

ФИЛОСОФИЯ НАУКИ

ПРОБЛЕМА САВАНТИЗМА СОВРЕМЕННЫХ НЕЙРОННЫХ СЕТЕЙ

А. В. Антипова

Московский государственный университет им. М. В. Ломоносова, г. Москва
annaantipova1415@gmail.com

Статья подготовлена при финансовой поддержке гранта МЦИС № 1/2022

Одной из основополагающих парадигм в исследованиях искусственного интеллекта является коннекционизм, который утверждает сетевую природу когнитивных феноменов. Соответственно, будущее как компьютерных технологий, так и когнитивного моделирования многие исследователи видят в развитии технологии нейронных сетей.

Однако ученые во многом опираются на модель «стандартного разума» не обращая внимание на патологические его варианты в духе Локка: дети, пожилые, неевропейские культуры или люди с ментальными расстройствами. Мы полагаем, что изучение именно таких «пограничных» случаев может быть более продуктивным в определении того, что является необходимым, а что случайным для человеческого мышления. К патологиям же в философии и когнитивных науках обращаются все чаще, хотя пока такие работы носят скорее описательный, а не теоретический характер, например книги Оливера Сакса [Сакс, 1986]. Как отмечают исследователи шизофрении Луи Сасс и Йозеф Парнасс, в психологических практиках имеется склонность упрощать происходящие с больными процессы, в то время как феноменологи обратили внимание на снятие некоторых «автоматических» процессов обработки информации и принятия решения [Парнас, Сасс, 2003]. Таким образом, мыслительные процессы качественно усложняются за счет снятия некоторых автоматизмов. Наиболее известным феноменом данного типа является савантизм – наличие выдающихся способностей в одной или нескольких областях знаний, контрастирующих с общей ограниченностью личности. Человек с синдромом саванта способен повторить несколько страниц текста, услышанного им всего один раз, безошибочно назвать результат умножения многозначных чисел, или сказать, на какой день недели придется 3 мая любого года. Чаще всего способности савантов проявляются в области музыки, математики, рисования, календарных расчетов и т. д. [Биджакджи и др., 2021].

Низкоуровневые процессы принципиально недоступны для интроспекции, поэтому саванты не могут объяснить источник своих способностей. Алан Снайдер полагает, что такие процессы намеренно ингибированы для сознания, а стратегия автоматизации ускоряет принятие

решений, особенно при предоставлении частичной информации о мире. [Снайдер, 2009]. Стоит отметить, что навыки савантов не связаны со способностью творчества, они зависят не от активного обучения, а от легкого считывания необработанной информации. Навыки не улучшаются с течением времени и исчезают с лечением болезни. Однако когнитивная стратегия савантов имеет некоторые преимущества для классов решения проблем, которые связаны с разрушением когнитивных искажений (саванты за счет концентрации на деталях не имеют биаса при решении задач) [Снайдер и др., 2003].

Таким образом, саванты имеют привилегированный доступ к низкоуровневым процессам нервной системы. Критерием высокоуровневости – низкоуровневости в перцептивных процессах является вектор обработки информации и опосредованность опытом. Процессы, влияющие на восприятие непосредственно через органы чувств, называются восходящими, а процессы, влияющие на восприятие посредством знаний, опыта, контекста являются концептуально управляемыми процессами [Чистопольская и др., 2019].

Мы полагаем, что развитие современных нейронных сетей находится на аналогичной савантной стадии. Нейронные сети обладают неограниченным доступом к «сенсорной» информации (все пиксели изображения, частоты в музыкальном произведении и т. п.) без top-down процессов. Характерно, что нейросети называют имитаторами в тех же областях, в которых саванты объявляются гениями: математика, рисование, музыка, моделирование и т. п. С применением савантных нейронных сетей возникают следующие сложности:

1. Интерпретируемость. Так же как саванты не могут объяснить собственные способности, выходной слой нейронной сети не располагает объяснениями, почему сеть определила изображение в класс «машины», а не «собаки». Нейронные сети используются как «черные ящики», что неприемлемо, когда речь идет о решении в суде или постановке диагноза.

2. Проблема генерализации. Нейросеть способна распознать кошку на изображении, но если в датасете нет изображения кота, повернутого на 90 градусов, то фотографии котов под разными углами могут значительно испортить качество нейронной сети.

Таким образом, мы полагаем, что проблематику сильного искусственного интеллекта можно свести к проблеме механизмов концептуализации данных, которую в свою очередь можно решить, обратившись к феномену людей-савантов. Мы также считаем необходимым проводить исследования, не опираясь на обучение нейронной сети естественным языкам, поскольку имеются статьи, подтверждающие наличие высокоуровневых категорий у врановых и некоторых млекопитающих [Korzeniewski, 2022]. Как следствие, мы предлагаем сосредоточиться на

исследованиях концептуальной сложности современных нейронных сетей [Szymanski, 2022] и сравнить концепции feature-engineering в нейронных сетях с высокоуровневым процессом категоризации в биологических нейронных сетях. Мы полагаем, что только подобные междисциплинарные исследования позволят нам приблизиться к пониманию универсальности высокоуровневых процессов.

Литература

Биджакджи М., Кексал М., Балоглу М. (2021). Клинический случай человека с синдромом саванта из Турции: когнитивные функции и календарный расчет // *Клиническая и специальная психология*. Том 10. №1. С. 1–14.

Парнас Й., Сасс Л. (2003). Шизофрения, сознание и самость. *Бюллетень по шизофрении*, № 29 (3). С. 427–446.

Чистопольская А. В., Лазарева Н. Ю., Маркина П. Н., Владимиров И. Ю. (2019) Представление о высокоуровневых и низкоуровневых процессах в когнитивной психологии. Теория изменения репрезентации С. Ольссона с позиции уровневого подхода // *Вестник ЯрГУ. Серия Гуманитарные науки*. № 3. С. 94–101.

Korzeniewski B. (2019). Philosophy of Conceptual Network. *Open Journal of Philosophy*. No 09 (02). Pp. 104–139

Sacks O. (1986). *The man who mistook his wife for a hat*. London, UK: Picador.

Snyder A. (2009). Explaining and inducing savant skills: privileged access to lower level, less-processed information. In: *Philos Trans R Soc Lond B Biol Sci*. Pp. 1399–1405.

Snyder A., Mulcahy E., Taylor J., Mitchell D., Sachdev P. & Gandevia S. C. (2003). Savant-like skills exposed in normal people by suppressing the left frontotemporal lobe. *J. Integr. Neurosci*. No 2. Pp. 149–158.

Szymanski L., McCane B., Atkinson C. (2022). Conceptual complexity of neural networks. *Neurocomputing*, No 469. Pp. 52–64.

НЕОДНОЗНАЧНОСТЬ ОПРЕДЕЛЕНИЯ ПОНЯТИЯ ПОЛЬЗЫ: СПОСОБЫ ЕЕ РАЗРЕШЕНИЯ

И. А. Гапаров

Самарский национальный исследовательский университет
имени академика С. П. Королёва, г. Самара
sarov-1@mail.ru

Статья подготовлена при финансовой поддержке РФФИ, проект № 20-011-00462 А «Философия техники Фридриха Дессауэра: эпистемология и антропология реалистской теории творчества».

Актуальность исследования обусловлена применением понятия пользы в научном знании с целью обоснования закономерностей естественного, технического, социального плана, в число которых входят: этические кодексы ИИ, правовые нормы, экономические законы и др. Однако неоднозначность, данная в форме «знания о незнании» [Дубровский, 1994], не позволяет отчетливо определить понятие пользы. Для представителей социально-гуманитарных наук проблема определения понятия пользы касается его содержания, т.к. под многообразием признаков они не в состоянии обнаружить наиболее важные. Данный аспект наглядно представлен как в филологических [Дронова, Яньчунь, 2022], так и в правовых исследованиях [Бентам, 1998]. Напротив, для представителей сфер естественного и технического познания следствием введения понятия пользы становится проблема формы, предстающая в виде перевода из одного кода в другой, возможностей ИИ и т.д.

В философском плане проблема определения понятия пользы выглядит запутаннее, т.к. главная задача философии заключена в устранении хаоса, в сведении многообразия в единство. «Польза» - многоаспектна, она охватывает разнообразные стороны жизни (бытовую, научную, религиозную, др.). Однако глубина содержания этого понятия неопределенна. Зачастую, мыслители, не отдавая себе отчета, прибегают к «пользе» только для того, чтобы связать свои мысли воедино, приобрести уверенность, подкрепить положение доводом, говоря «в пользу чего-либо», тогда как иные полагают в «пользе» запечатлеть существо дела. Однако, ещё Г. Риккерт в труде, посвященном определению «жизни», показал, что «слово, которое должно обозначать любую мыслимость, неизбежно теряет отчетливость значения» [Риккерт, 1998, с. 306]. «Польза» является одним из таких понятий.

Противоречие, возникающее в плане определения содержания понятия пользы, подводит к важному разграничительному моменту. Понятие «польза» либо необходимо, и потому является исходным основанием для построения системы, либо оно предстает в качестве одного из связующих

моментов. Если «польза» является исходным основанием, задающим все параметры системы, то она дана в системообразующем контексте. Если же она является логическим аспектом, то ее роль в построении системы завязана на синтезе опосредующих звеньев, а потому о ней можно вести речь только как о чисто вспомогательной компоненте. Разграничение между контекстами наиболее полно запечатлено в характере их систематичности, существование которой было вскрыто Н. Гартманом [Гартман, 2001].

Для метафизических систем не существует логических противоречий, т.к. их исходным основанием является построение такой упорядоченности и взаимосвязанности элементов, при которой одно вытекает из другого, следуя целесообразности телеологической направленности. Однако в таких условиях «польза» становится понятием, опосредующим или связывающим другие, т.е. чисто вспомогательным. Наиболее показательны системы И. Канта и Г. В. Ф. Гегеля. Кантом под понятием пользы понимается внешняя целесообразность, тогда как «объективная внутренняя целесообразность, то есть совершенство, уже ближе к предикату прекрасного» [Кант, 1994, с. 95]. Напротив, для Г. В. Ф. Гегеля «польза» – аспект реализации идеи, но не внешняя целесообразность, т.к. в последней не может быть учтен синтез. Он может быть запечатлен в «пользе», предполагаемой как «умозаключение, в котором субъективная цель смыкается с внешней ей объективностью через некоторый средний термин, который есть единство обеих» [Гегель, 1974, с. 396]. В целом, важно отметить, что от И. Канта до Г. В. Ф. Гегеля заметна трансформация содержания понятия пользы. Если у первого «польза» – это только форма выражения рассудочной способности, пригодная для построения суждений «здорового смысла», то у второго «польза» – форма выражения разумной способности, предстающая в качестве функции синтеза разрозненных аспектов бытия.

Метафизическим системам можно противопоставить утилитаристские, прагматические и технические. В них «польза» является системообразующим понятием. Она становится исходным основанием построения и существования вышепредставленных систем. Причиной их разграничения выступает предмет исследования, подразумевающий определение и актуализацию практических действий, соотносенных с корректно поставленной целью, т.е. связанной с удовлетворением потребностей материального плана. Отсюда следует, что для утилитаристов основная цель – это общественное благополучие. Исходным основанием его достижения выступает принцип полезности как ряд методов, направленных на регулирование уровня страданий и удовольствий. Однако учесть все интересы нельзя, поскольку «каждый должен считаться за одного, и никто не может претендовать на большее» [Милль, 2013, с. 231]. Напротив, прагматисты полагают «полезным» эффект, исходящий от идеи и данный в чувственном эмоциональном переживании [Джемс, 2011]. Но эффект должен

считаться позитивным, т.е. облегчающим жизнь и утверждающим человека в его существовании. Отсюда следует, что «польза» у прагматиков дана, с одной стороны, как способ утверждения чего-либо в действительности, а, с другой, как предикат, характеристика, свойственная идее, через посредство которой ее можно реализовать. Утилитаристский и прагматистский способы имеют свое завершение в техницистском, в соответствии с которым «польза» – в широком смысле, основа ценностного мировоззрения [Дессауэр, 2022], в узком – это способ повышения производительности труда [Энгельмейер, 2010]. В целом, под понятием пользы в техническом аспекте понимается сила или «некоторое изменение, нарушение естественного порядка вещей, влекущее в том числе и трансформацию горизонтов фантазии, которые в свою очередь показывают возможности для новых технических объектов» [Нестеров, 2017, с. 98].

При разграничении вспомогательного и системообразующих планов в определении понятия пользы, можно прийти к следующим выводам. В чисто вспомогательном контексте понятие «польза» определяется как недостаточное совершенство или субъективная цель, которые должны быть преодолены на более высокой стадии развития сообразно телеологической направленности. Выражением «пользы» может стать идея, отражающая сущность мира. Здесь «польза» не может быть дана в снятом виде (как реализованная цель), т.к. она подчинена высшим закономерностям, для которых человек с его прихотями является не более чем средство в планах мирового духа. Напротив, в рамках системообразующего подхода, «польза» – то, что задает направление всем человеческим поступкам или выступает в качестве высшей ценности, на основании которой возможна связь между идеальным и реальным аспектами бытия. Именно здесь человек волен решать, что он делает и для чего, подвергать свои поступки рефлексии.

Литература

Гартман, Н. (2001). *Систематическая философия в собственном изложении. Фауст и Заратустра*. СПб. Азбука.

Гегель, Г. В. Ф. (1974). *Энциклопедия философских наук. Т. 1. Наука логики*. М. Мысль.

Дессауэр, Ф. (2022). Человек и космос. Пер. с нем. А.Ю. Нестеров. Семиотические исследования. *Semiotic studies*. Т. 2. № 2. С. 25–49. DOI: 10.18287/2782-2966-2022-2-2-25-49

Джемс, У. (2011). *Прагматизм: новое название для некоторых старых методов мышления*. М. Изд-во ЛКИ.

Дубровский, Д. И. (1994). *Обман. Философско-психологический анализ*. М. Изд-во «РЭЙ».

Дронова, Л., Яньчунь, Л. (2022). Формирование понятия «польза» в русском языке. *Филология и человек*. № 2. С. 65–82 DOI: 10.14258/filichel(2022)2-05.

Кант, И. (1994). *Критика способности суждения*. М. Искусство.

Милль, Дж. Ст. (2013). *Утилитаризм*. Ростов-на-Дону. Донской издательский дом.

Нестеров, А. Ю. (2017). *Семиотические основания техники и технического сознания: монография*. Самара. Изд-во Самарской гуманитарной академии.

Риккерт, Г. (1998). *Философия жизни*. К. Ника-Центр.

Энгельмейер, П. К. (2010). *Теория творчества*. М. Либроком.

ПРОБЛЕМА СОЗНАНИЯ КАК ВТОРИЧНОЙ СТРУКТУРЫ РЕАЛЬНОСТИ

А. И. Геньш

Новосибирский государственный педагогический университет,
г. Новосибирск
aleksandrgensh19@gmail.com

Современные представления о сознании чаще всего охватывают его психологические свойства, такие как память, внимание, мышление. Этот подход в определенной мере можно назвать бихевиористским. Однако его роль довольно ограничена, когда речь идет о попытках понять специфику сознания как философской категории. При этом часто игнорируется другой подход, феноменологический. Он направлен на понимание природы цвета, тепла, прикосновения [Чалмерс, 2013.С. 19].

Все, из чего состоит жизнь каждого человека, - это его субъективная реальность. На данный момент феноменологические представления о сознании еще нуждаются в изучении. Человечеству практически не известны способы изучения этой стороны вопроса. Непонятно, что представляет собой субъективный опыт отдельного человека со стороны внешнего наблюдателя. Практически не существует ни одного физического способа проверить, что чувствует другой человек и есть ли у него вообще какое-либо внутреннее ощущение.

Также стоит учитывать, что последнее время стали появляться новые вопросы. Если относительно человеческого существа мы интуитивно догадываемся о наличии у него феноменологической части сознания, то с появлением все более развитых и думающих машин, задача становится намного более трудной. Нам не известно, обладают ли они тем же опытом, которым обладаем мы.

Если продолжить размышлять о наличии сознания в нашей вселенной, то можно привести определенную критику чисто физического подходу, при котором роль сознания не учитывается. Так, если представить две абсолютно идентичных вселенных, в одной из которых есть сознание, а в другой – нет, то сравнивая их, мы поймем, что сознание является вторичной структурой физического мира [Гаспарян, 2013].

Подобная мысль заставляет задуматься об общей роли сознания в нашей реальности. Появляется ли оно только у биологических форм жизни достигших определенного уровня познавательных способностей? Возникает вопрос об эволюционном возникновении сознания и самосознания, которые необходимо различать.

Так сознание состоит из ряда феноменологических чувств, которые не всегда могут быть осознаны, иначе бы мы столкнулись с проблемой

возникновения сознания у детей по мере взросления. В таком случае был бы не понятен механизм, при котором сознание появляется с возрастом, на каком этапе оно возникает? И каким образом информация превращается в сознание? Что такое информация?

Самосознание - намного более простое явление. Оно фиксируется, так как дети в определенном возрасте начинают отличать себя от других. Но это психологическое качество сознания, а не феноменологическое [Махова, 2011].

Для иллюстрации сказанного можно привести мыслительный эксперимент, предложенный Недом Блоком в 1978 году. Представим, что мы собрали 11 миллиардов человек и выдали им рации для общения, чтобы они могли взаимодействовать друг с другом. Так, каждый человек будет равен одному нейрону, что будет выглядеть как симуляция человеческого мозга. Здесь возникает вопрос, образуется ли сознание из подобной структуры, ответить на него, действительно, трудно.

Вернемся к искусственному интеллекту. Если сознание должно образовываться из-за 11 миллиардов людей и раций, то в случае создания подобной структуры из неживой материи, например, из кремнёвых микрочипов, у ИИ должен образовываться свой феноменологический опыт. Но где возникает подобный опыт? Ведь если начать убирать один за другим нейроном у человека хирургическим путем, то когда он лишится сознания? А если убирать один кремниевый микрочип за другим у искусственного интеллекта [Волков, 2017]? Все это сводится к выводу, что сознание – это полноценная часть вселенной, дублирующей структуры физической реальности.

Подобные теоретические выкладки являются доказательством концепции панпротопсихизма, название которой дал Дэвид Чалмерс. Эта доктрина утверждает, что каждый объект во вселенной имеет разум или бессознательную психику, и что все физические явления связаны психикой. Рассмотрим основные проблемы, с которой встречается данная доктрина. По мнению некоторых авторов, эти проблемы являются, по сути, вопросами комбинаторики [Гусев, 2021].

1. Каким образом отдельные части системы образуют единое общее сознание? Например, кварки и протоны имеют свои микрофеноменологические свойства. При этом не совсем понятно, что представляют собой их «ощущения». Вряд ли мы можем даже осознать, что чувствуют подобные микрочастицы, и, тем более, каким образом из них образуется полноценное мыслящее существо.

2. Интуитивно мы связываем каждый элемент системы с его сознательным опытом. Но это не позволяет образовывать множественное сознание, поскольку они не могут просто суммироваться.

3. Как микрокачества комбинируются так, чтобы получалось макрокачество? Микроопыт включает в себя микрокачества как некие аналоги макрокачеств. Однако это дает весьма ограниченное количество микрокачеств, которые должны порождать огромное количество ограничений нашего феноменологического опыта, что не является столь однозначным.

В заключении стоит отметить, что, несмотря на многообещающие выводы и возможности, который дает панпротопсихизм, он обладает множественными недостатками, которые не позволяют взглянуть на картину мира полноценно. Идеи наличия сознания как вторичной структуры реальности могли бы продвинуть понимание нами проблем искусственного интеллекта и менее развитых биологических форм жизни. Но они все еще остаются неидеальными теориями, которые требуют дальнейших улучшений и дополнений.

Литература

Волков, Д. Б. (2017) Психологический подход как решение проблемы тождества личности // *Философская мысль*. 2017. №2. С. 13-25.

Гаспарян, Д. Э. (2013) В защиту феноменального сознания: аргументы против физикализма в современной аналитической философии (часть I) // *Философские проблемы информационных технологий и киберпространства*. 2013. №2. С. 43–54.

Гусев, А. А. (2021) Панпсихизм в современной аналитической философии сознания // *Философская антропология*. 2021. Т.7. №2. С. 85-103.

Левин, С. М. (2016) Разновидности нейрошовинизма: от теории тождества до микрофункционализма // *Вестник Московского университета. Серия 7. Философия*. 2016. №3. С. 55-69.

Мохова, Ю. А. (2011) Психологическая сущность самосознания личности // *Омский научный вестник*. 2011. №6 (102). С. 166-169.

Чалмерс, Д. (2013) Осознающий ум. В поисках фундаментальной теории сознания. М: УРСС: Книжный дом «ЛИБРОКОМ», 2013. 512 с.

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ОБЩЕСТВЕННЫЕ ИНТЕРЕСЫ

Ю. Д. Данилов

Брестский государственный технический университет,
г. Брест, Республика Беларусь
ydd213@gmail.com

Активное развитие систем искусственного интеллекта (ИИ), призванных служить обществу, является не только важным направлением для научных исследований, но и вполне актуальной целью для разработки национальных и международных стратегий. На сегодняшний день ситуация в данном вопросе такова, что несмотря на декларируемые цели по улучшению общественного благополучия, имеется множество примеров, когда интеллектуальные системы наносят больше вреда, чем пользы. Одной из основных причин такого положения представляется несоответствие имеющихся разработок существующим общественным интересам.

На протяжении столетий категория «общественный интерес» была одним из основных понятий не только в политической и правовой теории, но и в практической деятельности по принятию социально значимых решений. Этот термин, как и некоторые другие, используемые в политической науке, не имеет универсального определения. Сейчас, в условиях широко развернувшейся дискуссии о роли искусственного интеллекта, можно говорить о том, что данная категория переживает настоящий ренессанс.

Наиболее часто, термин «общественный интерес» воспринимается как оборотная сторона «индивидуального», «частного» или «группового» интереса. Это связано с заложенными в суть понятия ценностями и декларируемыми им целями: счастья, всеобщего благополучия, гуманизма и всего прочего, что никак не связано с извлечением прибыли, накоплением богатства, рыночной активностью и т.п.


Весьма значительные противоречия в понимании общественного интереса обозначились в политической теории. Они касаются вопроса о том, как определить критерии, позволяющие четко понимать, что считать или не считать таковым. В книге «Общественный интерес и личные интересы» известная феминистка Вирджиния Хелд утверждает, что сущность данного понятия может быть достаточно просто объяснена с помощью какой-либо общепризнанной нормативной концепции. Она предлагает рассматривать, анализировать и оценивать общественный интерес в каждом конкретном случае, принимая и уравнивая все имеющиеся обоснования частных интересов [Held, 1970, с. 168]. Барри Бозман рассматривает проблему более широко. В работе «Общественные ценности и общественные интересы» он описывает уже несколько подходов к определению данного термина, классифицируя их на нормативные, консенсуалистские и процессуальные.

Нормативный интерес, по его мнению, предусматривает понимание общественных интересов как важнейшую цель, которую должны преследовать государственные должностные лица, принимающие важные политические решения. Консенсуалистский рассматривает общественные интересы как отражение потребности в чем-то большинства членов общества. Процессуальные автор разделяет на три разновидности:

- агрегативные, предполагающие, что данный конкретный интерес является величайшим благом для наибольшего числа граждан;
- плюралистические, предусматривающие конкуренцию существующих у людей потребностей и запросов;
- процедурные, основанные на возможном примирении частных интересов [Bozeman, 2007, с. 13].

В этом рассуждении Бозман во многом опирается на философа-прагматика Джона Дьюи, который, по его мнению, применил альтернативный и очень продуктивный взгляд на общественные интересы. Его суть заключается в том, что следование общественным интересам – это использование непредвзятости и разумных, справедливых процедур для максимального приближения к некоему идеалу.

В связи с активным распространением искусственного интеллекта и его проникновением в различные сферы человеческой жизнедеятельности, помимо политического контекста понимания общественных интересов, не меньшее значение имеет и юридический аспект. Юридический дискурс сегодня также не дает универсального, ясного и точного определения данного термина. Он рассматривает всю данную проблематику через историю отношений и взаимодействия между сложившимися общественными ценностями и конституционно закрепленными нормами, а также способами их регулирования. Ряд специалистов в области права соглашаются с идеей, что общественный интерес должен основываться на общих ценностях и нормах, но при этом политика не может считаться отвечающей общественным интересам, если она противоречит элементарным общественным потребностям, этическим традициям, ментальным установкам и т.д. В частности, такой точки зрения придерживается М. Фейнтак, который прямо утверждает, что общественные интересы граждан должны во все времена служить противовесом для власти [Feintuck, 2017].

Чтобы более определенно сформулировать требования к искусственному интеллекту, которые могут стать критериями его служения общественным интересам, требуется ответить на ряд ключевых вопросов  насколько ИИ:

- публичен и не ориентирован на получение прибыли;
- служит обеспечению равенства людей;
- обеспечивает процесс обсуждения и принятия совместных решений;
- способен соблюдать ключевые технические стандарты;

- открыт для общественной проверки.

Любая разрабатываемая сегодня система искусственного интеллекта, представляющая общественный интерес, нуждается в публичном обосновании, не ориентированном на получение прибыли. Чтобы она действительно служила этой цели необходимо доказать, что данные технологии создаются не просто ради инноваций или коммерческих выгод, а для служения общему благу людей. Это чрезвычайно важно, т.к. многие процессы, происходящие в обществе, по определению просто не имеют никаких оснований для автоматизации и не могут быть решены даже самыми совершенными техническими средствами.

Искусственный интеллект должен служить равенству и обеспечению прав человека (или, как минимум, не подрывать их). Это требование непосредственно связано с этическим принципом справедливости. Для общественных интересов важно избегать таких результатов действий искусственного интеллекта, которые прямо противоречат справедливости или провоцируют власть, других людей на принятие нежелательных решений. Технологии, заложенные в ИИ, должны успешно поддерживать равенство между всеми людьми, развивать гуманистические подходы к проектированию и способствовать продвижению инклюзивных принципов.

Интеллектуальная система должна быть открытой для общества в целом, потому что только при таких условиях у людей появится возможность проверить ее, проникнуться к ней доверием и использовать в дальнейшем для других проектов, представляющих общественный интерес. А это, в свою очередь, создаст условия для решения такой задачи, как установление равенства между поколениями, поскольку общественные интересы исторически находят себе подтверждение только в ретроспективе, на примерах жизни будущих поколений.

Создание новых ИИ-систем, претендующих на представление общественных интересов, требует коллективного совещательного участия как можно более широкого круга заинтересованных лиц. Никакая команда разработчиков, какой бы технически продвинутой, умелой, этической, социально ответственной она ни была, не может сама определить, что действительно отвечает широким общественным интересам. В какой-бы форме не проходил процесс обсуждения общественной значимости ИИ, важно чтобы он всем людям предоставлял возможность быть услышанными.

И, наконец, системы искусственного интеллекта, представляющие общественный интерес, должны соответствовать техническим стандартам безопасности и защиты. На данный момент уже можно вполне четко обозначить наиболее актуальные здесь проблемы: качество информации, конфиденциальность данных и их защищенность.

Важно отметить, что концепция искусственного интеллекта, рассматриваемая в контексте общественного интереса, поднимает все больше

и больше новых вопросов, требующих дальнейшего изучения. Представляется, что поиск ответов на них должен проходить в том числе и в рамках широких и открытых академических и общественных дебатов, основанных на принципах доступности, взаимной лояльности и доверия.

Литература:

Bozeman, B. (2007). Public Values: An Inventory. In *Administration & society, Journal, Magazine* [Электронный ресурс]. URL: <https://journals.sagepub.com/doi/abs/10.1177/0095399707300703> (дата обращения: 13.02.2022).

Held, V. (1970). *The public interest and individual interests*. New York. Basic Books – 240 p.

Feintuck, M. (2017). Regulating the Media Revolution: In Search of the Public Interest. In *The Journal of Information, Law and Technology (JILT)*. N. 3: [Электронный ресурс]. URL: https://warwick.ac.uk/fac/soc/law/elj/jilt/1997_3/feintuck/ (дата обращения: 13.02.2022).

ЯЗЫК СОВРЕМЕННОЙ НАУКИ КАК ФИЛОСОФСКАЯ ПРОБЛЕМА

И. А. Дедова

Поволжский государственный технологический университет,

г. Йошкар-Ола

iraira0906@mail.ru

Люди верят, что их разум повелевает словами. Но бывает и так, что слова обращают свою силу против разума
Ф. Бэкон

С момента своего появления философия стала заниматься анализом языка науки. Уже древние греки увидели, что слова как таковые не способствуют достижению истины, а уведут от нее. Парменид пришел к выводу, что сущность небытия и понятие небытия различны, ведь первого вообще не существует. Демокрит утверждал, что «атом» никак не дан нам в опыте. Софисты оставили вопрос о сущности вещей как принципиально неразрешимый. Платон попытался соединить сущность вещи с ее именем в «эйдосе», однако признал наличие пропасти между миром чувственных вещей и умозрительных идей. Аристотель попытался преодолеть эту пропасть, объяснив, что вещь и ее сущность едины, но сохранил дуализм формы и материи: мы или *знаем* нечто о сущности вещей, или обладаем *опытом* о реальном мире. Христианские теологи увеличили пропасть между божественной сущностью и миром, погрузившись в исследование слов как таковых.

В Новое время Фрэнсис Бэкон сформулировал требования к языку науки. В «старой» науке он увидел «суеверия» и «абстрактные формы» и предпринял попытку определить границы научных понятий для того, чтобы а) перестать воспринимать научные слова как общеизвестное и очевидное; б) исключить вымысел и фантазию [Бэкон, LXV, CXIX, LX]. Современная наука объявила себя объективным, доказательным и опытно проверяемым знанием. А философия науки выработала требования к научным понятиям, дефиниции которых должны а) отражать атрибутивное качество вещи; б) не быть тавтологичными; в) объяснять нечто неизвестное через известное; г) не содержать указание исключительно на более общий, родовой признак. Тезис должен быть ясным, а следствия тезиса должны подтверждаться опытным путем, что и должно в конечном итоге служить подтверждением истинности теории. Образцом в этом отношении объявляется классическая и неклассическая физика.

Однако наш анализ языка этих теорий показал, что принцип лексической полноты в них не соблюдается, так что научные утверждения почти не отличаются от религиозных догматов. В качестве одного из

примеров приведем объяснение теории относительности: *«В общей теории относительности постулируется, что гравитационные эффекты обусловлены не силовым взаимодействием, и не силовым взаимодействием тел и полей, а деформацией самого пространства-времени, в котором они находятся. Эта деформация связана, в частности, с присутствием массы-энергии»* [Википедия; 2022]. В этом высказывании: а) априори признается существование гравитации как таковой, хотя сущность этого явления не объясняется ни у Ньютона, ни у других авторов; б) отрицается силовое взаимодействие тел как причина гравитации (то есть одно ньютоновское понятие безоговорочно принимается, а другое – нет; в) причиной гравитации объявляется деформация пространства-времени, хотя затем в качестве причины пространственно-временной деформации объявляется масса-энергия. То есть вначале эта теория отделяет гравитацию от силы (которая у Ньютона есть порождение массы), а затем связывает ее с массой-энергией.

Нам могут возразить: необходимо использовать авторские определения понятий, например, самого создателя теории относительности. Однако при рассмотрении труда Эйнштейна мы не встретили ни одного точного определения терминов [Эйнштейн, 2000]. Нет объяснения сути теории и в учебных пособиях: в основном предлагается некая описательная характеристика: *«Теория гравитационного поля, построенная на основе принципа эквивалентности и общего принципа относительности, называется общей теорией относительности»* [Алексеев, 2015, с. 23].

Изучив дефиниции физических понятий «пространство», «масса», «сила», «поле», «материя», «энергия», мы не увидели указания ни на один существенный признак, что позволило бы считать их действительно научными, то есть претендующими на объективность, ясность и полноту. Обратившись к научным трудам, объясняющим теорию относительности, мы также не встретили точных определений. Так, например, Э. Тэйлор, заявив в 70-е годы о победе теории относительности в научных кругах, констатирует, что ее понятия являются чем-то самоочевидным, то есть, по сути, нарушает принцип науки, сформулированный Ф. Бэконом: *«Понимание теории относительности отныне не есть проблема обучения, а просто дело интуиции – практикуемого подхода к вещам»* [Тэйлор, С. 14].

Практически во всех определениях мы встречаем либо а) указание на более общие и абстрактные понятия – «физический объект», «некоторая физическая величина», «многообразие», «мера воздействия» без дальнейшего пояснения, чем данный объект отличается от других: *«Область пространства, где проявляют себя физические, достоверно зарегистрированные и точно измеренные силы, называется физическим полем»* [Санько, 2022]; *«Поля физические – это физические системы, обладающие бесконечно большим числом степеней свободы»* [Павлов, 2004]; *«Пространство – всеобщая форма существования материи»* [Введенский,

1955]; б) тавтологии: «Согласно современным представлениям поле является универсальной формой материи, к которой могут быть сведены как вещества, так и классические поля» [Википедия, 2022]; либо в) признание беспомощности в попытках объяснить сущность данного предмета: «Принято оперировать термином «масса» и использовать обозначение *m* без пояснений» [Википедия, 2022]; или: «Общепринятое определение силы отсутствует» [Википедия, 2022].

Лексическая неполнота свидетельствует о том, что сущность природы современная физика понимает не больше, чем физика Платона и Аристотеля, а в какой-то степени даже меньше. Разрыв между эмпирией, представленной в экспериментальных данных, и теорией, объясняющей сущность явлений, не преодолен. Причина этого, скорее всего, в неверно выбранном методологическом подходе современной науки, проанализировать и пересмотреть который должна именно философия. Основное ошибочное допущение, на наш взгляд, заключается в дифференцированности знания, потери целостности, дуалистичности – разрыве между духом и природой.

Литература

Алексеев С. О. (2015). *Введение в общую теорию относительности, ее современное развитие и приложения*: Учеб. Пособие / С.О. Алексеев, Е.А. Памятных, А.В. Урсулов и др. Екатеринбург.: Изд-во Урал. ун-та. 380 с.

Бэкон Ф. (1935). *Великое восстановление наук. Новый органон*. Lib.ru [Электронный ресурс] URL: <http://www.lib.ru/FILOSOF/BEKON/nauka2.txt> (дата обращения: 30.08.2022).

Введенский Б. А. (ред.). (1955). *Пространство*. Большая советская энциклопедия. М. Открытый текст. Электронное периодическое издание [Электронный ресурс]. URL: <http://opentextnn.ru/old/glossary/prostranstvo/index.html?id=5191> (дата обращения: 01.09.2022)

Масса (2022). *Википедия. Свободная энциклопедия* [Электронный ресурс]. URL: <https://ru.wikipedia.org/wiki/Масса> (дата обращения: 1.09.2022).

Общая теория относительности. (2022). *Википедия. Свободная энциклопедия* [Электронный ресурс]. URL: https://ru.wikipedia.org/wiki/Общая_теория_относительности (дата обращения: 30.08.2022):

Павлов В. П. (ред.). (2004). *Большая российская энциклопедия*. Поля физические [Электронный ресурс]. URL: <https://bigenc.ru/physics/text/3156418> (дата обращения: 01.09.2022).

Поле. (2022). *Википедия. Свободная энциклопедия* [Электронный ресурс]. URL: [https://ru.wikipedia.org/wiki/Поле-\(физика\)](https://ru.wikipedia.org/wiki/Поле-(физика)) (дата обращения: 1.09.2022).

Санько Н. Ф. (2022). *Астрономический словарь*. Поле физическое [Электронный ресурс]. URL:

<http://www.iki.rssi.ru/hend/Dictionary/Electromagnetic%20field.htm> (дата обращения: 1.09.2022).

Сила. (2022). *Википедия. Свободная энциклопедия* [Электронный ресурс]. URL: <https://ru.wikipedia.org/wiki/Сила> (дата обращения: 1.09.2022).

Тэйлор Э. Ф., Уилер Дж. А.(1971). *Физика пространства-времени*. М.: «Мир». 1971. 319 с.

Эйнштейн А. (2000). *Теория относительности*. Ижевск: НИЦ «Регулярная и хаотическая динамика». 224 с.

«СМЫСЛЫ» СИНЕРГЕТИКИ В СОЦИАЛЬНО-ГУМАНИТАРНОМ ПОЗНАНИИ

Е. С. Ермаков

Институт философии Национальной академии наук Беларуси, г. Минск
ermakovgeka@gmail.com

Расширение использования синергетики – как ее методологических принципов, так и теоретического инструментария – в исследовании социальности ставит проблему философско-методологической рефлексии последствий не только для социально-гуманитарных наук, но и собственно для синергетики, про которую можно говорить в следующих взаимосвязанных аспектах: как о науке, парадигме и научной картине мира.

Как наука синергетика возникла на стыке различных отраслей физики, химии и нелинейной математика, и как отдельное научное направление наследует и продолжает развивать идеи, заложенные и тектологии, теории катастроф, общей теории систем, нелинейного моделирования и кибернетик. Как парадигма синергетика включает в себя ряд методологических принципов, которые эксплицируют процессы самоорганизации в открытых, нелинейных, неравновесных, диссипативных и сложных системах любой онтологической природы.

Как заметил В. С. Степин, «синергетика по своим онтологическим и методологическим принципам предстает в качестве одной из важнейших экземпляров постнеклассической рациональности» [Степин, 2013, с. 51]. Постулируя глубокая связь между классической, неклассической и постнеклассической научной рациональностью в их эволюционной преемственности, В. С. Степин высказал мнение, что на постнеклассическом этапе развития рациональности происходит «осмысление ценностно-целевых структур научного познания и их социокультурной обусловленности» [Степин, 2017, с. 27]. Отдельным аспектом постнеклассической рациональности выступает интенция на междисциплинарное взаимодействие, что включает сближение социально-гуманитарного и естественнонаучного знания. В центре такого сближения находятся аксиологическая направленность и социокультурная обусловленность познания. Речь идет не только о переносе отдельных методов и концептов из одной науки в другую для решения отдельных проблем, но об изменении стратегий научного познания в целом.

Представленные три аспекта синергетики следует рассматривать как её взаимосвязанные аспекты, или «смыслы», которые используются социальными исследователями в зависимости от их методологической интенции. В частности, как наука синергетика в меньшей степени используется в социально-гуманитарных исследованиях как по причине

сложного теоретического аппарата, так и необходимости исследователей сталкиваться с математическим моделированием. В качестве примеров работы Л. И. Бородкина [Бородкин, 2016] в моделировании исторических процессов, а также исследования на стыке математического моделирования и философии А.В. Колесникова [Колесников, 2018].

Синергетика как парадигма требует от социального исследователя более глубокого осмысления ее методологических основ с одновременным пониманием того, что использование вместе с этим ее понятийного аппарата необходимо последовательно и со смысловой нагруженностью. В противном случае возникает эффект метафоричности понятий синергетики. Как отмечает Д. В. Винник, «популярность синергетики привела к ее проникновению в социальное знание, в результате чего многие его реципиенты оказались неспособны адекватно понять смысл моделей и стали использовать словарь произвольным образом» [Винник, 2016. с. 87].

В качестве примера использования синергетики как парадигмы в социально-гуманитарном познании следует рассматривать работу М. М. Маковского по исследованию динамических аспектов эволюции языка. Являясь неотъемлемой частью культуры, язык эволюционирует при изменении его частей – морфем, лексем и их значений для дальнейшего успешного своего функционирования [Маковский, 1992, с. 11]. В этом отношении данные части можно рассматривать как отдельные параметры порядка языка, которые, как указывает М. М. Маковский, наследуются и изменяются во время эволюционной динамики самого языка [Маковский, 1992, с. 12]. Таким образом, рассмотрение языка в качестве отдельной динамической системы не только позволяет подвергать его моделированию, но также содействует раскрытию генезиса и развития языка, связывая это, как продемонстрировал М. М. Маковский, с комбинаторным разнообразием и стохастической вариативностью самого языка [Маковский, 1988, с. 29].

Последний «смысл» синергетики – как научная картина мира – связан с философским осмыслением онтологических и эпистемологических ее основ и следствий. В качестве наиболее яркого примера можно привести работы В. П. Бранского в области фундаментальных основ социальной самоорганизации в целом. Согласно В. П. Бранскому, социальная синергетика «исследует *общие закономерности социальной самоорганизации*, то есть взаимодействия *социального порядка и социального хаоса* (курсив В. П. Бранского)» [Бранский, 1999, с. 117]. Для объяснения социальной самоорганизации и ее внутренних механизмов он обращается к концепту социального отбора, факторами чего выступают тезаурус, детектор и селектор [Бранский, 1999, с. 120]. Как отмечает автор, в отличие от самоорганизации в естественных системах, суть социальной самоорганизации сводится к чередованию двух противоположных процессов – иерархизации и деиерархизации, что соответствует усложнению и

упрощению систем [Бранский, 1999, с. 118]. Социальная система, таким образом, находится около двух аттракторов – простого, соответствующего усложнению системы, и странного, соответствующего системному упрощению. Через внутреннее взаимодействие своих факторов, социальный отбор, имеющий стохастический характер, обеспечивает «путь» системы к соответствующему аттрактору, тем самым обеспечивая самоорганизацию.

Представленная концепция В. П. Бранского представляет не только пример «вторжения» синергетики в изучение социальности как таковой, но и является еще образцом философской рефлексии теоретической и методологической роли синергетики на постнеклассическом этапе развития науки. С онтологической точки зрения синергетика как научная картина мира сориентировала познание на изучение сложных и динамичных человекомерных систем, где человек и как наблюдатель, и как элемент данных систем олицетворяет их динамику, но и в то же время сам на них влияет. С методологической же стоит признать не только определенные возможности методологии синергетики в исследовании социальных парадигмы и, соответственно, картина мира была связана с дальнейшим расширением использования ее методологических принципов, в том числе и в социально-гуманитарном познании.

Таким образом, синергетика в социально-гуманитарном познании на постнеклассическом этапе развития науки начала XX-го века являлась теоретико-методологической частью исследования социальности в целом. Однако расширение синергетики не только представляет угрозу как для ее самой, так и для философии и социально-гуманитарных наук, но и дает новый взгляд на исследование социальных процессов в онтологической и методологической проекции.

Литература

Бородкин, Л. И. (2016). *Моделирование исторических процессов. От реконструкции реальности к анализу альтернатив*. СПб. Алетейя.

Бранский, В. П. (1999). Социальная синергетика как постмодернистская философия истории. *Общественные науки и современность*. №6. С. 117–127.

Винник, Д. В. (2016). «Эпистемический искупитель»: свод приемов легитимации бессмыслицы. *Философия науки*. №3. С. 76–93.

Колесников, А. В. (2018). Хаос и самоорганизованная критичность в формировании концепции цифровой философии сознания. *Философия науки*. № 4. С. 120–136.

Маковский, М. М. (1988). *Лингвистическая комбинаторика: Опыт типологической стратификации языковых структур*. М. Наука.

Маковский, М. М. (1992). *Лингвистическая генетика: Проблемы онтогенеза слова в индоевропейских языках*. М. Наука.

Стёпин, В. С. (2013). Типы научной рациональности и синергетическая парадигма. *Сложность. Разум. Постнеклассика*. № 4. С. 45–59.

Стёпин, В. С. (2017). Типология научной рациональности и синергетика. *Філософія освіти*. №1. С. 6–29.

АРГУМЕНТ ИЛЛЮЗИИ ВОДОПАДА В СОВРЕМЕННОМ ДИСКУРСЕ О ТЕМПОРАЛЬНОСТИ СОЗНАНИЯ

А. С. Зайкова

Институт философии и права СО РАН, г. Новосибирск
Zaykova.a.s@gmail.com

Если взглянуть на быстро движущиеся облака, а потом перевести взгляд на дома или деревья, может показаться, что дома и деревья движутся. Это – пример эффекта «последствия движения», часто называемый также «иллюзией водопада», поскольку схожий эффект можно обнаружить при наблюдении за водопадом и неподвижным берегом рядом: после наблюдения за движением воды кажется, что берег начинает двигаться в обратную сторону. Впервые этот эффект был обнаружен ещё Аристотелем [Месяц, 2021] и после забыт на долгое время, но о нём вновь вспомнили в XIX веке, однако до сих пор причины его возникновения остаются под вопросом. В современном дискурсе этот эффект стал катализатором методологической дискуссии: действительно ли можно использовать факт существования этой иллюзии как аргумент в теоретическом споре или отдельных примеров работы некоторой сенсорной модальности недостаточно для утверждения о преимуществе некоторой теории?

Самая распространённая версия объяснения заключается в теории нейронной адаптации. Эта версия предполагает, что вычисление расположения объекта и зрительную адаптацию к движению производят различные нейронные потоки [Mather, Verstraten, & Anstis, 1998]. Впрочем, главный вопрос для данной иллюзии состоит не столько в том, почему это происходит, а, скорее, в том, что это значит для теории сознания, и, в частности, для теории темпоральности сознания.

С. Проссер рассматривает эту иллюзию как аргумент против большинства современных темпоральных моделей сознания [Prosser, 2017, p. 7]. Он использует классификацию темпоральных моделей Б. Дэйнтон, в рамках которой все существующие модели можно представить как одну из трёх основных моделей – или как их комбинацию. Первая из них, кинематографическая, предполагает, что время воспринимается как набор кадров. Ретенциональная модель утверждает, что мгновенное осознание вмещает недавний динамический опыт. Наконец, экстенциональная модель представляет наше восприятие как составленное из отдельных, последовательных или пересекающихся эпизодов осознания, соответствующих эпизодам недавнего динамического опыта.

С Проссер использует пример иллюзии водопада как случай, когда движение переживается, хотя само содержание опыта не содержит никаких изменений – кроме времени. Действительно, когда мы глядим на

неподвижный объект после того, как только что смотрели на движущийся поток, содержимое нашего опыта можно назвать неизменным: ни дома и деревья в примере из начала тезисов, ни берег реки в примере с водопадом не совершают никаких изменений или движений. Однако в этот момент наш опыт является динамическим. Этот факт, во-первых, вдохновляет С. Проссера на предложение модели динамических снимков, а во-вторых – и становится аргументом против иных темпоральных моделей сознания. С его точки зрения, кинематографическая модель предполагает существование только статичных кадров, следовательно, не может обеспечить восприятие движения для случая, когда движения нет. Экстенциональная модель, согласно его мнению, также не может объяснить этот эффект, поскольку в её рамках, по мнению С. Проссера, воспринимаемое в один момент времени не может определяться в иной момент. Более того: он убеждён, что все эти модели принципиально неразличимы [Prosser, 2017, p. 10], и аргумент против одной из модели, таким образом, становится аргументом против всех них.

Наиболее радикальную критику в адрес С. Проссера предложил К. МакКенна, который обвинил своего оппонента в несостоятельности методологических принципов. Он утверждал, что теории, которые пытаются объяснить наше восприятие времени, не могут быть обоснованы доказательствами единственной когнитивной функции [McKenna, 2021, p 44]. В какой-то мере его утверждение может быть справедливо: учитывая обилие когнитивных функций, теория, объясняющая единственный пример и доказываемая им же, в целом неконкурентноспособна. Однако, вообще говоря, теория динамических снимков несколько шире, чем объяснение иллюзии водопада, и, следовательно, критика К. МакКена в данном случае бьёт мимо цели.

С нашей точки зрения, мы должны критически оценивать применение отдельных примеров работы сенсорных модальностей и когнитивных функций в качестве аргументов против существующих моделей. В большинстве случаев, для любой из моделей мы можем придумать некоторую надстройку, которая способна объяснить существование той или иной иллюзии, не противореча сути теории. Поэтому для того, чтобы отдать предпочтение одной из моделей, одного примера, даже столь яркого, как иллюзия водопада, явно недостаточно. Однако нам следует признавать роль и корректность использования отдельных примеров работы сенсорных модальностей для создания новых теорий, понимая, что оценивать теории необходимо учитывая весь спектр известных когнитивных функций.

Литература

Месяц, С. В. (2021). Аристотель. О сне и бодрствовании (перевод и комментарий)». *Платоновские исследования*. № 14(1). С. 226–251.

Dainton, B. (2018) Temporal Consciousness. In: N. Zalta E.N. (ed). *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition) [Web]. URL: <https://plato.stanford.edu/archives/win2018/entries/consciousness-temporal/> (дата обращения: 21.05.2022).

McKenna, C. A. (2021) Don't Go Chasing Waterfalls: Motion Aftereffects and the Dynamic Snapshot Theory of Temporal Experience. *Rev.Phil.Psych.* № 12. P. 825–845.

Mather, G., Verstraten, F., Anstis, S. (1998). *The motion aftereffect: A modern perspective*. The MIT Press.

Prosser, S. (2017). Rethinking the specious present. In: Phillips I. (ed). *Routledge handbook of philosophy of temporal experience*. Oxford: Routledge. 146–156.

ЯЗЫК КАК ФОРМАЛЬНАЯ СИСТЕМА. ЭЛЕКТРОННЫЙ ПОВОРОТ И ЭЛЕКТРОННАЯ СЕМАНТИКА. РЕКУРСИВНАЯ ПАРАДИГМА

А. А. Кузнеченков

Самарский университет, г. Самара

ankuzn1970@yandex.ru

Теория познания предполагает семантику, и до тех пор,
пока мы не поймем последнюю, не стоит иметь дело с первой.

Альберто Коффа

В работе рассматривается проблема семантической неполноты моделей формальных систем. Проводится анализ результатов в области формальных систем и значимых результатов электронной эволюции, с целью их возможного синтеза. Делается предположение о том, что использование результатов электронной эволюции в исследованиях формальных систем позволит «наполнить» модели формальных систем семантикой, что по значимости для развития формальных систем, сопоставимо с лингвистическим поворотом. Делается попытка оценки возможности электронного поворота и перспектив появления электронной семантики. Особенностью исследования является то, что оно проводится в рамках рекурсивного подхода. Рекурсия рассматривается в качестве основного механизма формирования семантики в электронных моделях формальных систем. Так же ставится вопрос о развитии рекурсивной парадигмы и рекурсивного мышления в целом.

Для определения понятия формальной системы, состоящей из аксиом и правила вывода, Д. Гильбертом использовалось понятие «формальная аксиоматика» [Гильберт, Бернайс, 1979]. «Гильберту принадлежит, во первых, подчеркивание того, что строгая формализация теории предполагает полную абстракцию от смысла - результат такой формализации называется формальной системой, или формализмом (или иногда формальной теорией, или формальной математикой); во-вторых, ему принадлежит метод, делающей формальную систему в целом предметом изучения математической дисциплины, называемой метаматематикой или теорией доказательств» [Клини, 2009, с. 60]. «Если рассматривать картину полностью, то имеются три отдельные и отличные друг от друга «теории»: (а) содержательная (informal) теория, формализацией которой служит формальная система, (b) формальная система или предметная теория и (с) метатеория, в которой описывается и изучается эта формальная система.» [Там же, с. 63]. Как правило, в качестве метатеории используется формальная логика (исчисление высказываний, исчисление предикатов, и т.д.), основные

вопросы метатеории – доказуемость, непротиворечивость и полнота. Вызывает интерес подход Р. Смальяна к определению формальной системы. Так, Р. Смальян обращает внимание на то, что «Понятия формальной системы и механической операции тесно связаны: каждое может быть определено в терминах другого» [Смальян, 1981, с. 15] и дает *рекурсивное определение формальной системы* без использования понятий формальной логики с использованием понятия разрешающей процедуры [Там же]. Понятия «язык» и «формальная системы» в работе рассматриваются как эквивалентные и будут использоваться в зависимости от семантического контекста.

Понятие семантики, рассматривается в контексте классической логико-семантической традиции, сформировавшейся благодаря работам Б. Рассела, Г. Фреге, Л. Витгенштейна, Р. Карнапа, А. Тарского, и т.д. Значимыми для рассматриваемой проблематики являются: разграничение смысла и значения Г. Фреге [Фреге, 2018] и формирование теории формальных систем, получившей свое дальнейшее развитие в теории моделей. Значение теории моделей для развития семантической традиции заключается в том, что: 1) было формализовано соотношение синтаксиса и семантики и, 2) определено, что семантика находится в метасистеме (метаязыке) в качестве которого выступает, как правило формальная логика или разрешающая процедура.

В 1931 году были опубликованы теоремы К. Гёделя о неполноте формальных систем содержащих арифметику [Gödel, 1931], что, с одной стороны, ознаменовало начало кризиса в основаниях математики, с другой стороны, придало импульс исследованиям в области теории алгоритмов. Можно утверждать, что именно исследованиями в области теории алгоритмов содержатся предпосылки электронной эволюции. Для данного исследования, целью которого является попытка синтеза результатов исследований в области формальных систем и результатов электронной эволюции, важен факт установления общей природы семантических и алгоритмических исследований, при этом понятие формальной системы выступает в качестве связующего понятия. Рассмотрение языка в качестве формальной системы позволяет использовать как результаты полученные в семантическом направлении, так и в алгоритмическом.

В качестве значимых для данного исследования достижений в области формальных систем можно выделить: 1) аксиоматический метод как основной для построения формальных систем; 2) формирование логических формальных систем - исчисление высказываний и исчисление предикатов; 3) построение на основании логики формальных систем, имеющие в качестве моделей – арифметику (арифметика Пеано); геометрию (аксиоматика Гилберта); теорию множеств (аксиоматика Цермелло-Френкеля); 4) теорию моделей, как результат развития теории формальных систем. Особое место в

исследованиях в области формальных систем занимает Principia Mathematica [Уайтхед, Рассел, 2005], как попытка построения математического языка в целом. Отдельно стоит отметить о полученных результатах в отношении понятия «истина»: 1) установлена аксиоматическая относительность и аксиоматическая замкнутость понятия истина - в различных аксиоматиках одно и то же выражение может иметь различные истинностные значения; 2) семантическая концепция истины А. Тарского [Tarski, 1936] которая рассматривает истину как семантическое понятие, что позволяет рассматривать ее смысл и значение по Г. Фреге; 3) определена алгоритмическая и рекурсивная природа истины - «Определение истины можно очень просто получить из определения другого семантического понятия – понятия выполнимости», и далее, «Для определения понятия выполнимости нам лучше вновь обратиться к рекурсивной процедуре» [Ibid, p. 171]; 4) сформированы предпосылки для рассмотрения истины в качестве онтологического понятия.

В качестве значимых для данного исследования достижений в области электронной эволюции можно выделить: 1) факт того, что электронное пространство обладает значительно большими формализационными возможностями чем символы на плоскости. Средствами электронного пространства возможно зафиксировать структуру формального объекта и моделировать процессы, в том числе и рекурсивные; 2) рассматривая компьютерную программу как аналог формальной системы, можно отметить, что в процессе ее электронной реализации она существует на различных семантических уровнях – человекочитаемом, как текст, и - машиночитаемом, как число в двоичной (логической) системе исчисления; 3) машиночитаемый уровень электронной реализации модели формальной системы является доказательством ее логической природы; 4) факт существования формальных объектов в электронно реализованных моделях позволяет дать допустимый в целом ответ на вопрос об онтологии абстрактных объектов (они существуют в электронной модели).

Теория моделей позволяет рассматривать неполноту как *свойство модели* формальной системы. Предположение о том, что неполнота модели формальной системы связана со средством формализации основано на факте того, что модели формальных систем реализуются: 1) символами на плоскости; 2) в сознании человека. Пример естественного языка свидетельствует о том, что семантика языка в целом, с одной стороны, содержится во всем множестве текстов, с другой стороны, динамическая модель языка реализована в сознании человека и семантика естественного языка находится именно там – в сознании человека. Электронная эволюция изменяет общую ситуацию, и это изменение связано с появившейся возможностью электронной реализации моделей формальных систем. Таким образом, в рамках электронной эволюции происходит развитие средств

формализации, что для формальных систем означает возможность перехода от символической реализации модели к электронной. Существенным ограничением символического способа реализации является то, что символы на плоскости обладают ограниченными возможностями семантической формализации, а значит и невозможно реализовать в полном объеме разрешающую процедуру [Смальян, 1981], выполнимость [Tarski, 1936] и т.д. - для этого нужен человек, или электронный автомат. Электронная эволюция расширяет само понятие «формальное», включая в него модели формальных систем, и тем самым *делая их предметом изучения*, что важно.

«Четвертое царство» Фридриха Дессауэра [Дессауэр, 2017], как следствие активного развития философии техники, зафиксировало факт технического поворота в исследованиях. «Очевидно, что в XXI веке человечество вступает в состояние «третьей природы», выстраивая новую искусственную среду поверх искусственной среды, созданной в XIX и XX веках, и у философского сообщества пока нет обобщающих глобальных моделей, которые, могли бы описать взаимодействие природного и искусственного, показать глобальный горизонт прогноза» [Нестеров, 2016]. Результаты электронной эволюции позволяют сделать предположение о том, что в рамках технического направления исследований созданы предпосылки для «электронного поворота».

Ограничения возможности символического способа реализации моделей формальных систем возможно определить как общую проблему неполноты моделей формальных систем. Помимо общей проблемы, существует специальная, и она связана с рекурсивными расширениями в формальном языке, которые являются следствием общерекурсивной природы познания в целом. Примерами проявления второй проблемы служат логические парадоксы в естественном языке (парадокс Лжеца, парадокс Брэдоброя и т.д.), парадоксы в математике (парадокс Кантора, парадокс Рассела и собственно теоремы К. Гёделя о неполноте), имеющие рекурсивную природу. Значимым является то обстоятельство, что парадоксы являются логическими, а следовательно в контексте рассуждений - семантическими. Таким образом, в качестве решения общей проблемы неполноты моделей формальных систем рассматриваются электронное расширение моделей формальных систем, которые содержат электронную семантику.

В решении второй проблемы возникают сложности связанные с ее фундаментальностью. Устанавливаемая в процессе исследования неполнота модели формальной системы является свойством аксиоматики, логично предположить, что решением является изменение аксиоматики. В качестве расширения аксиоматики исследования предлагается рассмотреть рекурсивное расширение. Математическая наука выступала и выступает своего рода локомотивом в научных исследованиях. Понятие рекурсии занимало особое место в математических исследованиях и это связано

прежде всего с тем что это понятие в первую очередь отражает природу бесконечности и используется в основных математических понятиях таких как число и множество. Стоит предположить, что понятие рекурсии возникло в результате попытки формализации фундаментальных свойств объектов и процессов реального мира – объект состоит из элементов, которые являются объектами и т.д. Общая ситуация в основаниях математики всецело относится и к изученности рекурсивных процессов, которые относятся к фундаментальным, поскольку отражают возможность мира, человека, формальной системы изменять себя. Рекурсивная природа явлений окружающего мира проявилась в понятии аутопоэзиса в биологии, фактах того что социальное общество само себя регулирует, а право и математическая наука по сути являются саморазвивающимися. В рамках нашей работы интерес представляет дальнейшее изучение рекурсивных процессов человеческого мышления и их электронных моделей, исследование которых невозможно без формирования рекурсивного мышления, которое в свою очередь могло бы сформироваться на основании рекурсивной парадигмы в философии.

Исследование электронной трансформации языка как формальной системы на основе анализа результатов электронной эволюции в рамках рекурсивной парадигмы является позитивным и может привести к реализации семантики в электронных моделях. Электронные модели представляют собой следующий этап развития символьных формальных систем и более мощное средство формализации, включающее семантику, что позитивно. Рекурсивная парадигма может стать основой для создания подхода к исследованиям явлений имеющих бесконечную природу, каким и является познание.

Литература:

- Гильберт Д., Бернайс П. (1979). *Основания математики*. М. «Наука».
- Дессауэр Ф. (2017). *Спор о технике: монография*. Самара. Издательство Самарской гуманитарной академии.
- Клини С.К. (2009). *Введение в метаматематику*. М. Книжный дом «ЛИБРОКОМ».
- Нестеров А. Ю. (2016). Эпистемологические и онтологические проблемы философии техники: «четвёртое царство» Ф. Дессауэра. *Онтология проектирования*. Т. 6. № 3(21). С. 377-389. DOI 10.18287/2223-9537-2016-6-3-377-389. - EDN WMXBUB.
- Смальян Р. (1981). *Теория формальных систем*. М. «Наука».
- Уайтхед А., Рассел Б. (2005). *Основания математики*. Самара. Изд-во «Самарский университет».
- Фреге Г. (2018). *Логика и логическая семантика*. М. «ЛЕНЛАНД».

Gödel K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*. Band 38. Ausgabe 1. S. 173-198.

Tarski A. (1936). The Concept of Truth in Formalized Languages *Logic, Logic, Semantics, Metamathematics*. Oxford University Press. Pp. 152–278.

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ТРАНСГУМАНИЗМ: УТОПИЧЕСКОЕ И АНТИУТОПИЧЕСКОЕ СУЖДЕНИЯ О ЧЕЛОВЕЧЕСКОЙ СМЕРТНОСТИ И ЕЕ СВЯЗИ С ТЕХНОЛОГИЯМИ

Д. Д. Подкорытов

Тюменский индустриальный университет, г. Тюмень
Denis_podkorytov@internet.ru

Современный мир вступил в новую фазу информационной эпохи, связанную с цифровизацией всех аспектов социальной жизни человека. Цифровая среда проникает в повседневные практики и области профессиональной деятельности, меняя формы участия человека. Некоторые учёные предсказывают, что эволюция человечества рано или поздно должна будет перейти от Homo Sapiens к виду более утонченному и технологическому виду – Homo Digital [Чернов, 2020]. Новый открытия, производимые в области искусственного интеллекта, всё чаще вызывают эмоциональный коктейль из страха, надежды, возбуждения и безразличия. Однажды, в 2014 году известный физик-теоретик Стивен Хокинг в интервью телеканала ВСС заявил, что разработка полноценного искусственного интеллекта может означать конец человеческой расы [Вулластон, 2014]. Однако, некоторые ученые представляют более утопическое будущее, где будет стёрты ограничения продолжительности жизни, познания и эмоции, а на их замену придут бессмертие, сверхразум, счастье и удовольствие. Несмотря на наглядное расхождение во взглядах, стоит изучить, насколько современны описания потенциала искусственного интеллекта, раскрывающего тягу человека к трансгуманизму.

В перенасыщенном технологиями мире, где человеческие привязанности к технологиям уже быстро развиваются и по иронии судьбы ограничивают реалистичные человеческие взаимодействия, сохранение того, что делает нас людьми, кажется достойной ценностью, которую мы должны лелеять и продвигать. Однако стремительный технический прогресс вкуче с предсказаниями ученых указывают на будущее, в котором разделительная линия между роботами и людьми будет стираться, особенно по мере того, как люди начнут жить в роботах или они в нас.

Очевидно, что для некоторых компромисс между человеческим творчеством, духом и красотой и искусственно созданным интеллектом выглядит как плохая сделка и может привести к нашей гибели. Однако некоторые сторонники искусственного интеллекта не воспринимают это даже как компромисс с нашими идеалами или даже с тем, что значит быть человеком. Все разочарование ограничивается людьми, представляемых мир, населенный разумными машинами и людьми, покинувшими свои тела,

которые из-за технологического процесса будут обладать бессмертием и превосходными умственными способностями.

Рэй Курцвейл широко известный оптимист в области искусственного интеллекта, предсказывает, что в ближайшем будущем большинство всех болезней будет искоренено поскольку нанороботы оказываются высокоэффективными средствами здравоохранения, отличными от традиционных медицинских процедур. Также мы можем экспоненциально увеличить наш интеллект объединив нашу биологическую кору головного мозга с синтетическим веществом, которое в некоторой степени сможет загружать и обрабатывать информацию, взятую из облака. Человеческому мозгу сложно усваивать новые вещи, и ему трудно передавать эти знания дальше, однако, с помощью искусственного интеллекта эти ограничения могут быть разрушены и интеллект начнет принимать больше небιологическую форму. Эти улучшенные условия существенно подчеркиваются в перспективах трансгуманизма, который является условием, при котором люди будут обладать экстраординарными способностями и превзойдут присущий нам потенциал.

Драматизация антиутопического будущего можно увидеть в фантастических романах и фильмах, которые явным языком говорят о человеческой тревоги в этой области. В этих картинах изображена технологическая сложность сверхчеловеческих киборгов. Роботы спроектированные с учетом высокоразвитых чувств и интеллекта, чтобы выполнять единственную задачу защиты людей, в соответствии с тремя законами робототехники Айзека Азимова:

1. Робот не может причинить вред человеку или своим бездействием допустить, чтобы человеку был причинен вред.
2. Робот должен подчиняться приказам, отдаваемых ему людьми, за исключением, когда есть противоречие 1 закону.
3. Робот должен защищать своё собственное существование до тех пор, пока защита не противоречит 1 и 2 закону [Айзек Азимов, 1950].

Несмотря на то, что они кажутся четко определенными руководящими принципами, высокоинтеллектуальные роботы получают гораздо более глубокое понимание своей задачи и руководящих правил. Увидев разрушения, вызванные людьми в мире, роботы решают, что мы, скорее всего, будем в безопасности, если роботы возьмут управление на себя, поскольку люди склонны к ошибкам и иррациональности, которые могут нанести ущерб человечеству.

В отличие от этого мрачного будущего, изображаемого в антиутопической литературе, некоторые утописты верят, что с появлением полноценного искусственного интеллекта панацея от потребностей и ограничений человечества неизбежна. Пирамида потребностей Абрахама Маслоу рухнет, превратившись в плоский ландшафт постоянного комфорта и

процветания. С появлением высокоинтеллектуальных роботов-протезов людям со сломанными конечностями просто придется думать на более высоком уровне о том, что они хотят делать, и рецепторы этих роботов будут выполнять их запросы. Вместе с передовыми биотехнологическими усовершенствованиями смерть человеческой плоти ни в коем случае не будет самоцелью, а будет средством, с помощью которого мы достигнем бессмертия и превзойдем самих себя. Другие ученые, такие как Рэй Курцвейл, считают, что искусственный интеллект предоставляет нам выбор между ограничением или свободой. Утопическая перспектива, вызванная быстрым развитием сверхчеловеческого интеллекта, подпитала движение, известное как трансгуманизм.

Трансгуманизм, термин, обозначающий выход за пределы возможного для человека, постепенно прокладывает себе путь в будущее человечества благодаря недавним разработкам в области искусственного интеллекта. Однако стоит упомянуть, что в последнее время технологии претерпели изменения. В начале 20-ого века технологии стремились служить гуманистической цели, которая заключалась в том, чтобы сделать мир лучше и комфортнее для людей. Однако, пройдя через несколько лет социально-экономических сдвигов, технология постепенно взрастила свою трансгуманистическую цель – сделать люди лучше для всего мира. Для верующих в трансгуманистическую силу технологий человечество будет спасено божественной силой технологий, дарующей людям благословение бессмертия и сверхчеловеческих способностей. Для тех, кто выступает против гонки к сингулярности, которая представляет собой футуристическое состояние, в котором общество радикально меняется благодаря развитию сверхчеловеческого интеллекта, природа смерти и знание того, что жизнь коротка и эфемерна, это действительно то, что делает жизнь стоящей того, чтобы жить. Без определенного конца для всего и вся, все вещи могут потерять их значение. Большинство антиутопических прогнозов о сингулярности указывают на то, что переход от человека к человеческому дроиду, скорее всего, будет постепенным процессом. Они указывают на возможность социального неравенства, предрассудков и конфликтов между людьми и постлюдьми. Как только немногие люди достигают сверхума, они могут эгоистично стремиться застопорить трансформацию других, чтобы править и процветать. Наличие черт, которые превосходят то, что может сделать любой нынешний человек, особенно давнее желание человечества, и теперь технология может предоставить нам эту возможность [Луков, 2017].

На данный момент, изучив через призму научной литературы и киносредств возможные антиутопические и утопические результаты искусственного интеллекта, которые, в свою очередь, выявили соответствующие вопросы, присущие нашему пути, или нашей гонке в будущее, признает двойственную природу технологий. Раскрывая присущее

человеку стремление к трансцендентной жизни, полной блаженства и утопии, и в то же время признавая тревожную природу наших человеческих желаний, выдвигаются ряд соображений, связанных с принятием решений, которые должны вызвать критическое осмысление и дискуссия о будущем человеческой расы.

Литература

Айзек Азимов. (1950). *Я – робот*. Америка. Gnome Press.

Вулластон В. (2014). 'Artificial intelligence could spell the end of the human race': Stephen Hawking warns a robotic uprising will spiral out of our control. *MailOnline* [Электронный ресурс]. URL: <https://www.dailymail.co.uk/sciencetech/article-2857658/Artificial-intelligence-spell-end-human-race-Stephen-Hawking-warns-robotic-uprising-spiral-ontrol.html> (дата обращения: 15.09.2022).

Луков, В. А. (2017). Трансгуманизм. *Энциклопедия гуманитарных наук*. № 1. С. 245–252. DOI: 10.17805/zpu.2017.1.20

Чернов, И. В. (2020). Цифровизация как тенденция развития современного общества: специфика научного дискурса. *Гуманитарий Юга России*. Т. 10. № 1. С. 121–131.

КАК ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ ДЕМИСТИФИЦИРУЕТ ПОНЯТИЕ «ТВОРЧЕСТВО»

И.К. Ставровский

Институт философии НАН Беларуси, г. Минск
tutoriks@gmail.com

Многие склонны считать, что творчество – это одна из тех способностей человека, которая недоступна машине. Ведь одно дело слепо выполнять алгоритм, а совсем другое – создавать нечто принципиально новое и даже уникальное. Таким образом, изначально предполагается, что творчество не поддается алгоритмизации. Действительно, для создания чего-то нового может потребоваться наличие специальных знаний и навыков. Возможно даже есть некие правила, советы, рекомендации и т. д. Однако ничто из перечисленного само по себе еще не гарантирует успех, так как в творческом процессе неизбежно сохраняется нечто, нередуцируемое ни к сколь угодно высокому уровню мастерства, ни к способности воспроизводить сколь угодно сложный набор правил.

Данная позиция наивна и крайне уязвима для критики. Вполне очевидно, что здесь не хватает четкого и однозначного критерия для определения того, что является творчеством, а что нет. Кажется, что проблема решается более-менее произвольно на основе массы непроясненных посылок.

Скажем, если мы возьмем за точку отсчета понятия «новизны» или «оригинальности», которые часто используют для прояснение понятия «творчество», то мы никуда не продвинемся. Абсолютная оригинальность невозможна хотя бы потому, что любое творчество имеет ряд ограничений. К примеру, писатель подчиняется правилам языка, художник может использовать конечное количество цветов, а музыкант ограничен слышимым диапазоном звуков. Если же творец учитывает правила гармонии, композиции и т. д., то ограничений становится еще больше. В каком-то смысле произвольный набор звуков, символов или цветовых пятен может считаться даже более оригинальным. Но вряд ли многие назовут подобное творчеством.

Можно было бы попытаться связать понятие «творчество» с понятием «спонтанности», рассматривая его как необусловленное ничем иным, кроме самого себя. Предполагается, что творец создает нечто только ради самого процесса, не думая о славе, деньгах и т. д. Однако здесь мы имеем дело со своеобразной генетической ошибкой. По сути, утверждается, что произведение (независимо от его качества) обесценивается, если творец творил не ради творчества как такового. Подобное увязывание творчества с мотивацией творца приводит к парадоксальной ситуации, где мы оцениваем

не произведение, а того, кто его создал. Следовательно, даже самую лучшую картину мы должны объявить «неподлинным творчеством», если станет известно, что художник нарисовал ее ради денег. При этом если данный факт никогда не вскроется, то мы будем дальше считать картину шедевром. Такой подход выглядит крайне неудовлетворительным.

Теперь рассмотрим, как понятие творчества концептуализируется в рамках дискуссий об искусственном интеллекте. В 2003 году трио ученых Сельмер Брингсйорд, Пол Белло и Дэвид Ферруччи в качестве альтернативы тесту Тьюринга предложили оригинальный способ оценки разумности искусственного интеллекта, назвав его тестом Лавлейс [Bringsjord, Bello, Ferrucci, 2003]. Выглядит он следующим образом:

Человек **H** разработал искусственный интеллект **A**, создающий рассказы. Если **A** создает рассказ **o**, который не является ошибкой, и **H** не может объяснить, как **o** был создан, то **A** следует рассматривать как способный к творчеству. Важно, чтобы в распоряжении **A** была вся возможная информация об **H**. Успешным объяснением работы **A** будет представление четкого алгоритма его работы.

Однако проблема теста Лавлейс в том, что он содержит внутреннее противоречие. Полное знание о том, как устроена система уже подразумевает наличие знания о том, что и как она делает. Иными словами, решение данной задачи является вопросом времени и ресурсов. А наличие временных ограничений является произвольным и никак необоснованным условием.

Другой ученый Марк О. Ридл предложил улучшенную версию, тест Лавлейс 2.0 [Riedl, 2014]. Формулируется он следующим образом:

A создает артефакт **o** типа **t**. При этом **o** отвечает множеству условий **C**, где каждое условие $c_i \in C$ выразимо на естественном языке. **H** выбирает **t** и **C**, а затем оценивает, насколько созданный **o** им соответствует. Рефери **R** проверяет, чтобы сочетание **t** и **C** было возможным. Способность **A** создать **o**, отвечающее требованиям **t** и **C** является сильным индикатором интеллекта.

К сожалению, вторая версия теста решает проблемы первой лишь частично. Во-первых, мы должны довериться оценкам **H** и **R**. Во-вторых, мы вновь сталкиваемся с произвольным определением понятия «творчества», а именно: способность **A** создать **o**.

На этом этапе анализа становится очевидной ключевая проблема. Представление о неспособности машины к творчеству полностью основано непрямоугольности базового понятия. Пока последнее представляется загадочным или даже мистическим, то можно с полной уверенностью утверждать, что искусственный интеллект никогда не сможет заниматься творчеством. С другой стороны, попытка формализовать понятие «творчество» приводит нас к произвольным определениям и наборам критериев. Как отмечал специалист по психологии творчества Я. А.

Пономарев, сама идея поиска некоего алгоритма творчества противоречит смыслу этого понятия [Ильин, 2008, с. 22].

В действительности данная дискуссия неизбежно приводит к демистификации понятия «творчество». Без четкой дефиниции данное понятие не имеет смысла, так как мы даже не можем выделить объект его референции. Следовательно, осмысленное обсуждение становится невозможным. Введение произвольных критериев хотя само по себе неудовлетворительно, демонстрирует важную закономерность: однозначность в определении позволяет сразу перевести диалог в область создания алгоритма творчества. Проблема становится чисто технической.

Подобный подход вызывает недовольство у многих людей, так как даже постановка такого вопроса подразумевает устранение ореола вокруг понятия «творчества», а вместе с ним – исключительного онтологического статуса человеческого разума. А любые попытки сделать нечто подобное порождают своеобразную защитную реакцию. Ранее она проявлялась в обесценивании когнитивных способностей, которыми успешно овладевает искусственный интеллект. Например, в древности хорошая память считалась признаком высокого ума. Однако с появлением компьютеров она все более рассматривается как нечто механическое. Теперь развитие искусственного интеллекта может пошатнуть полумистический образ творчества. Это угрожает представлению об исключительности человека в гораздо большей степени, что порождает еще более сильную защитную реакцию.

Тем не менее описанная реакция не основана на логике и фактах, а скорее отражает иррациональное желание сохранить особый статус человека в онтологии. Вместо этого нам следует пересмотреть свою картину мира и принять тот факт, что человеческое сознание вообще и наша способность к творчеству в частности являются интересными, но не исключительными частями мира.

Литература

Ильин, Е. П. (2008). *Психология творчества, креативности, одаренности*. СПб. Питер.

Bringsjord, S., Bello, P., Ferrucci, D. A. (2003). Creativity, the Turing Test, and the (Better) Lovelace Test. *The Turing Test. The Elusive Standard of Artificial Intelligence*. Dordrecht. Springer Science+Business Media. Pp. 215–239.

Riedl, M. O. (2014). The Lovelace 2.0 Test of Artificial Creativity and Intelligence. *arXiv* [Электронный ресурс]. URL: <https://arxiv.org/pdf/1410.6142.pdf/> (дата обращения: 07.08.2022).