

МЕЧТАЮТ ЛИ АНДРОИДЫ ОБ АВТОБИОГРАФИИ?

Современное развитие идей и технологий искусственного интеллекта добавляет практического значения философским дискуссиям, касающимся вопросов личности и личного тождества. Цель статьи — продемонстрировать преимущество нарративного подхода для описания тождества кибернетических организмов. Для этого формулируется мысленный эксперимент по загрузке сознания ученого в тело андроида. В конце статьи автор рассуждает о потенциальной иерархии нарративных критериев.

Ключевые слова: тождество личности, нарративный подход, андроид, Волков.

Развитие искусственного интеллекта достигло небывалой скорости за последние годы. Один из актуальных примеров — это нейросеть Midjourney, способная по текстовому описанию нарисовать картину, персонажа и т.п. Но мало кто сейчас со всей серьезностью будет утверждать, что данный ИИ — самостоятельный творец, а не просто инструмент в руках человека. Чтобы он не был таковым, нейросеть должна уметь самостоятельно интерпретировать окружающую действительность и высказывать осмысленные суждения, например, посредством творчества. Midjourney можно будет назвать разумной, если она окажется способна творить на основе какого-то личного опыта, если у нейросети будет 'центральная абстракция' [1, с. 122], с позиции которой она станет оценивать вещи, явления и т.п.

То же самое можно сказать и про искусственный интеллект, инсталлированный в человеческую оболочку так называемых андроидов. Когда-нибудь людям придется всерьез задаться вопросом: имеет ли улучшенная версия, например, Амеки¹, разум? Может ли она называться личностью? Остро эта проблема встанет, когда андроид впервые совершит поступок, который можно будет приравнять к преступлениям с точки зрения права, или проступкам, за которые люди обычно чувствуют моральную ответственность. И так как отвечать за свои действия может лишь существо, осознающее их, в какой-то момент придется устанавливать разумность ИИ, чтобы решить: списывать его с эксплуатации, как дефектный прибор, или проводить над ним полноценный судебный процесс.

Однако проблема в том, что говорить об искусственном интеллекте с точки зрения его разумности, личности и ее тождества без особых проблем может лишь нарративизм. В этой статье я постараюсь доказать, что нарративная теория является наиболее удобной для рассуждений, касающихся обладания ИИ личностью, а также позволяет решить наибольшее количество проблем в контексте вопросов личного тождества такого объекта.

Кибернетические ощущения. Начну я с биологического подхода. Дело в том, что, если поставить этот подход перед ситуацией сложнее, чем наличие у андроида тела и мозга, близких к роботу, он стол-

кнется с проблемами и начнет противоречить сам себе².

Рассмотрим один из биологических подходов, а именно нейронный анимализм, утверждающий возможность сохранения личности при продолжении активности «исправно функционирующего мозга» [2, р. 63], выступающего источником различных чувственных переживаний и психологических активностей, которые и составляют собственно личность. Так, если рассматривать личность с привязкой к мозгу, который дает человеку возможность ощущать чувственные переживания, благодаря чему личность и выстраивается, то в обычном искусственном теле, неспособном обрабатывать раздражители в окружающей среде, личность содержаться не может. Но если взять тело, способное ощущать, то все становится не так однозначно. Представим, что некий Иван спроектировал искусственное тело, покрытое огромным количеством датчиков, способных определять различные раздражители, на которые реагирует человек (температура, гладкость поверхности и т.п.), и нейронный мозг, способный обрабатывать сигналы так же, как это делает биологический мозг человека. Добавим, что Иван дополнительно изобрел устройство, способное безошибочно просканировать человеческий мозг, нейронные связи в нем, позволяющее записать этот 'скан' в своеобразный программный код и моментально загрузить его в андроида с установленным в него ранее неорганическим мозгом. Далее Иван решает провести эксперимент над собой, надевает прибор на голову, запускает процесс... И просыпается в теле андроида. Поняв, что он осознает себя и ощущает окружающий мир кибернетическим телом, Иван на радостях встает в надежде увидеть свой оригинал, но с ужасом замечает, что человеческое тело лежит неподвижно. Проведя обследование, он понимает, что его человеческое тело под действием устройства впало в глубокую кому. И тут нейронный анимализм сталкивается с парадоксом. Родное тело Ивана, в котором содержится его 'родной' мозг, лежит перед ним и, согласно рассматриваемой точке зрения, его личность находится 'внутри'. Но в то же время, благодаря датчикам на искусственном теле, продолжилась непрерывность чувственных ощущений, значит, и личность осталась

прежней. Так, сторонник этого подхода должен утверждать, что Иван наблюдает со стороны и осознает самого себя, лежащего в коме, при этом он одновременно существует в другом, искусственном теле, но не является самим собой. Если же утверждать, что без сохранности биологического мозга, его ствола, нет сохранности разума и личности, так как организм, а значит, и животное существо, погибает [3], то в ответ можно представить следующую ситуацию. Раз Иван научился загружать свое сознание из человеческого тела в андроида, то он мог бы научиться и противоположному процессу: загружать себя туда и обратно бесконечное количество раз. Согласно этому примеру, выходит так, что Иван научился уничтожать разум и личность, а после 'возродить' их, что звучит абсурдно. И из-за того, что мысли об ИИ ставят нейронный анимализм с подобными парадоксами, я не считаю возможным для этой теории вступать в спор относительно андроидов. Остальные же биологические подходы никак не подразумевают сохранность личности Ивана в этом случае, поэтому отдельно рассматриваться не будут.

Следующим я разберу психологический подход. Про неолоккеевскую теорию о том, что тождество личности зависит от «непрерывности и связности между разными стадиями сознания личности» [4, с. 21], уточним, что она по сути является одним из «внутренних требований к содержанию биографической истории» [5, с. 25–26] к тождеству личности с точки зрения нарративного подхода, а именно «интелигибельностью траектории психической жизни» [5, с. 27]. Оба эти подхода постулируют, что действия, изменения и развитие личности должно быть логичным, прослеживаться эмпирически. Только если неолоккеевская теория полностью основана на данном принципе, то для нарративизма речь, скорее, идет лишь об одном из основных требований. То же самое касается и теории о q -воспоминаниях [6, с. 99–100], которая очень схожа с перспективой нарратива: q -воспоминания должны эмпирически прикрепляться к определенной личности, тем самым идентифицируя ее и превращаясь в обычные воспоминания. А что есть личные воспоминания человека, как не определенная перспектива, с которой он смотрит на происходящее вокруг [7, с. 387]? Они же и являются 'пржектором' [5, с. 24], что выхватывает из множества деталей именно те, которые волнуют определенную личность. И так как нарративная теория предполагает психологические качества личности, но при этом не налагает на нас требования точного описания содержания и принципов работы сознания, она более удобна для рассуждений об ИИ.

Теперь я хочу сконцентрировать внимание на внутренних и внешних принципах нарративного подхода, удобно систематизированных Б. Волковым [5]. Именно благодаря им нарративный подход выигрывает у остальных при обсуждении искусственного интеллекта и обходит парадоксы, создаваемые другими теориями. Так, если вернуться к примеру с Иваном, то в его непростой ситуации именно нарратив позволит решить проблему идентификации самости, существующей одновременно в двух состояниях, и сказать, что он остался личностью, хоть его разум и перешел в искусственную форму.

Начнем с внутренних требований к биографической истории. Во-первых, Иван все еще имеет определенную перспективу, с которой он смотрит на

происходящие события и оценивает их. Во-вторых, его автобиография остается интелигибельной. Можно эмпирически проследить, как Иван от предвкушения удачного эксперимента пришел к замешательству, когда увидел свое человеческое тело в коме, после к панике и затем к активному размышлению о том, как все исправить. В-третьих, видна и телеологическая направленность биографической истории Ивана. Допустим, целью всей его жизни было сделать себе имя в научном сообществе. Эта же цель могла остаться и после эксперимента, но отойти на время на второй план из-за возникших более насущных проблем. Она могла и интелигибельно, а значит, не противоречиво, измениться из-за них, если бы Иван забросил научную деятельность после пережитого шока.

Дополним пример рассуждениями о внешних требованиях. Возможность вербализации у Ивана присутствует, так как он явно способен осознать и организовать в жизненную историю свой опыт, иначе бы он не переживал из-за того, что случилось. Данное обстоятельство сохраняет правдоподобность нарратива. Имея воспоминания из человеческого тела, Иван может составить достоверную историю своей жизни, что будет устраивать знающих его людей.

Противопоставив биологический и психологический подход нарративному, я пришел к выводу: последний выигрывает у первых из-за того, что биологическая теория, сталкиваясь с ИИ, встречается с серьезными парадоксами, в то время как психологические теории могут быть 'переведены' в термины нарративизма. Учитывая это, можно сказать, что нарративный подход больше подходит для рассуждений о сознании искусственного интеллекта.

В следующей части статьи я намереваюсь сделать нарративную теорию еще более гибкой, чтобы она могла ответить на возникающие вопросы относительно проведенного мысленного эксперимента.

Кибернетический рассказ. Первая проблема — это решение того, какой из наших Иванов является продолжателем оригинального нарратива. То есть какого из них стоит считать 'настоящим'. Чтобы решить этот вопрос, я предлагаю обратиться к в корне отличающимся теориям о нарративе-фикции и нарративе-автобиографии, анти-реалистическому и реалистическому подходу соответственно. Я говорю о «метафизическом статусе героя нарратива» [8, с. 185].

Рассуждая ранее об Иване, я опирался на реалистический подход, на критерии, позволяющие признать автобиографию релевантной. Последними из них были внешние: возможность вербализации и правдоподобность нарратива. Говоря о вербализации/артикуляции, я имею в виду не только рассказ протагонистом своей истории от первого лица. Как писал Волков: «...дополнения истории могут и даже должны формулироваться другими личностями. Рассказы от третьего лица помогают включить в биографическую историю фрагменты, которые сам агент может не помнить, или отчет в которых он себе не отдает, а также обеспечить второе важное внешнее условие нарратива — правдоподобность» [9, с. 58].

Таким образом, утверждается тесная связь между внешними критериями. Работая в тандеме, они позволяют охарактеризовать определенную личность 'снаружи', идентифицировать ее. Нарративная фикция же как раз занимается характеристикой и идентификацией личности со стороны.

Можно сказать, что теорию о нарративной фикции в себя вбирает концепция нарратива-автобиографии. И раз фикция — это лишь часть от целого, то можно говорить о ее несамостоятельности. Чтобы более детально объяснить свою позицию, вернемся к мысленному эксперименту. Это поможет обосновать мне несостоятельность мыслей Э. Олсона и К. Витта о том, что успешное будущее нарративной теории возможно, только если ограничить ее вопросом характеристики [10] и выгодно выделить на фоне фикции нарратив-автобиографию.

Так, если применить обе рассматриваемые нарративистские теории к Ивану-в-коме и Ивану-андроиду, то выяснится, что первый имеет лишь прикрепленную к нему нарративную-фикцию. Хотя о нем и осталось представление у других людей, которые могут благодаря этому его идентифицировать и сказать, что он и есть 'настоящий' Иван — рядом с ним стоит Иван-андроид, который продолжает 'писать' связанную с ними обоими автобиографию. Он разумен и имеет при этом все качества известного нам Ивана. Если добавить к мысленному эксперименту жену ученого-изобретателя, допустим, она вошла в комнату в момент, когда Иван-андроид осознал себя, то она интуитивно будет более склонна считать андроида ее 'настоящим' мужем, особенно если он заговорит с ней и продемонстрирует свое прямое отношение к нарративу Ивана, которого она знала до эксперимента. Тело в коме с прикрепленной к нему фикцией окажется в еще более невыгодном положении, если внешность искусственной оболочки будет максимально приближена к биологическому телу ученого. Тогда, я уверен, интуиция жены и всех знакомых Ивана окажется решающим образом на стороне Ивана-андроида. Поэтому продолжателем оригинального нарратива здесь выступает именно киборг. Интуиция жены относительно тождественности Ивана в этом случае будет напоминать эксперимент с интуицией пересадки [11, с. 85], разве что здесь 'пересаживается' продолжатель оригинального нарратива, и спор идет между нарративистскими теориями, а не психологизмом и анимализмом.

Ранее я говорил, что теория о нарративе-автобиографии 'включает в себя' нарративную фикцию как свою составную часть, тем самым мы можем объединить два принципа, оставив из них наиболее фундаментальный. Не стоит упускать из виду и тот факт, что фикция все еще 'прикреплена' к ученому, находящемуся внутри искусственного тела. Просто в этом случае она не имеет решающего значения в определении сохранности 'настоящего' Ивана. И раз фикция содержится во внешних критериях релевантности нарратива-автобиографии и сама по себе оказывается несостоятельной, это позволяет выстроить некоторую иерархию нарративных критериев тождества личности. Она будет заключаться в том, что внутренние требования к нарративу играют более важную роль, чем внешние. Фикция может помочь личности быть собой в случаях, когда самостоятельное осуществление непрерывности нарратива невозможно.

Подобное сосуществование обеих концепций можно продемонстрировать на примере с младенцем. Личность новорожденного выстраивается только за счет окружающих, которые поддерживают прикрепленную к нему нарративную фикцию. Но постепенно, когда ребенок становится все сознательнее и сознательнее, он начинает приобре-

тать способность самостоятельно выстраивать автобиографию. Развиваясь, он заполучает личную перспективу, выраженную интеллигентность и телеологическую направленность. Это можно сравнить с тем, как младенцу помогают перемещаться, пока он не научится ходить сам, не станет более самостоятельным. Таким образом, со временем факт присутствия в теле ребенка личности становится более убедительным.

Фикция же в этом случае, во-первых, усиливается тем, что к вербализации нарратива окружающими добавляется его вербализация протагонистом и, во-вторых, отходит на второй план, уступая место полноценной автобиографии. Дальнейшая идентификация со стороны может понадобиться, если человек, например, попадет в вегетативное состояние. Тем самым нарратив продолжится, и личность сохранится благодаря фикции, выступившей в роли своеобразной заплатки в отсутствие 'основы' в виде полноценной автобиографии.

Таким образом, я делаю вывод: теории о нарративной-фикции и нарративе-автобиографии могут сосуществовать в иерархии, что позволяет нарративизму быть ещё более гибким и отвечать на возникающие вопросы, касающиеся ИИ и личности в целом.

Примечания

¹ Робот-андроид, созданный Engineered Arts в 2021 году, умеет копировать человеческую речь, может вести осмысленный диалог с пользователем.

² Перед началом аргументации необходимо прояснить, что в данной статье не обсуждаются вопросы наличия или отсутствия сознания у искусственного интеллекта, для удобства в своих рассуждениях я отгакиваюсь от позиции, что достаточно развитый ИИ не отличается от человеческого разума.

Библиографический список

1. Деннет Д. С. Почему каждый из нас является новеллистом // Вопросы философии. 2003. № 2. С. 121–130.
2. Licon J. A. Properly Functioning Brains and Personal Identity: An Argument for Neural Animalism // SATS: Northern European Journal of Philosophy. 2013. Vol. 14, Issue 1. P. 63–69. DOI: 10.1515/sats-2013-0004.
3. Sharpe K. W. Animalism and Person Essentialism // *Metaphysica*. 2015. Vol. 16, Issue 1. P. 53–72. DOI: 10.1515/mp-2015-0003.
4. Логинов Е. В, Мерцалов А. В, Салин А. С. [и др.]. Прологомены к проблеме тождества личности // Финиковый компот. 2018. № 13. С. 6–40. DOI: 10.24411/2587-9308-2018-00001.
5. Волков Д. Б. Нарративный подход как решение проблемы тождества // Вестник Санкт-Петербургского университета. Серия 17. Философия. Конфликтология. Культурология. Религиоведение. 2016. № 4. С. 21–32. DOI: 10.21638/11701/spbu17.2016.403.
6. Парфит Д. Тождество личности = Parfit D. Personal Identity / пер. с англ. Р. Л. Кочнева // Омский научный вестник. Сер. Общество. История. Современность. 2019. Т. 4, № 2. С. 94–107. DOI: 10.25206/2542-0488-2019-4-2-94-107.
7. Локк Дж. Опыт о человеческом разумении // Сочинения. В 3 т. Москва: Мысль, 1985. Т. 1. С. 78–620.
8. Левин С. М. Нарративный подход: фикция и реальность // Философский журнал. 2018. Т. 11. № 3. С. 184–187. DOI: 10.21146/2072-0726-2018-11-3-184-187.
9. Волков Д. Б. Как выжить, рассказывая историю. Нарративный подход к проблеме тождества личности // Финиковый Компот. 2018. № 13. С. 55–59. DOI: 10.24411/2587-9308-2018-00005.

10. Olson E., Witt K. Narrative and Persistence // Canadian Journal of Philosophy. 2019. Vol. 49, Issue 3. P. 419–434. DOI: 10.1080/00455091.2018.1486674

11. Парфит Д. Мы — нечеловеческие существа = Parfit D. We are not human beings / пер. с англ. У. В. Добронравовой // Омский научный вестник. Сер. Общество. История. Современность. 2020. Т. 5, № 4. С. 82–95. DOI: 10.25206/2542-0488-2020-5-4-82-95.

БЛИЖЕНСКИЙ Александр Дмитриевич, студент направления «Журналистика: журналистика в цифровом обществе: авторская и редакторская деятельность» Института социально-гуманитарных наук

Тюменского государственного университета, г. Тюмень.

Адрес для переписки: sblichenskiy@gmail.com

Для цитирования

Ближенский А. Д. Мечтают ли андроиды об автобиографии? // Омский научный вестник. Сер. Общество. История. Современность. 2023. Т. 8, № 2. С. 116–120. DOI: 10.25206/2542-0488-2023-8-2-116-120.

Статья поступила в редакцию 21.03.2023 г.

© А. Д. Ближенский

DO ANDROIDS DREAM OF AUTOBIOGRAPHY?

The modern development of ideas and technologies of artificial intelligence adds practical significance to philosophical debates concerning questions of identity and personal identity. The aim of this article is to demonstrate the advantage of the narrative approach for describing the identity of cybernetic organisms. For this purpose, a mental experiment is formulated to download the consciousness of a scientist into the body of an android. At the end of the article the author discusses a potential hierarchy of narrative criteria.

Keywords: identity of personality, narrative approach, android, Volkov.

References

1. Dennet S. D. Pochemu kazhdy iz nas yavlyayetsya novellistom [Why everyone is a novelist // Voprosy filosofii. *Philosophy Issues*. 2003. No. 2. P. 121–130. (In Russ.).
2. Licon J. A. Properly functioning brains and personal identity: an argument for neural animalism // *SATS: Northern European Journal of Philosophy*. 2013. Vol. 14, Issue 1. P. 63–69. DOI: 10.1515/sats-2013-0004. (In Engl.).
3. Sharpe K. W. Animalism and person essentialism // *Metaphysica*. 2015. Vol. 16, Issue 1. P. 53–72. DOI: 10.1515/mp-2015-0003. (In Engl.).
4. Loginov E. V, Mertsalov A. V, Salin A. S [et al.]. Prolegomeny k probleme tozhdestva lichnosti [Prolegomena to personal identity problem] // *Finikovyy kompot. Date Palm Compote*. 2018. No. 13. P. 6–40. DOI: 10.24411/2587-9308-2018-00001. (In Russ.).
5. Volkov D. B. Narrativnyy podkhod kak resheniye problemy tozhdestva [Narrative approach as the solution to the personal identity problem] // *Vestnik Sankt-Peterburgskogo universiteta. Seriya 17. Filosofiya. Konfliktologiya. Kul'turologiya. Religiovedeniye. Vestnik of Saint Petersburg University. Series 17. Philosophy. Conflict Studies. Culture Studies. Religious Studies*. 2016. No. 4. P. 21–32. DOI: 10.21638/11701/spbu17.2016.403. (In Russ.).
6. Parfit D. Tozhdestvo lichnosti [Personal identity] / trans. from Engl. R. L. Kochnev // *Omskiy nauchnyy vestnik. Ser. Obshchestvo. Istoriya. Sovremennost'. Omsk Scientific Bulletin. Series Society. History. Modernity*. 2019. Vol. 4, no. 2. P. 94–107. DOI: 10.25206/2542-0488-2019-4-2-94-107. (In Russ.).
7. Locke J. Opyt o chelovecheskom razumenii [An essay concerning human understanding] // *Sochineniya. V 3 t. [Works. In 3 vols.]*. Moscow, 1985. Vol. 1. P. 78–620. (In Russ.).
8. Levin S. M. Narrativnyy podkhod: fiktsiya i real'nost' [Narrative identity: fiction and reality] // *Filosofskiy zhurnal*.

The Philosophy Journal. 2018. Vol. 11, no. 3. P. 184–187. DOI: 10.21146/2072-0726-2018-11-3-184-187. (In Russ.).

9. Volkov D. B. Kak vyzhit', rasskazyvaya istoriyu. Narrativnyy podkhod k probleme tozhdestva lichnosti [How to survive telling a story. narrative approach to personal identity problem] // *Finikovyy Kompot. Date Palm Compote*. 2018. No. 13. P. 55–59. DOI: 10.24411/2587-9308-2018-00005. (In Russ.).

10. Olson E., Witt K. Narrative and persistence // *Canadian Journal of Philosophy*. 2019. Vol. 49, Issue 3. P. 419–434. DOI: 10.1080/00455091.2018.1486674. (In Engl.).

11. Parfit D. My — nechelovecheskiye sushchestva [We are not human beings] / trans. from Engl. U. V. Dobronravova // *Omskiy nauchnyy vestnik. Ser. Obshchestvo. Istoriya. Sovremennost'. Omsk Scientific Bulletin. Series Society. History. Modernity*. 2020. Vol. 5, no. 4. P. 82–95. DOI: 10.25206/2542-0488-2020-5-4-82-95. (In Russ.).

BLIZHENSKIY Aleksandr Dmitriyevich, Student of «Journalism: Journalism in a Digital Society: Author and Editorial Work» Faculty, Social Sciences and Humanities Institute, Tyumen State University, Tyumen.
Correspondence address: sblizhenskiy@gmail.com

For citations

Blizhenskiy A. D. Do androids dream of autobiography? // *Omsk Scientific Bulletin. Series Society. History. Modernity*. 2023. Vol. 8, no. 2. P. 116–120. DOI: 10.25206/2542-0488-2023-8-2-116-120.

Received March 21, 2023.
© A. D. Blizhenskiy