



Международный журнал информационных технологий и
энергоэффективности

Сайт журнала:

<http://www.openaccessscience.ru/index.php/ijcse/>



УДК 004.8

ФИЛОСОФИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Смаковский В.Н., научный руководитель: Малышева А.В.

*ФГБОУ ВО "КУЗБАССКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
ИМЕНИ Т.Ф. ГОРБАЧЕВА" (ФИЛИАЛ Г. ПРОКОПЬЕВСКЕ), Прокопьевск, Россия (653033,
42 42, Кемеровская область, город Прокопьевск, Ноградская ул., д. 19А), e-mail:
smakovskij.viktor@yandex.ru*

В статье рассматриваются ключевые проблемы философии искусственного интеллекта: определение границы между естественным и искусственным интеллектом, соотношение «слабого» и «сильного» искусственного интеллекта, перспективы развития искусственного интеллекта.

Ключевые слова: Искусственный интеллект, философия искусственного интеллекта; «сильный» искусственный интеллект; «слабый» искусственный интеллект.

THE PHILOSOPHY OF ARTIFICIAL INTELLIGENCE

Smakovsky V.N., scientific supervisor: Malysheva A.V.

*KUZBASS STATE TECHNICAL UNIVERSITY NAMED AFTER T.F. GORBACHEV
(PROKOPYEVSK BRANCH), Prokopyevsk, Russia (653033, 42 42, Kemerovo region, Prokopyevsk,
Nogradskaya str., 19A), e-mail: smakovskij.viktor@yandex.ru*

The article discusses the key problems of the philosophy of artificial intelligence: defining the boundary between natural and artificial intelligence, the ratio of "weak" and "strong" artificial intelligence, prospects for the development of artificial intelligence.

Keywords: Artificial intelligence, philosophy of artificial intelligence; "strong" artificial intelligence; "weak" artificial intelligence.

Философия достаточно активно изучает актуальные вопросы, в частности, проблемы искусственного интеллекта. Можно ли сделать копию сознание человека, и вообще, существует ли сознание? Что есть настоящий интеллект?

Аристотель считал, что разум - это свойство исключительно человеческое. Никто не сомневался до середины 20 века. В это время стали появляться первые ЭВМ (Электронно-вычислительная машина), или, как мы их сейчас называем - компьютеры.

В 1956 году, на Дартмутском семинаре, впервые вводится такой термин как "Искусственный интеллект", или сокращенно ИИ. Американский ученый, давший название этому термину, дал ему такое определение - процесс, в рамках которого машины выполняют те задачи, которые потребовались интеллектуальных усилий от человека - это искусственный интеллект. [1]

В 1980 году, Джон Серл - американский философ, в одной из своих работ вводит два термина "Сильный ИИ" и "Слабый ИИ".

“Слабый ИИ” - это машина, которая упрощает действия человека и является полезным инструментом. является просто полезным инструментом. В различных сферах деятельности возможно применение слабого искусственного интеллекта. Голосовые помощники, которые помогают пользователям найти необходимую информацию или участвуют в разговорах в качестве собеседника, являются одним из примеров. Мощные базы данных, которые обрабатывают информацию о тысячах симптомов и патологий, предоставляют удобный инструмент для быстрой и точной диагностики в медицине. В сельском хозяйстве применяется искусственный интеллект, который определяет комфортную среду обитания для растений или животных и изменяет ее в соответствии с временем года или периодом жизни. Бытовые устройства, такие как умные чайники, мультиварки и дома, также представляют собой примеры использования искусственного интеллекта на повседневном уровне.

“Сильный ИИ” - должен не просто решить интеллектуальные задачи, но и обладать сознанием в человеческом понимании. Вопрос о сознании человека всегда волновал философию.

Платон, под понятием “душа” - подразумевал именно сознание в метафизическом смысле. Христианские богословы утверждали, что душа - это неповторимое творение бога. Декарт вводит понятие “cogito” - что есть метафизическое сознание в совокупности с физическим телом.

Немецкий философ, логик, математик, Г. Лейбниц описал в своих трудах эксперимент, в котором просит попытаться представить душу человека, в виде механического объекта - мельницы. Если человек зайдет внутрь этой мельницы и начнет перечислять все детали, из которых состоит этот объект, то душу там не найти. В выводе, по мнению Лейбница, доказывает, что искусственно душу не создать.

В 1950 году знаменитый ученый Алан Тьюринг опубликовал в журнале статью, в которой был рассмотрен вопрос, актуальный и сегодня. Вопрос звучал так; “Может ли машина совершать действия неотличимые от осознанных?”. Так был создан “Тест Тьюринга”, в котором участник взаимодействует с машиной и другим человеком, используя только текстовые сообщения, не имея возможности видеть их. Основная цель машины заключается в том, чтобы ввести участника в заблуждение.[2]

Вернемся в 1980 году, уже к вышесказанному американскому философу Джону Серлу. В одной из своих работ он провел мысленный эксперимент, который назывался “Китайская комната”. Суть эксперимента заключалась в представлении коробки, в которой помещен человек не знающего китайского языка. К коробке подходит человек, представитель китайского народа и бросает иероглиф, в надежде получить ответ на свой запрос. Человек в коробке открывает книгу “Переводчик”, или некая книга с иероглифами - которая бы позволила бы к данному иероглифу найти ответ. Таким образом, следуя инструкции, человек в коробке может дать ответ на запрос, но не поняв смысл вопроса, а просто “технически” отвечая. Из этого эксперимента Серл сделал вывод, что ИИ - не может осознавать своих действий ровно также, как не может их осознавать человек в коробке.

Американский философ Дэниэл Деннет считает, что создать ИИ возможно, а человек сильно идеализирует “сознание”. По мнению Дэннета, сознание - это продукт человеческой эволюции, которая для нас, людей, является легкой формой описание окружающего мира. Он основывает эту позицию на своей теории “Интенциональной установки”. Интенциональность - это свойство сознания быть направленной на что - либо. Приведен пример на шахматах где

у человека стоит задача - выиграть. Таким образом, как описывает Дэннет, когда наступает наш ход, мы пытаемся просчитать следующие ходы противника. Чтобы это сделать с случаем компьютера, у нас есть несколько вариантов. Вычислить ход на уровне устройства - практически невозможно. В условиях реального времени, мы приписываем устройству интенциональность, а именно желание опасения и надежды, как мы обычно делаем в жизни. То есть мы, как бы одушевляем компьютер во время игры, у которого есть интеллект, присущее по мнению Платона, как сказано выше, человеку.[3]

Размышляя о проблемах, которые возникают на пути развития искусственного интеллекта, особое внимание следует уделить этическим аспектам. Одной из важнейших этических дилемм является вопрос о наличии свободной воли у искусственного интеллекта. Может ли ИИ быть способным соблюдать этические нормы наравне с человеком и понимать, что убийство ребенка или жестокость по отношению к животным являются безнравственными поступками? Возможно, в базу данных ИИ будут внесены законы страны, но что, если закон о не причинении вреда животным отсутствует? Сможет ли сильный искусственный интеллект самостоятельно принять решение о том, как поступить в такой ситуации?

Давайте представим ситуацию военного времени. [4] История знает много примеров солдат, которые, несмотря на то, что они были врагами, спасали жизни друг другу, оказавшись в беде. Но сможет ли сильный искусственный интеллект, который запрограммирован на устранение противника, пройти через систему и поступить подобным образом? Ведь его главная задача - уничтожение противника, и вопрос о спасении врага выходит за рамки его программы. Но может быть, в будущем ИИ сможет принимать такие решения, основываясь на этических принципах и моральных ценностях, подобно человеку? Это сложный вопрос, который требует дальнейших исследований и размышлений со стороны ученых и философов.

Вопрос о квалиа является одной из проблем построения искусственного интеллекта. Термин "квалиа" относится к субъективной реальности интеллекта и описывает его способность ощущать и переживать эмоции. Однако, разработка ИИ, способного настолько точно имитировать человеческие эмоции, вызывает определенные проблемы.

Одной из этих проблем является эффект "зловещей долины", который был предложен японским инженером М. Мори. Он заключается в том, что люди испытывают неприязнь к объектам или существам, обладающим человеческими чертами, но не идентичным им. Таким образом, ИИ, который слишком сильно похож на человека, может вызвать отторжение у людей.

Однако, несмотря на эти проблемы, разработка человекоподобных роботов и ИИ имеет огромный потенциал во многих сферах жизни. Они могут использоваться в медицине для помощи пациентам, в образовательных целях и даже в развлекательной индустрии. Важно найти баланс между созданием ИИ, обладающего достаточной степенью человечности, чтобы вызвать симпатию, но не настолько, чтобы вызвать неприязнь.

На пути создания сильного ИИ стоит одна из самых сложных задач - эмуляция квалиа, самого "человеческого" аспекта сознания. Роботы могут быть обучены реагировать советующим образом на выражение боли, но они не способны испытывать это болезненное ощущение на себе. В то время как человек может не только показать свою боль, но и почувствовать ее.

Слабый ИИ уже находит все новые сферы применения, что говорит о перспективах развития и дальнейшего продвижения в создании ИИ. Однако, вопрос о создании сильного ИИ остается под вопросом. Его достижение требует преодоления сложностей, связанных с эмуляцией квалиа и полным воспроизведением "человеческого" сознания.

Одной из актуальных проблем философии является обсуждение возможностей искусственного интеллекта. Воссоздание человеческого способа быть в единстве объективных и субъективных сторон ощущения, восприятия и представления - это не только эффективное запоминание любого теоретического материала и технологии обработки количественной информации. На данный момент между сильным и слабым ИИ существует огромная пропасть, но именно эту проблему нужно активно исследовать и разрешать. Новые технологии и исследования в области искусственного интеллекта могут привести к революционным изменениям в нашем понимании машины и человека. Необходимо учитывать как позитивные, так и негативные аспекты развития ИИ. Каким будет будущее человечества с развитием искусственного интеллекта? Какие проблемы возникнут? Возможно, с развитием ИИ мы сможем достичь новых уровней понимания и сотрудничества между машиной и человеком, но при этом необходимо помнить о сохранении человеческих ценностей и этических принципов.

Список литературы

1. Искусственный интеллект и кризис теорий сознания // Habr.com: сайт. –URL: <https://habr.com/ru/post/500732/> (дата обращения: 11.10.2023)
2. Философия искусственного интеллекта // Studopedia.ru: сайт. –URL: https://studopedia.ru/19_243552_VIIkriterii-otsenki-rabot.html (дата обращения: 11.10.2023)
3. Что такое философия искусственного интеллекта и зачем она нужна // Skillbox.ru: сайт. –URL: <https://skillbox.ru/media/code/filosofiya-iskusstvennogo-intellekta-chto-eto-takoe-i-zachem-ona-nuzhna/> (дата обращения: 11.10.2023)
4. Шляпников В.В. Киберфилософия как философия цифрового общества // Вестник Омского государственного педагогического университета. Гуманитарные исследования – 2020 - №4 – 3с.//Cyberleninka.ru: сайт. –URL: <https://cyberleninka.ru/article/n/kiberfilosofiya-kak-filosofiya-tsifrovogo-obschestva/viewer> (дата обращения: 11.10.2023)

References

1. Artificial intelligence and the crisis of theories of consciousness // Habr.com : website. –URL: <https://habr.com/ru/post/500732/> (date of request: 11.10.2023)
2. The philosophy of artificial intelligence // Studopedia.ru : website. –URL: https://studopedia.ru/19_243552_VIIkriterii-otsenki-rabot.html (date of application: 11.10.2023)
3. What is the philosophy of artificial intelligence and why is it needed // Skillbox.ru : website. – URL: <https://skillbox.ru/media/code/filosofiya-iskusstvennogo-intellekta-chto-eto-takoe-i-zachem-ona-nuzhna/> (date of request: 11.10.2023)
4. Shlyapnikov V.V. Cyberphilosophy as a philosophy of digital society // Bulletin of Omsk State Pedagogical University. Humanitarian studies – 2020 - No.4 – 3 p. // Cyberleninka.ru : website.

–URL: <https://cyberleninka.ru/article/n/kiberfilosofiya-kak-filosofiya-tsifrovogo-obschestva/viewer> (date of application: 11.10.2023)
