальную замену во всех ситуациях в случае сохранения архитектуры дискретного устройства.

Ключевые слова: алгоритм, алгоритмоцентризм, дискретный способ вычисления, аналоговый способ вычисления

DOES ALGORITHMOCENTRISM IMPOSE RESTRICTIONS ON THE FUNCTIONALITY OF ARTIFICIAL INTELLIGENCE?

V.A. Bazhanov

Ulyanovsk State University,
Ulyanovsk, Russia

Original article submitted: 11.08.24 Revision submitted: 12.09.24

Abstract. The article attempts to analyze the functionality of artificial intelligence (AI) from the point of view of the operating features of the algorithms that underlie it. The conception of algorithmocentrism is discussed, and it is shown that such boundaries, according to this conception, unconditionally exist, and they must be taken into account in order to more realistically comprehend the capabilities of AI. We claim that such limits exist, and they to be assessed in order to have more realistic picture representing AI capabilities. The latter compared with natural intelligence as well. The opinion is expressed that AI can and will be an effective tool, supplementing the natural one, but it is unlikely to replace it in any conceivable case of preserving the architecture of discrete device.

Keywords: algorithm, algorithmocentrism, discrete computation method, analog computation method

Алгоритмы известны и довольно активно используются с глубокой древности. История применения разного рода алгоритмов и алгоритмического мышления уходит в глубь веков [1]. Однако рождение и экспансия информационных технологий (включая шаги в направлении создания искусственного интеллекта – ИИ) заставили более пристально, чем раньше, взглянуть и проанализировать особенности работы алгоритмов, специфику и закономерности развития алгоритмического мышления. Такой анализ ставит серьезные эпистемологические вопросы, которые касаются природы и возможностей применения алгоритмов в разных сферах деятельности: что означает «мыслить алгоритмически»; каковы условия использования алгоритмов; допустимо ли рассуждать о границах и недостатках алгоритмического дискурса и алгоритмических процедур, тем более что всё чаще говорят уже и о нейронных алгоритмах [2]? Такого рода вопросы все настойчивее задаются и теми, кто непосредственно развивает информационные технологии, и теми, кто стремится осмыслить последствия их экспансии и применение этих технологий в самых разных сферах человеческой жизни [3]. Более того, ведущие мировые мыслители и ученые еще до момента прорыва в области ИИ выражали сомнения в том, достаточно ли серьезно человечество относится к достижениям в области создания ИИ [5].

Ключевые вопросы, связанные с возможностями алгоритмов и алгоритмического мышления, непосредственно затрагивают перспективы создания ИИ, в определенном смысле сравнимого с естественным интеллектом. И это вовсе не случайно, поскольку любой ИИ функционирует по некоторым алгоритмам.

Согласно общепринятому определению под алгоритмом понимаются процедуры поиска и решения некоторых задач в виде точных, последовательно выполняемых действий. Алгоритмы – вне зависимости от их вида (механические, вероятностные, циклические и т. д.) – должны быть эффективными, детерминированными, конечными и т. п. Все эти свойства могут быть реализованы на так называемых конечных автоматах, машинах Тьюринга, Поста посредством рекурсивных функций. Однако относительно работы едва ли не любого алгоритма можно сказать, что он представляет собой некоторое «бутылочное горлышко» (bottleneck), пропускной потенциал которого ограничен теми или иными факторами. Даже алгоритмы случайного перебора в информационных системах в действительности осуществляют не случайный, а псевдослучайный выбор (перебор).

Дело в том, что любой ныне действующий алгоритм является дискретным «устройством» и/или воспроизводится таковым. В случае цифровой репрезентации он может работать только с рациональными числами, тогда как всё множество действительных чисел принципиально ему недоступно. Между тем мозг человека нельзя сравнить с компьютером, который представляет собой дискретный автомат. Однако допустимо сравнение мозга, которому присуще мышление континуального типа, с аналоговой (вычислительной) машиной, которая по многим параметрам пока уступает «дискретным» компьютерам.

Действие алгоритмов сопряжено с множеством недостатков (особенно если речь идет о социально-политических или экономических процессах) [6]. Это и проблема воспроизводимости результатов вычислений [8], и проблема точности измерений (так называемая эпистемическая неопределенность), и проблема «плавающей запятой», точности репрезентации чисел, и необходимость

обращения к псевдослучайным генераторам чисел [9] и т. п. Известный тест Тьюринга в его классической форме по существу исключает апелляцию к «отелесненному» (embodied) субъекту, тогда как всё, что приписывается алгоритмам, приписывается «живым, т. е. отелесненным» субъектам, думающим и объективирующим знание значительно медленнее компьютера. Этот фактор, связанный с «отставанием» человека от результатов, выдаваемых компьютером, может быть решающим в те моменты, когда требуется незамедлительное важное решение, касающееся отдаленных перспектив. Более того, во многие алгоритмы незримо «впечатываются» имплицитные предпочтения и предвзятости (biases) субъекта, как, например, в случае расовых предрассудков в США, связанных с «каталогизацией» населения [10]. Нельзя также упускать из вида возможность алеаторической неопределенности. Вряд ли особенности действия алгоритмов в явном виде проявляются в модифицированных версиях теста Тьюринга и в тестах с аналогичными целями. «Большие языковые модели» (LLM) ChatGPT – 4, а также более продвинутые версии ChatGPT – 5, PresenAI, GFP-GAN, JADBio, Copy.ai, Notion.ai и т. д. далеко не столь совершенны, как обычно считают [11, 12]. Все модели такого рода вовлекают в свой функционал громадные массивы текстов, но они всегда задействуют тот их объем, который имел место на конкретный момент работы. Кроме того, как это ни парадоксально, эти модели испытывают значительные трудности и при соответствующих (даже арифметических) операциях с числовыми данными, поскольку как раз в вычислительных процедурах информационные технологии радикально превосходят естественный интеллект. Здесь ничего удивительного нет: если в текстах – умышленно или неумышленно – допущены арифметические ошибки или неточности, то они будут репрезентированы и в результатах работы LLM. Похожая ситуация наблюдается в простейших случаях дедуктивных рассуждений (типа простого категорического силлогизма, не говоря уж об индуктивных и тем более абдуктивных умозаключениях).

Кроме того, следует учитывать, что обучение ИИ (в частности, LLM) требует очень больших энергетических затрат со всеми сопутствующими финансовыми затратами и экологическими проблемами.

Всё это характеризуется как «алгоритмическая катастрофа», хотя такая характеристика и тяготеет к метафорическому выражению [12] (подробнее о концепции Ю. Хуэя говорится в [13]). Алгоритмизация приобретает особый статус в плане неопределенности функционала при поиске и выработке решения на базе нейросетей, в которых природа алгоритмических операций вообще может быть уподоблена «черному ящику»: внутренние процессы малодоступны для человеческого контроля и – во всяком случае пока – понимания.

Развитие ИИ (прежде всего «больших языковых моделей» типа GPT) поднимает нетривиальные вопросы о сохранении или модификации положений об авторских правах. Согласно статье № 1228 действующего ныне Гражданского кодекса РФ реальным автором интеллектуальной собственности является только лицо, творческий труд которого привел к тому, что она (интеллектуальная собственность) была создана. Поэтому в случае применения средств, предлагаемых LLM, честный и ответственный автор обязан сослаться на факт такого рода использования – аналогичным образом он в своем творчестве (работе) опирается на ресурсы Интернета.

Феномен алгоритмизации нашей деятельности, связанной с экспансие методов ИИ, требует обстоятельного анализа. Мы должны знать, на что способен и на что не способен ИИ. Только в этом случае опора на мощь и потенциал ИИ будет приводить к рациональным решениям.

Список литературы

1. A history of algorithms: from pebble to the microchip / Ed. J.-L. Chabert. – L.: N.Y.: Springer, 2015. – 524 p.
2. *Velickovic, P.* Neural algorithmic reasoning / P. Velickovic, G. Bendell // Patterns. – 2021. – Vol. 2. – Pp. 1–4.
3. *Beer D.* The tensions of algorithmic thinking: automation, intelligence, and the politics of knowing. Bristol: Bristol University Press, 2023. – 145 p.
4. *Padilla, L.M.* Limitations of algorithmic reasoning / L.M. Padilla. – L.: Our knowledge publ, 2023. – 52 p.
5. Transcendence looks at the implications of artificial intelligence – but are we taking AI seriously enough? / S. Hawking, S. Russell, M. Tegmark, F. Wilczek. – URL: https://www.independent.co.uk/news/science/stephen-hawking-transcendence-looks-at-the-implications-of-artificial-intelligence-but-are-we-taking-ai-seriously-enough-9313474.html (дата обращения 2024.04.25).
6. *Бажанов, В.А.* Искусственный интеллект, технологии Big Data (больших данных) и особенности современного политического процесса / В.А. Бажанов // Философия. Журнал ВШЭ. – 2023. – № 3. – С. 193–210.
7. *Филин, С.А.* Концепции знания и искусственного интеллекта применительно к инновационной сфере / С.А. Филин, А.Ж. Якушев. – Москва: РУСАЙНС, 2023. – 230 с.
8. *Бажанов, В.А.* Феномен воспроизводимости в фокусе эпистемологии и философии науки / В.А. Бажанов // Вопросы философии. – 2022. – № 5. – С. 25–35.
9. *Coveney, P.V.* When we can trust computers (and when we can't) / P.V. Coveney, R.R. Highfield // Philosophical Transactions A. – 2021. – Vol. A379. – Article 20200067.
10. *Obermeyer, Z.* Dissecting racial bias in an algorithm used to manage the health of population / Z. Obermeyer, B. Powers, C. Vogeli // Science. – 2019. – Vol. 366. – Pp. 447–453.
11. *Бажанов, В.А.* Можно ли всецело доверять искусственному интеллекту? Урок «Гёделя, Эшера, Баха» / В.А. Бажанов // Троицкий вариант. – 2023. – 25 июля. – № 383. – С. 12–13.
12. *Hui, Y.* Algorithmic catastrophe – the revenge of contingency / Y. Hui // Parrhesia. – 2015. – Vol. 23. – Pp. 122–143.
13. *Ивахненко, Е.Н.* Навстречу «новой эпистемологии»: рекурсивность и контингентность Юка Хуэя / Е.Н. Ивахненко // Эпистемология и философия науки. – 2022. – № 3. – С. 220–233.
14. *Rowbottom, D.P.* Does the no miracles argument apply to AI? / D.P. Rowbottom, W. Peden, A. Curtis-Trudel // Synthese. – 2024. – Vol. 203. – Article 173.

References

1. A history of algorithms: from pebble to the microchip / Ed. J.-L. Chabert. L.: N.Y.: Springer, 2015. 524 p.
2. *Velickovic P, Bendell G.* Neural algorithmic reasoning. Patterns. 2021. Vol. 2. Pp. 1–4.
3. *Beer D.* The tensions of algorithmic thinking: automation, intelligence, and the politics of knowing. Bristol: Bristol University Press, 2023. 145 p.

4. *Padilla LM.* Limitations of algorithmic reasoning. L.: Our knowledge publ, 2023. 52 p.

5. *Hawking S, Russell S, Tegmark M, Wilczek F.* Transcendence looks at the implications of artificial intelligence – but are we taking AI seriously enough? Available from: https://www.independent.co.uk/news/science/stephen-hawking-transcendence-looks-at-the-implications-of-artificial-intelligence-but-are-we-taking-ai-seriously-enough-9313474.html (accessed 2024.04.25).

6. *Bazhanov VA.* Artificial intelligence, Big Data technologies and features of the modern political process. Philosophy. HSE Journal. 2023;3:193-210.

7. *Filin SA, Yakushev AZh.* Concepts of knowledge and artificial intelligence in relation to the innovation sphere. Moscow: RUSAINS, 2023. 230 p. (In Russ.)

8. *Bazhanov VA.* The phenomenon of reproducibility in the focus of epistemology and philosophy of science. Problems of Philosophy. 2022;5:25-35. (In Russ.)

9. *Coveney PV, Highfield RR.* When we can trust computers (and when we can't). Philosophical Transactions A. 2021. Vol. A379. Article 20200067.

10. *Obermeyer Z, Powers B, Vogeli C.* Dissecting racial bias in an algorithm used to manage the health of the population. Science. 2019;366:447-453.

11. *Bazhanov VA.* Can artificial intelligence be completely trusted? Lesson from "Gödel, Escher, Bach". Troitskii variant. 2023. July 25. No. 383. Pp. 12–13. (In Russ.)

12. *Hui Y.* Algorithmic catastrophe – the revenge of contingency. Parrhesia. 2015. Vol. 23. Pp. 122–143.

13. *Ivakhnenko EN.* Towards a "new epistemology": Yuk Hui's recursiveness and contingency. Epistemology and Philosophy of Science. 2022. No. 3. Pp. 220–233. (In Russ.)

14. *Rowbottom DP, Peden W, Curtis-Trudel A.* Does the no miracles argument apply to AI? Synthese. 2024. Vol. 203. Article 173.

Информация об авторе

БАЖАНОВ Валентин Александрович – Заслуженный деятель науки РФ, доктор философских наук, профессор, заведующий кафедрой философии ФГБОУ ВО «Ульяновский государственный университет», г. Ульяновск, Россия; eLibrary SPIN: 2644-7587.
E-mail: vbazhanov@yandex.ru

Information about the author

BAZHANOV Valentin A. – Honored Scientist of the Russian Federation, Doctor of Philosophy, Professor, Head of the Department of Philosophy, Ulyanovsk State University, Ulyanovsk, Russia; eLibrary SPIN: 2644-7587.
E-mail: vbazhanov@yandex.ru