

УДК 172

DOI: 10.26456/vtphilos/2022.4.032

РЕШЕНИЕ ПРОБЛЕМЫ «СОЗНАНИЕ-ТЕЛО» И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Т.С. Георгиу

ФГБОУ ВО «Московский государственный областной педагогический университет», Московская обл., г. Мытищи

В статье исследуется дихотомия «сознание-тело» в контексте искусственного интеллекта (ИИ). Философы с древних времен пытаются решить эту серьезную проблему, но с появлением ИИ она стала еще более актуальной. В исследуемой дихотомии «сознание-тело» создание искусственного сознания становится центральной трудно разрешимой задачей для философии. Развитие ИИ продолжается быстрыми темпами, и вопрос моделирования мышления человека и построения искусственного сознания является совсем не гипотетическим. Рассматриваются главные мысленные эксперименты для решения данной проблемы, осуществляется философское осмысление и критический анализ разных подходов к исследованию означенной дихотомии. Особое внимание уделено вопросу «может ли машина мыслить и, следовательно, обладать сознанием?».

Ключевые слова: *искусственный интеллект, мысленный эксперимент, сознание, материя, разум, мышление.*

Введение

ИИ является самой обсуждаемой и противоречивой темой в современном обществе цифровизации. Возникающие проблемы в сфере ИИ затрагивают острые философские вопросы, связанные с человеческим разумом, мышлением и сознанием, которые приводят к серьезным этическим и экзистенциальным опасениям.

Проблема различия «сильного» и «слабого» искусственного интеллекта, трудно разрешимая дихотомия «сознание-тело» и проблема возможности создания «искусственного сознания» требуют особого философского осмысления. В данной статье предпринимается попытка исследовать соответствующие мысленные эксперименты и взаимоотношения между философскими теориями сознания и текущим состоянием ИИ.

Целью статьи является краткое описание основных мысленных экспериментов, касающихся вышеупомянутых проблем, и их критический анализ с точки зрения социальной философии. Следует отметить, что дихотомия «сознание-тело» существует гораздо дольше, чем компьютеры и ИИ. Философы с древних времен пытаются найти адекватный алгоритм ее решения, однако с появлением ИИ данная проблема стала еще более актуальной.

В данной работе рассматриваются разные вопросы: «Как работает разум и сознание?»; «Возможно ли, чтобы машины действовали разумно, как человек?»; «Возможно присвоить машинам ментальные свойства и даже сознание?». Данные проблемы становятся сложнее, когда развитие в сфере ИИ достигает такого уровня, что машине передается возможность принимать решения вместо человека. Еще сложнее вопрос о том, способна ли техническая система (робот, компьютер или другая автоматизированная техника), решая какие-то задачи, например спасения людей, брать на себя риски принятия решений, отвечать за эти решения перед людьми и другими техническими системами. Каким образом ИИ способен к машинному обучению в рамках этики и насколько общество готово принимать такое «машинное поведение», если допустить, что ИИ в принципе будет способен принимать разные решения как «осознающее» существо [8]? В исследуемой проблематике «сознание-тело» проблема искусственного сознания становится центральной трудной задачей для философии.

Определение терминов

Термин «материя» (от лат. *materia* – «вещество») вызывает огромный интерес среди философов и ученых с древних времен. Данное понятие играет важную роль в онтологии, натурфилософии, теории познания и в других науках. Материя в физике получает значение какого-либо объекта, который есть в Природе и в любом веществе, существование которого можно доказать экспериментальным путем. Под веществом обычно понимается любая совокупность атомов и молекул, находящихся в определенном агрегатном состоянии.

Однако в философском понимании данный термин принимает различные значения. Можно выделить следующие основные значения (понятия) в зависимости от разных философских подходов:

- 1) субстрат, «подлежащее», «то, из чего» (Аристотель) возникают и состоят вещи и Вселенная;
- 2) бесконечно делимый континуум, пространство, «то, в чем» (Платон), или протяжение (Декарт);
- 3) принцип индивидуации, т. е. условие множественности (Платон, Аристотель, Прокл, Лейбниц);
- 4) вещество, или тело, обладающее инертностью, т. е. массой, и непроницаемостью, упругостью или твердостью (античные стоики, новоевропейские материалисты) [1].

Мы же будем следовать современному определению материи как объективной реальности, данной познающему субъекту в виде ощущений, существующей вне и независимо от человеческого сознания и отражаемой им. Материей является бесконечное множество всех существующих явлений, объектов и систем, субстрат всех многообразных свойств, отношений, взаимодействий и форм движений.

Несмотря на тот факт, что на протяжении уже многих столетий исследуется проблема сознания в целом и термин «сознание» в частности, такой парадокс существует: общепринятого, понятного, логично структурированного и глубокого определения термина «сознание» до сих пор нет. В качестве всеобщего понимания и для целей данной работы используется следующее определение: сознание является одним из основных понятий философии, социологии и психологии, обозначающее состояние психической жизни индивида, выражающееся в субъективной переживаемости событий внешнего мира и жизни самого индивида и способность идеального воспроизведения действительности в мышлении. Эта высшая функция мозга, свойственная только человеку и тесно связанная с языком, заключающаяся в обобщенном и целенаправленном отражении действительности и внешнего мира, в предварительном мысленном построении действий и предвидении их результатов, в разумном регулировании и самоконтролировании поведения человека [5].

Следует отметить, что материя противопоставляется сознанию, духу, разуму, форме, идее, благу, Богу, актуальному бытию. Материя получает такое определение как противоположное сознанию, как подлинное, объективное, первичное бытие. Противостояние материализма и идеализма основывается на вышеупомянутом философском противопоставлении понятий материи и сознания [1].

История ИИ показала, что основатели и разработчики в этой сфере поставили задачу ответить на вопрос «Может ли машина мыслить и действовать разумно, как человек?», и если да, то «Возможно ли машине действовать автономно без вмешательства человека и достичь уровня наличия сознания?» [16]

Для решения этой задачи ученые сопоставляют функции мозга человека с компьютером и ИИ, как представлено в табл. 1.

Таблица 1. Сопоставление функций мозга человека с компьютером и ИИ

Функция	Человек	Компьютер и ИИ
Прием информации	Органы чувств	Устройства ввода
Хранение информации	Память	Устройства памяти
Обработка информации	Мозг	Процессор
Передача информации	Язык	Устройства вывода
Выполнение сложных задач	Нейронные сети мозга	Искусственные нейронные сети
Принятие решения	Разум, интуиция	Технология машинного обучения и др.

Конечно, такое сравнение является условным, не единственным и вызывает споры среди ученых. С одной стороны находятся те, кто утверждает, что невозможно моделировать все функции человеческого мозга, а с другой – те, кто уверен, что ИИ может даже превосходить интеллект человека.

Существуют два основных вида ИИ: слабый (англ. «narrow») и сильный (англ. «general», «strong»). Термин «сильный ИИ» был введен Джоном Сёрлом в работе «Разумы, мозги и программы» в 1980 г. Согласно Джону Сёрлу, компьютер оснащен сильным ИИ, и «это не просто инструмент для изучения разума; скорее, правильно запрограммированный компьютер на самом деле является разумом, в том смысле, что компьютерам, которым даны правильные программы, можно буквально сказать, чтобы они понимали и имели другие когнитивные функции» [17, р. 417].

Мысленные эксперименты

Мысленные эксперименты представляют собой специфический метод работы ученых и мыслителей. Они являются действенным вспомогательным ресурсом исследования и осмысления трудных философских проблем, когда традиционные методы исследования противоречат этическим нормам или ставят объект исследования в опасное положение. Проблема «сознание-тело» – одна из сложнейших в философии, и таким образом рассматриваются разные мысленные эксперименты в процессе размышлений о сознании, объективной реальности, ИИ и обществе.

«Мозг в колбе»

Мысленный эксперимент «Мозг в колбе» предложен философом Хилари Патнэмом в 1981 г. Главный тезис данного эксперимента состоит в том, что наше восприятие и воспринимаемая нами реальность является «объективной». Эксперимент ставит под вопрос истинность субъективного восприятия человека.

Патнэм предлагает осуществить следующий эксперимент: извлечь мозг человека из тела, поместить его в колбу и подключить к какому-то суперкомпьютеру так, чтобы тот может успешно симулировать нервные импульсы таким образом, как будто они приходят от органов чувств. Получается, что его человеческий опыт – это лишь результат воздействия электронных импульсов, поступающих из компьютера в нервную систему и далее в мозг. Вследствие этого возникает вопрос: может ли мозг определить, что он находится в виртуальной реальности и все его восприятия являются обманом? Данный эксперимент находится в русле рассуждений Рене Декарта, связанных с его мысленным экспериментом «Злой Демон» о радикальном сомнении в достоверности всех общепринятых знаний. Согласно мысленному эксперименту Декарта, поднимается вопрос: существует что-либо в мире, чего нельзя подвергнуть сомнениям? Декарт заявляет: что, если какой-то могущественный злой демон дурачит человека, предлагая его органам чувств иллюзии настолько достоверные, что человек ничего не может

заподозрить? Если это возможно, тогда человек не может отличить истину от лжи и он поверит явной лжи как реальности.

Согласно Патнэму, мозг в колбе никаким образом не может определить, что он не воспринимает «объективную реальность», а находится в колбе с симуляцией. Несмотря на отсутствие тела, мозг по-прежнему осознает себя существующим и постигающим окружающий мир, который считает реальным просто потому, что мышление является результатом физических процессов в мозге. Мозг продолжает мыслить по-прежнему и, следовательно, существует.

Из этого следует, что реальность каждого отдельного человека отличается от реальности другого, как в фильмах «Матрица», «Аватар» и «Шоу Трумана». Представим себе, что человек живет в таком мире и начинает понимать, что все вокруг симуляция, и весь ваш мир оказывается большим «обманом», где он (человек) исполняет главную роль, даже не подозревая об этом. Тогда какой смысл в такой жизни? Разумеется, если человек способен создать технологию, которая позволяет симулировать реальность искусственным способом таким образом, что невозможно обнаружить подмену, тогда перспективы нашей человеческой цивилизации весьма сомнительны. При таких обстоятельствах возникает фундаментальный онтологический вопрос, который может разрушить всю человеческую цивилизацию: «Как устроена реальность, каково место человека в этой реальности, и как мы вообще можем ее познавать?»

Мозг в Хьюстоне

Известный философ современности Дэниел Деннет в 1981 г. в статье «Где я?» (англ. «Where am I?») предложил очень интересный мысленный эксперимент в области философии сознания. В данном эксперименте Деннет представляет, как ученые отсоединили его (Деннета) мозг от тела и поместили в чан с питательным раствором, а затем перевезли в Хьюстон. Его тело продолжает исправно функционировать и взаимодействовать с мозгом, так как оно подключено при помощи специальных устройств и беспроводной связи к мозгу. Его мозг связан с его телом и способен контролировать свои задачи посредством сложной системы радиосвязи. По утверждению ученых, проводящих операцию, отдаление мозга от тела является всего лишь растяжением нервов, которое существенно не изменяет и не повреждает разум человека. Значит, расстояние между телом и мозгом не влияет на разум и личность человека. После успешной операции Деннет предлагает вопрос: «Где я?», если его мозг находится в совсем другом месте по отношению к его телу. И тогда вновь встает онтологический вопрос: «Где его личная идентичность?».

Далее его тело (отдаленное от мозга) участвует в миссии по обезвреживанию ядерной боеголовки. Однако происходит сбой соединения, и один за другим начинают отказывать все связи между телом и мозгом. Таким

образом вопрос «Где Деннет?» окажется очень сложным, поскольку невозможно определить, где на самом деле находится идентичность: в теле или в мозге?

Чтобы разрешить этот вопрос, философ предлагает дать своему мозгу имя Йорик, своему телу – имя Гамлет, а своей личности – имя Деннет, и далее проверяет истинность трех возможных ответов. Первый ответ гласит, что там, где Гамлет, там и Деннет. Сам философ утверждает, что этот тезис можно опровергнуть при помощи мысленных экспериментов по пересадке мозга. Например, если Том и Дик обменяются мозгами, тогда Том окажется в теле, которое раньше принадлежало Дику, и наоборот. Значит, ясно, что тело и мозг вполне могли бы расстаться друг с другом, чего нельзя сказать о личности и о своем мозге.

Второй ответ философа – там, где находятся Йорик, там и Деннет. Для размышления этого варианта философ выбрал некий юридический аргумент. Представьте себе, что Гамлет летит в Калифорнию, грабит банк, и его задерживают. Где будут судить Деннета – в Калифорнии, где произошло ограбление, или в Техасе, где находился его мозг (Йорик)? Этим противоречивым положением опровергается и второй ответ.

Третий ответ гласит, что Деннет там, где он сам считает для себя уместным, и этот ответ удовлетворяет философа больше всего. В любой данный момент человек имеет некую точку зрения, и местоположение точки зрения (внутренне определенное ее содержанием) и есть местоположение самого человека.

Суть этого эксперимента заключается в попытке доказать, что не стоит усложнять сентиментальными размышлениями тот факт, что сознание человека – продукт деятельности его мозга. Значит, можно создать человекоподобные машины, имеющие сознание как продукт программной эволюции машины. Деннет утверждает, что эволюция «слепа», как «слепы» и вычислительные машины. Таким образом, в ходе адаптации и эволюции машины приобретут когнитивные способности, разум и сознание, как все это появилось у человека в ходе собственной биологической эволюций.

«Китайская комната»

Самая известная критика теста Тьюринга – это мысленный эксперимент «Китайская комната», предложенный Джоном Сёрлом в 1980 г. Суть этого аргумента заключается в том, что понимание компьютера отличается от человеческого и что именно сознание, а не интеллект является существенным отличием высших млекопитающих, включая человека разумного. Машина не способна обладать сознанием, и моделирование человеческого сознания является нерешенной научной задачей. ИИ не может заменить человека, это невозможно [3; 11; 12, p. 139].

Сёрл предлагает представить такую ситуацию, что человека, не знающего китайский язык, т. е. самого себя, запирают в комнате, в которой находится книга с инструкциями. Человек не знает иероглифов китайского

языка, но у него есть инструкция на английском языке, которая говорит, какой кубик с китайскими иероглифами надо брать и какой кубик надо выдавать в ответ на определенные запросы, приходящие извне. Человек может использовать инструкцию, чтобы перевести фразы и отправить ответ на китайском языке, хотя сам он ни слова не говорит на китайском и вообще не знает китайский язык. Но следуя инструкции книги, выполняет все задачи правильно. Таким образом, он может убедить находящихся снаружи, что в совершенстве владеет китайским, просто из-за того, что правильно отвечает на заданные вопросы и задачи. Этот эксперимент был предложен с целью опровергнуть предположение, что компьютеры или другие виды ИИ могут мыслить и понимать так же, это делает как человек. То существо, которое понимает и сознательно выбирает свои действия – это человек и, следовательно, невозможно присвоить материальному объекту ментальные свойства мозга, в том числе и сознание [17]. Цифровые существа, компьютеры и агенты ИИ не понимают информацию, которая им дается, но в них может быть внедрена программа, которая прекрасно обрабатывает огромный массив данных, в результате быстрого вычисления выполняет заданные человеком задачи и таким образом создает видимость человеческого интеллекта.

«Комната Мэри»

Мысленный эксперимент «Комната Мэри» предложен философом Фрэнком Джексоном и был опубликован в статье «*Epiphenomenal Qualia*» журнала «*The Philosophical Quarterly*». Суть эксперимента состоит в том, что существует некий субъективный опыт, который называется «квалиа», и этот опыт не сводится к физическим процессам. В этой статье Джексон сообщает, что существует выдающийся нейрофизиолог Мэри, которая в течение всей своей карьеры исследует процесс восприятия цвета человеком. Она все знает про этот вопрос, она исследовала огромное количество научных работ, она писала много об этой теме и даже получила гранты на эти исследования. Но по какой-то причине она провела всю свою жизнь в одной комнате, где есть вещи только черно-белого цвета. Тогда возникает главный философский вопрос эксперимента, который сводится к тому, что произойдет, когда Мэри выйдет из своей комнаты и увидит реальные цвета: «Узнает ли она что-нибудь новое?» [13].

Аргумент заключается в том, что если Мэри узнает что-либо новое из опыта непосредственного восприятия цвета, то физикализм ложен. В результате данного эксперимента Джексон утверждает, что субъективный опыт не сводится к физическим процессам [14]. Подкреплением данного тезиса является закон достаточного основания Лейбница. Согласно закону достаточного основания ни одно явление не является истинным или действительным: «...ни одно явление не может оказаться истинным или действительным, ни одно утверждение справедливым – без достаточного основания, почему именно дело обстоит так, а не иначе, хотя эти основания в большинстве случаев вовсе не могут быть нам известны» [4, с. 455].

В итоге рассмотрения данного тезиса философ Джексон предположил, что кроме физических закономерностей и явлений существуют еще и нефизические. То есть ощущение красного цвета – это не только физический факт. В этом контексте представляется малоубедительным аргумент о создании ИИ подобного человеческого, судя по тому, что помимо объективного опыта существует и субъективный опыт, который полностью отсутствует у физических объектов. Главный недостаток машин – отсутствие «квалия».

«Каково это быть летучей мышью»

Этот мысленный эксперимент предложил в 1974 г. американский философ Томас Нагель. Суть его состоит в том, что самосознание является основным фактором, который делает проблему «сознание-тело» практически неразрешимой. Нагель выбрал летучую мышь, т. к. это достаточно близкое к человеку животное и в то же время крайне отличается от него, потому что видит мир вокруг себя с помощью эхолокации. Нагель спрашивает: «Что мы поймем и что означает быть летучей мышью?» То есть, как мир выглядит с точки зрения летучей мыши? Иными словами, каков ее внутренний мир и ее ментальный опыт? Вопрос риторический, потому что мы знаем, что не можем быть летучей мышью и, следовательно, не можем узнать, как это животное воспринимает окружающий мир при помощи эхолокации.

Нагель утверждает, что у каждого из нас есть уникальный опыт, и поэтому человеку невозможно узнать, как другой воспринимает мир [15]. Мужчины не могут узнать, что переживают женщины, просто потому, что никогда не могут пережить существование другого. Человеку в принципе недоступно знание того, что переживают другие, просто потому, что он никогда не может пережить опыт другого. И возникает главный онтологический вопрос: «Что значит быть любым из нас?»

Субъективный опыт, лежащий в основе сознания, не может рассматриваться объективно. Эксперимент затрагивает проблему соотношения объективного и субъективного – одну из фундаментальных проблем философии. Философ стремится опровергнуть тезис редуктивного физикализма о сознании. Согласно редуктивному физикализму, мысль является результатом физических процессов, которые происходят в клетках головного мозга, и их можно определить только изучая сам мозг человека. Каждому психическому состоянию, в том числе и сознанию, должно соответствовать определенное физическое состояние. Таким образом, сознание – продукт физических процессов.

Философ считает, что существуют вещи, недоступные человеческому познанию, и утверждает, что человек может верить в факты, которые своим умом ему невозможно постигнуть. Множество вещей существуют и будут существовать независимо от человека и его понимания, а сам человек может их понимать по-другому или вообще не воспринимать правильно. Приводим в пример ситуацию с бесконечными числами или с пределами

Вселенной. Человеку приходится признавать существование фактов, которые он не способен выразить или осознать. И в этом заключается главный тезис философа: каждое существо имеет свой собственный опыт, который означаем как субъективность опыта. Конечно, это не исключает вероятность того, что когда-нибудь человек сможет переступить через субъективность собственного опыта и выйти за пределы собственного сознания, и тогда люди смогут истинно ощутить себя другим существом, к примеру той же летучей мышью [9].

Если бы было возможно создать технологию, которая позволила бы нам трансформироваться в летучую мышь и обратно, опыт летучей мыши остался бы для нас недоступен, поскольку, превратившись в нее, мы смогли бы осознавать мир как летучая мышь, а превратившись обратно, не могли бы воспроизвести пережитый опыт, т. к. человеческое восприятие отличается от восприятия этого животного, и человек никак не может осознать тот опыт только потому, что человеческий мозг осознает как человек, а не как летучая мышь. Наш опыт ограничен тем материалом, который предоставлен нашему воображению. Следовательно, сознание невозможно описать как сумму материальных процессов в каком-то процессоре из-за ограничений возможности посчитать и с точностью воспроизвести субъективный опыт.

«Ф-зомби (p-zombie)»

Дэвид Чалмерс – автор мыслительного эксперимента «Философский зомби». Цель данного эксперимента – показать, что физическое и функциональное объяснения сознания ложны.

Согласно Чалмерсу, философским зомби является такое существо, которое обладает всеми функциональными свойствами человеческого организма, но не имеет субъективного опыта. Оно молекула в молекулу повторяет человека. Иными словами, внешне его никаким образом невозможно отличить от обычного человека. Есть все психические функции, как у прототипа – говорит, мыслит, но нет субъективной реальности, нет «квалиа» или феноменальных свойств. Например, знает, что такое боль, но ее не чувствует. Реагирует по-человечески на внешние стимулы, свет или звук, но лишено ощущений во внутреннем субъективном смысле. Зомби-двойник выглядит и ведет себя как человек, но у зомби-двойника отсутствует сознательный опыт. На самом деле двойник реагирует на все стимулы, но не имеет никаких переживаний относительно того, что такое красный цвет, закат солнца, добро и зло, несмотря на то что это существо может вести себя как человек и обладает человекоподобными реакциями. Этот двойник может получить всю необходимую информацию, но при этом не страдает и не обладает соответствующими чувствами. Суть эксперимента Чалмерса заключается в том, чтобы доказать, что субъективные аспекты сознания несводимы и нередуцируемы к функциональным фактам [11].

Придерживаясь данного положения, Чалмерс предлагает вообразить свой собственный дубликат «зомби – искусственную личность», организо-

ванную точно так же, как организован реальный Чалмерс, полагая, что робот функционально неотличим от реального Чалмерса. Между реальным Чалмерсом и его искусственным двойником разница состоит в том, что у реального Чалмерса – нейроны, а у «двойника» – чипы. Исходя из вышесказанного тезиса, Чалмерс утверждает, что «Чалмерс-зомби» не может пережить то, что переживает реальный Чалмерс и, следовательно, не обладает сознанием, поскольку у него внутри – пустота, т. к. ни в кремнии, ни в биохимии нет того, из чего рождается сознание.

В современном мире, где технологии и интеллектуальные устройства играют важную роль в повседневной жизни человека, «философский зомби» находится в фокусе внимания исследователей ИИ и становится объектом острой критики, и поддержки таких тематик, как искусственная жизнь, искусственная личность, искусственное общество и искусственное сознание.

Выводы

Возможен ли ИИ, т. е. «Возможно ли материальному объекту присвоить ментальные функции и процессы?» Ответить на данный вопрос весьма сложно, и проблема остается, увы, нерешенной. Это доказывает тот факт, что у всех вышеупомянутых экспериментов существуют сильные стороны, но одновременно есть и слабые стороны, которые не позволяют принимать одну или другую позицию как абсолютную истину. Следует отметить, что для диалектического осмысления необходимо принимать во внимание два основных противоположных направления. Таким образом, у философов появится возможность определиться с подходящими решениями по вопросам, связанным с ИИ.

Первый подход утверждает, что человеческая природа уникальна. Сознание присуще только человеку и не может быть присвоено никакому материальному объекту. Роджер Пенроуз в книге «Новый ум короля» аргументирует невозможность получения процесса мышления на основе формальных систем, человеческое сознание для него не является алгоритмическим [8]. Противоположную позицию занимают ученые, которые утверждают, что ИИ может развиваться в ходе эволюции, как развивался естественный человеческий интеллект, независимо от носителя, при этом неважно, является ли носитель материальным существом. Иными словами, ИИ возможен, и искусственное сознание можно конструировать, если человеку удастся исследовать и моделировать человеческий мозг. Кроме того, если принимать ту идею, которая гласит, что сознание является только биологическим казуальным процессом, и принципом функционирования искусственного интеллекта является причинность, тогда все действия такого осознанного интеллектуального существа будут результатами детерминированных функций. Этот конфликт естественной причинности с трансцендентальной свободой воли человека порождает серьезные философские проблемы в области онтологии и аксиологии.

Несмотря на эти нерешенные проблемы, современный ИИ создает видимость человеческого интеллекта, машина самостоятельно принимает решения без вмешательства человека, и имитация человеческого интеллекта делает его со временем неотличимым от естественного, а может быть, даже лучше, как утверждает Н. Бостром о сверхинтеллекте.

Важным моментом для любого осмысления в сфере проблемы «сознание-тело» является классификация ИИ по степени его интеграции. По этому вопросу существует несколько точек зрения, но нам необходимо сосредоточиться на уровне интеграции по степени выполнения интеллектуальных задач. Придерживаясь данного положения, можно выделить 5 основных видов ИИ:

- 1 Реагирующий – анализирует какую-то информацию, которую вводит человек, и делает вывод для выполнения определенной задачи; на этом цикл заканчивается.
- 2 Настраиваемый – в целом повторяет предыдущий, но его можно настраивать.
- 3 Автономный – машины, которые, получив однажды информацию и приняв на ее основании решение, помнят предыдущую информацию и предыдущее решение, которое они принимали и, следовательно, до некоторой степени способны к самообучению и могут действовать автономно.
- 4 Самосознающий – это интеллект, который осознает себя, у которого есть самосознание, который знает, кто он такой, у которого есть какие-то проблемы, требования, претензии, вообще желание заявить о себе миру.
- 5 Сверхинтеллект – это тот, у которого интеллект превосходит когнитивные и познавательные способности человека.

На сегодняшний день наука дошла до третьего уровня интеграции ИИ, остальные же пока являются элементами фантастики. Но мы осознаем, что до появления ЭВМ даже первые три категории были полем фантастики, а сегодня они – реальность. А это значит, что человек не застрахован ни от какого варианта развития технологии.

Впрочем, агенты ИИ даже без наличия сознания являются мощным инструментом с огромными возможностями, и если не будут предусмотрены правила безопасности и этические нормы, регулирующие их использование и существование, последствия будут не менее катастрофическими, чем при взрыве атомной бомбы или «успешном» клонировании человека.

Конечно, прогнозировать дорожную карту развития ИИ на среднесрочную перспективу невозможно. Таким образом, развитие технологий ИИ с целью ускорения технологического прогресса – лишь верхушка айсберга, поскольку сегодня мы не можем определить, что скрывается под водой, а именно – какие последствия можно ожидать для общества и всего человечества.

Сознание состоит из разных ментальных функций, и ученые до сих пор не знают наверняка, как оно функционирует. Вместе с тем, как доказали некоторые ученые, субъективный опыт не сводится к физическим процессам. В своем мысленном эксперименте «Китайская комната» Джон Сёрль показывает, что даже прохождение теста Тьюринга может не являться достаточным критерием наличия у машины подлинного процесса мышления, но факт, что машины в некоторых ситуациях действуют как моральные агенты, практически доказан.

Опасения самого различного характера высказываются нередко. К примеру, Г. Маклюэн развитие электронных средств интерпретирует как финальную «ампутацию» человеческого сознания [6].

Академик Н.Н. Моисеев в его книге «Судьба цивилизации. Путь Разума» предупреждает о серьезных рисках [7]. Согласно Моисееву, весьма вероятно такое развитие событий, связанных с ИИ: захват в личных интересах некоторой группой людей грандиозной информационной системы, функционирующей сегодня на планете. Это может привести к глобальному зомбированию человечества и к абсолютному информационному тоталитаризму.

Заключение

Подводя итог, можно отметить, что, пытаясь ответить на вопрос «Может ли машина мыслить и иметь сознание?», и несмотря на тот факт, что проблема является нерешенной, мы приходим к следующему выводу: если сознание окажется результатом химико-биологических процессов и, следовательно, наши действия определяет принцип казуальности, тогда ставится под вопрос наша идентичность как человека, а свобода воли человека вообще не может существовать. Поэтому автор принимает ту позицию, в соответствии с которой человеческая природа уникальна, а те функции, которые делают человека человеком, никоим образом невозможно воспроизвести в полной мере.

Однако необходимо принимать во внимание тот факт, что в современном обществе агентам ИИ уже позволено принимать решения вместо человека, а целью разработчиков, безусловно, является имитация всех функций человека, в том числе сознания. Машины, способные анализировать сложные данные для освоения и совершенствования конкретных навыков и имитировать сложные интеллектуальные функции человека, сегодня уже не редкость. Образы технологий ИИ, которые мы видели на экранах кинотеатров и телевизоров, вероятно, уже в ближайшем будущем не будут являться элементами фантастики, а станут нашей новой реальностью. Спрогнозировать дорожную карту развития ИИ на среднесрочную перспективу невозможно. Таким образом, развитие технологий ИИ в социальном контексте – лишь верхушка айсберга, но что же скрывается под водой, какие последствия можно ожидать для общества и, следовательно, для человечества, пока сложно сказать наверняка [2].

Если в будущем машины будут обладать (хотя бы в какой-то степени) сознанием, тогда определить, что именно делает человека человеком, а машину машиной, станет практически невозможно. На этом фоне с необходимостью возрождаются серьезные онтологические и аксиологические проблемы. Всем известно, что технологии могут быть использованы как во благо, так и во вред, и нет сомнений в том, что человечество столкнется с новыми вызовами, с фундаментальными вопросами о собственной идентичности и даже с вопросом о самом его существовании. Поэтому, если ИИ необходим современному обществу, то важно серьезно отнестись к главным этическим вопросам, связанным с ИИ, к вопросам о социальной природе нравственной деятельности, моральных отношений и морального сознания, отражающихся в категориях этики основного субъекта – человека, ведь искусственный интеллект ни в какой этике не нуждается. Этика нужна человечеству.

Список литературы

1. Бородай Т.Ю. МАТЕРИЯ [Электронный ресурс]. URL: <https://iphlib.ru/library/collection/newphilenc/document/HASH01a2e36ebf712afcc7be65a2> (дата обращения: 23.07.2022).
2. Георгиу Т.С. Философия автоматизации и искусственного интеллекта: от мифологического Талоса до будущих киборгов // Вестник Московского государственного областного университета. Серия: Философские науки. 2022. № 1. С. 68–75.
3. Душкин Р.В. Критика «Китайской комнаты» Дж. Сёрла с позиции гибридной модели построения искусственных когнитивных агентов // Сибирский философский журнал. 2020. № 2 (18). С. 30–47.
4. Лейбниц Г.В.Ф. Монадология // Антология мировой философии: в 4 т. // М.: Мысль, 1970. Т. 2. Европейская философия от эпохи Возрождения по эпоху Просвещения. С. 450–464.
5. Лекторский В.А. СОЗНАНИЕ [Электронный ресурс]. URL: <https://iphlib.ru/library/collection/newphilenc/document/HASH01089449209fc347f3553108?p.s=TextQuery> (дата обращения: 15.10.2022).
6. Маклюэн М. Понимание Медиа: Внешние расширения человека / пер. с англ. В. Николаева. М.: Жуковский: «КАНОН-пресс-Ц», «Кучково поле», 2003. 464 с.
7. Моисеев Н.Н. Судьба цивилизации. Путь Разума. М.: Языки русской культуры, 2000. 223 с.
8. Пенроуз Р. Новый ум короля: О компьютерах, мышлении и законах физики. / пер. с англ. В.О. Малышенко, 7-е изд. М.: Синергетика: от прошлого к будущему, 2019. 413 с.
9. Разин А.В. Этика искусственного интеллекта // Философия и общество. 2019. № 90 (1). С. 57–73.
10. Шерстобитова П.Д. Тело – разум (Т. Нагель. «Что значит быть летучей мышью?») // Молодой ученый. 2016. № 120 (16). С. 520–521.
11. Chalmers D. Conscious Mind: In Search of a Fundamental Theory. New York: Oxford University Press, 1996. 414 p.

12. Graziano M.S.A. Rethinking consciousness: a scientific theory of subjective experience. New York: W.W. Norton & Company, 2019. 224 p.
13. Jackson F. Epiphenomenal Qualia // The Philosophical Quarterly. 1982. № 127 (32). P. 127–136.
14. Jackson F. What Mary Didn't Know // The Journal of Philosophy. 1986. № 5 (83). P. 291–295.
15. Nagel T. Philosophical Review What Is It Like to Be a Bat? // The Philosophical Review. 1974. № 4 (83). P. 435–450.
16. Russell S., Norvig P. Artificial Intelligence. A Modern Approach. Harlow: Pearson Education Limited, 2021. 932 p.
17. Searle J.R. Minds, brains, and programs // Behavioral and brain sciences. 1980. № 3 (3). P. 417–457.

SOLUTION OF THE PROBLEM «MIND-BODY» AND ARTIFICIAL INTELLIGENCE

T.S. Georgiou

Moscow Region State Pedagogical University, Moscow region, Mytishchi

The article is devoted to the problem of the «mind-body» in the context of artificial intelligence (AI). Philosophers have been trying to solve this hard problem since ancient times, but the advent of AI, this problem has become notably very important. In the researched problematic «mind-body», the creation of artificial consciousness becomes the central difficult task for philosophy. The development of AI continues at a rapid pace, and the question of modeling human thinking and building artificial consciousness is not at all a hypothetical issue. For solving this problem are examined and studied the main mental experiments related to that. Philosophical comprehension and critical analysis of different approaches to the study of the «mind-body» problem is carried out. Particular attention is paid to the question «Can a machine think and therefore have consciousness?».

Keywords: *artificial intelligence, mental experiment, consciousness, matter, mind, thinking.*

Об авторе:

ГЕОРГИУ Такис Софокли – аспирант кафедры философии, ФГБОУ ВО «Московский государственный областной университет», Московская обл., г. Мытищи. E-mail: takis.georgiou.s@gmail.com

Author information:

GEORGIU Takis Sofokli – PhD student of Philosophy Dept., Moscow Region State Pedagogical University, Moscow region, Mytishchi. E-mail: takis.georgiou.s@gmail.com