

DOI: 10.26794/2226-7867-2019-9-4-52-56

УДК 004.8(045)

ФИЛОСОФСКИЕ ОСНОВАНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Махаматов Таир Махаматович, д-р филос. наук, профессор, профессор Департамента социологии, истории и философии, Финансовый университет, Москва, Россия
makhamatov.tair@mail.ru

Аннотация. В статье обосновывается тезис о том, что развитие технологии искусственного интеллекта тесно связано не только с открытиями в области естествознания, антропологии и медицины, но и с достижениями в сфере философии познания и когнитивных наук. Свой тезис автор обосновывает философско-эпистемологическим анализом проблем по совершенствованию нейронной сети как ядра современного искусственного интеллекта и делает вывод о том, что принципы функционирования нейронной сети соответствуют принципам сенсуализма Дж. Локка, априоризма И. Канта и др. Результаты сравнительного исследования позволили автору прийти к следующему заключению: совершенствование исследованных С. Хайкином, С. Расселом, П. Норвигом способностей нейронной сети («очевидность ответа», «классификация образов» и «достоверность принимаемого решения») возможно при опоре на гносеологические идеи Дж. Локка, использовании кантовских принципов («синтетическое единство апперцепции», «я мыслю») и поиске алгоритма формирования нейронной сетью способности создавать антиномии в искусственном интеллекте. Дальнейшее развитие искусственного интеллекта, основанного на нейронной сети, может опираться также на теорию познания Т. Гоббса, Р. Декарта, Б. Спинозы, Г.В.Ф. Гегеля и результаты современных когнитивных наук.

Ключевые слова: антиномии разума; априорные понятия; вторичные качества; нейронная сеть; первичные качества; сенсуализм; синтетическая апперцепция

PHILOSOPHY OF ARTIFICIAL INTELLIGENCE

Makhamatov T.M.,

Doctor of Philosophy, Professor, Department of Sociology, History and Philosophy, Financial University, Moscow
makhamatov.tair@mail.ru

Abstract. In the article, the author substantiates the thesis that the development of artificial intelligence technology is closely related not only to discoveries in the field of natural science, anthropology and medicine, but also achievements in the field of philosophy of knowledge and cognitive sciences. The author conducted a philosophico-epistemological analysis of the problems of improving the neural network as the core of modern artificial intelligence led to the conclusion that the principles of functioning of the neural network corresponding to such principles of the cognitive process discovered and studied in the philosophical concepts of New Time, such as J. Locke's apriorism, I. Kant's apriorism other. The results of the comparative study allowed the author to come to the following conclusion: the improvement of the abilities of the neural network studied by S. Haikin, S. Russell, P. Norvig ("evidence of the answer", "classification of images" and "reliability of the decision") is possible when relying on the epistemological ideas of J. Locke, using Kant's principles ("synthetic unity of apperception", "I think") and searching for the algorithm of neural network formation of the ability to create antinomies in artificial intelligence. Further development of artificial intelligence based on the neural network can also be based on the theory of cognition of T. Hobbes, R. Descartes, B. Spinoza, G. V.F. Hegel and the results of modern cognitive sciences.

Keywords: antinomies of mind; a priori concepts; secondary qualities; neural network; primary qualities; sensationalism; synthetic apperception

ВВЕДЕНИЕ

История промышленных революций и современные технологические революции демонстрируют, что возникновение и распространение определенного типа мышления (речь идет о «конструктивном мышлении, основанном на схемах»), явилось «важнейшим фактором процесса масштабного экономического роста» [1, с. 39]. Сравнительный анализ логико-эпистемологических проблем нейронной сети и раскрытых Р. Декартом, Т. Гоббсом, Дж. Локком и И. Кантом логических ступеней познавательного процесса показывает, что результаты философских исследований принципов мышления и в настоящее время остаются важнейшим фактором качественного и количественного роста искусственного интеллекта как основы развития производительных сил современных передовых экономик. Однако рассмотрение философско-эпистемологических принципов как теоретических факторов развития технологий до сих пор остается вне поля зрения исследователей искусственного интеллекта, о чем свидетельствует полное отсутствие в работах о нейронной сети упоминаний имен и трудов вышеназванных и современных философов [2].

В наше время обострение международной конкуренции обусловило переход к расширению практического применения *искусственного интеллекта, опирающегося на принципы конструктивистского мышления*. Как отмечают В.В. Вдовин, П.Г. Щедровицкий, «смысл конструктивизации в том, что дело расслаивается на явно оформленные деятельную и мыслительную части, причем деятельное мышление получает инструмент схематизации (чертеж) как форму увещивания и мыслительной имитации замысла и работ по его воплощению, а также — самого процесса „замаливания“, постепенно приобретающая форму конструктивного мышления...» [1, с. 43].

Основное назначение искусственного интеллекта — облегчить умственный труд человека. В последнее время развитие искусственного интеллекта перешло на качественно новый этап. Его совершенствование уже не ограничивается только универсальными формами умственной деятельности, изучением природы человеческого мозга, но основывается на логических принципах творческого, конструктивного мышления человека, памяти [3, 4], философско-эпистемологических способностей [5] и психологии познания и мышления [6, 7]. В своем фундаментальном труде «Нейронные сети» (Neural Networks) Саймон Хайкин пишет, что «предметная область нейронных сетей лежит на пересечении многих наук. Ее корни уходят в нейробиологию, ма-

тематику, статистику, физику, науку о компьютерах и инженерии» [2, с. 989]. То же самое подтверждают исследования Стюарта Рассела и Питера Норвига [8, с. 37–38]. Здесь, на наш взгляд, для дальнейшего развития искусственного интеллекта, особенно его современного ядра — нейронных сетей — было бы целесообразно использовать достижения и поставленные проблемы сенсуалистической теории познания Джона Локка, априоризма И. Канта, что до сих пор остается вне поля зрения многих исследователей искусственного интеллекта.

УЧЕНИЕ ДЖ. ЛОККА ОБ ИДЕЯХ ПЕРВИЧНЫХ И ВТОРИЧНЫХ КАЧЕСТВ И АКТУАЛЬНЫЕ ЗАДАЧИ НЕЙРОННЫХ СЕТЕЙ

Исследования механизма поэтапного функционирования человеческого мозга и нейронной сети как ядра искусственного интеллекта показывают, что первоначальный шаг познавательного процесса опирается на принципы сенсуализма. Согласно С. Хайкину, «большая часть усилий исследователей нейронных сетей была сфокусирована на задаче *распознавания образов*. Учитывая практическую важность этой задачи и ее повсеместную природу, а также тот факт, что нейронные сети исключительно хорошо подходят для решения задачи классификации, такая концентрация усилий ученых направлялась на поиск средств корректной классификации. Развивая это направление, стало возможным заложить основы *адаптивной классификации образов* (adaptive pattern classification). Однако мы достигли той точки, в которой системы классификации должны рассматриваться в более широком смысле, если мы хотим решать задачи классификации более сложной и интеллектуальной природы» [2, с. 990] (*курсив автора*).

На наш взгляд, решению этой задачи поможет сенсуалистическая теория познания Дж. Локка, согласно которой источником наших знаний являются чувственный опыт и рефлексия. Под рефлексией Дж. Локк подразумевает «то наблюдение, которому ум подвергает свою деятельность... вследствие чего в разуме возникают *идеи* этой деятельности» [7, с. 155]. Чувственный опыт дает разуму знания о таких качествах тел, как плотность, протяженность, форма, движение или покой и число. Эти качества Дж. Локк называет первичными и реальными. Идеи *этих* «качеств тел сходны с ними, и их прообразы действительно существуют в самих телах» [7, с. 186]. Такие «качества, как цвета, вкусы, звуки и т.д., которые на деле не играют никакой роли в самих вещах, но представляют собой силы, вызывающие в нас

различные ощущения *первичными качествами* вещей, т.е. объемом, формой, строением и движением их незаметных частиц», он называет вторичными качествами [7, с. 184]. Идеи, вызываемые в человеке этими качествами, вовсе не имеют сходства с телами. *Свет, тепло, белизна или холод реальны в телах «не более чем недомогание или боль — в манне. Уберите эти ощущения. Пусть глаза не видят света или цветов, пусть уши не слышат звуков, небо не ощущает вкуса, нос не обоняет — и все цвета, вкусы, запахи и звуки как особые идеи исчезнут, прекратят существование и сведутся к своим причинам, т.е. к объему, форме и движению частиц»* [7, с. 187].

Эти рассуждения очень близки к той задаче, о которой пишет С. Хайкин: «В контексте задачи классификации образов можно разработать нейронную сеть, собирающую информацию не только для определения конкретного класса, но и для увеличения *достоверности* (confidence) принимаемого решения. Впоследствии эта информация может использоваться для исключения сомнительных решений» [2, с. 35]. Задачи классификации, если отталкиваться от идей *первичных и вторичных качеств* Дж. Локка, тесно связаны с разработкой механизмов обнаружения взаимной обусловленности этих качеств предметов нейронными сетями искусственного интеллекта. Только в этом случае можно говорить, что, как пишет С. Хайкин, «использование нейронных сетей обеспечивает...очевидность ответа (evidential response)» [2, с. 35].

ЭПИСТЕМОЛОГИЧЕСКИЙ КОНСТРУКТИВИЗМ И. КАНТА И ПЕРСПЕКТИВЫ РАЗВИТИЯ НЕЙРОННЫХ СЕТЕЙ

Другая важнейшая проблема искусственного интеллекта, на наш взгляд, заключается в вопросе его *творческо-эпистемологического* конструктивизма мышления. Такие исследователи искусственного интеллекта и его современного ядра — нейронных сетей, как Дж. Баррат, М. Т. Джонс, Р. Курцвейг, Э. Бернхард, Д. Кэссин [9–12] и др., пишут о тенденции ускорения «эволюционного процесса» (за счет усложнения абстракции), создании саморазвивающегося искусственного интеллекта, который в скором будущем якобы превзойдет человека. Только немногие ученые, в числе которых С. Хайкин, пишут о границах возможности искусственного интеллекта, определяемых совокупным человеческим интеллектом. «Очень важно уяснить, что для создания компьютерной архитектуры, которая будет способна имитировать человеческий мозг

(если такое окажется возможным вообще), — пишет С. Хайкин, — придется пройти долгий и трудный путь» [2, с. 33].

В основании «компьютерной архитектуры, способной имитировать человеческий мозг» наряду с принципами сенсуализма Дж. Локка находится априоризм И. Канта. Программное обеспечение искусственного интеллекта, если говорить языком философии, по сути, является его априорной понятийной структурой. Его рамки определяют границы всех возможностей искусственного интеллекта, в том числе способности определения объекта оперативного контакта, т.е. способность *эпистемологического конструирования объекта*. «В структуру нейронных сетей должны быть встроены априорная информация и инварианты, что упрощает архитектуру сети и процесс ее обучения». Для выполнения такой задачи «необходимо понять, как разработать специализированную структуру, в которую встроена априорная информация. К сожалению, в настоящее время не существует четкого решения этой задачи» [2, с. 13, 62]. Здесь мы видим сходство с процессом образования «явления» в рассудке познающего субъекта, выступающего как результат соединения чувственных данных с априорными понятиями рассудка, о чем писал И. Кант в своей «Критике чистого разума».

Как известно, в наше время теория познания И. Канта, особенно его учение о явлении, вполне обоснованно рассматривается как эпистемологический конструктивизм [13]. Кантовское понятие «явление» с позиции логики научного познания *есть ни что иное как сконструированный посредством априорных понятий факт познания*. «Естественноиспытатели поняли, — пишет Кант, — что разум видит только то, что сам создает по собственному плану, что он с принципами своих суждений должен идти впереди, согласно постоянным законам, и заставлять природу отвечать на его вопросы, а не тащиться у нее словно на поводу, так как в противном случае наблюдения, произведенные случайно, без заранее составленного плана, не будут связаны необходимым законом, между тем как разум ищет такой закон и нуждается в нем» [13, с. 16]. Немецкий философ-неокантианец Э. Кассирер в «Феноменологии познания», в 3-м томе своего труда «Философия символических форм», анализируя кантовское понимание философии познания, особо выделяет эпистемологический конструктивизм И. Канта. Он пишет: «Знание не описывается ни как часть бытия, ни как его

отражение. Тем не менее ничуть не убывает его *соотнесенность* с бытием, которая, скорее, получает свое обоснование с новой точки зрения. *Функцией* знания оказывается построение и конституирование предмета — уже не абсолютного, но обусловленного именно этой функцией — как «явленного предмета». То, что мы называем «объективным» бытием, предметом опыта, возможно лишь при наличии предпосылаемого ему рассудка и его априорных объединяющих функций» [14, с. 14].

Одной из проблем нейронной сети, связанной с принципами «Критики чистого разума» И. Канта, является проблема формирования процесса обучения нейронной сети. Ее, на наш взгляд, следует рассматривать через призму двух гносеологических целей. **Первая цель:** спроектировать в структуре нейронной сети принцип «я мыслю», *т.е. самосознание* и его связь восприятиями-созерцаниями нейронной сети, который мог бы обеспечить их единство. Кант писал, что «все многообразное в созерцании имеет, следовательно, необходимое отношение к [суждению] “я мыслю” в том самом субъекте, в котором это многообразное находится». Принцип «я мыслю» обеспечивает единство и синтез всех представлений, а также «возможность априорного познания на основе этого единства» [13, с. 100].

Вторая цель: спроектировать в нейронной сети способность мыслить о «вещи в себе» и формировать антиномии, что означает выход за рамки действий в пространстве непосредственного восприятия и возможность творческого мышления.

Понятие «вещь в себе» выражает *попытку* Канта перейти к *эпистемологии теоретического знания* и демонстрирует наличие качественной глубокой разницы между чувственным и рациональным уровнями познания. Кант замечает, что кумулятивное накопление опытного знания *не может перевести* познание на рациональный уровень. Поэтому «вещь в себе» выступает как критерий *абсолютной* ограниченности наших чувственных восприятий, опыта.

В то же время кантовская «вещь в себе» как *мыслимая* сущность — есть показатель познавательных возможностей рационального, опережающего отражения теоретического мышления: «Что же касается предметов, которые мыслятся только разумом, и притом необходимо, но которые (по крайней мере, так, как их мыслит разум) вовсе не могут быть даны в опыте, то попытки мыслить их (ведь должны же они быть мыслимы) дадут нам затем превосходный критерий того, что мы считаем измененным мето-

дом мышления, а именно, что мы *a priori* познаем о вещах лишь то, что вложено в них нами самими» [13, с. 19].

По существу, кантовское *мышление* о вещах в себе, т.е. о сущностях, не данных в опыте, является *внеопытным, теоретическим познанием* внечувственных, теоретически конструированных объектов, что не связано с опытным, чувственным восприятием природы. Благодаря безусловности, «вещи в себе» Канта свободны от чувственной непосредственности, случайности и субъективности; они свою объективность и необходимость имеют в самих себе, т.е. они суть *causa sui*: «Так как мы познаем случайное только в опыте — здесь же речь идет о вещах, которые вовсе не должны быть предметами опыта, — то нам приходится выводить свое знание о них из того, что необходимо само по себе, — из чистых понятий о вещах вообще. Вот почему первый шаг, сделанный нами за пределы чувственно воспринимаемого мира, заставляет нас начинать свои новые знания с исследования абсолютно необходимой сущности и из ее понятий выводить понятия о всех вещах, поскольку они чисто умопостигаемые...» [13, с. 345]. Кантовские понятия «вещи в себе» и «антиномии чистого разума» следует рассматривать как философскую основу возможности разработки способностей нейронной сети *теоретического конструирования*. Нам представляется, что в принципе возможно обучить нейронную сеть создавать антиномии между своими способностями распознавать, классифицировать объекты в их трансформации и формировать чисто априорные, «умозрительные» схемы объектов и действий. Достижение этой задачи позволит переходить к разработке нейронной сети, обладающей способностью обучаться мыслить о развивающихся системах в рамках «структурной инвариантности» и максимизировать достоверность выводов и принимаемых решений.

ЗАКЛЮЧЕНИЕ

Анализ достижений и актуальных задач нейронных сетей как ядра современного искусственного интеллекта показывает, что многие проблемы, обсуждаемые в этой области, давно были исследованы в теориях познания Т. Гоббса, Р. Декарта, Б. Спинозы и особенно в трудах Дж. Локка и И. Канта. Рассмотрение достижений и перспективных задач дальнейших разработок в области нейронной сети показывает, что совершенствование ее «способностей имитировать человеческий мозг» повторяет эволюцию человеческого

познания от сенсуализма к рационализму. Поэтому обращение к идеям Т. Гоббса о номинализме в познании, трудам “*Cogito, ergo Sum*” Р. Декарта, “*Causa Sui*” Б. Спинозы и особенно сенсуалистической теории познания Дж. Локка, априоризму И. Канта, феноменологии духа Г.Ф.В. Гегеля и исследованиям современных когнитивных наук послужит философско-методологическим основанием решения многих проблем и задач развития искусственного интеллекта.

В представленной статье автор коснулся лишь наиболее близких к задачам нейронных сетей аспектов учений Джона Локка и Иммануила Канта. Более детальное и глубокое исследование проблем искусственного интеллекта в диалектической взаимосвязи с достижениями и проблемами гносеологии и эпистемологии даст плодотворные результаты как ученым-разработчикам нейронных сетей, так и философам, изучающим диалекту познавательного процесса.

СПИСОК ИСТОЧНИКОВ

1. Вдовин В.В., Щедровицкий П.Г. Конструктивное мышление: неучтенный фактор развития. *Вопросы философии*. 2018;(9):39–49.
2. Хайкин С. Нейронные сети: полный курс, 2-изд., испр. Пер. с англ. М.: Вильямс; 2006.
3. Hoskins A. Memory Ecologies. *Memory Studies*. 2016;(3):348–357.
4. Лекторский В.А. и др. Когнитивный подход. Монография. Лекторский В.А., ред. М.: Канон+ РООИ «Реабилитация»; 2008.
5. Harre R. Personal Being: A Theory for Individual Psychology. Harvard Univ. Press; 1984.
6. Harré R., & Gillett G. The discursive mind. London, England: Sage Publications; 1994.
7. Локк Дж. Опыт о человеческом разумении. Соч. в 3-х т.: Т. 1. Пер. с англ. М.: Мысль; 1985.
8. Рассел С., Норвиг П. Искусственный интеллект. Современный подход. 2-е изд. Пер. с англ. М.—СПб.: Диалектика; 2019.
9. Баррат Дж. Последнее изобретение человечества. Искусственный интеллект и конец эры Homo Sapiens. Пер. с англ. М.: Альпина нон-фикшн; 2015.
10. Джонс М.Т. Программирование искусственного интеллекта в приложениях. Пер. с англ. М.: ДМК Пресс; 2006.
11. Курцвейг Р. Эволюция разума. Пер. с англ. М.: ЭКСМО; 2015.
12. Barnard E. & Casasen D. Invariance and neuoral nets, IEEE Transactions. *Neural Networks*. 1991;2(5):498–508.
13. Кант И. Критика чистого разума. Пер. с нем. М.: Мысль; 1994.
14. Кассирер Э. Философия символических форм. Т. III: Феноменология познания. М.: Академический проект; 2011.

REFERENCES

1. Vdovin V.V., Schedrovitsky G.P. Constructive thinking: The unaccounted development factor. *Voprosy filosofii*. 2018;(9):39–49. (In Russ.).
2. Haykin S. Neural Networks: A Comprehensive Foundation. 2nd ed. Transl. from English. Moscow: Williams publishing house; 2006.
3. Hoskins A. Memory Ecologies. *Memory Studies*. 2016;(3):348–357.
4. Lektorsky V.A. et al. Cognitive approach. Monograph. Moscow: “Canon+”; 2008.
5. Harré R. Personal Being: A Theory for Individual Psychology. Harvard University Press; 1984.
6. Harré R., Gillett G. The discursive mind. London, England: Sage Publications; 1994.
7. Locke J. The experience of human understanding. Works in 3 vol. Vol. 1. Transl. from English. Moscow: Thought; 1985.
8. Russell S., Norvig P. Artificial intelligence. Modern approach. 2ne ed. Transl. from English. Moscow-St. Petersburg: Dialectics; 2019.
9. Barratt J. Our Final Invention: Artificial Intelligence and the End of the Human Era. Transl. from English. Moscow: Alpina non-fiction; 2015. (In Russ.).
10. Jones M.T. AI Application Programming. Transl. from Eng. Moscow: DMK Press; 2006. (In Russ.).
11. Kurzweil R. Evolution of mind. Transl. from Eng. Moscow: EKSMO, 2015.
12. Barnard E., Casasent D. Invariance and neural nets, IEEE Transactions. *Neural Networks*. 1991;2(5):498–508.
13. Kant I. Critique of pure reason. Transl. from German. Moscow: Thought; 1994. (In Russ.).
14. Cassirer E. Philosophy of symbolic forms. Vol. III: The phenomenology of knowledge. Transl. from German. Moscow: Akademicheskii Proekt; 2011. (In Russ.).