

ТЕСТ ТЬЮРИНГА И ПРОБЛЕМА КВАЛИА

Е. И. Гуринович

Белорусский государственный университет, г. Минск;

eugene.gurin@icloud.com;

науч. рук. – А. И. Зеленков, д-р филос. наук, проф.

Работа посвящена проблеме соотношения природы искусственного и естественного интеллектов. Обращаясь к тесту Тьюринга, являющегося в компьютерной науке (computer science) и теории информации (information theory) общепринятым критерием определения способности машины мыслить, можно обнаружить имплицитно содержащуюся в нем посылку об экстернатальности проявлений у какого-либо субъекта интеллекта и сознания. Это подтверждается позицией Г. Райла, утверждающего принципиальное отсутствие доступа к сознанию другого субъекта, кроме как посредством внешних его проявлений, к примеру, в диалоге с нами. Исходя из этого, делается заключение об объективном (не зависящем от природы интеллекта тестируемого субъекта) характере результатов теста и ставится под сомнение необходимость концептуального разделения интеллекта (и в целом – сознания) на искусственный и естественный (квалитативный).

Ключевые слова: искусственный интеллект; квалиа; философия сознания; аналитическая философия; тест Тьюринга.

Тема искусственного интеллекта, особенно в аспектах его «искусственности» и проблемы обретения им некоторой «натуральности», сегодня занимает достаточно значимое место как в массовой культуре (в кинофильмах «Я, робот», «Чаппи» и др., и, наиболее актуально, – в видеоигре «Detroit: Become Human»), так и в основательном научно-философском анализе. Оксфордский словарь определяет понятие «интеллект» (intelligence) как «способность обучаться, понимать и думать о вещах логическим образом» [1]. Стратегии же решения проблемы искусственности машинного интеллекта разделяются в массе своей на две общих стратегии:

- машины не могут «чувствовать» в том виде, в каком к этому способен человек (по причине отсутствия у них чувственного опыта, или «квалиа»);
- машины к этому способны, хоть и иным образом.

1. Подходя к решению данной проблемы, обратимся сначала к тесту Тьюринга: он вводится в работе «Может ли машина мыслить?», в ее первом разделе, – «Игра в имитацию». Уже само название раздела, в котором вводится тест с его классическими параметрами, как-бы намекает на то, что Тьюринга не беспокоит проблема искусственности, от которой он, и, в самом деле, уходит в самом начале главы, где он пишет: «Дело в том, что, если мы будем выяснять значение слов «машина» и «мыслить», исследуя, как эти слова употребляются обычно, нам трудно будет избе-

жать того вывода, что <...> ответ на вопрос «могут ли машины мыслить?» следует искать путем статистического обследования наподобие анкетного опроса <...> однако это нелепо» [3, с. 19]. Уклоняясь от такой чрезвычайно ресурсоемкой, и, по его мнению, бессмысленной, работы, Тьюринг и предлагает свой тест, суть которого можно кратко изложить по схеме:

- имеется три актора (мужчина (А), женщина (В) и задающий вопросы (С));
- актер С пространственно отделен от двух других участников, а подача вопросов актерам А и В производится дистанционно (во времена Тьюринга – с помощью печатной машинки, в идеале – телеграфа);
- цели теста (игры) для актора:
 - (мужчины) – состоит в том, чтобы задающий вопросы (С) пришел к неверному выводу о половой принадлежности как его самого (А), так и его «оппонента» (В);
 - (женщины) – состоит в том, чтобы доступными и разумными средствами помочь задающему вопросы (С) понять, кто есть кто;
 - (задающего вопросы) – состоит в том, чтобы понять, кто из двух неизвестных ему собеседников Х и У есть либо А (мужчина), либо В (женщина).

Очень часто смысл и содержание теста опускаются или искажаются, что приводит к неверным выводам такого характера, как, например, вывод о том, что цель теста в выявлении у машин способности мыслить посредством обнаружения у них некоторого наличествующего интеллекта в некотором субстанциальном (квалитативном) виде. Однако для Тьюринга это совсем не так, и, задаваясь вопросом о том, могут ли машины мыслить, он видит ответ и решение этой проблемы исключительно во внешних проявлениях этого свойства «мыслить», а именно, – в способности машины убедить человека в том, что она так же является человеком, используя для этого исключительно демонстрацию внешних признаков человечности и, как следствие, разумности. Признаки же эти, в сущности, состоят в способности вести аргументированный диалог, в демонстрации тех или иных «человеческих» повседневных речевых привычек (напр., употребление слов-паразитов), и в общей манере ведения диалога, а также в стратегиях аргументации, – все, в совокупности, в целях условной симуляции того, что собеседник актора С (задающего вопросы человека), – так же человек, того или иного пола.

2. В свою очередь, Гилберт Райл, британский аналитический философ, в своей работе [2] представил и обосновал позицию, утверждающую, что человеческое сознание (по сути, любое сознание вообще, а с

ним и интеллектуальные свойства) можно постичь исключительно и только посредством его внешних проявлений, преимущественно выражаемых в речевой коммуникации, но и не только. Как пишет Райл, – «Не существует прямой причинной связи (возможно, за исключением телепатии) между тем, что происходит в одном сознании, и тем, что происходит в другом. Только через посредство интересубъективно (публично) данного физического мира сознание одного человека способно что-то различать в сознании другого» [2, с. 23]. Таким образом, мы уверяемся в наличии сознания, а, следовательно, и интеллекта, у стороннего субъекта исключительно посредством внешних его проявлений, и исключительно посредством взаимодействия с данными проявлениями. Райл по данному вопросу заключает: «Люди <...> неисцелимо слепы и глухи к работе другого сознания и не в силах [прямо] воздействовать [непосредственно] на него» [2, с. 23].

И в самом деле, факт наличия сознания, равно как и интеллекта, у другого человека нам не доступен никаким иным опытным путем, кроме его условно внешних проявлений, к примеру, в диалоге с нами. Иными словами, чужое сознание для нашего собственного есть ни что иное как «black box»: мы можем сколь угодно препарировать и томографировать мозг человека, однако это совершенно (по крайней мере, пока) не дает нам никакого доступа к, если проводить параллель с П.К., интерфейсу этого сознания, к его личному внутреннему слуховому опыту, как, к примеру, пишет Райл: «Никто же не думает, что когда в моей голове вертится мелодия, то хирург может извлечь из моего черепа маленький оркестрик...» [2, с. 44], в отличие от той же функции удаленного доступа к рабочему столу стороннего П.К.

В сущности, из сказанного можно сделать вывод о том, что тест Тьюринга может быть использован не только для выявления некоторых интеллектуальных способностей у машины, посредством ее демонстрации возможностей к симуляции человеческого поведения в диалогичной форме и последующего убеждения собеседника-человека в собственной человечности, но и для определения интеллектуальности (и разумности) как таковой.

3. Одна из наиболее современных позиций, отстаивающих уникальность, натуральность, человеческого сознания и интеллекта была представлена американо-австралийским аналитическим философом Дэвидом Челмерсом [4], а именно в его концепции «квалиа». Челмерс определяет квалиа как «феноменальные качества», в числе наиболее известных – красность, как опытное переживание красного цвета. Сознательность же (consciousness) он определяет, как «...[возможность] говорить о том, как ково это – находиться в таком ментальном состоянии [сознательности].

То есть мы можем сказать, что ментальное состояние сознательно, если с ним связано качественное чувство, ассоциировано опытное качество» [4 с. 20]. Однако, наличие качественных чувств у субъекта, очевидно, подтвердить невозможно никак иначе, кроме как через получение от него описаний его личного феноменального опыта. Здесь, вспоминая Райла, мы понимаем, что подтвердить на опыте наличие качественного опыта сознания у того или иного субъекта изнутри этого самого субъекта мы никак не можем, и нам остается, как в тесте Тьюринга, полагаться лишь на его внешние признаки. Иного доступа к стороннему сознанию и его интеллектуальным способностям у нас попросту нет.

Таким образом, мы приходим к выводу следующего характера: либо квалиа, в виде некой субстанциальности, или как некоторого свойства материи, не существует вовсе, либо задаваться вопросом о ее существовании не имеет особого смысла, по крайней в рамках вопроса об искусственности машинного интеллекта. Говоря иначе, нам кажется вполне обоснованным утверждать то, что вопрос о наличии интеллекта у машины, будь он искусственен путем симуляции, или же условно натурален путем обретения качественных или квалиа-подобных свойств, уже и в текущую техническую эпоху, попросту снимается: как показывает опыт, независимо от уровня интеллекта субъекта с которым мы взаимодействуем, мы не имеем никакого эмпирического доступа к его сознанию, и, таким образом, мы не можем проверить его изнутри (а это единственный эмпирически-достоверный путь), дабы обнаружить (или не обнаружить, на примере концепта «философского зомби» (субъекта, имеющего полное соответствие на атомарном уровне другому человеку, но лишённого качественного опыта), предложенного самим Челмерсом) некоторые наличествующие у него сознание, качественный опыт и интеллектуальные свойства. Лишь то или иное самописание и демонстрация тех свойств, направленные сторонним субъектом вовне, являются для нас единственным доступным материалом для определения наличия интеллекта у этого субъекта и степени его развитости.

Следовательно, нет никакой принципиальной разницы между искусственным и натуральным (качественным, человеческим) интеллектами. Любой из них обнаруживается и исследуется нами исключительно по его внешним проявлениям, а получения прямого доступа к «интерфейсу» сознания носителей этих интеллектуальных способностей, по крайней мере, на данный момент, не представляется сколь-нибудь возможным: единственными признаками (с ходом технического прогресса постепенно устранимыми) для различения искусственного интеллекта от натурального для нас остается лишь несовершенство внешних проявлений первого. К примеру, такими несовершенствами с точки зрения чело-

века у машины, демонстрирующей интеллект, являются максимальная лаконичность и прямота фраз, отсутствие слов-паразитов в речи, очевидность схематичности построения диалога, а при личном общении с механизмом-роботом носителем этой формы интеллекта – механичность интонаций его голоса, и т. п. Следовательно, по крайней мере, в рамках решения задачи по созданию полноценного «искусственного» разума, во всем равном человеческому, а то и во всем его превосходящем, на наш взгляд, стоит заниматься не поисками возможностей по переносу в механическое устройство условной «частицы бога» в виде той самой кванта и качественного опыта, или попытками ее воспроизведения, а простой акселерацией тенденций по повышению плотности вычислительных процессов на единицу электро-технического пространства (следуя логике закона Мура), углублением способностей к обучению машин-носителей таких процессов, а также общим качественным прогрессом в робототехнической и компьютерной сфере.

Библиографические ссылки

1. Definition of intelligence noun from the Oxford Advanced Learner's Dictionary [Электронный ресурс]. URL: <https://www.oxfordlearnersdictionaries.com/definition/english/intelligence> (дата обращения: 16.05.2019).
2. Райл Г. Понятие сознания. М., 1999.
3. Тьюринг А. М. Может ли машина мыслить? М., 1960.
4. Челмерс Д. Сознательный ум : В поисках фундаментальной теории. М., 2013.