

О философском смысле понятия «сингулярность»

Научный руководитель – Шаронов Дмитрий Иванович

Ли Мэнян

Аспирант

Московский государственный университет имени М.В.Ломоносова, Философский факультет, Кафедра философии языка и коммуникации, Москва, Россия

E-mail: 1064739062@qq.com

Истоки концепции «сингулярность» восходят к Джону фон Нейману, пионеру компьютерных наук 1950-х годов. Фон Нейман предсказал, что «постоянно ускоряющийся прогресс технологий» приведет к «критической сингулярности в истории человечества». Под «сингулярность» подразумевается момент, когда искусственный интеллект, способный к самосовершенствованию, превосходит человеческий интеллект. Согласно взглядам Курцвейла на закон ускорения отдачи, технологическая сингулярность наступит примерно в середине XXI века. А другой известный сторонник сингулярности, Вернор Винчи, даже перенес ее на 2030 год. Всемирная шумиха вокруг ChatGPT, роботизированного инструмента для общения, разработанного американской лабораторией искусственного интеллекта OpenAI в 2023 году, вновь привлекла внимание к проблеме сингулярности, и кажется, что сингулярность, изменившая историю человечества, наступит раньше.

Сингулярность искусственного интеллекта претерпела трансформацию от пессимизма к оптимизму в западном обществе. В настоящее время существуют сторонники и противники сингулярности в вопросе о том, наступила она или нет. Вернор Виндж процитировал концепцию «интеллектуального взрыва» Гуда: сверхразумная машина, намного превосходящая человеческий интеллект, будет способна создавать более совершенные машины[n1]. Он считает, что человечеству нужно только изобрести первый суперкомпьютер, и наступит эра суперискусственного интеллекта. Знаменитый философ Чалмерс выдвинул философский аргумент в пользу появления сверхразума в области искусственного интеллекта. Игнорируя фактор времени, он считает, что сингулярность искусственного интеллекта возможна, поскольку после того, как люди создадут ИИ, равный их интеллекту, и ИИ+, превосходящий средний уровень человеческого интеллекта, ИИ+ создаст сверхразумный ИИ++, который будет еще более разумным, чем он сам, благодаря своей способности улучшать и усиливать себя[n2].

Скептики «сингулярности», представленные Дрейфусом, считают, что искусственный интеллект вообще не может развиваться в человеческий интеллект и обладать человеческим сознанием. В настоящее время большинство китайских ученых считают, что сингулярность искусственного интеллекта не наступит, и все еще существуют сомнения в том, прошел ли ChatGPT тест Тьюринга. Они считают, что ChatGPT не является человеком и не сможет заменить человека в будущем: способность к запоминанию знаний, которой обладает ChatGPT, намного превосходит человеческую. Однако интуитивные способности, которыми обладает человек, не могут быть достигнуты настоящим и будущим искусственным интеллектом.

Трансгуманизм - антропоцентрическая позиция, которая подразумевает, что человеческий интеллект является конечным мерилем искусственного интеллекта в трансгуманистической сингулярности искусственного интеллекта, поскольку только развитие интеллекта, подобного человеческому, может превратить искусственный интеллект в ориентированную на человека машинную аугментацию; в противном случае люди всегда будут жить

в тени того, как бороться с угрозами, исходящими от искусственного интеллекта. С технической точки зрения, проблема, стоящая перед трансгуманистами, ориентированными на сингулярность искусственного интеллекта, заключается в том, как можно соотнести человеческий и искусственный интеллект. Фактически, это также основная проблема, стоящая перед современным искусственным интеллектом. ChatGPT, несомненно, является самой умной машиной, существующей сегодня на планете, однако она все еще сталкивается с множеством проблем, таких как фактические ошибки. Основная причина этого заключается в том, что она еще не дошла до высокой степени редукции человеческого интеллекта, а скорее до высокой степени симуляции человеческого интеллекта в смысле явлений или поведения в форме машинного интеллекта (вычислительная мощность + база данных «глобального мозга» + алгоритм), но без возможности постичь весь процесс интеллекта. Основным препятствием, сдерживающим эту сингулярность искусственного интеллекта, является не технологическая проблема, а социальное и этическое ограничение в рамках антропоцентрической философской позиции.

В отличие от трансгуманизма, постгуманисты больше не признают субъективность человека или считают, что человек, по крайней мере, не единственный субъект, и придерживаются позиции нон-антропоцентризма. Нон-антропоцентрическое технологическое отношение означает, что понимание искусственного интеллекта отходит от ссылок на человеческий интеллект, подчеркивая, что человеческий интеллект может быть не единственной формой развитого интеллекта во Вселенной, но что могут существовать и другие формы интеллекта. Согласно этой философской позиции, приоритет человеческого интеллекта является в первую очередь объектом отрицания, таким образом, корреляционный барьер между человеческим интеллектом и машинным интеллектом, от которого страдает трансгуманистически ориентированный искусственный интеллект, перестанут быть труднопреодолимой проблемой, и идея создания искусственного интеллекта, независимого от человеческого интеллекта, скорее всего, будет поддержана. И как только такая программа будет реализована на практике, тогда независимо разработанный машинный интеллект сможет достичь сверхинтеллекта путем непрерывного самоусиления, тем самым достигнув сингулярности, поскольку ему не придется ограничиваться имитацией человеческого интеллекта, а когда сверхинтеллект будет реализован, он также сможет завершить «восходящий удар» по человеческому интеллекту[n3].

ChatGPT как представитель нового поколения искусственного интеллекта является крупным технологическим изменением, что же касается двух типов искусственного интеллекта сингулярности, который, вероятно, будет реализован первым, то сейчас нельзя определить, но можно с уверенностью сказать, что в случае внешних факторов для поддержки условий оба типа сингулярности в конечном итоге имеют возможность реализоваться. В это время, вероятно, человечество столкнется не только с конфликтом между людьми и машинами, но и с конфликтом между трансгуманистами и постгуманистами внутри человеческой расы, последний из которых, похоже, требует большей бдительности, чем другие, подобно борьбе между земными организациями трех тел и остальной частью человеческой расы до прихода цивилизации трех тел в «Трех телах».

Источники и литература

- 1) Chalmers D. J. (2010) The Singularity: A Philosophical Analysis. Journal of Consciousness Studies.p.7–65.
- 2) Vinge V. (1993) The Coming Technological Singularity: How to Survive in the Post-human Era. Science Fiction Criticism: An Anthology of Essential Writings. P.352–363.
- 3) Ник Бостром. (2015) Сверхинтеллект: дорожная карта, опасности и стратегии борьбы. Перевод: Бенвей Чжан, Юцин Чжан.