

ФИЗИЧЕСКИЕ, ФУНКЦИОНАЛЬНЫЕ И МЕНТАЛЬНЫЕ СОСТОЯНИЯ: ПРОБЛЕМА СООТНОШЕНИЯ

Д.В. Винник

Анализируется соотношение физических, функциональных и ментальных состояний. Ставится вопрос о физической и семантической природе функциональных состояний. Делается вывод, что если мы признаем тождество ментальных и функциональных состояний, утверждая при этом тождество функциональных состояний с состояниями физическими в каждом конкретном случае, мы вынуждены принять концепцию психофизической теории тождества свойств. В силу неприемлемости последней следует отказаться от тезиса о полном тождестве функциональных и ментальных состояний, что не запрещает рассматривать их как подобные состояния.

Ключевые слова: философия сознания, материализм, функционализм, коннективизм, семантика, синтаксис, когнитивные науки, экстернализм

Известна точка зрения, что не существует новых решений традиционных философских проблем. Пример такой концепции, как функционализм, демонстрирует ложность этой точки зрения. Функционализм предложил принципиально новый подход к интерпретации природы ментальных состояний.

Функционализм возник как реакция на редуктивный материализм, отождествлявший физические и ментальные свойства. Создатель функционализма Х. Патнэм оспорил подобное отождествление следующим образом: «С точки зрения функционализма, очевидно существует тождество, однако Смарт ориентировался на другой тип качеств мозга как части тождества. Согласно функционализму, мозг обладает качествами, которые по своему смыслу не являются *физическими*. Итак, что я имею в виду, говоря, что *мозг* имеет *нефизические* качества? Я имею в виду качества, *определяемые в терминах, не подразумевающих физику и химию мозга*. Если кажется странным, что физическая система может иметь нефизические качества, представим вычислительную машину. Вычислительная машина обладает множеством физических качеств. Она имеет определенную массу, она имеет определенное количество чипов или чего-то подобного. Она имеет экономические качества, например определенную цену; также она

имеет функциональные качества, например обладает определенной программой. Последняя разновидность качества является качеством *нефизическим в том смысле, что оно может быть реализовано системой независимо от того, какой композицией эта система является метафизически или онтологически*. Бестелесный дух может выполнять определенную программу, мозг может выполнять определенную программу, машина может выполнять определенную программу, и функциональная организация этих трех элементов – бестелесного духа, мозга и машины может быть совершенно одинаковой, даже если их материя, их состав кардинально различаются. Психологические качества проявляют такие же характеристики; одно и то же психологическое качество (например, гнев) может быть качеством тысяч различных биологических видов, которые могут иметь достаточно различную физику и химию (некоторые из этих видов могут быть инопланетными, и, возможно, роботы однажды проявят гнев). Предложение функционализма заключается в том, что, согласно наиболее правдоподобной “монистической” теории, созданной в XX в., наиболее правдоподобной теории, которая избегает трактовать Сознание и Материю как две различные субстанции или две различные области качеств, психологические качества тождественны функциональным качествам» [1].

Таким образом, Х. Патнэм скептически относится к теории психофизического тождества, однако поскольку его симпатии остаются на стороне этой теории, он считает, что в отношении машин Тьюринга можно сформулировать логический аналог вопроса о «тождестве сознания и мозга, который восприведет все аспекты этого вопроса» [2].

Отождествление ментальных состояний с состояниями функциональными приближает функционализм к материализму, однако вопрос о соотношении состояний функциональных и физических остается открытым. Существуют ли функциональные состояния независимо от физических? Очевидно, что это не так. Функциональные состояния могут быть в отношении супервентности или ковариативности с физическими состояниями, и это означает, что объект, обладающий неким функциональным состоянием, с необходимостью находится в некотором физическом состоянии. При этом одному и тому же ментальному состоянию могут соответствовать различные физические состояния.

Дело в том, что понятие функционального состояния все-таки является абстрактным, поскольку по своей сути оно представляет собой значение. Физические состояния, напротив, совершенно реальны. Физическим состояниям, например «вентиль открыт/закрыт», «заряд есть/нет», «свет поляризован/не поляризован» и т.п., можно приписывать разные

логические значения. Вот как описывает взаимоотношение этих двух типов свойств Дж. Серль: «Классы карбюраторов и термостатов определяются в терминах физических последствий. Именно поэтому никто не скажет, что карбюратор можно сделать и голубей. Класс же компьютеров определяется синтаксически в терминах приписывания нулей и единиц. Возможность различной реализации – это следствие того, что соотносимые свойства являются чисто синтаксическими, а не того, что один и тот же результат может быть получен с применением различных физических веществ. В данном случае единственная причастность физических свойств заключается в том, допускают ли они приписывание нулей и единиц и перехода между ними» [3].

Однако очевидно, что для цифрового компьютера в каждый взятый интервал времени, по крайней мере не больший времени переключения логического элемента, мы можем отследить все логические значения элементов. Все множество логических значений элементов можно описать как множество физических свойств физических элементов системы. Однако при том что мы признаем это, остается вопрос: обладают ли входящие в указанное множество физические состояния логическими значениями сами себе? Возможно несколько вариантов ответа на этот вопрос:

1) физические состояния обладают логическими значениями по своей природе;

2) физические состояния обладают логическими значениями только в рамках некоторой системы подобных же физических состояний, т.е. по отношению к множеству других подобных физических состояний;

3) множество физических состояний вычислительного устройства обладают внутренней семантической структурой, поскольку каузально зависят от физических состояний внешней среды. Отношение к физическим состояниям внешней среды и является интенциональностью системы;

4) физические состояния не обладают логическими значениями вообще, поскольку значения приписываются подлинно сознательным агентом на входе и интерпретируются на выходе;

5) существует объективная по содержанию примитивная синтаксическая структура, тождественно отражающая структуру множества физических состояний, однако ее семантическую интерпретацию может осуществлять только интенциональный субъект, приписывающий семантические значения, т.е. смысл в предложениях.

Утверждение (1) носит очевидно вульгарный редуccionистский характер, поскольку оно фактически состоит в том, что некоторому физическому свойству элемента, например «свет поляризован», внутренне при-

суще только одно значение, например 1, но никак не 0. Это явно абсурдное утверждение, ставящее во взаимно-однозначное соответствие логические значения и физические свойства, не допускающее иных интерпретаций. Это утверждение, вероятно, пришлось бы по вкусу сторонникам кибернетического креационизма, рассматривающим весь физический универсум как громадный цифровой компьютер, запрограммированный извне существом, приписавшим физическим состояниям эти значения.

Утверждение (2) более правдоподобно, однако оно оставляет открытым вопрос о природе системы как множества отношений между физическими свойствами. Является ли эта система отношений объективной? Ответ на этот вопрос завит от нашего понимания онтологической природы отношений. В рамках онтологического логицизма, как известно, отношения являются первичными онтологическими сущностями относительно свойств, и онтологический логицизм это утверждение, очевидно, вполне устроило бы. Однако для онтологического логицизма вопрос о соотношении функциональных и физических свойств вообще не является важным, поскольку физические свойства вторичны относительно фундаментальных логических отношений.

Преимущества этого холистического тезиса состоят в том, что такая система не требует «программиста» – агента, приписывающего значения. Она может быть естественной по своей природе. Эта система обладает примитивным синтаксисом, поскольку имеет объективное содержание как множество отношений между физическими свойствами. Это объективное содержание можно обнаружить на различных уровнях физических систем, так как физические свойства, которые могут выполнять роль логических значений, могут существовать на разных уровнях. Свойства более высоких уровней и являются теми свойствами, которые С. Александер называл эмерджентными.

Однако в рамках материалистически ориентированной философии утверждение (2) недостаточно сильное, поскольку оно в лучшем случае объясняет самые примитивные синтаксические свойства системы, но не способно объяснить свойства семантические, т.е. интенциональные, выражаемые в форме пропозициональных установок. Очевидно, что по своей природе это объективное содержание не является набором осмысленных предложений. Тем не менее нельзя полностью отрицать производную интенциональность у таких систем только на основании того, что мы выражаем направленность функционирования системы в предложениях о ее поведении. Система может быть объективно направлена на контроль какого-либо физического параметра или множества таких параметров.

Утверждение (3) характерно для большинства теоретических подходов, реализуемых в рамках так называемых «когнитивных наук», или «когнитивизма», а также для сильной версии теории искусственного интеллекта, согласно которой обладать ментальным состоянием – значит обладать компьютерной программой. Следует отметить, что «когнитивные науки» – это не класс научных дисциплин, а именно специфическая научная программа, включающая междисциплинарные исследования сознания и интеллекта, охватывающие философию, психологию, исследования искусственного интеллекта, нейробиологию (neuroscience), лингвистику и антропологию. Истоки этого направления исследований восходят к 1950-м годам, когда появилась возможность строить компьютерные модели и аналоги тех или иных психических функций.

Центральная гипотеза когнитивных наук состоит в том, что мышление может быть лучшим образом объяснено в терминах так называемых репрезентационных структур сознания и вычислительных процедур, обрабатывающих эти структуры. Большинство работ в области когнитивных наук базируются на предположении, что структура ментальных репрезентаций аналогична структурам компьютерных данных, а вычислительные процедуры разума сходны с вычислительными алгоритмами. Теоретики-когнитивисты предполагают, что психика содержит подобные ментальные репрезентации как логические пропозиции, правила, понятия, изображения и что мозг использует ментальные процедуры, такие как дедукция, поиск, сравнение и восстановление данных. Также мозгу приписываются функции сканирования и вращения изображений. Успехи в создании алгоритмов обработки оптической информации стали основанием для приписывания человеческому мозгу подобных же функций компьютерных программ, таких как распознавание, сканирование, увеличение/уменьшение, вращение и трансформация. Одной из ключевых философских концепций когнитивизма является репрезентационная теория сознания, в рамках которой была выдвинута гипотеза о существовании так называемого «языка мысли», более известного как «менталезе». Автором концепции является Дж. Фодор. Согласно Фодору, менталезе есть врожденный язык и основан он на репрезентациях, которые могут носить бессознательный характер и сходны с предложениями в логике предикатов.

Подобное приписывание человеческому мозгу работы с содержательными объектами не могло не вызвать серьезной критики. Одним из противников когнитивизма является Дж. Серль. Утверждение (4) достаточно адекватно описывает его понимание отношений между функциональными и ментальными состояниями. Это утверждение носит эли-

минативный характер по отношению к категории значения, поскольку отбирает у физических состояний право иметь какие-либо значения вообще. Принимая это положение, мы попадаем в объятия радикального эпистемического конструктивизма.

Этой точки зрения, судя по всему, придерживается Дж. Серль. Он обвиняет сильную концепцию искусственного интеллекта («когнитивизм» в его терминологии) в том, что она предполагает существование «гумункула» – подлинного интенционального объекта в самой системе. Согласно Серлю, это можно обнаружить в утверждениях исследователей-когнитивистов о том, что зрительная кора что-то вычисляет, переводит информацию из плоской картинки в трехмерное изображение. Поскольку трехмерное изображение существует только по отношению к наблюдателю, очевидно, что предполагается существование такого наблюдателя, не совпадающего с самим субъектом, частью которого является зрительный аппарат. «Короче говоря, – пишет Дж. Серль, – смысл, в котором когнитивная наука использует понятие “переработки информации”, находится на слишком высоком уровне абстракции, для того чтобы ухватить конкретную биологическую действительность внутренней интенциональности. Исходя из понимания когнитивной наукой слова “информация” мозг нельзя назвать устройством по переработке информации» [4]. Согласно Серлю, без гумункула не существует вычисления, а существует только замкнутая электронная цепь: «...Когда мы программируем компьютер на выполнение того же вычисления, приписывание вычислительной интерпретации зависит от нас, внешних гумункулов. Нет никакой интенциональной каузальности, внутренне присущей системе... Он не может следовать правилам, потому что у него нет интенционального содержания, внутренне присущего системе, функционирующей каузально, чтобы вызвать поведение» [5].

Одним из любопытных следствий, вытекающих из позиции этого философа, является то, что не существует никаких осмысленных бессознательных процессов, наличие которых предполагает гипотеза бессознательной универсальной грамматики. Никакой бессознательной универсальной грамматики просто не существует: «просто ложно допускать, что осмысленность входа и выхода предполагает совокупность осмысленных процессов между ними, и постулирование недоступных бессознательных процессов является нарушением принципа связи» [6]; «система может конвертировать осмысленный вход в осмысленный выход без следования всяким правилам, принципам, умозаключениям и прочим осмысленным связям между ними» [7]. Последнее утверждение есть по

своей сути выражение одного из основных тезисов коннективизма (connectionism), противопоставляющего себя функционализму, но фактически являющегося его специфической версией.

Коннективизм является радикальной концепцией искусственного интеллекта (ИИ), противопоставляющей себя классической вычислительной теории сознания (будем ее называть просто классической теорией). Рассмотрим основные идеи классической теории, а затем продемонстрируем специфику коннективизма. Классическая теория опирается на понятие логико-синтаксической формы ментальной репрезентации с целью определения степеней и сфер ментальных операций. Это понятие, однако, чуждо ортодоксальным коннективистам, пока они предполагают существование неатомарных ментальных репрезентаций. Классическая трактовка ментальных процессов основывается на двух идеях, каждая из которых восходит к классической теории вычислений. Вместе они объясняют, почему классическая точка зрения постулирует по крайней мере три различных уровня организации вычислительной системы: не только физический и семантический (уровень «знаний»), но и синтаксический.

Первая идея состоит в том, что существует возможность сконструировать языки, в которых определенные свойства синтаксических структур или формул систематически соотносятся с конкретными семантическими свойствами. В таком языке синтаксис формулы кодирует его значение, особенно те аспекты его значений, которые определяют его роль в логическом выводе. Все искусственные языки, используемые в логике, обладают этим качеством. Классицисты убеждены, что это решающее качество для языка мысли, если, конечно, вообще признается его существование.

Вторая идея, лежащая в основании классического подхода к ментальным процессам, заключается в том, что возможно изобрести машины, функции которых будут трансформациями символов, а операции будут чувствительными к синтаксической структуре символов, на основании которых такая машина действует. Это и есть классическая концепция компьютера, восходящая к машинам Тьюринга и фон Неймана.

Очевидно, что эти две идеи вполне совместимы друг с другом. Таким образом, в рамках классической модели ИИ компьютеры являются машинами, преобразующими символы согласно определенным логическим правилам. Существование сигнатуры дает нам основание приписывать вычислительным системам определенную ее синтаксисом семантику.

Концепция коннективизма полностью отвергает привычный подход, в котором основными сущностями являются символы и правила. Эта концепция была порождена идеологией параллельных вычислений и кон-

цепцией так называемых нейрокомпьютеров – вычислительных систем, построенных по аналогии с нейронными сетями, т.е. фактически моделирующих их поведение. В настоящее время существуют различные коннективистские модели, однако под коннективизмом будем подразумевать именно модель, основанную на параллельно распределенных вычислениях. Коннективисты заменили символы и правила на паттерны числовой активности групп элементов и паттерны весов групп связей между элементами. Понятно, что любой элемент или связь участвует во многих таких паттернах, которые могут перекрывать друг друга. Таким образом, все символическое коннективизм пытается элиминировать. Изменяя веса связей, экспериментаторы адаптируют сеть к внешней среде. Также в систему включаются алгоритмы обратной связи, позволяющие на основании статистических данных изменять вес собственных связей, т.е. алгоритмы самообучения [8].

Популярность коннективизма была обусловлена критическим отношением к отождествлению физических и функциональных состояний, трактуемых как состояния, обладающие семантическим значением. Вот как описывают причины успеха коннективизма П. Чёрчлэнд и Т. Сейновски: «Коннективизм... освобождает нас от того, что Хофстедтер в 1982 г. назвал “Булевым сном” – от ошибочного представления, что все когнитивные функции являются манипуляциями с символами по правилам логики. Точно таким же образом он освобождает нас и от *сна нейрофизиолога*, который является верой в то, что ответы на интересующие нас вопросы станут однажды очевидными в результате изучения мельчайших особенностей каждого нейрона (морфологии, физиологии, связей и т.п.). Коннективизм преподносит нам чрезвычайно важный урок, что качества системы не доступны на одном-единственном уровне системы» [9].

В самом деле, коннективизм моделирует ментальные или поведенческие явления, используя искусственные нейронные сети. Функционирование таких сетей описывается исключительно в статистических значениях работы элементов, из которых составлена система. Осмысленное поведение системы является лишь возможным следствием ее работы. В случае если мы наблюдаем осмысленное поведение, считается, что моделирование выполнено успешно, в противном случае – нет. Нейросетевая модель является семантически, впрочем, как и синтаксически, непрозрачным «черным ящиком», генерирующим осмысленное поведение по непонятным статистическим закономерностям.

Этой особенностью не преминули воспользоваться критики коннективизма, в частности Дж. Фодор и З. Пилишин: «Глубоко неверно в архи-

текстуре коннективизма следующее: поскольку он не признает ни семантических, ни синтаксических структур в ментальных репрезентациях, он невольно трактует их не как некое общее множество, а как список. Но списки сами по себе не имеют структуры, любой набор элементов является возможным списком. Согласно принципам коннективизма любой набор репрезентационных состояний (причинно связанных) является возможным сознанием. Таким образом, если принять посылки коннективизма, ничто не предохраняет умы от случайной несистематичности. Этот результат очевидно абсурден. Когнитивные способности проявляются в структурных кластерах. Их систематичность носит всеобъемлющий характер. Все доказательства говорят о том, что *умы в крапинку существовать не могут*. Этот аргумент был решающим в отношении коннективизма Хебба, Осгуда и Хула еще 20–30 лет назад. Сейчас вряд ли что-то существенно изменилось» [10].

Очевидно, что коннективистская концепция не может включать в себя представления о структурных отношениях между ментальными состояниями, поскольку она вообще не содержит сколько-нибудь интенционально осмысленного структурного описания. Все, с чем она имеет дело, – это статистические значения. Перспективы появления квантовых компьютеров и гипотеза квантовой природы сознания дают коннективистам дополнительные аргументы в пользу убеждения, что они находятся на правильном пути. Коннективизм очевидно предполагает эмерджентизм. Также коннективизмом оказывается слабая версия искусственного интеллекта, согласно которой ментальные процессы не являются программой, но могут быть симулированы вычислительно.

Утверждение (5), по сути, является неким компромиссом между утверждением (2) и утверждением (4). Эта точка зрения характерна для семантического экстернализма, основателем которого считается поздний Х. Патнэм. «В отличие от доктрины, которая известна нам начиная с XVII в., значения не находятся в голове» [11], – утверждает Патнэм. И далее: «Доктрина, согласно которой существуют ментальные репрезентации, с необходимостью относящиеся к вещам во внешнем мире, плоха не только с естественно-научной точки зрения, – это плохая феноменология и концептуальная путаница» [12].

Для демонстрации внешней природы значения Патнэм предложил мысленный эксперимент, известный как «Земля-Два». Согласно этому эксперименту, наше ментальное содержание не позволяет различить два феноменально идентичных объекта, при этом структурно различных, – для этого необходимо объективное знание о природе объектов, которое по

отношению к нашему опыту является внешним. Эта разновидность экстернализма делает акцент на ментальном содержании семантической природы, предполагая, что ментальное содержание не является супервентным по отношению к тому, что в происходит в голове. Существует интенциональное отношение между нашими мыслями и внешним состоянием дел – между понятиями и объектами.

Согласно функционализму, ментальное состояние определяется его функциональной ролью, которая включает в себя отношения между входами, выходами и другими ментальными состояниями. В стандартной формулировке функционализма входами и выходами, как правило, считаются объекты и информация, которые различаются относительно границ тела и не принадлежат к внешней среде. Н. Блок сформулировал радикальный функционализм «длинной руки», отвергающий это требование с целью принять экстернализм [13]. Для демонстрации своей концепции Блок предложил мысленный эксперимент, имеющий любопытные следствия. Он является своеобразным намеком на «китайскую комнату» Дж. Серля. Для начала Блок предлагает представить, что функциональная организация вашего мозга досконально известна как набор элементов и правил операций, производимых над ними. Это позволяет реализовать функциональную организацию вашего мозга на другой элементной базе, например последовательно заменить всю нервную ткань мозга на нечто другое, сохраняя связи, идущие от органов чувств к мышечной ткани, так чтобы поступающая извне информация и ваше поведение оставались прежними. С равным успехом, используя высокочастотную беспроводную связь с органами чувств и мышцами, вашу функциональную организацию можно реализовать и за пределами вашей черепной коробки. В качестве элементов этой системы теоретически может выступить китайский народ: «Предположим, что мы индоктринировали правительство Китая в функционализм и убедили китайских чиновников реализовать человеческий разум на час. Мы обеспечили всех китайцев (Китай выбран потому, что там как раз миллиард человек) двусторонней радиосвязью, чтобы связать их с другими людьми и с вашим телом... Эта система из миллиарда китайцев, связанных друг с другом по радио, играет роль внешнего “мозга”, связанного с вашим телом по радио. Ничего абсурдного в связи человека с его мозгом по радио нет» [14].

Экстернализм вызывает оптимизм у разного рода технократов-футуристов. Некоторые из них, например М. Эпштейн, даже предлагают новые понятия, такие как «синтеллект», «нейросоциум» или «церебрально открытое общество». «Синтеллект – соборный разум, соразум, который об-

разуется интеграцией индивидуальных сознаний через сеть электронных коммуникаций; высшая ступень в создании всечеловеческого нейроквантового мозга» [15]. «Нейросоциум – симбиоз мозга и социума; общество, которое непосредственно управляется мозговыми процессами и, в свою очередь, их контролирует. Это церебрально открытое общество (ЦОО), в котором мозговые процессы технически освоены, выведены наружу и прямо участвуют в информационных потоках и производственных процессах. Развитие современных информационных технологий ведет в перспективе к сращению всех линий коммуникаций: церебральных (между клетками мозга) и социокультурных (между ячейками общества). Со временем мозговые сигналы будут прямо передаваться по электронным сетям. Наряду с бывалым ускорением всех информационных процессов нейросоциум подвержен опасностям тоталитарного контроля. Мысли в принципе могут читаться, поэтому придется быть осторожным не только в словах. Церебрально открытое общество может потребовать от всех своих членов такой умственной аскезы, которой раньше предавались только монахи и йоги. Ментальная “корректность” или “гигиена” выработает привычку сурового “мыследержания” – дисциплины управления своими мыслями, воздержания от общественно опасных или греховных помыслов, как способ самозащиты в нейросоциуме. По логике антиутопии можно предположить, что гражданину нейросоциума будет позволено лишь иногда отключать свой высокоразвитый мозг от сигнальной панели, которая станет передавать малейшее возбуждение его нейронов в центральную диспетчерскую, планетарный или космический “нус”» [16].

Действительно, если принять положения функционализма, перспективы открываются фантастические начиная от создания протезов мозга и заканчивая полной заменой мозга на искусственный кибернетический аналог. Местонахождение этого кибернетического аналога является вопросом удобства, не более. Впрочем, если у одних эти следствия рождают нескрываемый оптимизм, у других они вызывают откровенное неприятие, если не ужас. Как справедливо отмечает Дж. Серль, если мы попытаемся серьезно воспринять идею о мозге как цифровом компьютере, то столкнемся с неприятным выводом о том, что можем из чего угодно создать систему, которая будет делать то же, что и мозг. С этой точки зрения, выражаясь языком вычисления, правомерно утверждать, что «мозг», способный работать как мой или ваш, «можно создать из кошек, мышей и сыра, или из прутков, или водопроводных труб, или еще чего-нибудь, с тем условием, что обе системы будут, как говорит Блок, “вычислительно эквивалентны”. Защитники

когнитивизма заявляют об этом выводе с полным и нескрываемым удовлетворением» [17].

Стоит сказать несколько слов о проблеме создания искусственного интеллекта. С точки зрения ментализма очевидно, что искусственный интеллект в полном смысле этого слова невозможен по определению, поскольку человеческий интеллект рассматривается как естественный и не сводимый к искусственному в принципе. Если мы принимаем положения механицизма, эта проблема становится практической. Однако это не означает, что она решаема теоретически, в том смысле, что мы сможем оценить степень приближения создаваемых вычислительных систем к порогу разумности.

Основной вопрос заключается в том, существуют ли некие формализуемые признаки интеллекта, которые мы способны вмонтировать в искусственную систему, и можем ли мы таким образом наделить ее разумом. Второй вопрос состоит в том, что если допустить, что такие признаки существуют, то имеем ли мы возможность вычленить их, исследуя естественный человеческий интеллект инструментарием физических и когнитивных наук и/или формальными логико-математическими средствами. Ответ далеко не очевиден, особенно относительно возможностей математической логики в данной сфере. Если применительно к человеческому разуму (аналог которого мы хотим воспроизвести технически) верна концепция классической теории вычислений, оперирующих с символами и правилами, то эта задача теоретически решаема. Впрочем, даже если классицисты не правы, есть подозрение, что их модели будут настолько хорошо симулировать искусственный интеллект, что смогут пройти тест Чёрча. Если же правы коннективисты, то это означает, что ясных критериев оценки приближения создаваемых систем к порогу разумности у нас нет, так как свойство разумности является по своей сути эмерджентным согласно коннективистскому пониманию проблемы. Известно, что эмерджентисты настаивают на том, что психические свойства появляются спонтанно и непредсказуемо, когда некая естественная система переходит на более высокий уровень сложности. Если принять это положение, то следует сделать вывод, что задача создания искусственного интеллекта не может быть решена радикальным рациональным путем «с нуля», – ее возможно решить только с помощью соразмерного наращивания сложности вычислительной базы и программного обеспечения в надежде, что в некий момент искусственная система наконец обретет сознание.

Если мы признаем тождество ментальных и функциональных состояний, утверждая при этом тождество функциональных состояний с состо-

яниями физическими в каждом конкретном случае, то принимая такую точку зрения, мы вынуждены принять концепцию психофизической теории тождества свойств. Последняя, как было продемонстрировано, неприемлема. Следовательно, нам следует отказаться от тезиса о полном тождестве функциональных и ментальных состояний, что не исключает запрещаает их как подобные состояния.

Примечания

1. *Patnam H.* Reason, truth and history. – Cambridge, 1997. – P. 19.
2. *Ibid.* – P. 24.
3. *Серль Дж.* Новое открытие сознания. – М.: Идея-пресс, 2002. – С. 201.
4. Там же. – С. 206.
5. Там же. – С. 201.
6. Там же. – С. 224.
7. Там же. – С. 225.
8. См.: *Smolensky P.* Computational models of mind // A Companion to the Philosophy of Mind. – Blackwall Published Ltd., 1993. – P. 176-200.
9. *Churchland P.S., Sejnowski T.J.* Neural representation and neural computation philosophical perspectives // Action Theory and Philosophy of Mind. – 1990. – V. 4. – P. 367.
10. *Fodor J., Pylyshin Z.* Excerpt from connectionism and cognitive architecture. – Blackwell Publisher Ltd., 2008. – P. 271.
11. *Patnam H.* Reason, truth and history. – P. 19.
12. *Ibid.* – P. 21.
13. См.: *Block N.* Inverted Earth // Philosophical Perspectives / Ed. By J.E. Tomberlin. – Atascadero, CA: Ridgeview Press, 1991.
14. *Block N.* An excerpt from troubles with functionalism // A Companion to the Philosophy of Mind. – P. 49.
15. *Эштетейн М.* Синтеллект // Проективный философский словарь. – СПб., 2003. – С. 358.
16. *Эштетейн М.* Нейросоциум // Проективный философский словарь. – С. 255–256.
17. *Серль Дж.* Новое открытие сознания. – С. 192.

Дата поступления 03.02.2010
Институт философии и права
СО РАН, г. Новосибирск
dvin@ngs.ru

Vinnik, D.V. Physical, functional and mental states: the problem of their correlation

The paper analyzes a correlation of physical, functional and mental states. The problem of physical and semantic character of functional states is raised. The following conclusion is made: if we accept identity of mental and functional states (and affirm identity of functional states and physical ones in each specific case) we have to accept the conception of property identity theory. Because it is inadmissible, we should reject the thesis of a complete identity of functional and mental states, but we may treat them as similar states.

Keywords: philosophy of mind