

**Применение концепции круга морали к анализу этических вопросов  
искусственного интеллекта.**

**Научный руководитель – Разин Александр Владимирович**

*Казюлин Вадим Сергеевич*

*Студент (бакалавр)*

Московский государственный университет имени М.В.Ломоносова, Философский  
факультет, Кафедра этики, Москва, Россия

*E-mail: dance.on.star@mail.ru*

В 20-ом веке перед философией по-настоящему остро встал вопрос отношений человека и других форм жизни. Стали очевидны огромные отрицательные эффекты, которые наша деятельность оказывает на экологию, а эксплуатация человеком животных стала беспрецедентна в своих масштабах. В то же время успехи биологии, генетики и психологии, в сочетании с падением авторитета религии, сделали очевидной неадекватность старых систем моральных ценностей в новых условиях, и необходимость создания нового этического учения, адекватного современным вызовам. Одним из крупнейших и самых влиятельных мыслителей, предпринявших попытку выработки этого нового учения, стал австралийский философ Питер Сингер. Его теория оказала огромное влияние на экологическое движение и движение за права животных, что, позволяет сказать: для своего времени его разработки были адекватны, пусть и достаточно спорны.

В 21-ом веке перед человечеством встают новые, ещё более серьёзные этические вызовы, чем в 20-ом. На наших глазах, создаётся нечто, чего природа никогда не знала, и не узнала бы без нашего вмешательства - искусственный интеллект. В контексте данной работы под “искусственным интеллектом” понимается в первую очередь сильный искусственный интеллект. Сейчас его возможности ограничены, но близится день, когда человечеству придётся ответить на вопрос, чем является искусственный интеллект по отношению к человеку, и какого морального статуса он заслуживает. В этом исследовании мы предпринимаем попытку проанализировать этическую систему Питера Сингера и выяснить, пригодна ли она для ответа на этические вызовы искусственного интеллекта.

Для начала мы рассматриваем этическую систему Сингера саму по себе. Для этого мы анализируем две его работы - “Animal Liberation” [3] и “The Expanding Circle” [4]. Свою работу Сингер относит к сфере социо-биологии, утверждая, что в основе общественной морали лежит присущий всем социальным животным альтруизм - родственный (kin altruism), взаимный (reciprocal altruism) и, в меньшей степени, групповой (group altruism). При этом, Сингер не считает, что мораль сводится лишь к биологии, видя в ней важный рациональный компонент, присущий людям, позволяющий нам выстраивать общие системы ценностей, важнейшим условием которых, является непредвзятость [4].

Исходя из этого, Сингер вводит важнейшую для нас концепцию - этический круг [4]. Вхождение в этот круг, являющийся своеобразной ценностной сферой, определяет моральный статус того или иного существа. По мнению Сингера, этот круг постоянно расширяется, уже вобрав в себя всё человечество, к которому скоро должны присоединиться ещё и животные. Растения и неорганическое вещество Сингер в этический круг не включает.

Мы приходим к выводу: для того, чтобы ответить на интересующий нас вопрос, нам необходимо выяснить, войдет ли в искусственный интеллект в этический круг Сингера. Чтобы ответить на этот новый вопрос, мы анализируем теорию Сингера далее, и приходим к выводу, что в ней можно выделить две главные категории, определяющие этический статус существа - разум (reason) [3] и способность испытывать страдания [4]. Люди обладают

обоими из этих свойств, животные, лишь одним. Таким образом, наш вопрос, разделяется на два: 1) Обладает ли искусственный интеллект разумом? 2) Способен ли искусственный интеллект испытывать страдания?

В поисках ответа на первый вопрос мы обращаемся сначала, к Рене Декарту, который, как мы считаем, является одним из первых философов, поставивших данный вопрос в “Рассуждении о методе” [1]. Ответ Декарта был отрицательным, однако многие не соглашались с ним, считая, что такой вывод - преждевременен и объясняется недостатком эмпирических знаний. После Декарта мы обращаемся к одному из первых специалистов в области искусственного интеллекта Алану Тьюрингу, который в своей работе “Computing Machinery and Intelligence” даёт утвердительный ответ на наш вопрос, говоря, что вычислительная машина может обладать разумом, неотличимым от человеческого [5].

Нахождение ответа на второй вопрос - намного более сложная задача. Для этого мы обращаемся к финскому учёному Аапо Хюваринену и его работе “Painful intelligence: What AI can tell us about human suffering”. Автор не даёт уверенного ответа на этот вопрос, утверждая, в схожем с Тьюрингов ключе, что искусственный интеллект может испытывать состояния в достаточной степени схожие со страданием [2].

Мы оцениваем оба полученных ответа как неудовлетворительные, так как они требуют принятия определённых допущений. По нашему мнению, сложность здесь заключается в фундаментальной невозможности нахождения какого-либо объективного ответа на данные вопросы, так как и разум, и страдания другого не могут быть эмпирически испытаны. В поиске этих ответов, мы вынуждены обращаться к суждениям о должном, а не о сущем.

Теория Сингера гласит, что человеческий разум вечно стремится к непредвзятости и обобщению, что и приводит к таким выводам, однако нам эта аргументация кажется недостаточной, потому что упирается в некие этические предубеждения, которые Сингер не замечает и не подвергает критике. Способность к страданию, которая для Сингера является условием вхождения в моральный круг, не является достаточным критерием сама по себе, так как остаётся без ответа вопрос о том, почему страдание нечеловеческих существ должно иметь значение для человека. Наконец, если мы поместим искусственный интеллект в систему Сингера, приняв в его отношении те же допущения, что принимаются по отношению к животным, моральный статус искусственного интеллекта окажется равен человеческому, а это представляется нам излишне смелым утверждением. В то же время, саму концепцию этического круга мы находим очень удачной эвристической конструкцией, хорошо подходящей для описания стоящих перед нами этических вызовов.

В итоге нашего исследования, мы приходим к выводу, что сама концепция этического круга является весьма конструктивной и полезной для изучения современных этических вызовов, но конкретная её интерпретация Питером Сингером не позволяет нам найти ответ эти вызовы, по причине наличия в ней необоснованных допущений, ведущих к неудобным для нашего общества выводам.

### Источники и литература

- 1) Декарт Р. Сочинения в 2 т.: Пер. с лат. и франц. Т. 1 / Сост., ред., вступ. ст. В. В. Соколова. // М.: Мысль, 1989.
- 2) Hyvärinen, Aapo. (2022). Painful intelligence: What AI can tell us about human suffering. 10.48550/arXiv.2205.15409. [https://www.researchgate.net/publication/360994220\\_Painful\\_intelligence\\_What\\_AI\\_can\\_tell\\_us\\_about\\_human\\_suffering](https://www.researchgate.net/publication/360994220_Painful_intelligence_What_AI_can_tell_us_about_human_suffering)
- 3) Singer Peter. Animal Liberation // NY, 2015.
- 4) Singer Peter. The Expanding Circle // Princeton University Press, 2011.

- 5) Turing A. Computing machinery and intelligence // Mind: журнал. — Oxford: Oxford University Press, 1950. No. 59. С. 433—460.