

НАУЧНЫЙ ПОТЕНЦИАЛ: РАБОТЫ МОЛОДЫХ УЧЕНЫХ

DOI: 10.17805/zpu.2016.4.26

«Простая» и «трудная» проблемы искусственного интеллекта в аналитической философии сознания

А. С. Гау

(Елабужский институт Казанского (Приволжского) федерального университета)

Данная статья посвящена рассмотрению двух основных аспектов проблематики искусственного интеллекта в современной аналитической философии сознания. Выделение из сферы метального двух ее аспектов — психологического и феноменального — позволяет, по аналогии с выделением двух аспектов соотношения телесного и ментального — говорить о существовании «простой» и «трудной» проблемы искусственного интеллекта.

Сущность «простой» проблемы искусственного интеллекта состоит в исследовании возможностей достижения ситуации возникновения психологического сознания на основе воспроизведения основ работы естественного человеческого мозга; сущность «трудной» — в изучении возможности переживания феноменальных ментальных состояний в рамках нейрокомпьютерной модели сознания.

В свою очередь, «трудная» проблема искусственного интеллекта предстает в двух аспектах. Первый из них связан с вопросом о возможности формализации работы мозга, предполагающей не только физическое воссоздание человеческого мозга, но и имеющей своим следствием возникновение феноменального сознания. Второй же аспект состоит в вопросе о том, можно ли вообще говорить о наличии сознания применительно к искусственной системе.

На основании представленного анализа делается вывод о том, что вопрос о возможности создания искусственного интеллекта является не столь существенным для исследования проблемы сознания. Сегодня философии необходимо «вынести за скобки» проблему возможности возникновения сознания в системе, манипулирующей физическими символами в соответствии с правилами, учитывающими их структуру, и сосредоточиться на исследовании свойств искусственного интеллекта безотносительно его технической составляющей. Это поставит нас перед одной из наиболее интересных проблем современной аналитической философии сознания, решение которой откроет совершенно новые перспективы в этой области.

Ключевые слова: аналитическая философия; искусственный интеллект; ментальные состояния; сознание; нейрокомпьютерная модель

ВВЕДЕНИЕ

Искусственный интеллект сегодня является предметом исследования самых разнообразных дисциплин как естественно-научной, так и гуманитарной направленности. Изучение возможностей формализации работы нейронов головного мозга человека, осуществляемое в современной нейронауке, дает основания для предположения о возможности возникновения ментальных состояний в такой системе и наделения

самой нейрокомпьютерной модели всеми теми свойствами, которые обычно приписываются человеческому сознанию: способности к обучению, переживанию и соответствующему выражению эмоций, даже творческих способностей. Однако это же порождает вопросы о том, каким статусом будут обладать подобные ментальные состояния, возникающие в искусственно созданной модели человеческого мозга: будет ли такое сознание обладать интенциональностью, будет ли его опыт иметь качественные и субъективные характеристики, наконец, будет ли оно вообще способно к переживанию феноменального опыта (Boden, 1973; Guttenplan, 1988; Harmer, 2010; Helman, 1987; Hochstein, 2016; Moor, 1978; Sanders, 1985; Searle, 1990; Williamson, 2003).

Поиск ответов на эти и подобные вопросы осуществляется в современной аналитической философии сознания, одной из наиболее значимых идей которой является положение о необходимости разведения двух аспектов ментального: психологического и феноменального. Согласно соответствующей точке зрения, все доступные нам феномены делятся лишь на два класса: те, к которым мы имеем доступ с позиции «первого лица», и те, к которым мы имеем доступ с позиции «третьего лица». При этом первые могут быть сведены в конечном итоге лишь к объяснению поведения, реакциям взаимодействия с окружающей средой, другими людьми и т. д. Эти феномены относятся к психологическому аспекту сознания, который ответственен за регуляцию физических и химических процессов, происходящих в человеческом организме и способных выражаться в ментальных проявлениях (таких как эмоции, адаптивные преобразования, поведение в ответ на сигналы, поступающие в мозг и т. п.). Вторая же группа феноменов — те, переживание которых происходит непосредственно, во всей чистоте нашего Я, относятся к феноменальному аспекту сознания и сводятся в конечном итоге к непосредственному опыту обращения к самому себе (тому самому cogito, которое остается в результате осуществления радикальной феноменологической редукции) (Чалмерс, 2013: 40; Weisberg, 2011: 539).

Данное положение имеет чрезвычайную значимость для построения основ фундаментальной теории сознания, возникновение которой способно дать ответы на вопросы о его возникновении, сущности и функциях. В то же время разделение ментальных состояний на феноменальные и психологические способно заложить основания для философских исследований *искусственного интеллекта*, который понимается нами как свойство искусственно созданных нейрокомпьютерных моделей человеческого мозга получать и перерабатывать исходящие сигналы в ментальные состояния.

Цель данной статьи состоит в исследовании философской сущности искусственного интеллекта в соответствии с выделяемыми в современной аналитической философии сознания двумя аспектами сферы ментального.

ИССЛЕДОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

В АНАЛИТИЧЕСКОЙ ФИЛОСОФИИ: «ПРОСТАЯ» И «ТРУДНАЯ» ПРОБЛЕМЫ

Действительно, на наш взгляд, благодаря разделению сферы ментального на психологическое и феноменальное, по аналогии с принятым в аналитической философии сознания выделением двух аспектов проблемы соотношения телесного и ментального, можно говорить о выделении «простой» и «трудной» проблемы искусственного интеллекта. В первом случае речь идет о достижении ситуации возникновения психологического сознания на основе воспроизведения основ работы естественного человеческого мозга; во втором — о возможности переживания феноменальных ментальных состояний в рамках нейрокомпьютерной модели сознания.

«Простота» первой из названных нами проблем объясняется тем, что сама по себе формализация работы нейронов головного мозга не представляется действительно трудной задачей: несмотря на то что сегодня мы еще не можем говорить об актуальном создании такой искусственной нейронной системы, которая имела бы своим следствием возникновение у машины, работающей на ее основе, сознания, решение данного вопроса лежит скорее в области науки, а не философии. Точное же воспроизведение естественных нейронных сетей, очевидно, повлечет за собой возникновение психологических аспектов сознания. Так, уже сегодня мы можем говорить о существовании машин, способных к обучению, переработке поступающих исходящих сигналов и осуществлению действий в соответствии с этими сигналами (иными словами, у таких машин уже присутствуют каузальные предпосылки сознательного поведения, что и составляет сущность психологического понятия ментального).

Вопрос же о возможности создания искусственного интеллекта, обладающего способностью к действительному переживанию опыта (то есть имеющего феноменальную составляющую сознания) предполагает два основных аспекта. Первый из них связан с вопросом о возможности формализации работы мозга, предполагающей не только предметное, физическое воссоздание человеческого мозга в его целостности, но и имеющей своим следствием возникновение феноменального сознания на основе *искусственной* нейронной системы. Второй же аспект заключается в вопросе о том, можно ли вообще говорить о наличии сознания, разума, интеллекта или интенциональности применительно к такого рода искусственной системе.

Как и в случае с «простой» проблемой искусственного интеллекта, первый аспект данной проблематики относится к сфере исключительно научных исследований, осуществляемых в данном направлении, а ответ на данный вопрос может быть получен (возможно, уже в обозримом будущем) опытным путем. В данном отношении центральной задачей, стоящей сегодня перед исследователями нейрокомпьютеров, является достижение так называемого эффективного параллелизма, сущность которого заключается в том, что производительность компьютера возрастает намного медленнее, чем количество простейших процессоров в искусственной нейронной системе. Для преодоления этой проблемы используются различные подходы, основным из которых является метод, при котором для различных типов задач строятся параллельные алгоритмы их решения с целью обеспечения возможности возникновения эффективного аппарата производства параллельных программ. Так или иначе, мы видим, что достижение эффективного параллелизма, способное повлечь возникновение сознания, является сегодня не философской, а технической задачей.

Что же касается второго аспекта «трудной» проблемы искусственного интеллекта, необходимо отметить, что в аналитической философии весьма распространена точка зрения, согласно которой сознание вообще не может возникнуть в какой бы то ни было формализованной системе. Иллюстрацией здесь может служить теоретический эксперимент, активно защищаемый уже в течение тридцати лет Дж. Серлом, который называется «китайская комната». Мысленный эксперимент, или модель, имеет своим назначением проиллюстрировать следующее положение: несмотря на то что ответы и вопросы осуществляющегося в ней диалога верно соответствуют друг другу, в «китайской комнате» все же не возникает никакого понимания. То есть до тех пор, «пока программа определяется в терминах вычислительных операций над формально определенными элементами, она не имеет связи с пониманием» (Searle, 1980: 421) и, следовательно, не обладает сознанием. Это объясняется тем, что формальные дей-

ствия сами по себе не имеют интенциональности, более того, они даже ничего не символизируют (то есть имеют лишь синтаксис, но не семантику). Интенциональность компьютера существует лишь в сознании тех, кто его программирует и использует.

Различие между мозгом и компьютером в этом случае состоит в том, что процессы, которые определяют нечто как компьютер, не зависят от типа аппаратной реализации, в то время как в мозгу наблюдается чрезвычайная специфичность анатомии и физиологии: «мозг не просто демонстрирует формальные процедуры или программы <...>, но и вызывает ментальные события благодаря специфическим нейробиологическим процессам» (Searle, 1990: 28). Применительно к осуществленному нами выше разведению двух аспектов ментального, мы можем перефразировать это следующим образом: в отличие от мозга, работа которого сопровождается возникновением как психологических, так и феноменальных ментальных состояний, нейрокомпьютерная модель может обладать лишь психологическим сознанием (то есть может осуществлять какие-либо действия в соответствии с заранее заданной программой, без действительного переживания ментального опыта).

ЗАКЛЮЧЕНИЕ

Вообще же, необходимо отметить, что для исследования проблемы сознания (как в плане изучения отдельных ее аспектов, так и в плане построения фундаментальной теории сознания), не столь существенным, на наш взгляд, оказывается вопрос о том, возможно ли вообще создать искусственный интеллект. Более того, сосредоточение на вопросе, поставленном именно так, лишает аналитическую философию сознания огромного количества интереснейших тем для исследований.

Действительно, на наш взгляд, в ходе исторического развития аналитической философии сознания произошло изменение вопроса о возможности создания искусственного интеллекта: вопрос «Может ли машина мыслить?» сменился на «Может ли мыслить машина, манипулирующая физическими символами в соответствии с правилами, учитывающими их структуру?»

Это замечание демонстрирует развитие философских исследований в области искусственного интеллекта, связанное с переходом от первоначального оптимизма лингвистического идеализма первой половины прошлого столетия к философскому реализму шестидесятых годов (в определенной степени занимающему ведущие позиции вплоть до настоящего времени), повлекшему за собой рассмотрение возможностей формализации работы естественных нейронных систем, то есть перенос интереса с первого на второй из указанных вопросов.

На наш взгляд, именно в различии этих двух форм вопросов заключается основная проблема, связанная с природой сознания в аспекте нейрокомпьютерного моделирования. Существующие сегодня мысленные эксперименты, такие как, например, «китайская комната», достаточно убедительно демонстрируют тот факт, что ответ на второй вопрос может быть только отрицательным. Однако первого вопроса они даже не затрагивают.

В то же время очевидным является и то, что от того, как именно поставлен вопрос, очень сильно зависит и ответ, который будет найден. Поэтому, на наш взгляд, основные проблемы, связанные с природой сознания вообще и искусственного интеллекта в частности должны лежать в сфере, которую намечает именно первый вопрос.

Это означает, кроме того, что необходимо вынести за скобки все проблемы, возникающие в связи с чисто формальными моделями мышления, и сосредоточиться на ги-

потезе о возможности существования искусственной системы, полностью соответствующей естественной связи нейронов головного мозга человека и, как следствие, способной производить сознательные акты, то есть мыслить.

Таким образом, на сегодняшний день ответ на вопрос о том, возможно ли возникновение сознания в искусственно созданной нейрокомпьютерной системе, скорее всего, будет отрицательным, поэтому дальнейшее исследование свойств такой системы покажется бессмысленным. Однако если заключить в скобки вопрос о возможности конструирования мыслящей машины, и, в рамках мысленного эксперимента, задать вопрос о том, каковы будут свойства уже существующего искусственного интеллекта и в чем будут его отличия от естественного, мы окажемся перед одной из наиболее интересных проблем современной аналитической философии сознания, решение которой откроет перед нами совершенно новые перспективы в области исследования сферы ментального.

СПИСОК ЛИТЕРАТУРЫ

- Чалмерс, Д. (2013) Сознający ум: В поисках фундаментальной теории : пер. с англ. М. : Книжный дом «ЛИБРОКОМ». 509 с.
- Boden, M. A. (1973) How artificial is artificial intelligence? // The British Journal for the Philosophy of Science. Vol 24. №1. P. 61–72.
- Guttenplan, S. (1988) Thinking Machines: The Evolution of Artificial Intelligence // Journal of Design History. № . P. 151–152. DOI: 10.1093/jdh/1.2.151
- Harmer, J. (2010) All in the Mind? // The Cambridge Quarterly. №39. P. 385–391. DOI: 10.1093/camqtly/bfq026
- Helman, D. H. (1987) Realism and Antirealism in Artificial Intelligence // The British Journal for the Philosophy of Science. Vol 38. №1. P. 19–26.
- Hochstein, E. (2016) Categorizing the Mental // The Philosophical Quarterly. №66. P. 745–759. DOI:10.1093/pq/pqw001
- Moor, J. H. (1978) Three Myths of Computer Science // The British Journal for the Philosophy of Science. Vol 29. №3. P. 213–222.
- Sanders, J. T. (1985) Experience, Memory and Intelligence // The Monist. Vol 68. №4. P. 507–521.
- Searle, J. R. (1990) Is the brain's mind a computer program? // Scientific American. №3. P. 26–31.
- Searle, J. R. (1980) Minds, Brains, and Programs // The Behavioral and Brain Sciences. №3. P. 417–457.
- Weisberg, J. (2011) Subjective Consciousness: A Self-Representational Theory, by Uriah Kriegel // Mind. №120. P. 538–542. DOI: 10.1093/mind/fzr042
- Williamson, J. (2003) Abduction, Reason, and Science: Processes of Discovery and Explanation // The British Journal for the Philosophy of Science. №54. P. 353–358.

Дата поступления: 07.10.2016 г.

“EASY” AND “HARD” PROBLEMS OF ARTIFICIAL INTELLIGENCE IN ANALYTIC PHILOSOPHY OF MIND

A. S. GAU

(ELABUGA INSTITUTE, KAZAN (VOLGA REGION) FEDERAL UNIVERSITY)

This article examines two aspects of the problem of artificial intelligence in contemporary analytic philosophy of mind. Contemporary researchers distinguish between two aspects of the mental — psychological and phenomenal. By analogy with the two aspects of the mind-body correlation, they allow us to talk of the «easy» and «hard» problems in artificial intelligence.

In its essence, the «easy» problem of artificial intelligence lies in studying the possibilities of achieving psychological mental states based on the reconstruction of the human brain. The essence of

the «hard» problem is to examine the possibilities of phenomenal mental experience within neural computational modeling.

In its turn, the latter problem appears in two aspects. The first asks whether it is possible to create an artificial mind which would both physically reconstruct the human brain and, as its effect, achieve the stage of phenomenal consciousness. The second aspect is whether we can actually talk of consciousness in an artificial system.

The article concludes that the possibility of artificial intelligence is not essential for study of the problem of consciousness. Contemporary philosophy needs to shift the focus from the issue of the possibility of consciousness in the system manipulating physical symbols in accordance with structure-compliant rules to studying artificial intelligence regardless its technical features. It will allow us to face one of the most interesting problems in contemporary analytic philosophy of mind. Providing a solution to this problem will open up new prospects in this field.

Keywords: analytic philosophy; artificial intelligence; mental states; consciousness; neural computational modeling

REFERENCES

- Chalmers, D. (2013) *Soznaiushchii um: V poiskakh fundamental'noi teorii*. Moscow, Knizhnyi dom «LIBROKOM». 509 p. (In Russ.)
- Boden, M. A. (1973) How artificial is artificial intelligence? *The British Journal for the Philosophy of Science*, vol. 24, no. 1, pp. 61–72.
- Guttenplan, S. (1988) Thinking Machines: The Evolution of Artificial Intelligence. *Journal of Design History*, no. 1, pp. 151–152. DOI: 10.1093/jdh/1.2.151
- Harmer, J. (2010) All in the Mind? *The Cambridge Quarterly*, no. 39, pp. 385–391. DOI: 10.1093/camqtly/bfq026
- Helman, D. H. (1987) Realism and Antirealism in Artificial Intelligence. *The British Journal for the Philosophy of Science*, vol. 38, no. 1, pp. 19–26.
- Hochstein, E. (2016) Categorizing the Mental. *The Philosophical Quarterly*, no. 66, pp. 745–759. DOI:10.1093/pq/pqw001
- Moor, J. H. (1978) Three Myths of Computer Science. *The British Journal for the Philosophy of Science*, vol. 29, no. 3, pp. 213–222.
- Sanders, J. T. (1985) Experience, Memory and Intelligence. *The Monist*, vol. 68, no. 4, pp. 507–521.
- Searle, J. R. (1990) Is the brain's mind a computer program? *Scientific American*, no. 3, pp. 26–31.
- Searle, J. R. (1980) Minds, Brains, and Programs. *The Behavioral and Brain Sciences*, no. 3, pp. 417–457.
- Weisberg, J. (2011) Subjective Consciousness: A Self-Representational Theory, by Uriah Kriegel. *Mind*, no. 120, pp. 538–542. DOI: 10.1093/mind/fzr042
- Williamson, J. (2003) Abduction, Reason, and Science: Processes of Discovery and Explanation. *The British Journal for the Philosophy of Science*, no. 54, pp. 353–358.

Submission date: 07.10.2016.

Гау Анастасия Сергеевна — аспирант кафедры философии и социологии, специалист научного отдела Елабужского института Казанского (Приволжского) федерального университета. Адрес: 423604, Россия, г. Елабуга, ул. Казанская, д. 89. Тел.: +7 (85557) 7-03-32, +7 (987) 228-32-65. Эл. адрес: anastasia.gau@gmail.com. Научный руководитель — канд. филос. наук, доц. Е. В. Громов.

Gau Anastasia Sergeevna, Postgraduate student, Department of philosophy and sociology; Specialist, Research department, Elabuga Institute of Kazan (Volga Region) Federal University. Postal address: 89 Kazanskaya st., Elabuga, Russian Federation 423604. Tel.: +7 (85557) 7-03-32, +7 (987) 228-32-65. E-mail: anastasia.gau@gmail.com. Research adviser: E. V. Gromov, Candidate of Philosophy, Associate Professor.