

УДК 519.7



М.Ф. Бондаренко¹, Н.П. Кругликова², Н.Е. Русакова³,
Ю.П. Шабанов-Кушнаренко⁴

¹⁻⁴ ХНУРЭ, г. Харьков, Украина

ПРОБЛЕМЫ МОДЕЛИРОВАНИЯ СУБЪЕКТИВНЫХ СОСТОЯНИЙ

В статье обсуждаются новые обстоятельства в области интроспекции, которые возникают в процессе развития теории интеллекта. Рассматривается природа интроспекции и корректируется постановка одной из задач теории интеллекта с целью приведения ее решения в соответствие с интроспективными данными на примере изучения механизма цветового зрения человека. Сформулирован принцип тождества, согласно которому субъективное и соответствующее ему объективное — это один и тот же предмет или процесс, рассматриваемый с разных сторон.

ИНТРОСПЕКЦИЯ, ТЕОРИЯ ЦВЕТОВОГО ЗРЕНИЯ, СУБЪЕКТИВНЫЕ СОСТОЯНИЯ, ПРИНЦИП ТОЖДЕСТВА

Введение

О значении для теории интеллекта интроспекции и получаемых с ее помощью сведений о психофизических процессах и субъективных состояниях испытуемого уже говорилось в работе [1]. Однако, в процессе дальнейшего развития теории интеллекта в этой области обнаруживаются новые важные обстоятельства. Некоторые из них обсуждаются в этой статье. По мере продвижения вглубь механизма интеллекта исследователь обнаруживает, что сложность формулировок постулатов, характеризующих физически наблюдаемое интеллектуальное поведение испытуемого, катастрофически растет. Если же эти постулаты записывать не в объективных, а в субъективных терминах, тогда, как свидетельствует уже накопленный опыт исследований, сложность формулировок постулатов непомерно не растет и всегда остается в приемлемых пределах.

Необходимость записи постулатов в субъективных терминах существенно усложняет работу исследователя. Как и прежде, он обязан на каждом шаге своей деятельности обеспечить безупречную объективность изучения интеллекта. Это означает, что все термины, используемые для формулировки очередного постулата, он должен однозначно выразить через понятия, опирающиеся исключительно на сведения о поведении испытуемого, получаемые в объективном эксперименте. Связь этих терминов, задаваемая постулатом, должна подтверждаться физическими опытами. Но теперь исследователь дополнительно должен относительно каждого вновь вводимого понятия и каждой связи доказывать, что они согласуются также и с интроспективным свидетельством испытуемого. Чтобы этого достичь, испытуемый должен тщательно сличить все детали введенного объективного понятия с элементами своего субъективного опыта. Кроме того, он должен убедиться в том, что его субъективные состояния связаны друг с другом именно так, как этого требует формулировка постулата. Результат сличения будет

удовлетворительным лишь в том случае, если испытуемый при максимуме внимания и старания не сможет обнаружить отличия между данными своего субъективного опыта и содержанием логической конструкции объективных понятий и составленного из них постулата.

Если же такого тщательного сличения объективного и субъективного не делать, то очень скоро, по мере продвижения вглубь механизма интеллекта, объективно вводимые понятия перестанут соответствовать субъективным состояниям испытуемого. Вследствие этого исследователь рано или поздно почувствует зыбкость почвы, на которой он строит здание теории интеллекта, и утратит возможность опираться на субъективный опыт испытуемого при постановке новых задач исследования. Отказавшись от услуг логмана-интроспекции, исследователь потеряет ориентировку и утонет в сложности психики человека, не зная, в какую сторону двигаться дальше, чтобы обеспечить успех физического изучения интеллекта. Таким образом, интроспекция нужна для того, чтобы не заблудиться в лабиринте возможных постановок задач физического исследования интеллекта. Только немногие пути в этом лабиринте ведут к успеху, остальные — в тупики непреодолимой сложности формулировок свойств поведения испытуемого. Для преодоления этой сложности найден пока единственный прием — свертывание объективной формулировки постулатов в субъективную.

Хотелось бы лучше узнать, что собой представляет интроспекция — этот уникальный источник сведений о субъективных состояниях, от которого столь сильно зависит успех исследований человеческого интеллекта. Кажется, что в механизме интроспекции есть что-то физически невозможное. С одной стороны, нельзя усомниться в том, что субъект действительно ощущает, чувствует, переживает. С другой, не удается ясно представить принцип действия прибора, способного генерировать субъективные состояния. Очень трудно человеку

поверить в то, что чувствовать может не только он сам, но также и система искусственного интеллекта. Столь же трудно ему представить наличие субъективных состояний у других людей. Однако нельзя не согласиться и с тем, что у каждого человека есть свой внутренний мир. Перед лицом этих трудностей и противоречий легко прийти к мысли, что чувствует не тело, а отличная от него душа. Но и это не помогает делу. Говорить так — значит просто откладывать объяснение феномена интроспекции: ведь всегда можно спросить, а как же тогда удастся чувствовать душе. Можно ответить: сверхъестественным образом. Но тот, кто так утверждает, просто уклоняется от научного объяснения факта наличия субъективных состояний у людей.

1. О природе интроспекции

Нам представляется, что не следует относиться слишком серьезно к тем противоречиям (несомненно, мнимым), которые возникают при попытке проникнуть в структуру механизма интроспекции методом самонаблюдения, и делать из них выводы нигилистического или мистического характера. Многочисленные факты говорят о том, что сведения, доставляемые нашим внутренним зрением, будучи ложно истолкованы, часто ведут к ошибкам. Примерами этого могут служить разнообразные обманы чувств. Особенно же осторожно надо относиться к свидетельствам интроспекции о самой себе. Когда кажется, что показания наших чувств противоречат прочно установленным законам физики, следует решительно стать на сторону физики и считать, что такие показания нуждаются в более удачной интерпретации. Например, испытуемый помещает левую руку в горячую воду, а правую — в холодную и держит их там несколько минут. Затем окунает обе руки в теплую воду. Теперь он ясно чувствует, что одна и та же вода и холодна (по показаниям левой руки) и горяча (по показаниям правой). Однако с физической точки зрения — это абсурд, и поэтому при оценке температуры воды такие показания наших чувств мы просто игнорируем, вопреки их кажущейся очевидности и непреложности.

Но можно ли полагаться на интроспективные данные при формулировке задач теории интеллекта, если они нас обманывают? Мы полагаем, что можно, только делать это надо достаточно осторожно и осмотрительно, по принципу «доверяй, но проверяй». Взятые сами по себе чувства никогда не лгут. Если под действием воды одной и той же температуры левой рукой я чувствую, что вода холодна, а правой — что она горяча, то этот факт моего сознания непреложен. Рассматриваемый вне связи с вызвавшим его внешним воздействием он не может противоречить каким бы то ни было физическим данным. Факт чувствования не может быть оспорен.

Противоречие между объективным и субъективным появляется лишь тогда, когда человек пытается судить о воспринимаемом внешнем физическом объекте по своему субъективному состоянию (в нашем примере — о температуре воды по ощущению тепла или холода). Но так ли уж непреодолимо это противоречие? Когда физик видит, что его термометры показывают различную температуру одного и того же тела, он не спешит объявлять о наличии неразрешимого противоречия или сверхъестественных эффектов. Быть может, термометры имеют разные шкалы. С подобной меркой можно подойти и к оценке разнобоя в показаниях температурных анализаторов человека. Известно, что температурный анализатор человека (как впрочем и все другие его анализаторы) снабжен механизмом адаптации, меняющим масштаб шкалы в зависимости от уровня измеряемой температуры.

2. Корректировка постановки задач

Рассмотрим пример корректировки постановки одной из задач теории интеллекта с целью приведения ее решения в соответствие с интроспективными данными. Речь идет о задаче изучения механизма цветового зрения человека, описанной в работе [1]. Испытуемый воспринимает два произвольно выбранных исследователем световых излучения на расположенных рядом небольших тест-полях. Он сличает цвета излучений и устанавливает их равенство или неравенство. Оказывается, что результат сличения цветов ни от каких побочных факторов не зависит. Не зависит он, в частности, и от фона, окружающего тест-поля. Убедившись в рефлексивности, симметричности, транзитивности, аддитивности, трехмерности и непрерывности предиката, реализуемого испытуемым, исследователь дедуктивно выводит из этих свойств-постулатов модель цветового зрения, то есть конкретный вид преобразования излучения в его цвет.

Все сделано безупречно, и теперь, казалось бы, можно с помощью этого преобразования предсказать заранее, будут ли два произвольно взятые световые излучения восприниматься испытуемым в виде одинаковых цветов или нет вне зависимости от того, в каких условиях они ему предъявлены. Рекомендация теории проста: если физическая характеристика цветов совпадет (а именно: должны совпасть две тройки чисел — координаты этих цветов), то цвета будут одинаковыми, если же не совпадет, значит, они должны оказаться различными. Однако существует простое интроспективное наблюдение, которое противоречит такого рода предсказаниям. Оно демонстрируется следующим «опровергающим экспериментом». Сформируем на небольшом тест-поле некоторое цветное излучение и окружим его обширным синим фоном. Подбе-

рем спектр излучения тест-поля таким, чтобы оно воспринималось испытуемым в виде серого цвета. Далее, сохраняя неизменным излучение тест-поля, заменим синий фон на желтый. Испытуемый видит, что цвет тест-поля после замены фона меняется, превращаясь из серого в голубой. Читатель может легко убедиться в справедливости этого наблюдения. Итак, данные интроспекции явно противоречат выводам физической теории, когда она утверждает, что цвет однозначно определяется вызвавшим его световым излучением.

Означает ли это, что физическая теория цвета ошибочна? Отнюдь нет. Ведь в ней под термином «цвет» понимается объективно определяемое понятие, а именно — класс всех световых излучений, равных по цвету в экспериментах со сравнением двух рядом расположенных полей, а это не совсем тот цвет, который мы субъективно переживаем. Если взять два световых излучения, равных по цвету на окруженных каким-либо фоном соседних полях, то при любой перемене фона цветовое равенство сохранится, но сами цвета могут меняться. Значит субъективно переживаемый испытуемым цвет каждого тест-поля определяется не только породившим его излучением, но также и фоном, окружающим поля. Получается, что субъективный цвет излучения — это не класс всех метамерных излучений, а нечто такое, что при фиксированном фоне связано с ним взаимно однозначной зависимостью. Причем сама эта зависимость меняется с изменением фона. Итак, физическая теория цвета верна, но она отвечает не совсем на тот вопрос, который мотивирован данными интроспекции. Она ведет не совсем туда, куда указывает содержимое нашего внутреннего мира. То, что называется в колориметрии цветом, не всегда можно интерпретировать как психологическую реакцию испытуемого на световое излучение. Таким образом, колориметрическое понятие цвета не всегда согласуется с ощущением цвета. Объективное и субъективное в данном случае вступают в противоречие друг с другом, следовательно, постановка задачи о цвете нуждается в усовершенствовании, если целью является познание природы цвета как субъективного состояния человека.

Как же изменить формулировку задачи, чтобы ее решение давало не только описание поведения испытуемого при сравнении двух световых излучений по цвету, но также и описание преобразования светового излучения в его цвет, субъективно переживаемый испытуемым? Простейший ответ состоит в том, чтобы раз и навсегда зафиксировать фон, окружающий тест-поля. Например, все опыты делать только на черном фоне. Тогда будет устранена причина (изменение фона), которая нарушает взаимно однозначное соответствие между цветом физическим и

цветом психологическим. Такая постановка задачи выглядит не вполне естественной по физическим соображениям: незачем фиксировать фон, если он не влияет на результаты эксперимента. Неудовлетворительна она и по психологическим соображениям, так как не снимает ранее выдвинутого возражения: существует такая ситуация, при которой упомянутая выше модель цветового зрения дает результат, расходящийся с субъективным опытом испытуемого. Чтобы исправить положение, нужно фон включать в характеристику входного сигнала наряду со спектром предъявляемого излучения. Тогда опыты можно будет поставить таким образом, чтобы охватить ими и «опровергающий эксперимент».

Задача изучения цветового зрения человека формулируется следующим образом. Берем маленькое тест-поле (теперь уже не два соседних поля, а вместо них только одно), на котором испытуемому предъявляются произвольные световые излучения. Тест-поле окружено некоторым фоном, который заполняет всю остальную часть поля зрения. Фон может выбираться произвольно. Испытуемый в течение некоторого времени наблюдает на тест-поле цвет предъявленного излучения в присутствии заданного фона. Затем излучение на тест-поле и фон, его окружающий, произвольно меняются, и испытуемый наблюдает новый цвет на тест-поле, сравнивая его по памяти с исходным цветом. В случае равенства цветов испытуемый должен отреагировать положительным ответом, в случае неравенства — отрицательным. Ясно, что такие опыты включают в себя, как частный случай, и ранее рассмотренный «опровергающий эксперимент». Его мы получим, если для первого и второго предъявлений на тест-поле используем одно и то же излучение, фоны же должны быть взяты разными.

На основе этих опытов должна быть построена скорректированная модель цветового зрения человека в виде зависимости субъективного цвета x от порождающих его светового излучения X и фона Y . Обозначим зависимость колориметрического цвета x от излучения X , получаемую при решении задачи в исходной постановке, символом f , тогда $f(X) = x$. В новой постановке решение задачи будет иметь вид $h = g(f(X), Y)$. Здесь g — зависимость субъективного цвета x от колориметрического цвета $x = f(X)$ и фона Y . Теперь h — это психологический цвет. Мы использовали колориметрический цвет $f(X)$ вместо излучения X в роли первого аргумента функции g на том основании, что на любом фоне субъективный цвет x не меняется при замене излучения X на любое метамерное ему излучение. Об этом свидетельствует упоминавшийся ранее факт, что равенство колориметрических цветов сохраняется при любом изменении фона, окружающего тест-поля.

Приведенный выше пример корректировки одной из психофизических задач теории интеллекта наглядно демонстрирует, как под натиском интроспективных данных исследователь интеллекта вынужден расширять постановку решаемой им физической задачи. При фиксированном фоне колориметрическое решение задачи о цветовом зрении человека безупречно как с физической, так и с психологической точек зрения. Но как только ограничение на фиксацию фона снимается, психологические данные сразу же вступают в противоречие с колориметрической моделью цветового зрения. Чтобы это противоречие устранить, придется сам фон включать в характеристику входного сигнала, предъявляемого испытуемому. При этом происходит расширение постановки исходной физической задачи. Важно отметить, что после решения задачи в новой, более широкой, постановке ранее полученные результаты не обесцениваются, поскольку старая модель вливается в новую, становясь ее частью. Нечто подобное происходит и в физике: при появлении более общей теории старая теория обычно не сходит со сцены, но по-прежнему успешно используется в рамках, очерченных новой теорией. Итак, мы видим, что интроспекция снабжает исследователя ценной информацией, которую тот может использовать для прокладывания маршрута движения вглубь механизма интеллекта при его познании. Математические модели различных сторон механизма интеллекта человека строит и обосновывает физика. На вопрос же, какие именно стороны интеллекта надо изучать на данном этапе его исследования, ответ дает психология.

Создаваемые физикой теории относительноны в том смысле, что новые теории уточняют и обобщают уже имеющиеся теории. И нет ни одной такой физической теории, которая была бы застрахована от пересмотра. Видимо, подобное положение со временем возникнет и в психофизической части теории интеллекта. Во всяком случае, уже сейчас ясно, что теория цветового зрения, учитывающая фон, может быть еще более обобщена. Об этом свидетельствует наличие таких интроспективных данных, которые вступают с ней в противоречие. Оказывается, что благодаря цветовой адаптации, на субъективный цвет влияет не только фон, окружающий тест-поле, но и то, что видел испытуемый на тест-поле и вокруг него до момента предъявления светового стимула, порождающего этот цвет. Субъективный цвет несколько меняется также под действием звуковых, осязательных и других ощущений. Цвет можно изменить, надавливая на глазное яблоко, а также воздействуя на сетчатку глаза или на кору затылочных долей мозга испытуемого электрическим током. Кроме того, при исследовании цветового зрения

разных лиц даже при одинаковой постановке задачи могут получаться различные модели (индивидуальные различия функций спектральной чувствительности зрения, зрение дальтоников (дихроматов различных типов, ахроматов, цветаномалов)). Так что перед исследователем открываются поистине неограниченные возможности для дальнейшего обобщения и усовершенствования модели цветового зрения. То же относится и к моделям любых других психофизических процессов.

3. Проблема природы субъективных состояний

Выше на конкретном примере исследования цветовых ощущений был продемонстрирован относительный характер физических знаний о субъективных состояниях испытуемого. В процессе углубления наших знаний о цветовом ощущении в нем открываются все новые и новые свойства. Вместе с тем, сама природа цвета, а именно — его способность быть синим, желтым и т.п. (и в этом качестве непосредственно переживаться испытуемым) всегда остается за пределами физической характеристики цвета. Мы чувствуем цвет таким, каков он есть на самом деле, мы не сомневаемся в том, что цвет, взятый сам по себе (как «вещь в себе»), именно таков, каким он является перед нашим сознанием. И оказывается, что эта непосредственно воспринимаемая нами сущность цвета (например, его синева) не выражается в объективно наблюдаемых свойствах цвета.

В связи со сказанным напрашивается еще одна аналогия теории интеллекта с физикой. Широко известен афоризм: «Материя исчезла, остались одни уравнения». Здесь имеется в виду, что любая физическая теория описывает лишь свойства предметов, но не их субстанцию, природу, не сами предметы как «вещи в себе». Исследователь интеллекта тоже может с полным основанием сказать: «Ощущение исчезло, остались одни уравнения». Ощущение как «вещь в себе», как субъективное состояние испытуемого, реально им переживаемое, не извлекается из объективно наблюдаемого поведения испытуемого. Физик имеет доступ лишь к свойствам ощущений, которые он выражает в виде уравнений. Само же ощущение испытуемого, как его личное переживание, находится вне пределов описательных возможностей физики.

Но вправе ли мы при таком положении дела допускать возможность создания полноценного искусственного разума на базе достижений теории интеллекта? Вопрос этот далеко не праздный, ибо от него зависит сфера практического применения теории интеллекта. Если бы описание природы субъективных состояний человека было возможно, тогда можно надеяться, что искусственное воспроизведение полученных моделей приведет к тому,

что техническое устройство будет на самом деле переживать свои состояния в том виде, в каком их переживает человек, отдавая себе отчет о них. Если же это принципиально недостижимо, тогда не исключено, что такое устройство будет лишь чисто внешне имитировать интеллектуальную деятельность человека, оставаясь, по существу, всего лишь бесчувственным, бездушным автоматом.

Человек субъективно переживает каждый цвет (например, синий) как нечто совершенно конкретное, вполне определенное, абсолютное, как бы данное извне, существующее само по себе. Переживаемый цвет — это некая субстанция, она обладает характерным качеством синевы, причем это не имя синевы, а она сама. Эта синева предметна, она переживается принудительно, ее нельзя изменить усилием воли, она не зависит от точки зрения субъекта, от того, чем она сопровождается, с чем сравнивается. Каждый человек неизбежно верит, что другие люди переживают цвета точно такими же, какими переживает их он сам, хотя проверить, так ли это на самом деле, объективными методами он не может. Субъективный цвет (как и все другие психологические состояния человека) очень похож на Кантову «вещь в себе», природа которой тоже недоступна для физического описания.

В противоположность цвету субъективному цвет колориметрический есть всего лишь математическая абстракция, он задается с точностью до обозначений, его конкретное представление зависит от воли исследователя. Колориметрический синий цвет — это всего лишь конкретная тройка чисел, он не субстанционален, в нем нет субъективно переживаемого качества синевы. В одной системе обозначений он будет представлен одной тройкой чисел, в другой — совсем иной тройкой чисел. Если бы субъективно переживаемые цвета вдруг поменялись своими качествами (например, синий цвет стал зеленым, зеленый — желтым и т.д.), то физик, изучающий лишь внешние реакции испытуемого, не смог бы заметить этой подмены. Он продолжал бы называть синим тот цвет, который уже изменился на зеленый, характеризуя его тройкой чисел, предназначавшейся ранее для описания синего цвета.

Синева, как органически присущее субъективному цвету качество, не попадает в число характеристик, фигурирующих в колориметрическом цвете, она ускользает от математического описания. Поэтому даже в том случае, когда построенные для двух испытуемых математические модели цветового зрения оказываются совершенно одинаковыми (в колориметрическом варианте для этого требуется совпадение всех трех функций спектральной чувствительности зрения), нет достаточных оснований для вывода из этого факта, что такие испытуемые

реагируют одинаковыми субъективными цветами на идентичные световые излучения. Теоретически не исключено (хотя это кажется невероятным с точки зрения здравого смысла), что существуют люди, реагирующие на световые излучения ощущениями, которые для других людей являются слуховыми или обонятельными, или же такими ощущениями, которые другим людям совершенно неизвестны.

Итак, мы видим, что субъективное качество переживания цвета не улавливается колориметрической моделью цветового зрения человека. Очевидно, что оно не может быть выражено никакой другой, более общей или более совершенной, моделью. Ясно также, что этот вывод в равной мере относится и ко всем другим субъективным состояниям человека. Выходит, что в каждом субъективном состоянии человека есть нечто непознаваемое, причем как раз то, что придает им качество субъективности, что обеспечивает причастность этих состояний к нашему собственному «я».

Не подрывает ли этот вывод ценности теории интеллекта как средства искусственного воспроизведения субъективной стороны человеческого разума? Если мы не в состоянии описать синеву ощущения такой, какой она есть сама по себе, то как же тогда можно надеяться, что удастся изготовить искусственно полноценное ощущение цвета, которое автоматически действующее мыслящее устройство могло бы субъективно переживать? Мы можем ответить на эти вопросы следующим образом. Приходится признать, что субстанция ощущения, которая делает его субъективным переживанием, не может быть описана в научных терминах. Тем не менее, это ограничение не является препятствием для искусственного воспроизведения феномена личного переживания ощущений и вообще — любых субъективных состояний. Попытаемся обосновать последнее утверждение.

Предположим на минутку, что по достаточно обширным математическим описаниям теории интеллекта (которые, как мы надеемся, она разрабатывает в будущем) построена полноценная в смысле разумности физических реакций копия испытуемого в виде некоторого действующего технического устройства. Судя по результатам наблюдений за внешними действиями такой копии испытуемого, в ней обитает некий дух, личность, субъект. Последний утверждает, что он переживает различные психологические состояния, в том числе и ощущения цвета с характерными для них качествами синевы, желтизны и т.п. Так ли это на самом деле или нет, внешний наблюдатель проверить не может. Похожи ли ощущения, переживаемые копией испытуемого (если таковые вообще возникают), на ощущения самого испытуемого, никто сказать не в состоянии.

В копии испытуемого можно обследовать то место в механизме ее устройства, в котором разыгрываются образы световых излучений, воспринимаемых ею. Найти это место не составит особого труда, поскольку подразумевается, что копия испытуемого создана людьми, а, следовательно, ее устройство и функционирование полностью им известны. Отыскать же точное место возникновения зрительных образов в организме испытуемого было бы делом гораздо более сложным и вряд ли выполнимым при современном уровне знаний об анатомии и физиологии мозга человека. Именно по этой причине мы обращаемся в данном рассуждении не непосредственно к испытуемому, а к его гипотетической искусственно созданной копии.

Зададимся вопросом: какая материальная субстанция служит носителем образа светового излучения, воспроизведенного в копии испытуемого? Отвечая на него, мы можем сказать, что это, к примеру, — магнитное поле, которое в каждой своей точке характеризуется направлением и плотностью силовых линий и т.п. Но все это — не более чем слова, коды чисел и другие символы. Никакие описания не смогут охарактеризовать носитель образа таким, каков он есть сам по себе, т.е. как «вещь в себе». Если мы перейдем к другим единицам измерения или к иной координатной системе, то получим иное описание того же самого физического объекта. Как видим, свойства физического объекта с успехом описываются, однако сам объект, его природа, субстанция ускользает от описания точно так же, как ранее ускользала от описания природа субъективного ощущения. Получается, что физический объект, как и ощущение, можно описать только с точностью до обозначений.

4. Принцип тождества

Таким образом, природа физических процессов как «вещей в себе» и природа соответствующих им субъективных состояний человека, рассматриваемых как существующие сами по себе, сходна в том смысле, что и та и другая не поддается научному описанию. По этой причине мы не знаем, чем же природа объективного и соответствующего ему субъективного отличается, и, весьма вероятно, что знать этого не можем. Никто не сомневается в том, что субъективное и соответствующее ему объективное состояния существуют реально, но сверх этого ничего об их природе сказать не удастся. Вместе с тем, в силу психофизического параллелизма, все физически наблюдаемые свойства субъективных и соответствующих им объективных состояний совпадают. Следовательно, все, что известно (а возможно, и все, что можно узнать) о физических процессах и соответствующих им субъективных состояниях,

идентично. Ничто не мешает нам сделать еще один шаг и предположить, что не только свойства, но и природа, субстанция физических и соответствующих им психических состояний совпадают. Такое утверждение недоказуемо и неопровержимо, но оно существенно упрощает картину мира. Основываясь на сказанном, мы считаем возможным сформулировать следующий **принцип тождества**: *природа физических процессов — носителей субъективных состояний и природа самих субъективных состояний одна и та же.*

Согласно принципу тождества, субъективное и соответствующее ему объективное — это две стороны одной и той же медали (лучше было бы сказать: две стороны одной плоскости), это один и тот же предмет или процесс, рассматриваемый с разных сторон. На самом деле есть только один предмет, но до сих пор предполагалось, что их два. Возьмем, к примеру, субъективное зрительное ощущение, которое разыгрывается в поле зрения. Его можно рассматривать как поле точек, в каждой из которых имеется некоторый цвет, меняющийся во времени. Этому ощущению в организме человека соответствует (по современным представлениям — в коре затылочных долей головного мозга) объективное поле физических состояний некоторой природы, меняющееся во времени синхронно со зрительным ощущением. Принцип тождества в данном случае гласит, что в случае равенства цветов в произвольно взятых двух точках поля зрения должны совпадать также и физические состояния в соответствующих точках коры головного мозга. Если же цвета не совпадают, то обязательно должны различаться и соответствующие им физические состояния.

Если принять принцип тождества, то самые темные и трудные (они же и ключевые) проблемы теории интеллекта легко и естественно решаются. Будет ли машина — копия испытуемого, созданная по математическим описаниям теории интеллекта, обладать собственной психикой, субъективно переживать свои внутренние состояния? Да, будет, потому что воспроизведенные в ней физические состояния, играющие роль образов физических процессов, как раз и являются, согласно принципу тождества, такими субъективными переживаниями. Можно ли субъективные состояния человека изучать физическими методами? Да, можно, поскольку, согласно принципу тождества, любое субъективное состояние есть также и состояние физическое. Не будут ли ущербными описания субъективных состояний человека, даваемые теорией интеллекта, из-за того, что природа субъективности психологических состояний ими не улавливается? Нет, не будут, поскольку, согласно принципу тождества, описать природу субъективности переживания — это то же

самое, что узнать, каков физический объект сам по себе. Последнего физика достичь не может, но из-за этого никакого ущерба она не терпит.

Мы видим, что принцип тождества играет роль «подпорки» при решении теорией интеллекта важнейших проблем. В связи с этим он заслуживает самого тщательного обсуждения. Прежде всего возникает вопрос, можно ли принцип тождества обосновать методами физики. Ответ отрицателен уже хотя бы потому, что в принципе тождества идет речь о физических процессах и субъективных состояниях как о «вещах в себе». И те и другие с этой стороны не представляется возможным наблюдать физически. Нельзя обосновать физическими методами тождество природы субъективных и соответствующих им объективных состояний, поскольку эту природу невозможно воспринять физическими приборами. Однако в этом отрицательном ответе содержится и нечто положительное: если тождество природы субъективного и соответствующего ему объективного физическими методами не может быть доказано, то оно, вместе с тем, не может быть ими и опровергнуто. Поэтому авторы могут не опасаться за судьбу сформулированного ими «принципа тождества», ибо последний неуязвим для критики со стороны физики. Вместе с тем, принцип тождества, как мы только что видели, не является бесполезным утверждением. Видимо, не все важные для физики утверждения могут быть ею доказаны. Некоторые из них приходится принимать как априорные. Принцип тождества относится к числу именно таких утверждений.

Как же могло случиться, что принцип тождества, будучи, по существу, искусственно придуманным утверждением, не имеющим опоры в опыте, все же помогает решать важные проблемы теории интеллекта? В науке такого рода принципы не являются новостью. Физики их хорошо знают и с пользой применяют в своих исследованиях. Например, таким принципом является так называемая «брита Оккама». Формулируется он следующим образом: «сущностей не следует умножать без необходимости». Иными словами, если в какой-нибудь науке все можно объяснить без допущения той или иной гипотетической сущности, то и не следует ее вводить. Известный философ, логик и исследователь интеллекта Бертран Рассел пишет: «Я лично убедился в необычайной плодотворности этого принципа» [2, с. 491].

К принципу тождества можно прийти, применяя «бриту Оккама». Есть две сущности: психическое состояние и соответствующее ему физическое состояние. Общепринято, что любое психическое состояние всегда сопровождается соответствующим ему физическим состоянием (так называемый

«психофизический параллелизм»). Никем не оспаривается, что носителем каждого психического состояния служит свое особое состояние физическое. Психическое всегда неразлучно с физическим как тень с пешеходом в солнечный день. Принцип тождества утверждает лишь то, что в данном случае нет надобности вводить две разные сущности, достаточно использовать одну из них, поскольку сущности эти невозможно ни различить, ни разделить. История науки знает примеры слияния двух сущностей в одну. Так, в древние времена планета Венера принималась за два разных объекта: Вesper — вечерняя звезда и Люцифер — утренняя звезда. Потребовались длительное время и большие интеллектуальные усилия астрономов, чтобы объединить эти два понятия в одно.

5. Применение принципа тождества

Действенность принципа тождества основана на том, что он устраняет воображаемую пропасть между психическим и физическим, которая возникла из-за «необоснованного» разделения состояний на психические и физические. Употребление в предыдущей фразе слова «необоснованный» без кавычек было бы несправедливостью по отношению к тем, кто когда-то ввел такое разделение. Разве в свое время было необоснованным введение двух названий для одной и той же (как это мы знаем теперь) планеты Венеры? Это была совершенно естественная мера предосторожности в условиях дефицита знаний об этих, как казалось тогда, различных объектах.

Введение двух разных имен для одного и того же предмета — еще не ошибка. Ничто не мешает отождествить содержание этих имен впоследствии, когда для этого накопится достаточный объем знаний. Разделение состояний на психические и физические было вполне естественным на начальных стадиях развития науки об интеллекте. Но теперь, когда о свойствах психических и физических состояний известно уже достаточно много, и, тем не менее, не возникает необходимости в различении этих понятий, вполне созрели условия для провозглашения тождества этих двух «сущностей». Чтобы воображаемая пропасть между психическим и соответствующим ему физическим не отравляла в будущем жизнь специалистам по теории интеллекта, пропасть эту надо ликвидировать. А чтобы сделать это, необходимо принять принцип тождества.

Формулировался ли принцип тождества теми, кто ранее занимался исследованием интеллекта? Наши поиски в этом направлении не дали ясного положительного ответа. Мы можем привести только несколько высказываний двух классиков естествознания, которые с большой натяжкой можно рассматривать как вклад в провозглашение принципа тождества. Шредингер в послесловии к

книге [3, с. 122] развивает взгляд, согласно которому «Я», взятое в самом широком значении этого слова — т.е. каждый сознательный разум, когда-либо говоривший или чувствовавший «я», представляет собой не что иное, как субъект, могущий управлять «движением атомов» согласно законам природы».

Казалось бы, это высказывание не имеет ничего общего с принципом тождества. Однако далее Шредингер пишет: «Само по себе это представление не ново. Насколько мне известно, наиболее ранние упоминания о нем насчитывают уже, по крайней мере, 2500 лет, если не больше. Начиная с древних великих Упанишад, представление о том, что Атман = Брахман (т.е. личная индивидуальная душа равна вездесущей, всепостигающей, вечной душе), ... считалось квинтэссенцией глубочайшего прозрения в то, что происходит в мире» [3, с. 123]. Если под словом «Атман» понимать природу субъективных состояний, а под словом «Брахман» — природу соответствующих им физических состояний, то мы получим принцип тождества. Однако Шредингер, как нам представляется, вкладывает в эти слова иной смысл. Это тем более вероятно, что в другом месте этой же книги Шредингер отрицает возможность естественнонаучного объяснения психофизического параллелизма. Он пишет: «На мой взгляд, природа этого параллелизма лежит в стороне от области естественных наук и, весьма возможно, за пределами человеческого понимания» [3, с. 18]. Между тем, принцип тождества как раз и объясняет существование этого параллелизма. Рассел весьма близок к тому, чтобы признать в субъективном восприятии физический объект. Он отличает восприятие от знания о нем. Рассел пишет: «Суть голого явления — просто определенные цветные пятна. Они ассоциируются с образами осязания, они могут вызвать слова и могут стать источником воспоминаний. Психический объект восприятия, поскольку он наполнен образами осязания, становится «объектом», о котором предполагают, что он физический; наполненный же словами и воспоминаниями, он становится «восприятием», которое является частью «субъекта» и считается духовным. Психический объект восприятия — это именно явление; он не бывает ни истинным, ни ложным. Заполненный словами, он есть суждение, и способен быть истинным или ложным. Это суждение я называю «суждением восприятия». Предложение «знание есть восприятие» следует истолковать как означающее «знание есть суждение восприятия». Только в этой форме оно способно в грамматическом отношении быть правильным» [2, с. 174].

В другом месте этой же книги Рассел спорит с утверждением о нефизичности восприятий: «Не доказано, например, что цвета, существенно отличные

от тех, которые мы видим, не могут существовать, будучи невидимы. По физиологическим основаниям мы можем полагать, что этого не происходит, но такое основание эмпирично; что касается логики, то нет оснований отрицать наличие цветов там, где нет глаз или мозга» [2, с. 671]. Рассел ставит под сомнение утверждение, что «ничто не может быть одновременно и духовным и материальным» [2, с. 672]. Он добавляет: «Быть наблюдаемым или быть воспринимаемым объектом означает просто воздействовать определенным образом, и нет логических оснований отрицать за всеми событиями способности к воздействиям такого рода» [2, с. 672].

Термин «принцип тождества» широко используется в философской литературе и связывается с именем Шеллинга, который, пытаясь преодолеть дуализм кантовской и фихтевской систем, выдвинул новый исходный принцип монистической философии — абсолютное тождество субъективного и объективного, идеального и реального. Принцип тождества мышления и бытия лежит также в основе гегелевской системы. В настоящее время идеи тождества мышления и бытия проповедуют отдельные школы неотомизма. Марксистская философия обосновывает свой монизм, исходя из идей материального единства мира. Ленин пишет: «...противоположность материи и сознания имеет абсолютное значение только в пределах очень ограниченной области: в данном случае исключительно в пределах основного гносеологического вопроса о том, что признать первичным и что вторичным. За этими пределами относительность данного противоположения несомненна» [4, с. 133]; «За этими пределами оперировать с противоположностью материи и духа, физического и психического, как с абсолютной противоположностью было бы громадной ошибкой» [4, с. 231].

Выводы

Отправляясь от принципа тождества, можно ответить на множество интересных вопросов. Возникают ли субъективные состояния в организме человека, или же они существуют в каком-то ином смысле, нежели физические состояния? Ответ однозначен: как физические процессы субъективные состояния, естественно, следует искать только в организме человека. Можно ли и как отыскать физический процесс, совпадающий с данным субъективным состоянием? Можно, для этого надо иметь полное математическое описание данного субъективного состояния (включая и все его связи с другими субъективными состояниями) и полное математическое описание физического состояния (включая и все его связи с другими физическими состояниями), о которых предполагается, что они соответствуют друг другу. Если эти математические

описания с точностью до обозначений совпадают друг с другом, значит, мы имеем дело с одним и тем же объектом. В противном случае — это разные объекты.

Можно ли непосредственно воспринять произвольно выбранное физическое состояние таким, каким оно есть на самом деле, т.е. как «вещь в себе»? Можно, для этого надо данное физическое состояние использовать в роли материального носителя субъективного состояния. Будучи подсоединен к этому носителю, механизм разума (неважно чьего) непосредственно воспримет данное физическое состояние в виде некоторого субъективного состояния. Природа этого субъективного состояния не зависит от воспринимающего разума, поскольку, согласно принципу тождества, она совпадает с природой соответствующего физического состояния.

Какова природа произвольно взятой «вещи в себе»? Она такова, какой психологически воспринимает ее разум, использующий эту «вещь в себе» в роли носителя субъективного состояния. И никакой иной она быть не может. Бесполезно искать природу «вещи в себе» еще где-то. Если б нашлась иная природа, отличная от первой, то мы имели бы две различные природы одного и того же физического состояния, иными словами, — не одну, а две «вещи в себе». Если бы мы не воспринимали физические процессы непосредственно в виде субъективных состояний, то никогда б не узнали ни о существовании самих себя, ни о существовании окружающего нас внешнего мира. Любые два разума будут воспринимать одинаковые физические состояния как идентичные субъективные состояния, если первые являются носителями последних. Если два субъекта имеют, к примеру, физически неотличимые механизмы зрения, то с необходимостью они будут реагировать на одинаковые световые излучения одинаковыми ощущениями цвета.

Находится ли субъективное состояние за воспринимающим аппаратом, реагирующим на физический его носитель, или перед ним? Перед ним, т.к. воспринимающий аппарат анализирует физическое состояние, которое, согласно принципу тождества, совпадает с соответствующим ему субъективным состоянием. Субъективное состояние (оно же и физическое) объективно существует вне зависимости от того, воспринимается ли оно каким-либо разумом или нет. Воспринимающий аппарат разума играет роль окна, из которого открывается вид на

находящиеся перед ним предметы. Если допустить, что субъективное состояние находится за воспринимающим аппаратом и является его продукцией, то получаем противоречие в виде регресса в бесконечность: потребуется еще один прибор для восприятия полученного субъективного состояния и т.д. Утверждать, что субъективное состояние есть не сам физический процесс, а лишь образ этого физического процесса — носителя этого состояния, столь же нелепо, как говорить, что измеренная длина стола характеризует не сам стол, а лишь образ стола.

Авторы должны сознаться, что последовательное применение принципа тождества приводит к весьма непривычным представлениям. Однако создается впечатление, что разработчику теории интеллекта такие представления навязываются принудительно вне зависимости от того, нравятся они ему или нет. Здесь очень многое требует еще осмысления и дальнейшей проработки.

Список литературы: 1. Шабанов-Кушнарченко, Ю.П. Начала теории интеллекта. Проблемы и перспективы [Текст]/ Ю.П. Шабанов-Кушнарченко. — Рукопись деп. в ВИНТИ 28.06.82, №3324 — 82. Деп. — 210 с. 2. Рассел, Б. История западной философии [Текст]/ Б. Рассел — М.: ИЛ, 1959. — 932 с. 3. Шредингер, Э. Что такое жизнь с точки зрения физики? [Текст]/ Э. Шредингер — М.: ГИИЛ, 1947. 4. Ленин, В.И. Материализм и эмпириокритицизм [Текст]/ В.И. Ленин — М.: Госполитиздат, 1967.

Поступила в редколлегию 17.02.2011.

УДК 519.7

Проблеми моделювання суб'єктивних станів / М.Ф. Бондаренко, Н.П. Кругликова, Н.Е. Русакова, Ю.П. Шабанов-Кушнарченко // Біоніка інтелекту: наук.-техн. журнал. — 2011. — № 2 (76). — С. 24-32.

У статті розглядається природа інтроспекції та коректується постановка одної із задач теорії інтелекту з метою приведення її рішення у відповідність з інтроспективними даними на прикладі вивчення механізму теорії кольору.

Бібліогр.: 4 найм.

UDC 519.7

Problems of design of the subjective states / M.F. Bondarenko, N.P. Kruglikova, N.E. Rusakova, Yu.P. Shabanov-Kushnarenko // Bionics of Intelligence: Sci. Mag. — 2011. — № 2 (76). — P. 24-32.

In the article nature of introspection is examined and raising of one of tasks of theory of intellect is corrected with the purpose of bringing its decision over in accordance with introspective information.

Ref.: 4 items.