

УДК 167.1+ 004.8+004.05

DOI: 10.15372/PS20220308

Д.В. Винник**О ВОЗМОЖНОСТИ ГИБРИДНОГО
ИСКУССТВЕННОГО РАЗУМА***

В статье рассматривается сильный искусственный интеллект как возможность создания искусственной личности. Утверждается, что радикальный коннекционизм ложен и неприемлем для этих целей. Мозг имеет гибридную вычислительную природу: помимо механизмов самообучения, он содержит алгоритмические процедуры. Выдвигается эволюционный аргумент против эффективности нейросетевого моделирования. Также обсуждается вопрос о необходимости верификации гипотезы квантового мозга для прояснения физических ограничений моделирования. Ставится вопрос о приемлемости феноменологического знания. Утверждается, что тринитарная архитектура личности Аристотеля отражает фундаментальную структуру личности на феноменальном уровне. Предлагается трактовка эмоций как аналоговой формы репрезентации. Делается вывод, что если искусственная личность как адекватная модель может быть построена, то она может быть создана как гибридная мультиагентная система. Отдельные модули и целые уровни могут быть аналоговыми, другие – цифровыми; одни системы могут самообучаться, другие – функционировать согласно алгоритмическим правилам и конкретным математическим функциям.

Ключевые слова: искусственный интеллект; коннекционизм; когнитивные функции; эмерджентизм; перцептрон; эмоции; разум; рефлексия; интенциональность; самосознание; сознание

D.V. Vinnik**ON THE POSSIBILITY OF HYBRID ARTIFICIAL
INTELLIGENCE**

An article considers strong artificial intelligence as an opportunity to create an artificial personality. It is argued that radical connectionism is false and unacceptable for these purposes. The brain has a hybrid computational nature – in addition to self-learning circuits, it contains algorithmic procedures. An evolutionary argument is put forward against the effectiveness of neural network modeling. The question is raised about the need to verify the quantum brain hypothesis in order to clarify the physical limitations of model-

* Статья публикуется в авторской редакции.

ing. The question is raised about the acceptability of phenomenological knowledge. It is argued that Aristotle's trinitarian personality architecture reflects the fundamental structure of personality at a phenomenal level. The interpretation of emotions as an analogue form of representation is proposed. It is concluded that if an artificial personality as an adequate model can be built, then it can be created as a hybrid multi-agent system. Individual modules and entire levels may be analog, others digital; some systems can self-learn, others can function according to algorithmic rules and specific mathematical functions.

Keywords: artificial intelligence; connectionism; cognitive functions; cognitivism; emergentism; perceptron; emotions; mind; reflection; intentionality; self-awareness; consciousness.

Об идее искусственной личности

Вокруг понятия сильного искусственного интеллекта имеет место множество споров. В самом деле, сам вопрос о том, возможно ли техническими средствами можно создать полноценный искусственный разум, вторгается в парадоксальную область оснований теории вычислений и в сферы подлинно метафизические. Пожалуй, наиболее содержательно этот вопрос изложен в книге В.В. Целищева «Алгоритмизация мышления: Геделевский аргумент» [10]. Если отстраниться от логико-философской и мета-математической проблематики в сторону более прагматической, то понятие искусственного интеллекта можно осмыслить как возможность создания искусственной личности, обладающей соответствующими квазипсихическим атрибутами: аналогами способности к ощущениям, экспрессивному оцениванию (эмоциям), суждениям. Если в этот список включить аналоги самосознания и творческого мышления, то можно говорить и о сильном искусственном интеллекте. Это означает, что понятия сильного искина и искусственной личности можно рассматривать как тождественные. Споры о возможности создании полноценной искусственной личности ведутся с привлечением аргументов из самых разнообразных дисциплин, начиная от respectable эволюционной биологии [25] и заканчивая самой вульгарной метафизикой экстерналистского извода.

Несмотря на ажиотаж вокруг тематики искина, оптимизм относительно скорых перспектив его создания питается многочисленными спекуляциями. Создание человекоподобного робота, обладающей подлинным самосознанием и, следовательно, – личностью (как той сущности, что обладает этим свойством), в обозримой перспективе не следует рассматривать всерьез, даже если разделять концепцию «механизма» или функционализм как его наиболее из-

вестную форму [10, с. 9]. Тех, кто отрицает принципиальную возможность создания или даже эмерджентности разума (спонтанного зарождения нового качества, как результата эволюционного скачка сложности субстрата) в сложных кибернетических системах, как правило относят к сторонникам слабого искина. Последние обычно допускают воспроизводство отчужденных когнитивных функций без проникновения в сущность того, что мы считаем разумностью и человечностью.

О некоторых физических ограничениях моделирования мозга

Следует отметить, что крайне редко в полемике о возможности сильного искина обсуждается проблематика физических ограничений, накладываемых нейронной архитектурой мозга на реализацию тех или иных свойств, которые обычно трактуются как вычислительные. Отметим, что радикальные философские коннекционисты, например, Дж. Серль отрицают вычислительную природу мозга в том смысле, что самообучение нейронной сети не подпадает под строгое мета-математическое понятие вычисления. Схожей точки зрения придерживается и Р. Пенроуз, трактующий сознание как тонкий физический процесс квантово-механической природы о котором в терминах алгоритмов говорить не приходится. К настоящему времени в научной литературе по биофизике нервной деятельности накопился ряд нерешённых вопросов о функциональном значении некоторых специфических феноменов. Это такие феномены, как:

- 1) «эфаптическая» или «несинаптическая», т.е. полевая передачи информации [12, р. 108];
- 2) квантовое туннелирование на клеточных мембранах [24];
- 3) устойчивая квантовая когерентность биомолекул [18];
- 4) нейронные лавины как механизма амплификации квантовых эффектов [13] до макроуровня и, в конечном счете, – до поведения [21].

Почти все перечисленные феномены были предсказаны теоретически и установлены экспериментально. В самом деле, существование нейронных лавин очевидным образом следует из теории хаоса (синергетики). В нейронной сети неизбежно (даже чисто статисти-

чески) должны иметь место отдельные нейроны, статическое или динамическое положение которых таково, что его возбуждение или торможение способно запускать каскад подобных процессов в миллионах и даже миллиардах других нейронов. Вопрос о квантовом туннелировании на клеточных мембранах напрашивается исходя из наших знаний о работе ионных насосов и самой размерности процессов. К слову, размеры синаптических щелей настолько велики, что туннелирование на них не может быть частым событием, чтобы иметь функциональное значение. Идея о квантовой когерентности биомолекул является органичной для теоретической физико-химии. Несколько особняком стоит феномен несинаптической передачи, который был экспериментально открыт еще в середине XIX века, но и её природа вполне прозрачна с точки зрения основ электродинамики.

В отношении каждого из этих феноменов разумно задать вопрос, имеют ли они значение для работы мозга? Если ответы хотя бы на 2-й и 4-й вопрос будут позитивными, это будет означать, что гипотеза квантового мозга истинна. Если это так, что теории параллельных вычислений, являющейся математическим основанием коннекционизма, для полноты объяснения работы естественных нейронных сетей недостаточно. Из этого, в свою очередь, должно следовать, что функционализм ложен. Если квантовые эффекты неустранимо встроены в причинно-следственные связи нормальной работы мозга, нельзя исключать, что феноменальные свойства сознания действительно являются не алгоритмическим результатом или даже просто странным эпифеноменом, а неким вполне конкретным квантовым макро-феноменом функционирования нервной ткани (по аналогии с макро-феноменами сверхтекучести, сверхпроводимости и когеренции) [5].

Эволюционные ограничения нейросетевого моделирования

Допустим, что коннекционизм и эмерджентизм истинны в самых сильных формулировках. Это должно означать, что усложняющаяся искусственная вычислительная сеть может преодолеть порог сложности и стать сопоставимой (по количеству вычислительных узлов и связей между ними) с мозгом живого существа. Достаточное ли подобным количественным образом понимаемая

сложность для спонтанного для внешнего наблюдателя проявления (эмерджентности) нового диалектического качества, которую мы именуем сознанием или ментальностью? Классическая философская традиция склоняют понимать сознание именно как атрибут такой крайне специфической формы природного единства, как личность, а не как относительно автономную сущность любых сложных вычислительных систем. Можно задать и еще более общий вопрос. В самом деле, есть ли в универсуме форм сознаний сознание как свойство некоей целостности, которое при этом не является личностью? Этот вопрос настолько же интересен, насколько метафизичен. Основатель весьма экзотической и даже архаичной для XX века трансцендентальной феноменологии Эдмунд Гуссерль, очевидно, допускал подобную формы сознания. В одном из фрагментов своих «Идей...» он утверждает, что, «безусловно мыслимо и бесплотное, а также, сколь парадоксально это бы не звучало, и бездушное, не одушевляющее человеческую телесность сознание, т.е. такой поток переживания, в котором не конституировались бы интенциональные единства опытного постижения – тело, душа, эмпирический «я»-субъект. Для всех этих понятий опыта «не было бы никакой опоры, они бы были лишены всяческой значимости» [7, с.123].

Даже если между естественным и искусственным типами интеллекта нет различия на онтологическом уровне, это не означает, что воспроизводство техническими средствами такой сокровенной сущности как разум, есть задача конструктивного типа. Очевидно, что любая конкретная психика, человека или животного, есть результат эволюции нервной системы позвоночных в течение десятков и даже сотен миллионов лет, если смотреть на эволюцию с точки зрения больших временных масштабов. Любую конкретную психику, т.е. любую личность (в отношении психики животных также есть смысл применять понятие личности) с этой точки зрения следует рассматривать как некоторую комбинацию психических свойств в самом широком смысле, – начиная от скорости реакции на раздражители, обусловленные конституцией нервной системы и заканчивая развитостью высших психических функций, которые ранее было принято именовать когнитивными. К сожалению, на сегодняшний день вследствие увлечения «когнитивными науками» понятие когнитивного оторвалось от своего латинского корня, претерпело значительную инфляцию и означает просто любую психическую функцию. Даже простые ощущения трактуются как когнитивные функции. Каж-

дая комбинация психических свойств является результатом настройки параметров торможения/возбуждения миллиардов нейронов живых существ в триллионах итераций естественного отбора (жизнях особей, участвующих в эволюции). Если обобщить этот взгляд до категоричного уровня форм сознания, то можно утверждать, что известные нам формы сознания есть всего лишь одна из комбинаций в многомерном универсуме психических способностей.

Исключительно сомнительно, что в обозримом будущем человеческие вычислительные ресурсы будут способны получить (эмпирическим или полу-эмпирическим путем) синаптическую матрицу, необходимую для самой простой модели самой примитивной личности. Технические средства человечества к моделированию эволюции весьма далеки от объекта моделирования по степени аппроксимации. Особенно это касается высших форм организации нервной системы. Не следует переоценивать успехи в техническом воспроизводстве отдельных психических функций, вроде распознавания образов. Универсум психических функций весьма велик, они значительно варьируются как по форме, так и по сложности реализации. Следует обратить внимание на очень простую количественную оценку, – дистанция между количеством искусственных нейронов в существующих нейросетях и количеством нейронов в самых примитивных мозгах составляет много порядков. Какое количество нейронов следует признать минимально необходимым для существования самой примитивной личности? Знания зоологии и самое оптимистическое менталистское толкование зоопсихологии, приписывающее существование личностей, например, у насекомых, склоняют к оценке в шестой порядок. Иными словами, самой простой личности нужны миллионы нейронов. Именно таков порядок ганглий у насекомых. Однако, не стоит забывать, что для искусственных нейронов это число должно быть значительно большим, поскольку нейрон такого типа настолько просто, что обладает только одним типом состояния (вес от 0 до 1, моделирующий спайк). Это состояние есть результат простого интегрирования подобных же весов от соседних нейронов. Живые нейроны характеризуются множеством физико-химических свойств, которые обуславливают спайк. Однако, и по содержанию самого спайка далеко нет полного согласия: какова его роль, – аналоговая или цифровая. Если цифровая, то внутренняя структура спайка как формы сигнала не имеет значения. Если аналоговая, то это не так. Каким типом машины является мозг:

цифровой или аналоговой? На этот вопрос нет окончательного ответа, но пока перевес на стороне цифровой гипотезы [26, p.1].

О ложности радикального коннекционизма

Если встать на радикальные позиции коннекционизма, вдохновленного потрясающими успехами перцептроники, то следует признать, что всякие реальные представления об иерархии психических функций совершенно бесполезны, поскольку, с точки зрения этой психофизической теории, структура информации на выходе любой нейронной сети, (в данном случае, – поведение и иные ментальные феномены), есть не более чем результат аппроксимации синаптических параметров при взаимодействии с внешней средой. Эта структура никоим образом не отображает внутреннюю структуру самой вычислительной сети. Иными словами, знание о способе представления информации на выходе не пригодно для создания модели самообучающейся вычислительной системы. Информация может иметь понятный нам смысл только на входе и на выходе, внутри системы может быть только статистический хаос миллионов аппроксимирующихся параметров.

Несмотря на острую полемику, сегодня ясно, что даже для более-менее правдоподобного объяснения работы мозга радикальный коннекционизм не годится, поскольку эмпирические данные говорят о том, что мозг есть гибридная система, которая, помимо механизмов самообучения, содержит алгоритмические процедуры. Имеет значение и сама архитектура мозга, которая содержит алгоритмические правила на «железном» уровне, реализуемая разными слоями. Например, подобным образом устроена наша зрительная кора, которую можно описать как многослойный перцептрон. Известно, что каждый слой исполняет отдельную функцию. Эти функции экспериментально установлены. Самые нижние слои распознают линии и углы: т.н. геометрические примитивы. Более высокие слои ответственны за создание геометрических фигур: овалов, прямоугольников и треугольников. На самых верхних уровнях распознаются сложные эстетические объекты и человеческие лица. Обратим внимание на любопытный факт ригидности высших уровней, – они гораздо устойчивей к искажениям и реакции их нейронов менее зависят от нюансов перспективы, реагируя на более широкую апертуру. Именно бионическое моделирование архитектуры зрительной коры

привело к успеху К. Фукусиму. В 1980 году он нейросеть для распознавания образов на основе конкурентного обучения, получившую название «когнитрон» [19].

Следует отметить, что исследования механизмов зрения оказались очень продуктивными, они привнесли множество аргументов против коннекционистов в пользу их противников, известных как «когнитивисты» или сторонники алгоритмической трактовки природы ментальных состояний. П. Бреслов и Дж. Кован обратили внимание на типичность визуальных геометрических галлюцинаций и сделали выявление их нейрофизиологического механизма предметом своей работы. Если надавить на закрытые глаза или подвергнуться воздействию психофармакологических средств, то можно увидеть совершенно характерные для всех людей в таких обстоятельствах геометрические фигуры. Это многогранники, логарифмические спирали, матричные текстуры, решетки и туннели. Их изображения приводятся в работе. Эти геометрически структуры вполне просты, их математическая форма известна давно и не составляет никакой тайны. Авторы исследования настаивают, что им удалось обнаружить в зрительной коре те устойчивые источники сигналов, которые генерируют эти фосфены и галлюцинозные изображения. Что характерно, эти сигналы также имеют вполне понятную математическую форму, имеющую много общего с теми математическими функциями, которые способны «рисовать» подобные структуры. Очевидно, что эти сигналы работают в фоновом режиме, прописывая каркас и текстуру нашего зрительного поля, которые в нормальных психофизических условиях не доступны сознанию. Только при специфических обстоятельствах нарушения работы мозга этот скрытый «технический» уровень синтеза зрительных образов становится доступен сознанию [14, p. 299].

В свете сказанного обратим внимание на уникальный случай американца Джейсона Паджетта, пережившего тяжелейшее сотрясение мозга. После контузии сама структура зрительного восприятия Паджетта изменилось самым радикальным образом, – образы объектов распались на отдельные визуальные фрагменты; только движущиеся предметы сохраняли относительную целостность и позволяли и ориентироваться в окружающем пространстве. Крайне любопытно, что окружности Д. Паджетт воспринимает исключительно в виде многоугольников с большим количеством граней. Это вполне соответствует известным представлениям евклидовой гео-

метрии о природе окружности. Подобное восприятие носит для него тотальный характер: струи воды, облака, лужи, радуга – все это предстает мозаикой из различных пикселей. Особенно сложные и завораживающие структуры этот человек видит на границах раздела сред. Наиболее привычным подобным примером для него являются края облаков, подсвеченные солнцем. Тридцать лет назад Паджетт начал рисовать эти структуры с помощью линейки и циркуля. Как стало ясно впоследствии, часто он изображал вполне известные фрактальные структуры. Б. Брогард выдвинул гипотезу, что в результате механического повреждения зрительной коры сознанию Паджетта стали доступны процессы обработки зрительной информации нижних слоев этой коры. Крайне правдоподобно объяснение, согласно которому он стал созерцать некую опорную структуру своего зрительного поля. Что в данном случае имеет значение, – это то, что она оказалась не хаотичной, напротив – высокоупорядоченной и даже вписывающейся в понятные нам математические представления [15, p.19].

Применимость феноменологических представлений

Тех, кто отрицает принципиальную возможность создания робота с подлинным самосознанием или отказывает возможности эмерджентности разума в сложных кибернетических системах, фактически отрицают возможность создания сильного искусственного интеллекта. Сторонники слабого искусственного интеллекта обычно допускают отчуждение когнитивных функций на внешние технические устройства, т.е. их воспроизводство или моделирование. Это не одно и то же. Когда способность к арифметическому счету реализуется на калькуляторе, то это следует считать воспроизводством, поскольку калькулятор справляется с функцией счета ничуть не хуже и даже лучше человека. В этом смысле, калькулятор является сортом искусственного интеллекта, только очень узким – воспроизводящим одну единственную умственную способность. Когда речь идет о способности формальной системы прогнозировать поведение конкретного человека, имитировать стиль его письма и т.п., корректнее говорить о модели. Если реализации когнитивных функций и модели возможны, означает ли это, что

Возможно моделирование личности как некой целостности известных когнитивных функций? А.Ю. Алексеев выделяет четыре разных понятия искусственной личности: 1) имитация; 2) модель;

3) репродукция естественной человеческой личности; 4) креация – создание «сверхличности» [2, с. 157]. Относительно модельного подхода, который автор называет «когнитивно-модульным», компьютерная система должна включать в свой состав блок «псевдосознания» [2, с. 159]. Действительно, именно создание модели личности как действующего физического аналога можно трактовать как целостность, максимальным образом приближающуюся к тому, что мы понимаем под личностью природной.

Моделирование личности должно опираться на некие знания о структуре той сущности, которую мы называем личностью. Какими знаниями мы обладаем? Это некие знания феноменологического характера, почерпнутые из интроспекции и рефлексии, а также особенности поведения. Важно иметь в виду, что знание о формах поведения вторично, поскольку оно не содержит явной информации об ощущениях и высших когнитивных функциях. *Самым важным в этом свете представляется вопрос, является ли феноменологическое знание об иерархии когнитивных функций приемлемым для конструирования искусственной личности?* Если да, то является ли оно необходимым?

Следует признать, что эти знания в этой области весьма скудные и относительные. В пределах одной только психологии (как феноменологической дисциплины) существует огромное количество классификаций ментальных состояний. Нет особого смысла анализировать все из них. Большинство являются модификациями и расширениями неких предельно общих фундаментальных способов деления психофизических понятий. Большинство из них, за исключением самых экзотических и маргинальных, в той или иной степени восходят к аристотелевской классификации психических способностей по условным частям души.

О пригодности психической архитектуры Аристотеля

Как известно, Аристотель он различил три вида души (растительную, животную и разумную) с присущими им способностями. При этом их можно поделить на два класса: смертные виды (растительная и животная) и бессмертный вид (разумный). В человеческом существе все три вида души присутствуют как условные части единой души, являющейся энтелехий тела [4, с. 435]. Части условны, поскольку душа является неделимой сущностью. С современной

точки зрения это скорее когнитивные аспекты (способности или функции) личности, филогенетически доставшиеся в наследие от животных и растений. Функции растительной части души – питание и размножение. Функции животной части – ощущения и эмоции, стремления и моторика. Функция разумной – мышление.

Разумная часть души способна отделяться от тела после его смерти и предаваться единственному виду деятельности, к которому предназначена – мышлению. А поскольку связь с эмпирической реальностью разорвана вместе с гибелью тела, разумная часть души будет занята мышлением о мышлении, т.е. бесконечной рефлексией. Таким образом в пределе она сольется с божественным сверхинтеллектом, – аристотелевским «нусом». Поскольку разумная часть души не «смешана» с материальным миром, с внешней средой, то все воспоминания о пребывании в теле вскоре будут забыты, как несущественные. Иными словами, разумная часть души трактуется как некая автономная метаструктура, делающая человека деятельным разумным существом. Строго говоря, растительные функции души нельзя отнести к психическим в современном смысле слова, – в них нет места феноменальному опыту или сознанию. Вот как об этом писал В. Асмус: «Между двумя крайними звеньями – душой разумной и душой неразумной – Аристотель помещает – как среднее звено – способность ощущения. Он причисляет ее то к разумному, то к неразумному началу... Аристотель не отводит особого, отдельного места способности организмов ощущать удовольствие и страдание... В аристотелевском разборе состава психических явлений единство основного принципа деления не нарушено: все чувствования Аристотель считал одним моментом развития второй силы души – животной. Там, где имеют место ощущения, пояснял он, является удовольствие и страдание, а где есть удовольствие и страдание, необходимо возникает и желание» [3, с. 60]. Аристотелевская тринитарная структура души в несколько видоизмененном виде сохранилась в общей психологической теории до сих пор, – она включает: сенсорно-перцептивный, эмоционально-экспрессивный и сознательно-волевой уровни психической организации. Это знание прошло проверку временем и может претендовать на объективность.

А.Ю. Алексеев, судя по всему, отвечает на поставленный вопрос о применимости знания об иерархии психических функций положительно. Он предлагает «типовую» трехуровневую архитектуру когнитивно-компьютерной системы: 1) уровень коннекциони-

стских образов (паттернов), осуществляющий перцептивную обработку данных; 2) уровень первичных репрезентаций, переводящий восприятия в дискретные представления и суждения; 3) уровень вторичных репрезентаций, на котором осуществляется представление представлений (моделирование других моделей представления знаний и моделирование собственной модели) [1, с. 163]. По сути предлагается вполне узнаваемая тринитарная структура из перцептивного, первичного обобщающего и оценочного, а также рефлексивного (мета-психологического) уровней психической организации. Моделирование первого уровня средствами самообучающейся перцептроники идет очень успешно. Моделирование второго уровня более проблематично, но инструментарии формализации и представления знания также вполне успешны. Эмоции можно рассматривать как первичные аналоговые репрезентации. Моделирование метамодельного уровня, воспроизводства аналога рефлексии наталкивается на серьезные семантические и метаматематические проблемы.

Аналоговая форма эмоций

В течение веков не утихает полемика по поводу границ между тремя уровнями психической организации. Особенно полемичен статус эмоциональных состояний. Одни философы их относят их к перцептивному уровню, другие к уровню самосознания. В качестве общего признака для подобного рода классификации выступает их логическая форма. Первые обращают внимание на то, что перцептивные и эмоциональные состояния являются качественными. Это значит, что их содержание выражается с помощью предикатов первого порядка. Например, лиловость и грусть суть простые качества. Вторые настаивают, что эмоции обычно явлены нам не как первичные чувственные данные, но вмонтированы в страхи и желания, т.е. в состояния интенциональные по своему типу. Логическая форма интенциональных состояний такова, что их содержание не может быть выражено одноместными предикатами, – это высказывания об отношении к конкретным объектам, так и к различному содержанию (выражающемуся в отдельном суждении). Важно, что отношение может различаться модальностями: как простым поведенческими интенциями (стремление и избегание как корреляты желания и страха), так и более сложными эпистемическими (вера,

убеждение, сомнение, знание). Известно, что такие базовые эмоции, как удовольствие и неудовольствие сопровождают состояния желания и страха и, обычно напрямую отождествляются с ними.

Этот взгляд на природу эмоций восходит Д. Юму, который сводил классы ментальных состояний к двум: впечатлениям (т.е. ощущениям) и идеям. Эмоции Юм причислял к содержанию рефлексивных актов: «идея удовольствия или страдания, возвращаясь в душу, производит новые впечатления – желание и отвращение, надежду и страх, которые, собственно, могут быть названы *впечатлениями рефлексии*, так как извлечены из последней» [11, с. 64]. Юм утверждал, что исследование наших ощущений касается скорее анатомов и естественников, чем моралистов, а «аффекты, желания и эмоции возникают по большей части из идей» [11, с. 65].

Феноменологическая гетерогенность ментальных состояний

Указанная разница между качественным и интенциональным послужила причиной выдвижения Я. Кимом так называемой *дизъюнктивной проблемы* природы ментального: «Благодаря чему наши боли и убеждения имеют нечто общее, что они попали под одну категорию “ментальных феноменов”? Они, конечно же, удовлетворяют дизъюнктивному свойству “качественное или интенциональное”, но это подобно попытке обнаружить общность между красным и круглым, аргументируя, что обе вещи, как красная, так и круглая удовлетворяют критерию “красный или круглый”» [20, р. 23].

В качестве решения проблемы был выдвинут критерий «эпистемической точки зрения от первого лица». Такой точки зрения придерживался Р. Рорти: «Мой ответ на вопрос – “Почему мы склонны сваливать в одну кучу интенциональное и феноменальное?” – заключается в том, что Декарт использовал для заполнения пропасти между ними понятие “непоправимо познанное”» [9, с. 51]. Однако, с точки зрения добротной гносеологии, способ наблюдения является скверным кандидатом выступать в качестве свойства наблюдаемого.

С точки зрения формальной логики деление понятия осуществляется не по одному основанию. Если воспринять эту проблему не как спекулятивную, а с избыточной онтологической серьезностью,

можно допустить, что природа психического носит двойственный характер. Даже более того, – то, что мы называем ментальным, суть не одна, а две психики различной природы в одном теле. Это не настолько экзотическое допущение, как может показаться. Судя по всему, Аристотель придерживался подобного взгляда, утверждая, что растительная и животная части души произошли от соответствующих видов, а разумная часть есть некая автономная надприродная сущность. В некоторых концепциях разница между чувственным и рациональным познанием носит настолько радикальный характер, что, фактически требует допущения сожительства в одной голове разных типов умов.

Если отстраниться от столь смелых онтологических допущений о существовании трансцендентальных паразитов сознания и прочих агентов имматериального разума, подселившихся в психику приматов, проблему можно осмыслить как вопрос о соотношении феноменальной и структурной дескрипций сознания. Судя по всему, качественное и интенциональное на самом деле отражает разные контуры психической организации, с одной стороны и, разные аспекты конкретных ментальных состояний, с другой.

Обсуждаемая проблема имеет практическое измерение. В настоящее время успешная детекция содержания сознания возможна только по отношению к ощущениям и эмоциям. Уверенная дешифровка интенционального содержания возможна только на основании полноценной функционалистской теории, которой пока не существует, поскольку не известен способ церебрального кодирования семантики.

Судя по всему, феноменальная разнородность ментального содержания отражает реальный гибридный характер естественного интеллекта. Возможно, что это знание может быть использовано для создания интеллекта искусственного.

Теоретическое предложение Дэвида Юма состоит в том, что эмоциональные состояния следует понимать форму оценки любого содержания, безотносительно к теоретико-познавательному типу этого содержания, – конкретно-перцептивному или когнитивно-абстрактному. Эта форма аналоговая по своей сущности, т.е. она представляет интенсивность некоего параметра с помощью непрерывной шкалы. Кстати, создатель квантовой электродинамики Ф. Дайсон утверждает, что мозг – аналоговая машина, поскольку информация эмоций и понимания представлена в аналоговом виде [17]. Обратим

внимание, что психиатрии известны паталогические состояния психики, которые сопровождаются не только отчуждением (человек осознает некие эмоции, но воспринимает их как не свои, они не беспокоят его), но и катастрофическим вырождением эмоциональных состояний. Такой формой является состояние деперсонализации [8, с. 430]. Наверняка эти вырожденные случаи говорят нам что-то о природе эмоций, а именно об их относительной автономности по сравнению с привычным интегративным взглядом. Это означает, что они могут быть представлены как самостоятельный процесс, а значит – подлежат самостоятельному моделированию.

В той или иной степени, аристотелевская иерархия психических функций соотносится со знанием нейрофизиологии: известно, какие церебральные поля отвечают за классы ощущений; имеют место знания о субстрате эмоциональных состояний; высшие когнитивные функции связываются с работой неокортекса и лобных долей. Судя по всему, феноменологическая иерархия в самом общем виде отражает иерархию переработки входящей информации. Поскольку когнитивисты правы, психофизическое знание для нужд моделирования личности приемлемо. Однако, пригодны ли для этих целей феноменологические классификации ментальных свойств и состояний в отрыве от знаний об их физической реализации, которых нам критически недостает? Для имитации личности с целью пройти тест Тьюринга – это знание необходимо. Но есть ли в нем смысл для модерирования личности в рамках антропоморфного робототехнического устройства? Этот ответ следует считать отрицательным.

Если искусственная личность как адекватная модель может быть построена, то она может быть создана как гибридная мультиагентная система. Отдельные модули и целые уровни могут быть аналоговыми, другие – цифровыми; одни системы могут самообучаться, другие – функционировать согласно алгоритмическим правилам и конкретным математическим функциям. Моделирование чувственного восприятия невозможно без самообучения. Первичных репрезентации не могут быть реализованы без инсталляции формальных систем представления знаний (т.н. «онтологий»). Моделирование эмоций представляется с помощью физической аналоговой

формы, которая может быть использована для обобщений в форме суждений как вторичных репрезентаций. В любом случае, добротную искусственную личность без встроенных механизмов обучения и самообучения сделать невозможно, – она должна вырасти в конкретной среде.

Литература

1. *Алексеев А.Ю.* Трудности проекта искусственной личности // Искусственное общество. 2008. Т. 3. № 1. – URL: <https://artsoc.jes.su/su207751800000077-3-1/> (дата обращения 22.02.2022).
2. *Алексеев А.Ю.* Когнитотехнологические проекты искусственной личности // Гуманитарное знание и вызовы времени. – М., СПб.: Центр гуманитарных инициатив; Университетская книга, 2014. С. 156–174.
3. *Асмус В.* Трактат о душе // Аристотель. Соч. в 4-х т. М.: Мысль, 1976. Т. 1. С. 50–62.
4. *Аристотель.* О душе // Аристотель. Соч. в 4-х т. М.: Мысль, 1976. Т. 1. С. 435.
5. *Винник Д.В.* Квантовые свойства в физической организации мозга: амплификация или нивелировка? // Философия науки. 2020. № 1. С. 96–118.
6. *Винник Д.В.* Осознание самосознания как «нулевой уровень» разума // Философия науки 2015. № 4. С. 76–96.
7. *Гуссерль Э.* Идеи к чистой феноменологии и феноменологической философии. М.: Дом интеллектуальной книги, 1999.
8. *Каплан Г.И., Сэдок Б. Дж.* Клиническая психиатрия: из синописа по психиатрии: в 2-х т. Т. 1. – М.: Медицина, 1998.
9. *Рорти Р.* Философия и зеркало природы. – Новосибирск: Изд-во Новосибир. ун-та, 1997. 320 с.
10. *Целищев В.В.* Алгоритмизация мышления: Геделевский аргумент. Изд. 2-е, испр. – М.: ЛЕНАНД.
11. *Юм. Д.* Трактат о человеческой природе. Сочинения в 2-х т. М.: Канон, 1996. Т. 1. С. 64–65.
12. *Arvanitaki A.* Effects Evoked in an Axon by the Activity of a Contiguous One // *Journal of Neuropsychology*, 1942. Vol. 5. No.2. P. 89–108. – URL: <https://www.physiology.org/doi/abs/10.1152/jn.1942.5.2.89> (date of access 07.07.2022).
13. *Beggs M.* Neuronal avalanche / Scholarpedia – URL: http://www.scholarpedia.org/article/Neuronal_avalanche (date of access 07.07.2022).
14. *Bressloff C., Cowan D., Golubitsky M., Thomas J., Wiener C.* Geometric Visual Hallucinations, Euclidean Symmetry and the Functional Architecture of Striate Cortex // *Philosophical Transactions: Biological Sciences*. – 2001. No. 1407. P. 299–330.
15. *Brogaard B.* A Case of Acquired Savant Syndrome and Synesthesia Following a Brutal Assault. 2011 // URL: <https://drive.google.com/file/d/0B0GEjtSycjTKNDU4ZmVhNjktNDk2OC00MjBhLTk5ZmQtYzBhYTRkM2ZlNmU4/view> (date of access 07.07.2022).
16. *Crane T.* Conceptual and Non-Conceptual Content / *The Routledge Encyclopedia of Philosophy*. – Taylor and Francis, 1998. – URL: <https://www.rep.routledge.com/articles/thematic/content-non-conceptual/v-1> (date of access 07.07.2022)

17. *Dyson F.* Are Brains Analogue or Digital? – URL: <https://www.youtube.com/watch?v=JLT6omWrv1w> (date of access 07.07.2022)
18. *Chenu A., Scholes G. D.* Coherence in energy transfer and photosynthesis // Annual Review of Physical Chemistry. – 2015. № 66. P. 69–96.
19. *Fukushima K.* Neocognitron: A hierarchical neural network capable for visual pattern recognition / Neural networks, 1988. Vol. 1, No. 2. P. 119–130.
20. *Kim J.* Philosophy of mind. – Colorado: Westview Press, 1998.
21. *Koch C., Hepp K.* Quantum mechanics in the brain. – URL – <https://www.nature.com/articles/440611a> (date of access 07.07.2022)
22. *Plenio M.B., Huelga S.F.* Dephasing-assisted transport: quantum networks and biomolecules, *The Journal of Physics*, 2008, Vol. 10. – URL: <https://iopscience.iop.org/article/10.1088/1367-2630/10/11/113019/meta/> (date of access 07.07.2022).
23. *Shulman C., Bostrom N.* How Hard is Artificial Intelligence? Evolutionary Arguments and Selection Effects // Journal of Consciousness studies. 2012. Vol. 19, No. 7–8. P. 103–130.
24. *Tee J., Taylor D. P.* Is Information in the Brain Represented in Continuous or Discrete Form? – URL: <https://arxiv.org/ftp/arxiv/papers/1805/1805.01631.pdf> (date of access 07.07.2022).

References

1. *Alekseev A.Y.* Trudnosti proekta iskusstvennoj lichnosti” [Difficulties of artificial personality project] // Iskusstvennye obshchestva [Artificial societies] 2008. Vol.3. No.1. – URL: <https://artsoc.jes.su/s207751800000077-3-1/> (date of access 07.07.2022) (In Russ.)
2. *Alekseev A.Y.* Kognitotekhnologicheskie proekty iskusstvennoj lichnosti” [Cognitive projects of artificial personality] / Gumanitarnoe znanie i vyzovy vremeni [Social knowledge and challenge of time] – Moscow, Sankt Petersburg: Centr gumanitarnykh iniciativ; Universitetskaya kniga, 2014. P.156-174. (In Russ.)
3. *Asmus V.* Traktat o dushe // Aristotel'. Soch. v 4-h tomah. Vol.1. – Moscow: Mysl', 1976. P. 50-62. (In Russ.)
4. *Aristotel* O dushe [De anima] / Aristotel. Soch. v 4-h tomah. [Essays in 4 volumes]. Vol.1. Moscow: Mysl', 1976. (In Russ.)
5. *Vinnik D.V.* Kvantovie svoystva v fisicheskoy organizacii mozga: amplificaciya ili nivelirovka? [Quantum properties in the physics of the brain: amplification or leveling?] *Filosofiya nauki* [Philosophy of science]. 2020. №.1. P. 96-118. (In Russ.)
6. *Vinnik D.V.* Osoznanie samosoznaniya kak «nulevoj uroven» razuma [Awareness of self-awareness as a «zero level» of intelligence] // *Filosofiya nauki* [Philosophy of science] 2015. №4. P. 76-96. (In Russ.)
7. *Husserl E.* Idei k chistoj fenomenologii i fenomenologicheskoy filosofii [Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie] – Moscow: Dom intellektual'noj knigi, 1999. (In Russ.)
8. *Kaplan G.I., Sadock B. J.* *Klinicheskaya psixiatriya: iz sinopsisa po psixiatrii: v 2-h tomah* [Clinical psychiatry: from a synopsis on psychiatry in 2 volumes] Vol.1. – Moscow: Medicina, 1998. (In Russ.)
9. *Rorty R.* *Filosofiya i zerkalo prirody* [Philosophy and the Mirror of Nature] – Novosibirsk: Izd-vo Novosib. un-ta, 1997. (In Russ.)
10. *Tselishchev V.V.* *Algoritmizaciya myshleniya: Gedelevskij argument. [Algorithmic thinking: Gödel's argument]* Izd. 2-e, ispr. Moscow: LENAND, 2021. (In Russ.)

11. *Hume D.* Traktat o chelovecheskoj prirode [Treatise of human nature] / Sochineniya v 2-h tomah [Essays in 2 volumes]. Vol.1. – Moscow: Kanon, 1996. P.64–65. (In Russ.)
12. *Arvanitaki A.* Effects Evoked in an Axon by the Activity of a Contiguous One // *Journal of Neuropsychology*, 1942. Vol. 5. № 2. P. 89–108. – URL: <https://www.physiology.org/doi/abs/10.1152/jn.1942.5.2.89> (date of access 07.07.2022)
13. *Beggs M.* Neuronal avalanche / Scholarpedia – URL: http://www.scholarpedia.org/article/Neuronal_avalanche (date of access 07.07.2022)
14. *Bressloff C., Cowan D., Golubitsky M., Thomas J., Wiener C.* Geometric Visual Hallucinations, Euclidean Symmetry and the Functional Architecture of Striate Cortex // *Philosophical Transactions: Biological Sciences*. – 2001. №. 1407, P. 299–330.
15. *Brogaard B.* A Case of Acquired Savant Syndrome and Synesthesia Following a Brutal Assault. 2011 // URL: <https://drive.google.com/file/d/0B0GEjtSycjTKNDU4ZmVhNjktNDk2OC00MjBhLTk5ZmQtYzBhYTRkM2ZlNmU4/view> (date of access 07.07.2022)
16. *Crane T.* Conceptual and Non-Conceptual Content / *The Routledge Encyclopedia of Philosophy*. – Taylor and Francis, 1998. – URL: <https://www.rep.routledge.com/articles/thematic/content-non-conceptual/v-1> (date of access 07.07.2022)
17. *Dyson F.* Are Brains Analogue or Digital? – URL: <https://www.youtube.com/watch?v=JLT6omWrvlw> (date of access 07.07.2022)
18. *Chenu A., Scholes G. D.* Coherence in energy transfer and photosynthesis // *Annual Review of Physical Chemistry*. – 2015. № 66. P. 69–96.
19. *Fukushima K.* Neocognitron: A hierarchical neural network capable for visual pattern recognition / *Neural networks*, 1988. Vol.1. № 2. P.119–130.
20. *Kim J.* Philosophy of mind. – Colorado: Westview Press, 1998.
21. *Koch C., Hepp K.* Quantum mechanics in the brain. – URL – <https://www.nature.com/articles/440611a> (date of access 07.07.2022)
22. *Plenio M.B., Huelga S.F.* Dephasing-assisted transport: quantum networks and biomolecules, *The Journal of Physics*, 2008, Vol. 10. – URL: <https://iopscience.iop.org/article/10.1088/1367-2630/10/11/113019/meta/> (date of access 07.07.2022)
23. *Shulman C., Bostrom N.* How Hard is Artificial Intelligence? Evolutionary Arguments and Selection Effects // *Journal of Consciousness studies*. 2012. Vol. 19. № 7–8. P. 103–130.
24. *Tee J., Taylor D.P.* Is Information in the Brain Represented in Continuous or Discrete Form? – URL: <https://arxiv.org/ftp/arxiv/papers/1805/1805.01631.pdf> (date of access 07.07.2022)

Сведения об авторе

Винник Дмитрий Владимирович – доктор философских наук, Финансовый университет при правительстве Российской Федерации, профессор Департамента гуманитарных наук (125167, Москва, пр-кт Ленинградский, д. 49/2, e-mail: dvinnik@fa.ru)

Information about the author

Vinnik, Dmitry Vladimirovich – Doctor of Science (Philosophy), Financial University under the Government of the Russian Federation, Department of social studies, professor (49 Leningradsky Prospekt, 125993, Moscow, Russia, e-mail: dvinnik@fa.ru)

Дата поступления 11.09.2022