

УДК: 141: 16: 004.5
DOI: 10.15372/PS20250509
EDN: OAZLRQ

А.С. Зайкова

**О СОИЗМЕРИМОСТИ ВОПРОСОВ
«МОЖЕТ ЛИ МАШИНА МЫСЛИТЬ?»
И «МОЖЕТ ЛИ МАШИНА ПОБЕДИТЬ
В ИМИТАЦИОННОЙ ИГРЕ?»**

В статье проводится критический анализ соизмеримости двух центральных вопросов в философии искусственного интеллекта: «Может ли машина мыслить?» и «Может ли машина победить в имитационной игре?». Показано, что, несмотря на прагматическую замену, предложенную Аланом Тьюрингом в 1950 году, вопросы концептуально и онтологически различны. На основе анализа аргументов Д. Деннета, Дж. Сёрла, Д. Чалмерса и др. демонстрируется, что соизмеримость возможна лишь в ограниченном, индуктивном смысле, но не в логическом или онтологическом. Подчеркивается необходимость различения поведенческой симуляции и подлинного мышления.

Ключевые слова: соизмеримость, тест Тьюринга, машинное мышление, сознание, имитационная игра, философия ИИ, эпистемология, поведенческий критерий.

A.S. Zaykova

**ON THE COMMENSURABILITY OF THE QUESTIONS
“CAN A MACHINE THINK?”
AND “CAN A MACHINE WIN THE IMITATION GAME?”**

This paper offers a critical analysis of the commensurability between two foundational questions in the philosophy of artificial intelligence: “Can a machine think?” and “Can a machine win the imitation game?”

It is argued that, despite Alan Turing's pragmatic substitution proposed in 1950, these questions differ in ontological status and conceptual framework. Drawing on the works of D. Dennett, J. Searle, D. Chalmers, and others, the paper demonstrates that commensurability is possible only in an epistemologically limited, inductive sense—but not in a logical or ontological one. The study emphasizes the necessity of distinguishing behavioral simulation from genuine thinking, thereby avoiding the category mistake often associated with interpreting the Turing Test as proof of machine consciousness.

Keywords: commensurability, Turing Test, machine thinking, consciousness, imitation game, philosophy of AI, epistemology, behavioral criterion.

1. Введение

Понятие соизмеримости в эпистемологии и философии науки чаще всего понимается как возможность сравнения, перевода или сопоставления двух теорий, концепций или вопросов на основании общей онтологии, методологии или семантической базы. Такое понимание было выработано во многом благодаря Т. Куру и его работе «Структура научных революций» [8], где он ввёл идею несоизмеримости научных парадигм, утверждая, что смена парадигм влечёт за собой трансформацию самого языка науки, критериев значимости и даже восприятия данных. П. Фейерабенд в своих работах развил эту идею, показав, что между конкурирующими теориями часто отсутствует нейтральный метаязык, позволяющий их объективно сравнить [6].

В этом контексте соизмеримость предполагает наличие общего концептуального пространства, в котором можно оценивать истинность, адекватность или объяснительную силу альтернативных подходов без необходимости радикального перевода или пересмотра базовых предпосылок.

В общем случае, мы можем выделить четыре основные критерии соизмеримости, выведенные из классических споров о сопоставимости теорий:

- (1) они направлены на один и тот же онтологический объект или свойство (например, мышление как когнитивный феномен);
- (2) они встроены в одну и ту же концептуальную рамку, не требующую радикального переопределения ключевых понятий;

(3) они используют совместимые эпистемологические критерии (например, оба полагаются на поведенческие данные или оба требуют доступа к внутренним состояниям);

(4) их ответы логически связаны: положительный (или отрицательный) ответ на один из них влечёт за собой определённый статус другого.

Если все четыре условия выполняются, мы, как правило, можем говорить о полной соизмеримости; если только одно или несколько – о частичной или ограниченной соизмеримости. В последнем случае вопросы могут быть связаны прагматически или индуктивно. Именно такой ограниченный тип связи, как будет показано далее, существует между вопросами «*Может ли машина мыслить?*» и «*Может ли машина победить в имитационной игре?*». Настоящая статья ставит своей задачей не просто констатировать различие этих вопросов, но систематически проанализировать степень и характер их соизмеримости через призму указанных четырёх критериев.

2. «Может ли машина мыслить?»

Традиционный вопрос «*Может ли машина мыслить?*» апеллирует к концептам сознания, намерения, субъективного опыта, что, применительно к машинам (техносубъектам, интеллектуальным агентам, нейросетям, роботам, большим языковым моделям и др.), выглядит как откровенное лукавство. Даже сама постановка вопроса применительно к современным интеллектуальным агентам подвергается критике, не столько из-за технических возможностей машин, сколько из-за неопределённости самого понятия «мышление». В классической философии мышление связывается с наличием внутреннего мира, интенциональности и способности к рефлексии – качествами, которые невозможно или почти невозможно приписать вычислительным системам без радикального переосмысления этих понятий.

Кроме того, этот вопрос близок к понятию агентности. В философии действия и этике агентность связывается с автономией воли, моральной ответственностью и свободой выбора, что ставит под сомнение возможность корректно рассуждать об агентности искусственных интеллектуальных систем. В научном сообществе обсуж-

дается необходимость рассмотрения новых видов агентности (интеллектуальной, рациональной, социальной и др.) [1], однако такие обсуждения направлены на решение иных научных задач, к примеру, для описания функциональных ролей ИИ в социальных системах.

В любом случае, задавая вопрос о мышлении машин, мы сталкиваемся либо с невозможностью приписать им мышление в строгом смысле, либо с необходимостью радикального пересмотра самого понятия мышления. В этом отношении вопрос оказывается скорее постановкой проблемы пределов понятий, нежели вопросом о фактических возможностях вычислительных систем. Он выявляет напряжение между традиционной антропоцентричной концепцией разума и новыми формами когнитивной активности, возникающими в искусственных интеллектуальных системах.

Одним из сторонников второго подхода (пересмотра понятия мышления) можно назвать Д. Деннета. Он утверждает, что приписывание ума другому агенту – это стратегия интерпретации, а не метафизическое утверждение [5]. С этой точки зрения, если машина демонстрирует последовательное, адаптивное и целенаправленное поведение, рационально приписывать ей «верования» и «желания», даже если мы не знаем, есть ли у неё субъективный опыт. Однако такая позиция не устраняет онтологическую пропасть между симуляцией и реализацией мышления; она лишь предлагает прагматический способ взаимодействия с когнитивно сложными системами.

3. «Может ли машина победить в имитационной игре?»

Идея замены вопроса о мышлении на вопрос об имитации была впервые предложена А. Тьюрингом в его знаменитой статье «Вычислительные машины и разум» («Computing Machinery and Intelligence»), опубликованной в 1950 году в журнале *Mind*. Тьюринг начинает статью с признания трудности исходного вопроса: «Я предлагаю рассмотреть вопрос: „Могут ли машины мыслить?“ Следовало бы начать с определений значений слов „машина“ и „мыслить“» [12, р. 433]. Признавая, что дебаты о значении слова «мыслить» ведут в тупик, он предлагает «заменить вопрос другим, который тесно с ним связан и выражен относительно недвусмысленными словами [Там же].

Этот новый вопрос основан на модифицированной версии игры, которую Тьюринг назвал «имитационной». В оригинальной версии игры участвуют три человека: мужчина (А), женщина (В) и судья (С), который должен по переписке определить, кто из А и В – женщина. Тьюринг предлагает заменить одного из игроков машиной. Если судья не может надёжно отличить машину от человека, машина считается прошедшей тест.

Целью Тьюринга было не дать определение мышления, а предложить операционализируемый критерий, пригодный для эмпирической проверки. Он писал: «Исходный вопрос „Могут ли машины мыслить?“, по-моему, слишком лишён смысла, чтобы заслуживать обсуждения» [12, р. 442]. Вместо этого он сосредоточился на поведенческой эквивалентности: если машина ведёт себя так, как мы ожидали бы от мыслящего существа, то нет рациональных оснований отрицать её способность мыслить.

Тем не менее, Тьюринг не утверждал прямо, что победа в имитационной игре доказывает наличие сознания или понимания. Он рассматривал тест как прагматическую замену вопросу, на который мы всё равно не способны дать ответ. Однако для обоснования правомерности замены одного вопроса другим он использовал ряд аргументов, основным из которых является методологический: в отсутствие прямого доступа к внутреннему опыту (даже у других людей) мы вынуждены полагаться на внешние проявления. Таким образом, имитационная игра становится инструментом для обхода метафизических трудностей и перехода к эмпирически проверяемым гипотезам.

Особый интерес в этом контексте представляет ответно-зависимая интерпретация теста, развитая Д. Праудфут. Согласно её позиции, наиболее правильной интерпретацией теста Тьюринга соответствует следующее суждение:

«х является разумным (или мыслит), если в неограниченной компьютерной игре-имитации-человека х кажется разумным среднему опрашиваемому» / «субъекту» [10, р. 397–398].

Эта интерпретация подчёркивает социокультурный аспект атрибуции разумности и рассматривает успешное прохождение теста как результат восприятия, а не объективного факта. Если система кажется разумной среднему субъекту, она таковой признаётся, независимо от её внутренней структуры.

Подобный подход, хоть и не является широко признанным, сближает имитационную игру с более широкими дискуссиями о социальной агентности, где решающим оказывается не то, «обладает ли» субъект внутренней автономией, но то, как он воспринимается другими [2]. Сама возможность машины пройти тест фиксирует её социальное присутствие и способность занимать в коммуникации ту же позицию, что и человек. Однако это социальное признание не обязательно влечёт за собой признание когнитивных характеристик машины или субъективности её опыта.

4. Соизмеримость вопросов

Для анализа соизмеримости вопросов «*Может ли машина мыслить?*» и «*Может ли машина победить в имитационной игре?*» рассмотрим наиболее яркие позиции в философии сознания и искусственного интеллекта.

В прикладной когнитивной науке и робототехнике часто используется градуалистский подход: вместо утверждения, что что-то «мыслит» или «не мыслит», вводятся уровни когнитивной сложности. В этом контексте успешное прохождение имитационной игры может рассматриваться как маркер высокого уровня когнитивной симуляции, даже если он не гарантирует наличия сознания. Такой подход, однако, требует чёткого разграничения между симуляцией, репликацией и реализацией когнитивных процессов [9].

Ещё более радикальной является позиция Д. Чалмерса, согласно которой мы должны различать «лёгкие» и «трудные» проблемы сознания [4]. Поведенческие задачи, такие как имитационная игра, относятся к «лёгким» проблемам: они касаются функций, которые в принципе могут быть воспроизведены вычислительно. Однако «трудная» проблема – почему и как возникает субъективный опыт – остаётся за пределами поведенческого анализа. С этой точки зрения, даже идеально прошедшая тест машина может быть философским зомби – существом, идентичным человеку по поведению, но лишённым сознания. Это означает, что поведенческий критерий логически не влечёт наличия мышления, и, следовательно, вопросы остаются несоизмеримыми даже в ослабленном смысле.

Хотя аргументы Дж. Сёрла [11] и Н. Блока [3] демонстрируют принципиальную возможность поведенческой имитации без наличия понимания или сознания, некоторые исследователи настаивают на эпистемологической ценности теста Тьюринга даже при отсутствии онтологической гарантии. К примеру, Р. Френч в статье «Тест Тьюринга: первые пятьдесят лет» признаёт, что прохождение теста не доказывает наличие мышления в строгом смысле, однако подчёркивает: «Прохождение теста Тьюринга не доказало бы, что машина мыслит, но предоставило бы очень веские основания полагать, что это так» [7, р. 118]. Френч утверждает, что критики, такие как Сёрл, ошибочно требуют дедуктивной достоверности от теста, в то время как в науке преобладают индуктивные обобщения, основанные на наблюдаемых коррелятах. С его точки зрения, устойчивая способность машины вести осмысленный, контекстуально адаптивный диалог – особенно в расширенной версии теста, включающей эмоциональные, социальные и сенсомоторные аспекты – создаёт высокую степень правдоподобия наличия когнитивных способностей, сопоставимых с человеческими. Эта позиция не отождествляет имитацию с мышлением, но признаёт её как сильный эпистемологический индикатор в условиях недоступности прямого доступа к внутреннему опыту.

Теперь, когда мы упомянули наиболее яркие позиции, перейдём к оценке соизмеримости по четырём критериям, сформулированным во введении.

(1) Онтологическая направленность. Вопрос «Может ли машина мыслить?» обращён к мышлению как к внутреннему, интенциональному состоянию, связанному с сознанием и субъективным опытом. Вопрос об имитационной игре, напротив, касается исключительно поведенческой компетентности в лингвистическом взаимодействии. Хотя Тьюринг предполагает, что поведенческая неотличимость может служить основанием для приписывания мышления, сама игра не предполагает и не требует наличия внутреннего мира. Следовательно, онтологические объекты различны: первый вопрос касается сущности, второй – функции.

(2) Концептуальная рамка. Вопрос о мышлении предполагает традиционную философскую концепцию разума, восходящую к картезианской модели субъекта, где мышление неразрывно связано с рефлексией и внутренней автономией. Имитационная игра,

напротив, опирается на бихевиористскую и функционалистскую парадигму, где значение определяется через использование, а разум – через внешнее поведение. Эти концептуальные рамки конфликтуют: одна требует внутреннего наблюдателя, другая отвергает его как ненаучного.

(3) Эпистемологические критерии. Вопрос о мышлении сам по себе не предлагает операциональных средств проверки и остаётся открытым для метафизических интерпретаций. Вопрос об имитационной игре предлагает чёткий поведенческий критерий. Однако оба вопроса сталкиваются с общей эпистемологической проблемой – «проблемой других умов»: как приписать мышление тому, чей внутренний мир принципиально недоступен? В этом узком смысле критерий частично выполняется: оба вопроса разделяют одну и ту же эпистемологическую трудность и опираются на аналогичную логику индуктивного вывода от поведения к внутреннему состоянию.

(4) Логическая связь ответов. Критика Сёрла и Блока показывает, что успех в имитационной игре не влечёт логически наличия мышления. Но и обратное утверждение также не является бесспорным: для прохождения теста недостаточно, чтобы некое существо мыслило, также необходимо, чтобы оно обладало коммуникативными навыками и возможностями к поддержанию текстового диалога. Таким образом, между вопросами существует неполная асимметричная логическая связь: мышление влечёт возможность имитации (которая реализуема только при добавлении коммуникативных функций), но имитация не влечёт необходимость мышления.

Из четырёх предложенных условий соизмеримости частично выполняются два: критерий совместимости эпистемологических оснований (3) и критерий асимметричной логической связи (4). Это позволяет говорить об ограниченном, эпистемологически опосредованном характере соизмеримости исследуемых вопросов. Вопросы связаны через общую проблему атрибуции мышления и асимметричную импликацию от мышления к поведению, тем не менее, их нельзя назвать ни тождественными, ни взаимозаменяемыми. Такая частичная соизмеримость достаточна для прагматического использования теста Тьюринга как индуктивного инструмента, но недостаточна для онтологического отождествления имитации и мышления.

5. Заключение

Проведенный анализ позволяет сделать вывод о несоизмеримости двух вопросов. Конечно, между ними существует асимметричная связь, но она обусловлена не логикой, а условиями человеческого познания. Успешная имитация не требует наличия мышления, но, в условиях принципиальной недоступности внутренних состояний других агентов – включая как людей или животных, так и машин – она повышает правдоподобие наличия когнитивных способностей. Эта связь, однако, остаётся строго индуктивной и вероятностной: она опирается на аналогию и прагматическую полезность, а не на логическое следование или онтологическую необходимость.

Кроме того, хотя поведенческая имитация не является достаточным условием мышления, мышление является достаточным условием для возможности имитации при подключении коммуникативных функций. Любой мыслящий субъект, обладающий необходимыми навыками, в принципе способен участвовать в имитационной игре и проходить её. Эта асимметрия указывает на то, что второй вопрос является необходимым, но не достаточным следствием первого. Таким образом, соизмеримость ограничена направлением: только от мышления к имитации; но не от имитации к мышлению.

Следовательно, даже в ослабленной форме полная соизмеримость не достигается. Вопросы остаются разнородными по своей эпистемологической природе: один принадлежит метафизике сознания, другой – методологии поведенческой оценки. Их сопоставление возможно лишь при явном указании на принимаемые методологические и метафизические допущения – например, на готовность отождествлять функциональную эквивалентность с когнитивной тождественностью.

Признание этого факта не отменяет ценности теста Тьюринга как инструмента. Он остаётся важным эталоном для оценки лингвистической компетентности, социальной адаптивности и коммуникативной сложности ИИ-систем. Однако его онтологическая переинтерпретация, подразумевающая подмену поведения сознанием, а симуляции – мышлением, ведёт к категориальной ошибке, особенно опасной сейчас, в эпоху генеративных моделей, способ-

ных создавать иллюзию понимания без его наличия. Различение между «кажется, что мыслит» и «действительно мыслит» является необходимым условием как для теоретических, так и для практических исследований в области искусственного интеллекта.

Литература

1. *Зайкова, А. С.* Агентность систем искусственного интеллекта: интеллектуальная, социальная, моральная // *Философия науки*. 2024. № 5(104). С. 89–97. <https://doi.org/10.15372/PS20240507>
2. *Зайкова А.С.* Можно ли относиться к роботам и к иным техносубъектам как к животным: к вопросу о социальной агентности // *Семиотические исследования*. 2025. 5(3). С. 16–23.
3. *Block, N.* Psychologism and behaviorism // *The Philosophical Review*. 1981 90(1). P. 5–43.
4. *Chalmers, D. J.* (Facing up to the problem of consciousness // *Journal of Consciousness Studies*. 1995. No. 2(3). P. 200–219.
5. *Dennett, D. C.* The intentional stance. 1987. MIT Press.
6. *Feyerabend, P.* Against method: Outline of an anarchistic theory of knowledge. New Left Books. 1975
7. *French, R. M.* The Turing Test: The first fifty years // *Trends in Cognitive Sciences*. 2000. № 4(3). P. 115–122.
8. *Kuhn, T. S.* The Structure of Scientific Revolutions. University of Chicago Press. 1962.
9. *Preston, J.* The philosophy of artificial intelligence. // E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2013 ed.). URL: <https://plato.stanford.edu/archives/fall2013/entries/artificial-intelligence/>
10. *Proudfoot D.* Rethinking Turing’s Test and the Philosophical Implications // *Minds and Machines*. 2020. 30. P. 487–512.
11. *Searle, J. R.* Minds, brains, and programs // *Behavioral and Brain Sciences*. 1980. 3(3). P. 417–457.
12. *Turing, A. M.* Computing machinery and intelligence // *Mind*. 1950. 59(236). P. 433–460.

References

1. *Zaykova, A. S.* (2024). Agency of artificial intelligence systems: Intellectual, social, moral. *Philosophy of Science*, S5(104), 89–97. (In Russ.).
2. *Zaykova, A. S.* (2025). Can We Treat Robots and Other Technosubjects as Animals? On the Question of Social Agency. *Semiotic Studies*, 5(3), 16–23. (In Russ.).

3. *Block, N.* (1981). Psychologism and behaviorism. *The Philosophical Review*, 90(1), 5–43.
4. *Chalmers, D. J.* (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2(3), 200–219.
5. *Dennett, D. C.* (1987). *The intentional stance*. MIT Press.
6. *Feyerabend, P.* (1975). *Against method: Outline of an anarchistic theory of knowledge*. New Left Books.
7. *French, R. M.* (2000). The Turing Test: The first fifty years. *Trends in Cognitive Sciences*, 4(3), 115–122.
8. *Kuhn, T. S.* (1962). *The Structure of Scientific Revolutions*. University of Chicago Press.
9. *Preston, J.* (2013). The philosophy of artificial intelligence. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2013 ed.). URL: <https://plato.stanford.edu/archives/fall2013/entries/artificial-intelligence/>
10. *Proudfoot D.* (2020). Rethinking Turing's Test and the Philosophical Implications. *Minds and Machines* 30: 487–512.
11. *Searle, J. R.* (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–457.
12. *Turing, A. M.* (1950). Computing machinery and intelligence. *Mind*, 59(236), 433–460.

Информация об авторе

Зайкова Алина Сергеевна – научный сотрудник Института философии и права СО РАН (630090, Новосибирск, ул. Николаева, 8).
Zaykova.a.s@gmail.com

Information about the author

Zaykova Alina Sergeevna – researcher of Institute of Philosophy and Law, Siberian Branch of the Russian Academy of Sciences (8, Nikolaev st., Novosibirsk, 630090, Russia).
Zaykova.a.s@gmail.com

Дата поступления 02.03.2025
Принята к печати 11.12.2025