

ЭНАКТИВИЗМ И ПРОБЛЕМА СОЗНАНИЯ*

Иванов Дмитрий Валерьевич – кандидат философских наук, старший научный сотрудник. Институт философии РАН. Российская Федерация, 109240, г. Москва, ул. Гончарная, д. 12, стр. 1; e-mail: ivdmitry@mail.ru



Статья посвящена энактивистскому подходу к проблеме сознания, проблеме натуралистического объяснения феноменальных аспектов нашего опыта. Согласно классической когнитивной науке, мы можем объяснить все ментальные состояния как функциональные, репрезентационные состояния. Многие философы не согласны с таким взглядом. Они демонстрируют, что феноменальные качества сознательных состояний не могут быть поняты в терминах ментальных репрезентаций. Современные дебаты относительно природы феноменальных качеств являются дебатами между представителями репрезентационизма и анти-репрезентационизма. Аргументы, предложенные анти-репрезентационистами, демонстрируют неудовлетворительность классического репрезентационистского подхода. Однако это не означает, что мы должны принять существование квалиа – особых нерепрезентационных феноменальных свойств. Вполне возможно защитить репрезентационизм, переосмыслив природу ментальных репрезентаций. В этой статье исследуется трансформация понятия ментальной репрезентации в когнитивной науке за последние несколько десятилетий. В работе демонстрируется, что понятие репрезентации-в-действии, обсуждаемое в энактивистских теориях, может помочь нам найти основания для натуралистического понимания сознательного опыта.

Ключевые слова: философия сознания, философия когнитивной науки, энактивизм, проблема сознания, феноменальное сознание, квалиа, ментальная репрезентация

ENACTIVISM AND THE PROBLEM OF CONSCIOUSNESS

Dmitry Ivanov – PhD in Philosophy, senior research fellow. Institute of philosophy, Russian Academy of Sciences. 12/1 Goncharnaya St., Moscow, 109240, Russian Federation; e-mail: ivdmitry@mail.ru

The paper deals with the enactivist approach to the problem of consciousness. The problem of consciousness is the problem of naturalistic explanation of phenomenal aspects of our experience. According to classical cognitive science, we can explain all mental states as functional, representational states. Many philosophers disagree with this view. They demonstrate that phenomenal qualities of conscious states cannot be understood in terms of mental representations. Contemporary debates about the nature of phenomenal qualities are the debates between representationalists and anti-representationalists. The arguments proposed by anti-representationalists demonstrate the insufficiency of classical representational approach. But it doesn't mean that we should accept the existence of qualia – special non-representational phenomenal properties. It is possible to defend representationalism by reconsidering the nature of mental representations. This article examines the transformation of the concept of mental representation in cognitive science over the last few decades. It demonstrates that the notion of representation in action discussed in enactivist theories can help us to provide the foundation for naturalistic understanding of conscious experience.

* Статья подготовлена при поддержке РГНФ, проект № 16-03-00373а «Экстерналистская парадигма в философии сознания и когнитивной науке».



Keywords: philosophy of mind, philosophy of cognitive science, enactivism, the problem of consciousness, phenomenal consciousness, qualia, mental representation

Для того чтобы решить психофизическую проблему, мы должны объяснить феноменальную природу сознания. Подобное объяснение предполагает демонстрацию логической выводимости утверждений о феноменальных аспектах сознания из утверждений о естественнонаучных фактах. Однако в последние несколько десятилетий многие философы указывали на то, что выработка подобного объяснения представляет собой особую сложную проблему. Ведущий современный философ сознания Дэвид Чалмерс обозначил данную проблему термином «трудная проблема сознания». По мнению ряда философов, такое объяснение в принципе невозможно предложить.

Одним из первых к обсуждению проблемы объяснения сознания обратился философ Томас Нагель. В статье 1974 г. «Каково быть летучей мышью?» Нагель отмечает следующее [Nagel, 1974]. Знание о данностях сознания, которое доступно субъекту сознания, является знанием, полученным из уникальной перспективы – перспективы первого лица. Только субъект определенного сознательного состояния знает, каково это – пребывать в данном состоянии. Наличие этой перспективы является существенным свойством сознания. Именно она конституирует нашу субъективность. Знание о сознании, которое лежит в основе любой естественнонаучной теории, является знанием, полученным из перспективы третьего лица. Это – знание объективных, intersubjectively наблюдаемых фактов (поведенческих, физических, функциональных). Пытаясь выработать натуралистическое объяснение сознания, мы прежде всего опираемся на последний тип знания, мы стремимся редуцировать субъективные аспекты восприятия к каким-либо объективным фактам, тем самым неизбежно упуская наиболее существенные аспекты сознательного опыта. Например, изучая то, как летучие мыши ориентируются в мире посредством эхолокации, мы можем узнать все факты, касающиеся их поведения и нейрофизиологии, однако знание того, каково это – воспринимать объекты подобным образом, будет по-прежнему для нас недоступно. Как отмечает другой философ, Джозеф Левин, в нашем объяснении сознательного опыта другого существа всегда будет наличествовать провал (explanatory gap) [Levin, 1986].

Аргументы, предложенные Т. Нагелем, Дж. Левиным, Ф. Джексоном, разработавшим интересный мысленный эксперимент «Мэри в черно-белой комнате» [Jackson 1982], можно отнести к группе эпистемологических аргументов. В них делается акцент на познавательной ситуации, в которой мы оказываемся, пытаясь объяснить сознание. Существует также группа метафизических аргументов. Все



современные аргументы этой группы по сути являются вариациями картезианского аргумента в пользу различия души и тела. В них из некоторой мыслимой ситуации делаются выводы об онтологической природе феноменальных аспектов сознания.

Наиболее обсуждаемым аргументом в последние два десятилетия был аргумент от мыслимости зомби, разработанный Д. Чалмерсом [Chalmers, 1996]. Если Декарт пытался представить сознание без тела, то Чалмерс, напротив, предлагает нам помыслить существование тел без феноменального сознания. Он просит нас представить существование зомби, наших полных физических и функциональных двойников, обладающих всеми психическими процессами, которые есть у нас, но лишенных феноменальных качеств, которыми окрашивается наша психическая жизнь. Подобно Декарту из мыслимости такой ситуации Чалмерс делает вывод о ее логической возможности и о нетождественности феноменальных качеств каким-либо физическим или функциональным характеристикам организма.

Эпистемологические и метафизические аргументы, в которых демонстрируется нередуцируемость феноменальных качеств сознательного опыта к физическим или функциональным свойствам, убедили многих философов занять антинатуралистическую позицию по вопросу о природе сознания. По сути, эти аргументы способствовали возрождению дуализма в философии сознания. Например, свою позицию Д. Чалмерс обозначает как натуралистический дуализм. Однако современный дуализм отличается от классического дуализма субстанций. Этот вид дуализма является дуализмом свойств. Представители данного вида дуализма принимают единую физикалистскую онтологию, однако эта онтология допускает, что в мире помимо физических свойств и объектов существуют особые ментальные качества, несводимые к физическим свойствам. Наиболее популярными вариантами дуализма свойств являются эпифеноменализм и панпсихизм. С точки зрения эпифеноменализма, хотя возникновение ментальных качеств каузально обусловлено физическими процессами, однако сами эти качества каузально не воздействуют на физические объекты. Согласно панпсихизму, ментальные свойства следует рассматривать не как возникающие на каком-то этапе развития физических систем, а как фундаментальные свойства всех объектов, существующих в универсуме.

В современной философии сознания для обозначения феноменальных качеств, нередуцируемых к физическим и функциональным свойствам, используется латинский термин «квалиа». Однако многие философы, оппоненты антинатурализма, не согласны с тем, что для понимания феноменальных аспектов сознательного опыта необходимо вводить такие сущности, как квалиа. По их мнению, данные аспекты могут быть проинтерпретированы с помощью понятия ментальной репрезентации, обозначающего нашу способность определенным



образом представлять какую-либо ситуацию. По сути, современная полемика по вопросу о природе феноменальных аспектов сознательного опыта это – дискуссия между сторонниками репрезентационизма и анти-репрезентационизма. Представители анти-репрезентационизма справедливо указывали, что классическое понятие ментальной репрезентации, проинтерпретированное, как правило, в качестве внутреннего функционального состояния, определенным образом представляющего некоторое положение дел в мире, неспособно учесть все аспекты сознательного опыта. Однако из этого не следует, что мы должны допускать существование квалиа. Возможно, нам необходимо пересмотреть понятие ментальной репрезентации.

Долгое время это понятие играло главную роль в когнитивной науке. Репрезентационизм являлся методологическим основанием классической когнитивной науки. Однако в последние несколько десятилетий подобное положение дел стало меняться. В когнитивной науке появились подходы, пересматривающие природу ментальных репрезентаций. Одним из таких подходов является энактивизм.

Появление энактивистских теорий в 90-х годах прошлого века и активное их развитие в 2000-х гг. можно рассматривать как третий этап в развитии когнитивной науки. На первом этапе развития когнитивной науки, на этапе ее возникновения в середине прошлого века, сфера когнитивных процессов понималась как функциональная система, осуществляющая особые вычислительные операции над внутренними репрезентациями. Когнитивная наука возникает как оппозиция доминировавшему в первой половине двадцатого века бихевиористскому подходу в психологии. Как известно, бихевиоризм стремился показать, что ментальные состояния могут быть сведены к определенным поведенческим реакциям либо к поведенческим диспозициям. Существенным недостатком такого взгляда было игнорирование того факта, что некоторые ментальные состояния являются скорее причинами определенного поведения, чем самим поведением. Благодаря когнитивной революции 1950-х годов ученые начали рассматривать ментальные состояния не как фикции, которые должны быть элиминированы, а как внутренние репрезентационные состояния, которые опосредуют воспринимаемые стимулы и поведенческие ответы.

Особое влияние на развитие когнитивной науки оказали исследования искусственного интеллекта. Как отмечает Энди Кларк, три момента оказались определяющими для развития этих исследований и когнитивной науки в целом [Clark, 2001]. Во-первых, особая роль в понимании когнитивных процессов закрепляется за формальной логикой. Формальная логика представляет собой систему простых символов, правил образования из них более сложных символов и синтаксических правил вывода, позволяющих от одних символов переходить



дить к другим. Обращаясь к формальной логике, ранние ученые-когнитивисты предлагали интерпретировать когнитивные процессы как систему оперирования особыми символическими объектами по синтаксическим правилам. Например, Саймон и Ньюэлл рассматривали когнитивную сферу как физико-символьную систему, т. е. систему, содержащую физические символы, стоящие в отношении денотации к физическим объектам внешнего мира, а также организованные по соответствующим правилам процессы оперирования данными символами [Newell, Simon, 1981]. Преимущество подобного понимания когнитивных процессов заключается в том, оно позволяет объяснить семантическую составляющую интеллектуальной деятельности исключительно синтаксическим образом. Семантический момент когнитивных процессов определяется отношением денотации символических объектов к объектам внешнего мира и тем, какую роль эти объекты играют в системе синтаксических связей.

Вторым моментом, повлиявшим на формирование когнитивной науки, была формализация вычислительных процессов Тьюрингом. Введение понятия «машина Тьюринга», обозначающего устройство, состоящее из ленты, хранящей символичные данные, и считывающей головки, способной двигаться вдоль ленты, считывать и записывать на нее новые символы, позволило продемонстрировать возможность физического устройства, которое, следуя простым синтаксическим правилам, способно решить любую четко специфицированную проблему. Третьим моментом, определившим развитие и исследований искусственного интеллекта, и когнитивной науки, было изобретение конкретных машин с архитектурой фон Неймана, позволяющих реализовать формализованные ранее идеи вычислительных процессов.

Как отмечает Кларк, притягательность символического подхода в когнитивной науке определялась не только тем, что он позволял ученым претендовать на объяснение природы интеллектуальной деятельности, но и тем, что он обещал дать объяснение предсказательной силе нашего повседневного менталистского дискурса. По мнению Джерри Фодора, наш менталистский дискурс позволяет нам успешно предсказывать и объяснять поведение агента, наделенного убеждениями и желаниями, поскольку он воплощает принципы работы внутренних когнитивных процессов. Нашим убеждениям, желаниям и другим ментальным состояниям, описываемым обыденным языком, соответствуют внутренние символические состояния, обладающие каузальной силой. Наши обыденные предсказания поведения интеллектуального агента работают, поскольку им соответствуют реальные каузальные процессы в мозге. Фодор так пишет об этом: «В первом приближении, думать “сейчас пойдет дождь, поэтому я пойду внутрь” значит обладать образцом ментальной репрезентации, означающей я



пойду внутрь, причинно обусловленным, в определенном смысле, образом ментальной репрезентации, означающей *сейчас пойдет дождь*» [Fodor, 1987, p. 17].

Резюмируя, можно представить классическую когнитивную науку следующим образом.

1. Когнитивные процессы существенным образом предполагают репрезентационные и вычислительные процессы. Репрезентации, в самом общем смысле, являются символическими структурами, а вычислительные процессы, в общем смысле, являются правилами (подобными правилам логики) манипуляции этими символическими структурами.

2. Хорошей моделью для понимания функционирования системы сознание-мозга является компьютер – т. е. машина, базирующаяся на тех же логических основаниях, что и машина Тьюринга, и на фоннеймановской архитектуре цифровых компьютеров. Подобные машины идеально подходят для манипулирования символами в соответствии с определенными правилами. Компьютерная метафора предполагает, что система сознание-мозг на процессуально-информационном уровне может быть понята как вид цифрового компьютера. Проблемой для когнитивной психологии является определение той программы, которая выполняется нашими мозгами [Churchland, Sejnowsky, 1990, p. 138].

Второй этап развития когнитивной науки следует связывать с коннекционистским поворотом, произошедшим в 80-х годах прошлого века. В результате этого поворота внутренние репрезентации перестали рассматриваться многими учеными как особые символические объекты, а вычислительные процессы стали пониматься как параллельно распределенные.

Основной недостаток компьютерной метафоры, на которую опирается классическая когнитивная наука, заключается в том, что наш мозг во многих существенных аспектах не подобен компьютерам, основанным на архитектуре фон Неймана. Выход из строя одного из элементов, определяющих работу компьютера, или изменение единственного символа в программном коде могут привести к отказу работы всей системы. Такого мы не наблюдаем в случае с мозгом. Даже значительные повреждения мозга не обязательно приводят к прекращению его работы. Нарушение работы какого-либо отдела мозга вполне может быть компенсировано другими его частями. Мозг представляет собой не набор дискретных элементов, последовательно выполняющих специфические функции (центральный процессор, память и т. д.), а сеть нейронов, характеризующуюся параллельной обработкой информации, при которой в возбуждение приходит вся сеть целиком. Выпадение из работы какого-либо участка сети приводит к изменению паттернов ее возбуждения, но не выводит ее це-



ликом из строя. Как кажется, рассмотренный в эволюционном плане, параллельный способ обработки информации в большей степени соответствует задаче организма выдавать оперативные ответы (т. е. в ситуации нехватки времени) на сложные вызовы окружающей среды, чем последовательное решение адаптационных проблем.

Объясняя природу когнитивных процессов, коннекционисты пытаются прежде всего смоделировать работу мозга, используя при этом понятие искусственных нейронных сетей. Искусственная нейронная сеть является простейшей моделью мозга. Она состоит из единиц, или узлов, являющихся аналогами нейронов, и связей между ними, подобных синаптическим связям. Существуют три группы единиц нейронной сети. Единицы, имеющиеся на входе, ответственные за первоначальную обработку информации, являются аналогами сенсорных нейронов. Единицы на выходе подобны нейронам, отвечающим за моторные реакции. Третья группа обозначается как группа скрытых нейронов, которые выполняют функцию хранения и переработки информации. Каждая единица на входе получает от внешней среды некоторое значение возбудительной или тормозящей активности, которое передается скрытым единицам, с которыми она связана. Значение активности скрытых единиц зависит от получаемых значений активности от входных единиц и других скрытых единиц. Это значение скрытые единицы передают единицам на выходе. Суммарная активность всей сети зависит не только от значений активности, передаваемых единицами, но и от весов, которыми обладают связи между единицами. По сути, вес означает здесь силу связи между единицами. От него зависит, какое значение активности будет передано.

Воспользовавшись цитатой из работы Фодора и Пылишина, можно следующим образом резюмировать основные моменты коннекционизма.

Коннекционистские системы представляют собой сети, состоящие из огромного числа простых, тесно связанных друг с другом «элементов». Относительно элементов сетей и их связей обычно принимаются следующие допущения: 1) каждый элемент получает на входе сети некоторое значение активности (возбудительной, тормозящей или и той, и другой); 2) с помощью элементов сети ищется суммарная активность, а состояние сети меняется как функция (как правило, пороговая) от этой суммарной активности; 3) предполагается, что каждая связь может менять значение активности, которую она передаст; при этом модуляция активности является функцией от ее внутреннего (но подверженного модификациям) свойства, именуемого «весом». Таким образом, активность на входных нитях сети представляет собой некоторую нелинейную функцию от состояния активности ее источников; 4) в целом поведение сети можно описать с помощью функциональной зависимости от начального состояния активности элементов сети и весов по всем их связям; сами эти веса, и только они, образуют память системы [Фодор, Пылишин, 1996, с. 232].



По сути, коннекционизм является вариантом такой философской теории, как элиминативный материализм. Представители элиминативизма, к которым могут быть отнесены такие философы, как Селларс, Куайн, Фейерабенд, Рорти, Пол и Патриция Черчленды, исходят из того, что наши повседневные представления о природе психических процессов формируют определенную теорию, которую элиминативисты обозначают термином «народная психология». Подобно любой другой теории народная психология может оказаться ложной теорией. Согласно элиминативистам именно таковой она и является. В отличие от редуктивного материализма, представители которого принимают существование объектов народной психологии, но стремятся редуцировать их к материальным процессам, элиминативисты полагают, что неверной является сама онтология народной психологии. Ее объекты, такие как мысли, эмоции, желания и т. д., используемые для объяснения нашего когнитивного поведения, должны не объясняться посредством редукции, а элиминироваться подобно тому, как были элиминированы такие сущности, как флогистон или, скажем, *elán vital*.

Коннекционизм подобен элиминативному материализму в том, что он отрицает существование внутренних ментальных репрезентаций, понимаемых как особые символические объекты, к которым якобы отсылает повседневный язык описания ментальных состояний. С точки зрения коннекционизма понимание процессов обработки информации и механизмов репрезентации, имеющих в нейронных сетях, не должно осуществляться в терминах, навязываемых нам повседневным языком. Как полагают элиминативисты, объяснение когнитивных процессов должно опираться на совершенно иную онтологию, отличную от онтологии, представляемой народной психологией. Коннекционисты не отказываются от понятия репрезентации. Однако репрезентации, наличествующие в нейронных сетях, являются субсимволическими. Репрезентационные состояния нейронных сетей не являются теми ментальными состояниями, которые мы обозначаем, используя язык народной психологии.

Третий этап развития когнитивной науки следует связывать с появлением энактивистских теорий. Подобно коннекционистам представители энактивизма отказываются от классического понимания ментальных репрезентаций как внутренних состояний, представляющих определенным образом мир. Однако энактивисты идут дальше в пересмотре оснований когнитивной науки.

Как уже отмечалось, репрезентационизм являлся одним из ключевых методологических принципов классической когнитивной науки. Другим не менее важным принципом данной науки был компьютерционизм – представление о том, что когнитивные процессы являются особым рода вычислительными операциями с ментальными репрезентациями. Нужно отметить, что, хотя коннекционизм и критикует



тот вариант репрезентационизма, на котором основывается классическая когнитивная наука, он по-прежнему остается в рамках компьютерционистского подхода.

Компьютерционизм является вариантом функционализма – философской теории, согласно которой ментальные состояния есть функциональные состояния мозга. Одно из главных положений функционализма заключается в том, что функциональные состояния инварианты по отношению к тому физическому носителю, на котором они реализованы. Это значит, что ментальные состояния, будучи функциональными состояниями, в определенном смысле не зависят от своей физической реализации. Неважно, с какой физической системой мы имеем дело – это может быть мозг, компьютер или какой-либо иной объект, – если эта система находится в определенном функциональном состоянии, то она обладает соответствующим ментальным состоянием. Именно против этой идеи выступает энактивизм.

Согласно энактивистским представлениям, для понимания природы ментальных состояний нам необходимо учитывать уровень имплементации этих состояний, т. е. их воплощенность (*embodiment*) в конкретном организме, деятельностно (*enactive*) встроенном (*embedded*) в окружающую среду. Вот как об этом пишут Варела, Томпсон и Рош в своей книге, которая была одной из первых работ, предлагавших энактивистский подход.

Используя термин *воплощенный*, мы стремимся выделить два момента: первое, когнитивные процессы зависят от того вида опыта, который происходит в результате обладания телом с различными сенсорно-моторными навыками, и, второе, эти индивидуальные сенсорно-моторные навыки сами являются встроенными в более обширный биологический, психологический и культурный контекст. Используя термин *действие*, мы стремимся подчеркнуть еще раз, что сенсорные и моторные процессы, восприятие и действие, фундаментальным образом неотделимы друг от друга в контексте живой когнитивной системы [Varela, Thompson, Rosch, 1991, p. 172–173].

Эти авторы подчеркивают, что восприятие и действие представляют собой петлю. Содержание восприятия существенным образом зависит от характера деятельности живого организма, в свою очередь, те действия, которые совершает существо, зависят от того, как оно воспринимает мир. И, конечно же, то, как организм воспринимает окружающую среду и взаимодействует с ней, определяется его конкретной телесной организацией. Таким образом, согласно энактивистам, когнитивные процессы не могут быть поняты как абстрактная функциональная (компьютерционистская) система, лишенная конкретного телесного воплощения.



На формирование энактивизма повлияли идеи, выдвинутые в теории динамических систем и в исследованиях в области робототехники. Теория динамических систем в когнитивной науке призвана описывать сферу когнитивных процессов как динамическую систему, т. е. систему, развивающуюся во времени. Особенностью такой системы является то, что она включает в себя как части не только агента действия, но и окружающую среду, с которой взаимодействует этот агент. Процессы, протекающие в этой системе, могут быть охарактеризованы как достигающие определенного уровня сложности, нелинейные, эмерджентные и самоорганизующиеся. Данная теория возникает в когнитивной науке как альтернатива компьютерционизму и репрезентационизму. Сторонники этого подхода предложили множество различных примеров того, как системы могут взаимодействовать с окружающей средой, не осуществляя особого рода вычислений, базирующихся на каких-либо внутренних репрезентациях.

Например, Тим ван Гелдер предлагает нам рассмотреть работу центробежного регулятора Уатта [van Gelder, 1995]. Эта машина призвана поддерживать определенную частоту вращения вала какого-либо регулируемого объекта, например турбины. В свою очередь вал приводится в движение паровым двигателем. В качестве датчиков, позволяющих регулировать вращение вала, используются два или более груза, закрепленных на рычагах на оси вала, которые, разлетаясь под воздействием центробежной силы, приводят в движение муфту,двигающуюся вверх или вниз по оси вала. Движение муфты посредством рычага открывает или закрывает клапан, регулирующий поступление пара. Таким образом, увеличение частоты вращения вала приводит к тому, что грузы под действием центробежной силы расходятся от оси вращения и закрывают заслонку, прекращая поступление пара в двигатель. Уменьшение частоты вращения вала приводит к уменьшению центробежной силы и приближению грузов к оси вращения, что, в свою очередь, открывает заслонку поступления пара в двигатель.

Машина Уатта представляет собой не-компьютерционистскую, не-репрезентационистскую контролирующую систему. Попытки описать ее как осуществляющую вычисления и определенным образом репрезентирующую состояния парового двигателя будут неудовлетворительны по следующим причинам. Посмотрим, как в сравнении с машиной Уатта работала бы физическая реализация машины Тьюринга. Для регулирования поступления пара в двигатель компьютеру нужно было бы проделать ряд операций: измерить скорость вращения вала; сопоставить данные об этой скорости с извлеченными из памяти данными о желаемой скорости; при выявлении расхождения в данных измерить давление пара, высчитать требуемое изменение этого давления и уровень открытия заслонки; привести заслонку в



соответствующее положение. Операции, осуществляемые подобным устройством, являются манипуляциями с определенными репрезентационными объектами. Кроме того, эти операции являются дискретными, последовательными действиями. Ничего подобного мы не наблюдаем в работе машины Уатта. Эта машина является примером континуального, одновременного каузального взаимодействия тесным образом сопряженных частей, представляющих единое целое, – контролирующего устройства и того, что оно контролирует. Анализируя подобные примеры, представители теории динамических систем в когнитивной науке, как, собственно, и энактивисты, рассматривают сферу когнитивных процессов как подобную скорее машине Уатта, чем машине Тьюринга.

Идеи, близкие по духу теории динамических систем, высказывал австралийский робототехник Р. Брукс. Он знаменит разработкой поведенчески ориентированных (behavior-based) роботов. Развивая свои идеи, Брукс критиковал понимание интеллектуальных систем, предлагаемое классическими исследованиями в области искусственного интеллекта. Согласно этому пониманию, архитектура роботов должна включать в себя прежде всего следующие элементы: восприятие, моделирование, планирование и действие. Иначе говоря, будучи классической вычислительной машиной, такой робот должен быть способен составлять репрезентации внешнего мира и, оперируя с ними, осуществлять определенные действия. Первые роботы, разрабатываемые по этой модели в 60-х годах прошлого века, например, Шейки (Shakey), с трудом справлялись с простейшими задачами.

Инновационность подхода Брукса к проблеме построения роботов заключалась в отказе от обозначенной классической архитектуры. Вместо того чтобы рассматривать робота как систему, в которой поступающая на сенсоры информация последовательно передается в виде определенных репрезентаций от модуля к модулю, проходя стадии моделирования, планирования и пр., до тех пор пока система не перейдет к действию, Брукс предложил архитектуру роботов, предполагающую непосредственную связь сенсорного восприятия с действием. По сути, роботы Брукса представляют собой «набор конкурирующих поведенческих реакций» [цит. по: Shapiro, 2011, p. 140], из взаимодействия которых складывается согласованный паттерн движений робота. Например, простейший робот включал в себя уровень (layer) поиска, направляющий робота в произвольном направлении каждые десять секунд; уровень остановки, включающийся, когда сенсоры регистрировали препятствие на пути; уровень исследования, направляющий робота в сторону зарегистрированного сенсорами объекта, с которым необходимо провести определенные манипуляции. Как пишет Брукс: «Каждый производящий определенные действия уровень связывает непосредственным образом восприятие с



действием. Это только наблюдатель за данным Созданием является тем, кто приписывает некую централизованную репрезентацию или централизованный контроль. Само Создание такового не имеет» [цит. по: Shapiro, 2011, p. 139].

По мысли Брукса, эти роботы являются примером ситуационных, телесно-воплощенных интеллектуальных систем. Эти роботы «имеют тела и воспринимают мир непосредственно – их действия являются частью динамического взаимодействия с миром, и эти действия оказывают незамедлительное влияние на восприятие робота». Они «не имеют дело с абстрактными описаниями, только с тем, что находится «здесь» и «сейчас» в их окружении и что оказывает непосредственное влияние на поведение системы» [цит. по: Shapiro, 2011, p. 139]. В определенном смысле классические роботы никогда не взаимодействуют с миром. Они работают прежде всего с репрезентациями, моделями мира, но, как известно, «карта не есть территория». Роботы же Брукса взаимодействуют с миром непосредственно. Перефразируя цитату Кожибского, можно сказать, что для них сама осваиваемая территория и является собственной картой, или, как пишет Брукс, для этих созданий «мир выступает как собственная модель».

Опираясь на изложенные идеи, энактивисты пытаются показать, что если мы стремимся сохранить понятие репрезентации в когнитивной науке, то нам следует его мыслить только как репрезентацию-в-действии. Ментальные репрезентации не должны пониматься как дискретные внутренние символические объекты, являющиеся подобно фотографическим образам пассивными отражениями определенного положения дел в мире. Понятие репрезентации-в-действии довольно сильно отличается от данного классического понятия ментальной репрезентации. Оно обозначает динамическое состояние всего организма, его способность определенным образом действовать, отвечать на вызовы окружающей среды, а не просто способность мозга представлять некое положение дел.

На мой взгляд, наибольшим потенциалом, способствующим решению проблемы сознания с позиции репрезентационизма, обладает такой вариант энактивизма, как теория сенсорно-моторного знания, предложенная А. Ноэ. Ноэ следующим образом характеризует энактивистский подход: «Центральное утверждение того, что я называю энактивистским подходом, заключается в том, что наша способность восприятия не только зависит, но и конституируется нашим владением определенного рода сенсорно-моторным знанием» [Noë, 2004, p. 2]. Как подчеркивает Ноэ, это знание является знанием-как, а не пропозициональным знанием-что. Это имплицитное практическое знание того, как изменяются сенсорные стимулы в результате разных способов движения. Будучи имплицитным практическим знанием, оно по сути является определенного рода телесным навыком. Эти навыки



связаны со способностью существа к самодвижению. В свою очередь движение существ зависит от способности определенным образом воспринимать самих себя. Одним из следствий тезиса, принимаемого Ноэ, является отрицание того, что «восприятие является процессом в мозге, где перцептивная система конструирует внутренние репрезентации мира» [Ноэ, 2004, р. 2]. По мнению Ноэ, восприятие является определенной активностью всего живого существа.

Рассмотрение восприятия как практического знания зависимости изменения сенсорных стимулов от движения позволяет решить проблеме перцептивного присутствия. Эта проблема связана со скептицизмом относительно нашего знания того, как именно мы воспринимаем мир. Согласно классическому подходу, наше сознание представляется как своего рода картинка, мгновенный фотографический снимок, на котором присутствуют все детали воспринимаемого, например, визуально мира. Однако, как показывают скептики, опираясь на различные экспериментальные данные, такое представление о собственном сознательном опыте является иллюзией. Например, известно о существовании слепого пятна – области на сетчатке глаза, нечувствительной к свету, из которой выходит зрительный нерв. Несмотря на наличие слепого пятна, визуальное пространство не воспринимается нами как неполное. Можно пытаться объяснить это тем, что мозг каким-то образом достраивает картинку, заполняя пробелы. Однако, как пытался показать Деннет, у нас нет оснований полагать, что где-то в мозге осуществляются подобные операции заполнения пробелов [Dennett, 1991]. Дело в том, что отсутствие информации о какой-либо части воспринимаемого мира не тождественно информации об отсутствии этой части. Вполне возможно, что мы не замечаем разрывов в визуальном пространстве именно в силу отсутствия информации. Полагать же, что у нас имеется информация об отсутствии чего-либо и что эта информация каким-то образом обрабатывается где-то в мозге, где достраивается полноценная картинка, означает совершать ошибку гомункулуса. Ошибка гомункулуса предполагает, что мы допускаем существование материальной или функциональной области мозга, ответственной за реконструкцию полноценной репрезентации воспринимаемого окружения, изоморфной самому этому окружению. Без эмпирических свидетельств существования такой области подобный ход рассуждения, по мнению Ноэ, недопустим.

Избегая ошибки гомункулуса, Деннет объясняет иллюзию сознания следующим образом. Например, в ситуации восприятия обоев с повторяющимся образом Мэрилин Монро наш мозг, восприняв лишь несколько образов, формирует убеждение о наличии на стене сотни образов актрисы. Однако мы полагаем, что у нас есть не просто убеждение, а именно восприятие всей сотни образов. Именно в этом заключается иллюзия сознания. На самом деле, мы не воспринимаем все образы Монро, мы лишь думаем, что воспринимаем.



Соглашаясь со скептиками в том, что наш сознательный опыт не является континуальным, изоморфным воспринимаемому миру, Ноэ тем не менее не согласен с тем, что мы подвержены иллюзии сознания. В обычной жизни мы не думаем, что наше восприятие является полным и безупречным (что мы воспринимаем все детали), таким, как его рисует классическая пиктографическая теория. Скорее, это философы, в том числе скептики, полагают, что мы должны думать подобным образом. Однако у нас все же есть ощущение восприятия мира как присутствующего в качестве чего-то целого. В этом и заключается проблема перцептивного присутствия (мира, воспринимаемых объектов и т. д.). Ноэ формулирует ее в виде вопроса: «Каким образом нам кажется, что мир будто бы представлен перед нами визуально во всех своих деталях, без того чтобы нам казалось, что мы *видим* все эти детали?» [Ноэ, 2004, р. 60].

Ответ Ноэ заключается в том, что мир представлен нам виртуальным образом. Это значит, что феноменологическое ощущения восприятия целостного объекта связано с нашей способностью иметь *доступ* к недостающим частям воспринимаемого объекта. В свою очередь, это означает, что у нас имеется сенсорно-моторное знание того, каким образом нам следует двигаться, чтобы компенсировать недостаток информации о воспринимаемом объекте. Например, мы видим помидор как нечто целое, хотя и воспринимаем только определенную его сторону, поскольку мы понимаем, какого рода движения позволят нам увидеть его заднюю часть. Заимствовав термин виртуального присутствия объекта у Минского, Ноэ сравнивает перцептивное присутствие воспринимаемого мира с нашей работой с каким-либо сайтом. Хотя сайт и локализован на удаленном сервере, у нас всегда есть доступ к его данным, что создает впечатление присутствия сайта на нашем компьютере.

Важным аспектом такого решения является представление о том, что содержание перцептивного опыта в определенном смысле локализовано в самом мире. Это утверждение вполне согласуется с идеями, озвученными в теории поведенчески ориентированных роботов Брукса. Как отмечал Брукс, нет необходимости моделировать роботов как создающих модель окружающей среды. В определенном смысле сам мир является лучшей моделью. Ноэ эту мысль иллюстрирует следующим образом. Существует два способа, как вы можете ориентироваться в городе. Например, если вам нужно добраться до центра города, где возвышается башня, вы можете воспользоваться картой, содержание которой представляет город. Другой способ добраться до центра – это двигаться по улицам города таким образом, чтобы вам всегда был виден шпиль башни. Ваше движение будет определяться самим городом, выступающим собственной моделью, и вашей сенсорно-моторной способностью двигаться таким образом, чтобы при этом вы получали больше доступа к сенсорному стимулу, представленному башней.



Здесь важно добавить, что частью содержания перцептивного опыта также оказываемся мы сами или, как уточняет Ноэ, «наш пункт наблюдения или телесная локализация». Ноэ пишет: «Восприятие, таким образом, является направленным на мир и само-направленным» [Ноэ, 2004, р. 68]. Такая характеристика восприятия связана с двухаспектностью содержания перцептивного опыта. В содержании восприятия мы можем выделить фактуальное измерение и перспективное измерение: в опыте восприятия нам дано то, какие вещи, и то, как они выглядят для нас. Мы воспринимаем тарелку как круглую (фактуальное измерение), однако, рассмотренная с определенной позиции, она выглядит для нас как эллиптическая (перспективное измерение). В свете этого замечания сенсорно-моторное знание можно определить как имплицитное практическое знание того, какого рода движения будут способствовать изменению перспективного измерения опыта с целью получения представления о самом объекте. «Мы воспринимаем в опыте тарелку как круглую именно потому, что сталкиваемся с эллиптическим видом именно отсюда, и мы понимаем трансформации, которые бы претерпела эллиптически кажущаяся форма (аспект) по мере того, как мы двигаемся. Это означает, что мы понимаем, что ее эллиптическая внешность зависит от нашего пространственного отношения к ней, отношения, которое изменяется движением» [Ноэ, 2004, р. 78].

Важно подчеркнуть, что двухаспектность содержания восприятия не означает, что восприятие состоит из двух отдельных частей: фактуального восприятия объекта и каких-то ментальных, феноменальных сущностей – видимостей объекта. В движении мы воспринимаем единый объект, но как обладающий определенными свойствами, которые видятся из определенной перспективы. Эти свойства Ноэ обозначает как П-свойства. Они не являются какими-то особыми ментальными свойствами, подобными квалиа или чувственным данным. Ноэ характеризует их следующим образом: «П-свойства – видимая форма и размер объектов – являются вполне “реальными” или “объективными”. Действительно, отношение П-формы и П-размера к форме и размеру может быть описано с помощью точных математических законов (например, закона линейной перспективы). Важно, для того чтобы охарактеризовать П-свойства, не нужны отсылки к ощущениям и чувствам. <...> Тем не менее, П-свойства являются реляционными. Конкретно, они зависят от отношения между телом воспринимающего субъекта и воспринимаемым объектом. <...> Если существует разделение между сознанием и миром (в картезианском смысле, раздел между ментальным внутренним пространством и нементальной внешней средой), то П-свойства выступают определенно на стороне мира» [Ноэ, 2004, р. 83].



Подобная интерпретация содержания перцептивного опыта позволяет избежать недостатков в понимании природы ментальных репрезентаций, присущих классической когнитивной науке, на которые указывали анти-репрезентационисты. В теории Ноэ ментальные состояния мыслятся не как внутренние символические состояния мозга, а как репрезентации-в-действии, характеристики всего живого организма, деятельностно включенного в окружающую среду. Будучи натуралистической, его концепция тем не менее избегает крайностей редукционизма физикалистских и функционалистских подходов, против которых были направлены аргументы анти-репрезентационистов. Как мы видели, содержание восприятия, о котором говорил Ноэ, является характеристикой взаимодействия мира и организма. Такое понимание содержания позволяет зафиксировать перспективный характер нашего опыта, о котором писал Нагель, и объяснить квалиа не как особые ментальные сущности, а как объективные реляционные свойства репрезентируемых объектов. Все вместе это делает сенсорно-моторный энактивизм хорошим основанием для натуралистического решения проблемы сознания.

Список литературы

- Фодор, Пылишин 1996 – *Фодор Дж., Пылишин З.* Коннекционизм и когнитивная структура: критический обзор // *Язык и интеллект* / Пер. Г.Е. Крейдлина. М.: Прогресс, 1996. С. 230–313.
- Chalmers 1996 – *Chalmers D.* *Conscious Mind.* Oxford: Oxford University Press, 1996. 414 p.
- Churchland, Sejnowsky 1990 – *Churchland P.S., Sejnowsky T.* *Neural Representation and Neural Computation* // *Mind and Cognition* / Ed. by W. Lycan. Oxford: Blackwell, 1990. P. 133–152.
- Clark 2001 – *Clark A.* *Mindware: An Introduction to the Philosophy of Cognitive Science.* N.Y.: Oxford University Press, 2001. 210 p.
- Dennett 1991 – *Dennett D. C.* *Consciousness Explained.* Boston: Little, Brown and Company, 1991. 511 p.
- Fodor 1987 – *Fodor J.* *Psychosemantics: The Problem of Meaning in the Philosophy of Mind.* Cambridge (MA): MIT Press, 1987. 171 p.
- Jackson 1982 – *Jackson F.* *Epiphenomenal Qualia* // *Philosophical Quarterly.* 1982. No. 32. P. 127–136.
- Levin 1986 – *Levine J.* *Materialism and qualia: the explanatory gap* // *Pacific Philosophical Quarterly.* 1986. No. 64. P. 354–361.
- Nagel 1974 – *Nagel T.* *What is it like to be a bat?* // *Philosophical Review.* 1974. No. 83. P. 435–50.
- Newell, Simon 1981 – *Newell A., Simon H.* *Computer science as empirical inquiry: Symbols and search* // *Communications of the Association for Computing Machinery.* 1981. No. 19. P. 113–126.



Noë 2004 – Noë A. *Action in Perception*. Cambridge (MA): MIT Press, 2004. 296 p.

Shapiro 2011 – Shapiro L. *Embodied Cognition*. N.Y.: Routledge, 2011. 237 p.

van Gelder 1995 – van Gelder T. What might cognition be if not computation? // *Journal of Philosophy*. 1995. Vol. 7. No. 92. P. 345–381.

Varela, Thompson, Rosch 1991 – Varela F., Thompson J.E., Rosch E. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge (MA): MIT Press, 1991. 328 p.

References

Chalmers D. *Conscious Mind*. Oxford: Oxford University Press, 1996. 414 p.
Churchland P.S., Sejnowsky T. Neural Representation and Neural Computation. *Mind and Cognition*. Oxford: Blackwell, 1990, pp. 133–152.

Clark A. *Mindware: An Introduction to the Philosophy of Cognitive Science*. New York: Oxford University Press, 2001. 210 p.

Dennett D. C. *Consciousness Explained*. Boston: Little, Brown and Company, 1991. 511 p.

Fodor J. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press, 1987. 171 p.

Fodor J., Pylishin Z. Коннекционизм и когнитивная структура: критический обзор. [Connectionism and Cognitive Architecture: A Critical Analysis]. In: *Язык и интеллект. [Language and intelligence]*. Moscow: Progress Publ., 1996, pp. 230–313. (In Russian)

Jackson F. Epiphenomenal Qualia. *Philosophical Quarterly*, 1982, vol. 32, no. 127, pp. 127–136.

Levine J. Materialism and qualia: the explanatory gap. *Pacific Philosophical Quarterly*, 1986, vol. 6, pp. 354–361.

Nagel T. What is it like to be a bat? *Philosophical Review*, 1974, vol. 83, pp. 435–450.

Newell A., Simon H. Computer science as empirical inquiry: Symbols and search. *Communications of the Association for Computing Machinery*, 1981, vol. 19, pp. 113–126.

Noë A. *Action in Perception*. Cambridge, MA: MIT Press, 2004. 296 p.

Shapiro L. *Embodied Cognition*. New York: Routledge, 2011. 237 p.

van Gelder T. What might cognition be if not computation? *Journal of Philosophy*, 1995, vol. 92, no. 7, pp. 345–381.

Varela F., Thompson J.E., Rosch E. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press, 1991. 328 p.