

Самосознание и искусственный интеллект: взгляд современной философии

Научный руководитель – Новиков Владимир Иванович

Солдатенкова Кристина Владимировна

Аспирант

Московский государственный университет имени М.В.Ломоносова, Факультет журналистики, Кафедра литературно-художественной критики и публицистики, Москва, Россия

E-mail: vasina@live.ru

Научное направление по созданию и изучению искусственного интеллекта (ИИ) начало формироваться в 50-ых гг. XX века. Под искусственным интеллектом понимается способность цифрового компьютера или управляемого компьютером робота к характерным для человека интеллектуальным процессам (способности рассуждать, обобщать, учиться на опыте и пр.). В 1980-м году американский философ Дж. Серл ввел понятия слабого и сильного ИИ, которые до сих пор активно используются и являются предметом дискуссий в научном сообществе. Ключевое отличие сильного ИИ заключается в его способности к самосознанию, тогда как слабый способен лишь помогать человеку в решении его повседневных задач. Именно с проблемой создания сильного искусственного интеллекта тесно связана современная философия, поскольку только она может ответить на вопросы о природе самосознания, квалиа, интенциональности.

Примечательно, что сам Серл скептически относится к возможности создания сильного ИИ. Свои сомнения он подтвердил в 1980-м году посредством проведения мысленного эксперимента под названием «Китайская комната». Его скептицизм поддерживают философы Д. Деннет и Дж. Лукас.

Основной аргумент данного круга ученых заключается в том, что искусственный интеллект является синтаксической системой, а человек благодаря сознанию - семантической. Также они считают, что формальный объект (искусственный интеллект) не может автоматически приобрести каузальных свойств физического объекта (естественного интеллекта). Кроме того, обладание самосознанием и рефлексивным мышлением рассматривается ими как когнитивная способность, возникшая у людей и животных в результате эволюции. Особи, которые могли испытывать эмоции и контролировать свое поведение, получали преимущество в эволюционной гонке. Таким образом, способность к самосознанию у искусственного интеллекта отрицается по причине отсутствия возможности моделирования процесса эволюции.

Тем не менее, идея создания сильного ИИ обрела немало сторонников. В настоящее время ее поддерживают философы нового поколения Н. Бостром, Д. Чалмерс, Э. Юдковский, которые тесно сотрудничают с крупными компаниями, занимающимися разработкой систем искусственного интеллекта.

Так, сторонники сильного ИИ утверждают, что ему не удастся избежать эволюционного развития и борьбы за гегемонию, вследствие чего им и будет обретена способность к самосознанию. Более того, Н. Бостром даже опасается, что в процессе эволюции способный самостоятельно мыслить и принимать решения искусственный интеллект может представлять угрозу для существования человечества. Поэтому, по его мнению, сперва необходимо разработать философскую концепцию, которая даст понимание того, каким образом интегрировать искусственный интеллект в существующую цивилизацию. В связи с этим возникает одна из основных задач философии в области сильного ИИ - обозначить

те представления, ценности и цели, которые программисты будут внедрять в разрабатываемые системы.

Стоит отметить, что сторонники сильного ИИ сходятся во мнении о том, что для его создания потребуется еще не меньше нескольких десятков лет. Согласно теории, сформулированной в 1980-ых гг. и получившей название «парадокс Моравека», наделить искусственный интеллект способностью совершать сенсомоторные операции значительно сложнее, чем обучить его высококогнитивным процессам. Например, в настоящее время программистам не составляет труда «научить» систему ИИ игре в шахматы, но при этом наделить ее бессознательными навыками восприятия и эмоциональности годовалого ребенка пока даже не представляется возможным. По этой причине проблема создания сильного ИИ находится в проблемной области философской науки, перед которой стоит задача выяснить природу самосознания, рефлексивного мышления, личного субъективного опыта и других процессов, которые неразрывно связаны с понятием человека и личности.

Таким образом, вопрос о возможности создания искусственного интеллекта, обладающего самосознанием, является одним из наиболее актуальных и спорных в современной философии. Чтобы программисты смогли разработать сильный ИИ, философам предстоит выяснить принцип функционирования человеческого сознания, а также определить основные моральные ориентиры и ценности для будущих систем. В связи с этим нельзя не согласиться со словами физика Д. Дойча о том, что «проблема искусственного интеллекта - это проблема философии, а не компьютерных наук и нейрофизиологии, т.к. именно философия поможет разработать машины, которые думают, как люди» [2].

Источники и литература

- 1) Бостром Н. Искусственный интеллект. Этапы. Угрозы. Стратегии. М., 2016.
- 2) Deutsch D. Philosophy will be the key that unlocks artificial intelligence // The Guardian. 2012.
- 3) Kak S. Artificial Intelligence, Consciousness and the Self // Medium. 2018.