

УДК 1(091)

DOI: 10.17223/1998863X/43/2

В.И. Суховой

МЕНТАЛЬНЫЕ РЕПРЕЗЕНТАЦИИ, ВЫЧИСЛИМОСТЬ И ИНТЕНЦИОНАЛЬНОСТЬ: ПОДХОД ДЖ. ФОДОРА

На суд читателя выносится классическая модель устройства человеческой психики Джерри Фодора. Показывается, какое место в ней занимает понятие ментальной репрезентации, как оно взаимосвязано с идеей вычислимости и насколько такая модель способна разрешить проблему интенциональности. Оцениваются перспективы данной теории с точки зрения современного состояния философии сознания и когнитивной науки.

Ключевые слова: *ментальные репрезентации, вычислимость, интенциональность, классическая модель.*

Введение

Джерри Алан Фодор – современный американский философ и когнитивист, оказавший огромное влияние на развитие как философии сознания, так и когнитивной науки, – наиболее известен широкому кругу людей благодаря своей концепции «языка мысли». Эта теория встретила широкий отклик среди философов и когнитивных ученых. Ее популярность во многом объясняется доминированием в западной мысли 1960–1980-х гг. так называемого классического подхода, связанного с именами таких исследователей, как Алан Тьюринг, Алонзо Черч, Ноам Хомский и др. Сформировавшись в рамках этого подхода, Фодор развил его, разработав оригинальную модель психики. В последние десятилетия появилось несколько альтернатив классическому подходу, но несмотря на это, он сохраняет привлекательность для исследователей. Не в последнюю очередь это произошло благодаря Джерри Фодору и его интерпретации этого подхода.

В классической модели ключевое место занимают понятия ментальной репрезентации и вычислимости, которые служат основой в объяснении и моделировании психических процессов. Однако, как отмечают критики классического подхода, он не отдает должного таким принципиально важным особенностям психических явлений, как интенциональность, осознаваемость, субъективность и т.п. И если «трудная проблема сознания» действительно не включается сторонниками классического подхода в число вопросов, на которые можно получить ответ в ходе научного изучения психики, то к интенциональности они относятся иначе – для них этот аспект может и должен быть учтен в теоретическом объяснении психики. В этой статье я хотел бы рассмотреть, что представляет собой классическая модель психики в интерпретации Фодора и какое решение проблемы интенциональности он предлагает.

Классическая когнитивная модель

Классическая идея, идущая от Тьюринга, состоит в том, что уровень вычислений независим от уровня имплементации, т.е. независим от физической системы, на которой он реализован. Вычисления определяются за счет про-

граммы, задающей рамки возможных действий. Программа должна оперировать символами, т.е. синтаксическими, формальными объектами, взаимодействия между которыми происходят исключительно на основании правил, управляющих сочетаниями этих символов. Неважность содержания символов для процесса вычислений можно продемонстрировать на примере, предложенном русским лингвистом Л. Щербой: «Глѡкая кѹздра штѣко будланѹла бѡкра и курдѣчит бокрѣнка». Каждое слово в данном предложении бессмысленно, тем не менее мы с полным основанием можем утверждать, что предложение является грамматически правильным. Нам понятно, что в нем играет роль существительного, что – глагола, что является отглагольным наречием, что – прилагательным, относящимся к существительному, и т.д. и что все части речи связаны между собой грамматически верно. Из этого примера видно, что можно строить правильные сочетания символов без учета их содержания.

Как известно, бихевиоризм и гибсоновский реализм обходятся без понятия ментального уровня, но классический подход с его компьютерной метафорой возвращает в психологию и философию ментализм. Бихевиоризм строит свои обобщения только на основе связей между стимулами на входе и поведением на выходе. Но с точки зрения когнитивизма входные данные, поступающие в организм, запускают определенное ментальное состояние, которое, в свою очередь, может вызвать другое ментальное состояние, никак не отразившись на поведении. Это хорошо видно на примере машины Тьюринга. Еще Хилари Патнэм в знаменитой статье «Сознание и машины» 1960 г. показал, что уровень «логических состояний» машины Тьюринга отличается от ее физических состояний.

В частности, «логическое описание» машины Тьюринга не содержит никаких данных о *физической природе* этих «состояний» – и даже о физической природе машины в целом. (Неизвестно, состоит ли она из электронных реле, картона, клерков, сидящих за пультами, или чего-либо еще.) Иными словами, любая данная «машина Тьюринга» — это *абстрактная* машина, которая может иметь практически неограниченное число различных физических реализаций [1. С. 35].

Но машина Тьюринга или просто компьютер помимо логических или программных состояний обладает еще и структурными состояниями, характеризующимися устройством микросхем, хотя для понимания работы программы знание микросхем не имеет никакого значения. В соответствии с этой идеей классический подход предполагает, что устройство человека имеет два уровня: нейрофизиологический, соответствующий структурному описанию компьютера, и ментальный, соответствующий описанию компьютерных программ.

Это означает, что некоторые ментальные состояния должны иметь символичный характер, т.е. быть символами, по определенным правилам участвующими в различных взаимодействиях в психике. В связи с этим возникает закономерный вопрос – почему только некоторые состояния? Дело в том, что, к примеру, состояние боли мы также можем назвать ментальным, но оно не обладает символической природой, в отличие от идеи боли или суждения о боли. Два последних ментальных состояния являются репрезентациями потому, что отображают (репрезентируют) определенные состояния окружающего мира или организма, и эти символичные, т.е. дискурсивные, состояния философы называют концептами. Тем самым концепт *человек* является мен-

тальным состоянием, репрезентирующим всех возможных людей, а концепт *боль* – все возможные болевые ощущения.

Классическая модель: две интерпретации

С момента своего появления классический подход был представлен в двух различных вариантах. Так, согласно интерпретативистской трактовке Патнэма мы не можем выделить как отдельный уровень психологических законов, поскольку ментальные состояния «не образуют такую же причинно-замкнутую систему, какая имеет место в случае ‘конфигураций’ машины Тьюринга» [1. С. 37]. Это означает, что мы можем описывать человеческое поведение в виде реализации определенной программы или в виде совокупности логических состояний машины Тьюринга, но делаем это из соображений удобства, поскольку в реальности человек не следует никаким внутренним программам. Вместе с тем в соответствии с тезисом Черча с помощью машинной таблицы можно описать все то, что может быть строго представлено в виде множества шагов и тем самым может быть смоделировано на компьютере. А отсюда, отмечает Джон Серл, следует тривиальный вывод о том, что операции в мозге человека могут быть смоделированы на компьютере. Но таким же образом на компьютере можно смоделировать и систему погоды, и поведение нью-йоркской биржи, и движение планет в Солнечной системе и т.п. [2. Р. 87]. Но такой интерпретативистский подход не продвигает нас в понимании природы психики.

Фодора такая половинчатая позиция не устраивает, и он наряду с другими исследователями придерживается той точки зрения, что уровень человеческой психики *действительно* является аналогом программного уровня компьютера, что поведение человека *действительно* формируется на основании вычислений, происходящих в его психике, и что некоторые ментальные состояния *действительно* имеют символичный характер. В пользу этого у него есть ряд аргументов. С одной стороны, Фодор подчеркивает, что ментальные состояния должны быть продуктивными. Идея продуктивности берет свое начало в теории американского лингвиста Ноама Хомского, где она применяется к естественному языку, но Фодор распространяет ее на ментальные репрезентации. Продуктивность предполагает, что при наличии определенного конечного набора слов или концептов и правил, регулирующих их взаимные сочетания, можно построить неограниченное число всевозможных предложений или суждений. Так, имея в своем распоряжении три слова (концепта) «А», «В» и «С» и правило сочетания конъюнкцию, мы можем составлять все новые предложения, просто добавляя одно из слов (концептов) в конец имеющегося предложения: “А & В & С”, “А & В & С & А” и т.д. Таким образом, длина возможных предложений потенциально бесконечна и зависит от количества рабочей памяти: увеличивая память, мы можем увеличить количество возможных предложений.

Отсюда следует, что применительно к психике компьютерная метафора не означает простое предоставление наилучшего описания определенной программы, потому что такое описание всегда дается *post factum* и не учитывает продуктивности ментальных состояний. А это означает, что ментальные состояния должны быть аналогичны «вычислительным состояниям автомата, нежели его машинно-табличным состояниям» [3. Р. 58]. Тем самым компьютерная метафора предполагает, что психику не просто можно представить в

виде описания определенной компьютерной программы и смоделировать на компьютере, но что психика *действительно* является своего рода компьютером и различные ментальные операции имеют вычислительный характер.

Вычислимость и психика

Остановимся подробнее на характеристике вычислимости. Как следует понимать вычислительные процессы? Фодор выдвигает два требования: 1) вычислительные процессы являются *символическими*, т.е. они определяются на множестве репрезентаций; 2) они являются *формальными*, т.е. применяются к репрезентациям исключительно на основе их синтаксиса [4. Р. 226]. Но не всякий формальный процесс, по мнению Фодора, является синтаксическим, поскольку процесс определяется как синтаксический через его противопоставление семантике, т.е. через указание на то, что этот процесс не является семантическим, а значит, содержание символа или репрезентации не имеет для него никакого значения [Ibid. Р. 227]. Ведь к образам также приложимы формальные операции, они могут взаимодействовать друг с другом по определенным правилам, но в этом взаимодействии решающее значение будет иметь семантика образа, поскольку именно она является ключевым моментом, отличающим один образ от другого. Тем самым в вычислительной теории психики, как ее понимает Фодор, утверждается, что «две мысли отличаются по содержанию, только если они могут быть соотнесены с формально различающимися репрезентациями» [Ibid. Р. 227]. Это означает, что мысли, одинаковые по форме, не могут иметь различное содержание. Если, к примеру, в естественном языке имена собственные «Мадонна» (мать Христа) и «Мадонна» (певица) имеют одинаковую форму, но разное содержание и потому являются омонимами, то они обязаны иметь разные по форме ментальные репрезентации. Отсюда следует, что репрезентация не может существовать без соответствующей структуры, т.е. не обладая отдельным вычислительным уровнем, представленным в человеческой психике. И как указывает Фодор, операции с символами не могут существовать без самих символов, а точнее, «не может быть внутренней репрезентации без внутреннего языка» [5. Р. 55]. Этим последним не может быть наш естественный язык, поскольку тогда бы пришлось предположить, что младенцы, не овладевшие языком, не способны репрезентировать для себя окружающий мир. К тому же как бы они смогли овладеть языком при том условии, что он систематично организован, а нерепрезентативная психика младенцев – нет?

Систематичность является одним из важнейших свойств естественного языка (наряду с продуктивностью), она предполагает, что при наличии предложения «Дима любит Асю» всегда можно построить предложение «Ася любит Диму». По словам Брайана Маклафлина, «идея о систематичности мысли заключается в том, что любое существо, способное иметь определенную мысль, способно также иметь семейство связанных с ней мыслей. Это способность обладать мыслями, которые не являются в принципе точечными (punctate)», но собраны в кластеры [6. Р. 32]. Таким образом, если у нас есть предложение «Дима любит Асю», мы не можем не быть способными построить предложение «Ася любит Диму», поскольку все составные части обоих предложений одинаковы, различно только отношение между ними. Но если психика младенцев не обладает репрезентативной природой, значит, ей не свойственна систематичность, но как в таком случае может несистематиче-

ская психика обучиться систематическому языку, т.е. как может язык выражать пропозиции, которые психика не способна ввести? [7. Р. 26].

Подобного же рода соображения можно высказать и в отношении высших животных. По мнению Фодора, существует гомогенность между ментальными способностями невербальных организмов и обычных представителей человеческого общества, которую нельзя объяснить, не признав, что психология невербальных организмов гомогенна с нашей психологией. Вычислимость предполагает язык репрезентаций, но не предполагает, что он должен быть языком коммуникации, т.е. естественным языком [5. Р. 56–57].

Психика как репрезентативная система

Подобные рассуждения приводят нас к вопросу, уже затронутому ранее: не будет ли описание психики (или организма в целом) как репрезентативной вычислительной системы просто удобным предсказательным механизмом? А это, в свою очередь, ставит вопрос о том, что значит быть репрезентативной вычислительной системой?

С одной стороны, этот термин имеет максимально широкое применение и репрезентацией можно считать практически все что угодно. К примеру, «рисунки» на песке, оставленные ветром, репрезентируют движение воздушных масс, а следы, оставленные на земле человеком, репрезентируют форму его стопы. Но такое максимально широкое понятие репрезентации малопригодно для объяснения природы человеческой психики. Как отмечает когнитивист Зенон Пыльшин, не любой информационный процесс можно назвать репрезентацией. Нужно различать информационные процессы, в которых содержание информационного состояния играет роль в объяснении, и такие, в которых оно (если даже и присутствует) не играет никакой роли. Если мы не получаем преимущества в объяснении при указании, *что* именно репрезентирует данное конкретное состояние, тогда для нас не имеет смысла рассматривать это состояние как репрезентацию [8. Р. 74]. Что же в таком случае делает репрезентацию не просто исследовательским инструментом, но тем механизмом, который присущ самому организму? По мнению Пыльшина (и Фодор с ним солидарен в этом), для того, чтобы репрезентация была присуща самому организму, необходимо, чтобы характеристики данного конкретного информационного состояния *определяли поведение организма посредством кодируемых в этом состоянии свойств*. Другими словами, необходимо, чтобы *свойства информационного состояния функционировали для репрезентации организму свойств мира* [Ibid. Р. 74]. Это означает, что организм действует в мире в соответствии с тем, *как* этот мир ему репрезентирован. Планеты не репрезентируют законы Кеплера для того, чтобы двигаться по орбитам в Солнечной системе, но человек репрезентирует окружающий мир для того, чтобы эффективно в нем действовать.

Из вышесказанного видно, что позиция Фодора достаточно радикальна в том смысле, что для него компьютерная аналогия не является просто удобной метафорой, человек действительно обладает репрезентативной системой, состоящей из концептов. Но если это так, то как можно объяснить появление такой репрезентативной системы в процессе эволюции? Где в эволюционном развитии произошел скачок от нерепрезентативных систем к репрезентативной системе? Теоретически наличие такого скачка не является чем-то экстраординарным, например, его можно связать с появлением нервной системы или сложной

нервной системы (развитого головного мозга), способной имплементировать достаточное количество сложных программ. Если обратиться к развитию вычислительной техники, то и здесь можно наблюдать похожий скачок при увеличении ее сложности, к примеру при переходе от карманного калькулятора к современному сверхмощному компьютеру. Если калькулятор пользуется небольшим количеством простейших программ в силу небольшого объема памяти, то в современном компьютере объем памяти позволяет существовать разветвленной иерархии программ. Разве нельзя то же самое признать и в отношении животных, у которых при наличии простейшей нервной системы используются простейшие программы, а там, где есть очень сложная нервная система, существует и сложная иерархия программ. По аналогии с компьютерами зависимость количества программ от объема рабочей памяти становится важным требованием и при классическом подходе к психике.

Но если и в психике существует определенная иерархическая градация программ, то каким образом они будут представлены организму? Если определить наиболее сложную программу как язык мысли, то означает ли это, что она должна быть представлена для всего организма в целом?¹ В этой связи философ Дэниел Деннет и некоторые другие авторы критикуют Фодора за то, что его концепция предполагает наличие гомункулуса, внешнего по отношению к репрезентациям агента. В компьютере, утверждает Деннет, репрезентации функционируют как маленькие подрядчики и рабочие, каждый из которых выполняет определенную функцию. Поэтому, согласно Деннету, чем сложнее организм, тем больше соответствующих «рабочих» и «подрядчиков» должна содержать его «компьютерная модель». А это означает, что не может существовать единого уровня репрезентаций, который был бы представлен организму как целому. Наличие же такого уровня, по мнению Деннета, свидетельствовало бы о том, что наша психика является картезианским театром, на сцене которого происходит представление репрезентаций, наблюдаемое гомункулусом-зрителем. На языке компьютерной метафоры это означает, что это мы, интерпретаторы, выделяем уровень логических состояний машины, а для самого компьютера его как такового не существует. Таким образом, Деннет полагает, что не существует никакого онтологически реального репрезентативного уровня, где наши концепты были бы эксплицитно представлены, поскольку допущение такого уровня приводит к несуразным выводам [9. Р. 99–102].

На первый взгляд аргументация Деннета выглядит довольно убедительной, но в ней не учитывается то важное обстоятельство, что градация программ не означает их обязательной представленности организму как целому. Вполне можно допустить, что более примитивные и базовые программы будут представлены лишь субличностной части организма, в то время как наиболее сложные – даны организму как целому. Такой наиболее сложной программой как раз и может являться язык мысли. Кроме того, как будет показано ниже, отказ от гипотезы языка мысли делает классическую модель уязвимой перед аргументом «Китайская комната» Джона Серла.

¹ Представленность всему организму в целом означает, что она касается не только его субличностной части; к примеру, человек не контролирует во время ходьбы детали походки (активацию и взаимодействие всех мышц ноги и других частей организма), но управляет ею в целом. Представленность всему организму в целом можно сравнить с сознательным контролем человека над его действиями и реакциями, хотя привязка к сознанию не должна пониматься как необходимая в данном случае.

Синтаксис и семантика

Говоря о компьютерной аналогии, программах и символах, мы не касались их содержания, однако использование символов неизбежно предполагает, что они определенным образом проинтерпретированы, т.е. имеют определенное значение. Вместе с тем вычислительный подход в принципе не требует брать в расчет влияние внешнего мира на наши репрезентации. Тогда возникает вполне закономерный вопрос, репрезентациями чего являются в таком случае наши концепты? Рационалистическая психология предполагает, что большинство наших репрезентаций являются врожденными, а это позволяет допустить, что врожденной является не только их форма, но и содержание. Это означает, что вопрос о соотношении репрезентаций с внешним миром автоматически снимается. Организм человека снабжен органами чувств, сообщающими информацию о внешнем мире нашему мозгу. Но, как отмечает Фодор, «пока мы рассматриваем ментальные процессы как исключительно вычислительные, влияние информации из окружающего мира на эти процессы исчерпывается формальным характером того, что органы-посредники (*oracles*) пишут на ленте» (имеется в виду лента машины Тьюринга) [4. Р. 231]. Поэтому не имеет значения, является ли истинным то, что пишут эти органы-посредники, т.е. действительно ли они отражают состояние окружающей среды. Мы можем предположить, что они что-то отражают, но безотносительно к тому, что же именно. Эти положения равносильны признанию *методологического солипсизма*¹. Как отмечал Патнэм, для менталистских психологов со времен Декарта не имело особого значения, существуют ли вещи и индивиды за рамками сознания. По сути, они могут и не существовать, но тогда, по мнению Патнэма, описание ментальных состояний, в которых находится человек, будет выглядеть довольно странно. Ведь пребывание в ментальном состоянии ревности к у означает, по крайней мере, что у должен существовать [11. Р. 170–171]. Но Фодор полагает, что «если ментальные процессы являются формальными, тогда они имеют доступ только к формальным свойствам соответствующих репрезентаций окружающей среды, предоставляемых чувствами» [4. Р. 231]. Иными словами, они не имеют доступа к семантическим свойствам репрезентаций, таким как *быть истинными, иметь референты* или *быть репрезентацией* окружающей среды [Ibid. Р. 231]. Отсюда следует, что для того чтобы быть в состоянии ревности к у, совсем не обязательно, чтобы у существовал; главное, чтобы наши органы чувств передавали информацию о его наличии и у нас в психике имелся соответствующий концепт.

Вместе с тем согласно каузальной теории референции в случае, когда я думаю, к примеру, что Месси забил 55 голов, моя мысль имеет каузальное отношение к тому, как я и Месси встроены в окружающий мир, и поэтому только натуралистическая (в противовес рационалистической) психология может адекватно объяснить это обстоятельство, поскольку здесь мы вступаем в область отношений между миром и организмом [4. Р. 233].

¹ Этот термин использовал в свое время еще Рудольф Карнап («Логическое построение мира»). Карнап стремился реконструировать эмпирическое знание с соответствующих позиций, т.е. с феноменалистической перспективы, без возможных онтологических обязательств и некоторых эпистемологических амбиций, чтобы показать взаимозависимую и структурную природу нашей системы эмпирических понятий, выражающую единство и объективность [10].

Синтаксис и «Китайская комната». Проблема интенциональности

Более того, находясь в рамках методологического солипсизма, мы становимся легкой добычей для критики со стороны Джона Серла. Его знаменитый аргумент «Китайская комната» нацелен на то, чтобы показать несостоятельность вычислительного подхода к пониманию человеческой психики. Смысл этого аргумента в том, что человек, находящийся в комнате с окошком, получает через него запросы на китайском языке в виде наборов символов, которые он совершенно не понимает. Но в его распоряжении есть инструкции, объясняющие, каким образом формировать ответы на поступающие запросы. Человек производит манипуляции с символами, следуя инструкциям, и формирует другие наборы символов, служащие ответами на поступающие запросы. Но человек в комнате не понимает и сформированные ответы – он их создает, просто манипулируя символами в соответствии с инструкциями. Однако по формальным признакам его деятельность полностью отвечает требованиям носителя китайского языка: на все поставленные вопросы, просьбы, требования и т.п. он предоставляет адекватные ответы¹. Китайские выражения этого человека обладают семантикой, недоступной для него самого. Тем самым Серл хочет сказать, что вычислительная теория не способна объяснить устройство человеческого ума, поскольку синтаксис не способен объяснить семантику, значение или содержание символов, сформированных в ответ на запросы в китайской комнате, недоступно для понимания находящегося там человека. А значит, никакая программа, поскольку она имеет дело только с формальными характеристиками символов, не способна объяснить природу интенциональности.

Хотя понятие интенциональности, по мнению Серла и др., тесно связано с понятиями понимания и осознанности, не все современные аналитические философы признают эту тесную связь². Серл на примере «Китайской комнаты» стремится показать, что существуют два вида интенциональности: внутренняя (*intrinsic*) и интенциональность, приписываемая по аналогии (*observer-relative*). Так, когда мы приписываем интенциональность термостатам, мы в

¹ Изначально этот аргумент нацелен против теста Тьюринга. Смысл последнего состоит в том, что если программа будет давать на вопросы человека адекватные ответы, то мы вправе говорить о наличии у нее разума. Сам тест был задуман таким образом, что человек общается через компьютер, не зная, кто ему отвечает – реальный человек или программа. Пользователь не видит своего собеседника, все, что он имеет, – это его ответы и вопросы на выходе.

² Аргумент «Китайская комната» был призван показать, что машина не может обладать сознанием, поскольку она не обладает истинной интенциональностью. Но даже Серл, для которого понятия сознания и интенциональности имеют почти синонимическое значение, не отождествляет их. Поскольку для него интенциональные феномены должны обладать условиями выполнимости, т.е., чтобы понять, чем является какое-либо интенциональное состояние, нужно знать, при каких условиях оно является истинным. Серл в этой связи отмечает, что «некоторые наиболее важные логические особенности (*features*) интенциональности находятся за рамками феноменологии, поскольку они не обладают непосредственной феноменологической реальностью» [12. Р. 115]. Деннет в этой связи также отмечает, что мы, люди, часто не ведаем, как мы смогли осознать и понять новые вещи. Однако, по мнению Деннета, отличительным признаком понимания служит способность применять наши уроки к новым материалам, новым проблемам. «Идея о том, что индивидуальный организм обладает мобильными (*portable*), улучшающими собственное устройство умениями, является ядром нашего обыденного (*folk*) представления о процессе понимания. Она не зависит от предположения о сознательном опыте, несмотря на то, что оно (это предположение) является знакомой декорацией, идеологическим преувеличением по отношению к изначальному понятию» [13].

этом случае не считаем, что они обладают ментальной жизнью; мы приписываем им интенциональность по аналогии с собой, потому что спроектировали их для того, чтобы они служили определенным целям, чтобы осуществляли функции на основании нашей интенциональности. Истинная же интенциональность присуща только людям. Весь функционализм, по мнению Серла, строится на неспособности видеть это различие. Функциональные атрибуты всегда зависимы от наблюдателя, и потому не существует внутренних (intrinsic) функций в том смысле, в каком понимается внутренняя интенциональность [14. Р. 451–452], что и должен продемонстрировать аргумент «Китайская комната».

Возвращаясь к обсуждению позиции Фодора, сразу отмечу, что для него понятие осознанности не имеет никакого отношения к проблеме интенциональности¹. Фодор связывает интенциональность с семантикой, поскольку в рамках вычислительной теории психики проблема интенциональности – это в первую очередь проблема семантической ментальных репрезентаций [16. Р. 28]. Вместе с тем при натуралистическом подходе к психологии содержание ментальных репрезентаций должно быть локализовано во внешнем мире. Каким же образом понятие интенциональности соотносится с вычислительной теорией психики, которая является краеугольным камнем в теории Фодора?

Согласно Фодору интенциональные состояния нельзя считать логическими конструкциями телодвижений или состояниями нервной системы, ибо в первом случае мы будем иметь позицию бихевиоризма, а в последнем – теории психофизического тождества². И тогда или у нас будут проблемы с ментальной каузацией, поскольку бихевиористы игнорируют ментальные состояния, или нам придется отрицать номологическую возможность существования рациональных машин или существ с другой анатомией мозга, а это будет означать *несостоятельность вычислительной теории психики* [17. Р. 67]. Отказываясь от методологического солипсизма, Фодор заявляет, что интенциональным содержанием обладает синтаксически структурированная ментальная репрезентация объекта внешнего мира. Таким образом, интенциональность присутствует там, где есть репрезентации, *имеющие каузальную связь с репрезентируемыми ими объектами в мире*. Причем только объект в действительном мире может иметь каузальную связь с ментальной репрезентацией; возможные объекты, замечает Фодор, не обладают каузальной силой и не могут повлиять на содержание концепта [Ibid. Р. 141].

Здесь можно видеть, как блокируется критика функционализма, высказанная Серлом. Конечно, если мы понимаем символы в чисто формальном ключе, непонятно, что придает им содержание и откуда появляется у них интенциональность. Но символы, указывает Фодор, нужно брать семантически интерпретированными, т.е. находящимися в каузальной связи с объектами внешнего мира. Это означает, что и компьютерная программа будет обладать интенциональным содержанием, если мы установим связи между ее симво-

¹ Так, Фодор отмечает: «Том Нагель как-то написал, что именно сознание делает философию сознания (philosophy of mind) такой трудной. Это почти верно. По сути, это интенциональность делает философию сознания такой трудной, а сознание делает ее невозможной» [15. Р. 22].

² Это теория, согласно которой ваши ментальные состояния идентичны состояниям вашей нервной системы. Представители этого направления – Д. Армстронг, Дж. Сمارт, Д. Папино и др.

лами и внешним миром. Так, Зенон Пыльшин, позиция которого очень близка к позиции Фодора, пишет о том, что пока мы работаем с компьютерными программами в привычном смысле, у нас часто возникает возможность множественной интерпретации символов – то как чисел, то как слов или описаний. Но если мы снабдим компьютер органами-посредниками и позволим ему свободно взаимодействовать с естественной и лингвистической средой, то наши возможности интерпретации сильно сократятся [14. Р. 443], т.е. при установлении каузальной связи между символами и миром через органы-посредники возможность множественной интерпретации символов программ сузится настолько, что останется разве что люфт, обозначенный куайновским тезисом неопределенности перевода¹.

Заключение

Итак, привлекательность классической модели зиждется на том факте, что она свободна от привязки к определенной архитектуре носителя психики: будет ли это мозг или компьютер, мы всегда можем выделить уровень программ, который и характеризует психику. Работа этого уровня основана на идее вычислимости, т.е. переход из одного ментального состояния в другое является вычислительной операцией. А для ее характеристики имеет значение только форма наших состояний, мы можем не учитывать их содержание. Данные аспекты хорошо объясняют в первую очередь работу таких видов высокоуровневого когнитивного поведения, как язык, комплексное планирование и дедуктивное рассуждение.

И хотя сторонники классического подхода, отстаиваемого Фодором, апеллируют к идее Тьюринга и компьютерам, следует иметь в виду, что данный подход выходит за рамки простой и изящной идеи Тьюринга о вычислимости и имеет дополнительные предпосылки в своем основании. Во-первых, он опирается на предположение о концептах как фундаментальных кирпичиках, благодаря которым и происходят вычислительные операции. Во-вторых, в нем используется немного позже разработанная Фодором идея модулярности психики, которая означает, скажем, отличие процесса восприятия от познания и мышления. Отсюда следует, что релевантные для организма в целом вычисления происходят в когнитивном аппарате, в котором и хранятся концепты.

Эти две идеи взаимодополняют друг друга, поскольку из идеи вычислимости самой по себе еще не следует ответ на вопрос, чем являются минимальные элементы, участвующие в вычислениях. Классическая модель предполагает, что на этапе процесса восприятия также совершаются вычислительные операции, особенностью которых является инкапсулированность, т.е. непроницаемость информации из высших уровней или из других модулей на вычислительные процедуры в самом модуле восприятия. Уже просчитанная информация из модулей восприятия поступает в принципиально

¹ В.О. Куайн утверждал, что никаких эмпирических данных будет недостаточно для точного перевода с одного языка на другой, иллюстрируя это на примере радикального перевода с абсолютно неизвестного исследователю языка туземцев на его родной язык. Так, даже в случае остенсивного указания на кролика и произнесения слова на туземном языке «гавагай» у исследователя остается свобода перевести его не только как «кролик», но и как «неотъемлемая часть кролика», «временной срез кролика» и т.д.

немодулярную когнитивную систему, активируя соответствующие концепты. Так, чирикание воробья и визуальная информация о воробье (т.е. внешний облик воробья) поступают из аудиального и зрительного модулей в когнитивную систему, активируя уже имеющийся в ней концепт *воробей*.

Учитывая сказанное, правомерно задаться вопросом: почему релевантные для организма в целом вычисления происходят именно на уровне концептов, которые сами являются целостными и далее не разлагаемыми единицами? Ведь и на уровне модулей происходят вычисления, в которых участвуют другие единицы. Собственно, именно эта идея имеет свои корни в рационалистической психологии, идущей от Декарта и утверждающей врожденность многих наших понятий¹. Противоположная же идея – о том, что важна только интерпретация входных данных нашей системой восприятия, формирующей наши представления и понятия, – соответствует эмпиристской точке зрения, имеющей своим источником философию Локка.

Против эмпиристской идеи вполне уместно возразить, что она не может исключать наличия у субъекта врожденных свойств. К примеру, чтобы распознать глубину или контрастность, нам надо иметь соответствующие рецепторы, способные их распознавать; мы, скажем, не способны распознавать свет или звуковые колебания определенного спектра или диапазона, например инфракрасное излучение или ультразвук. Абсолютный эмпиризм в принципе невозможен, поскольку при взаимодействии окружающей среды и воспринимающего субъекта последний уже обладает определенной структурой.

Естественно задаться вопросом, можно ли считать адекватным решением для этой проблемы рационализм? Ведь помимо классической идеи есть немало альтернативных подходов (тот же коннекционизм), представители которых не признают наличия каких-либо врожденных или встроенных концептов, полагая, что наши представления формируются на основании распознаваемых рецепторами раздражителей. Но, что важно, и сам классический вычислительный подход может обходиться без идеи встроенных представлений или концептов, свидетельством чего может служить теория Дэниеля Деннета.

На мой взгляд, преимущество подхода Фодора состоит в том, что он предлагает изящное решение проблемы ментальной репрезентации. Данное понятие давно уже является своеобразным камнем преткновения в философии сознания. К примеру, Джон Серл в своей теории психики использует понятие ментальной репрезентации, не основанное на идее вычислимости. Он отрицает вычислительный характер психики, поэтому ментальные состояния понимаются им не-синтаксически. С точки зрения Серла, все, что реально имеет значение для соответствующего ментального состояния, это его содержание. Но как в таком случае без обращения к синтаксису мы могли бы адекватно определять это состояние? Выход из этого затруднения как раз и предлагает Фодор: путем трактовки ментального состояния как ментальной репрезентации, обладающей формальными и содержательными характеристиками. Такая репрезентация – концепт – определяется за счет каузальной связи с репрезентируемым ею в мире объектом. И благодаря этому шагу вычислительный характер психики хорошо сочетается с натуралистическим ее

¹ Строго говоря, это более современный (относящийся к Новому времени) вариант рационализма, потому что корни его мы найдем уже у Платона.

объяснением: свое содержание она получает через органы чувств от реальных объектов в мире. Остается еще вопрос о характере этой каузальной связи, но в классическом подходе он, по сути, приобретает техническое значение. Кроме того, признание вычислительной природы психики хорошо объясняет мыслительную деятельность человека, которая не может быть простой цепочкой ассоциаций, но является деятельностью, имеющей формальные характеристики: на основании посылок делаются определенные заключения.

Вдобавок классический подход, представленный Фодором, прекрасно инкорпорирует в себя понятие интенциональности, не пытаясь его свести к чему-либо или вовсе отбросить. Так, определенная ментальная репрезентация обладает интенциональностью в том случае, если между ней и репрезентируемым объектом устанавливается каузальная связь. Одно из преимуществ такого объяснения состоит в том, что интенциональность не понимается исключительно как биологический феномен. Она может существовать также и у роботов и компьютеров, если установлена каузальная связь между их символами и объектами мира.

В итоге можно признать, что классическая модель, отстаиваемая Фодором, несмотря на свои недостатки, неплохо справляется с объяснением высших когнитивных функций человека, способна дать удовлетворительное решение проблемы ментальных репрезентаций и интенциональности, а также не идет в разрез с представлениями современной науки и здравым смыслом.

Литература

1. Патнэм Х. Сознание и машины // Патнэм Х. Философия сознания. М., 1999.
2. Searle J. Is the Brain Digital Computer // Searle J. Philosophy in a New Century. Selected Essays. Cambridge : University Press, 2008. P. 86–107.
3. Block N., Fodor A.J. What psychological states are not? // Block N. Consciousness, Function, and Representation. Collected Papers. Vol. 1 : A Bradford Book. Cambridge, Massachusetts : The MIT Press, 2007. P. 45–63.
4. Fodor J. Methodological Solipsism Considered as a Research Strategy in a Cognitive Psychology // Jerry A. Fodor Representations. Philosophical Essays on the Foundations of Cognitive Science. Harvester Press, 1981. P. 225–257.
5. Fodor J. Language of thought. Tomas Y. Crowell Company, Inc. New York, 1975.
6. McLaughlin B.P. Can an ICS Architecture Meet the Systematicity and Productivity Challenges? // Paco Calvo and John Symons (ed.) The architecture of cognition: rethinking Fodor and Pylyshyn's systematicity challenge. Cambridge, Massachusetts : The MIT Press. 2014. P. 31–77.
7. Fodor J. Concepts. Where Cognitive Science Went Wrong. Oxford : Clarendon Press, 1998.
8. Pylyshyn Z. Things and Places: How the Mind Connects with the World. Cambridge, Massachusetts : The MIT Press, 2007. 256 p.
9. Dennett D. A Cure for the Common Code? // Dennett D.C. Brainstorms. Philosophical Essays on Mind and Psychology. 1981. P. 90–108.
10. Uebel T. Vienna Circle // Stanford Encyclopedia of Philosophy. First published Wed Jun 28, 2006. URL: <https://plato.stanford.edu/entries/vienna-circle/#RedFouTwoCriParReb> (substantive revision Wed Feb 17, 2016).
11. Патнэм Х. Значение «значения» // Патнэм Х. Философия сознания. М., 1999.
12. Searle J. The Phenomenological Illusion // Searle J. Philosophy in a New Century. Selected Essays. Cambridge : University Press. 2008. P. 107–137.
13. Dennett D.C. From Bacteria to Bach and Back. The Evolution of Minds. W.W. Norton & Co., 2017.
14. Searle J.R. Minds, Brains and Programs // The Behavioral and Brain Sciences. 1980. № 3. P. 417–457.
15. Fodor J. Language of Thought. Revisited. Oxford : Oxford University Press, 2008.

16. Fodor J.A. Fodor's Guide to Mental Representation: The Intelligent Auntie's Vade-Mecum // Fodor J.A. Theory of Content and Others Essays. Cambridge, Massachusetts : The MIT Press, 1990. (Third print.)

17. Fodor J.A. Psychosemantics: The Problem of Meaning in the Philosophy of Mind. Cambridge, MA : MIT Press, 1987.

Vitalii I. Sukhoyi, Higher School of Economics (Moscow, Russian Federation).

E-mail: witaliy.suhovjy@gmail.com

Vestnik Tomskogo gosudarstvennogo universiteta. Filosofiya. Sotsiologiya. Politologiya – Tomsk State University Journal of Philosophy, Sociology and Political Science. 2018. 43. pp. 19–32.

DOI: 10.17223/1998863X/43/2

MENTAL REPRESENTATIONS, COMPUTABILITY AND INTENTIONALITY: JERRY FODOR'S APPROACH

Keywords: mental representations; computability; intentionality; classical model.

The paper is devoted to Jerry Fodor's classical computational theory of mind. Fodor is one of the brightest proponents of the theory, the one who developed it during almost all his research career. The paper discusses the major features of classical computationalism which starts from the Turing idea. The classical approach describes the mind as an independent program level. There is no great necessity in the knowledge of brain neurobiology, because we can describe the mind as an autonomous level, equal to the program level of the computer. The main point of the computational theory of mind is that the brain is a real analog of the computer, and the mind is a real analog of the program level: it is not a system which we could only describe as a set of programs, but a system which lacks the corresponding ontological status. The main feature of the work of this level is that syntactically driven computations are made over the mind's symbols – concepts. Computations are independent of the meaning of concepts; and the only thing the system should 'know' is the formal rules which govern the behaviour of concepts. The major advantage of a classical approach is that it could explain intentionality via causal links between objects in the world and symbols in the mind. It is shown that this explanation of intentionality can block John Searle's well-known Chinese Room argument. One of the strongest sides of a classical model is that it could explain intentionality, on the one hand, without appealing to biology terms and, on the other hand, without relativising or eliminating it. Another advantage of computationalism is that it uses such notions as mental states and understands mental states as representations. Of course, the computational theory of mind lost many of its proponents just because of the rising of opposing models – connectionism, for instance. And there has always been a strong eliminativism movement in philosophy of mind which finds additional power within connectionism, but classical computationalism has its own advantages because it saves and uses the notion of mental representations and intentionality and comports with our common-sense view.

References

1. Putnem, H. (1999a) *Filosofiya soznaniya* [Philosophy of Consciousness]. Translated from English by L. Makeeva, A. Nikiforov. Moscow: Dom intellektual'noy knigi.
2. Searle, J. (2008a) *Philosophy in a New Century. Selected Essays*. Cambridge: University Press. 2008. pp. 86–107.
3. Block, N. & Fodor, A.J. (2007) What psychological states are not? In: Block, N. *Consciousness, Function, and Representation*. Collected Papers. Vol.1. Cambridge, Massachusetts: The MIT Press. pp. 45–63.
4. Fodor, J. (1981) *Representations. Philosophical Essays on the Foundations of Cognitive Science*. Harvester Press. pp. 225–257.
5. Fodor, J. (1975) *Language of thought*. Tomas Y. Crowell Company, Inc. New York.
6. McLaughlin, B.P. (2014) Can an ICS Architecture Meet the Systematicity and Productivity Challenges? In: Calvo, P. & Symons, J. (eds) *The architecture of cognition: rethinking Fodor and Pylyshyn's systematicity challenge*. Cambridge, Massachusetts: The MIT Press. pp. 31–77.
7. Fodor, J. (1998) *Concepts. Where Cognitive Science Went Wrong*. Oxford: Clarendon Press.
8. Pylyshyn, Z. (2007) *Things and Places: How the Mind Connects with the World*. Cambridge, Massachusetts: The MIT Press.
9. Dennett, D. (1981) *Brainstorms. Philosophical Essays on Mind and Psychology*. Cambridge, Massachusetts: The MIT Press. pp. 90–108.

-
10. Uebel, T. (2006) Vienna Circle. In: Zalta, E.N. (ed.) *Stanford Encyclopedia of Philosophy*. [Online] Available from: <https://plato.stanford.edu/entries/vienna-circle/#RedFouTwoCriParReb>. (Accessed: February 17, 2016).
 11. Putnem, H. (1999b) *Filosofiya soznaniya* [Philosophy of Consciousness]. Translated from English by L. Makeeva, A. Nikiforov. Moscow: Dom intellektual'noy knigi.
 12. Searle, J. (2008b) *Philosophy in a New Century. Selected Essays*. Cambridge: University Press. 2008. pp. 107–137.
 13. Dennett, D.C. (2017) *From Bacteria to Bach and Back. The Evolution of Minds*. W.W. Norton & Co.
 14. Searle, J.R. (1980) Minds, Brains and Programs. *The Behavioral and Brain Sciences*. 3. pp. 417–457.
 15. Fodor, J. (2008) *Language of Thought*. Revisited. Oxford: Oxford University Press.
 16. Fodor, J. (1990) *Theory of Content and Others Essays*. Cambridge, Massachusetts: The MIT Press
 17. Fodor, J. (1987) *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: The MIT Press.