

Разум, истина и история.

Хилари Патнем

1981 год

1 Мозги в бочке

Муравей ползет по песку и оставляет на нем след. Совершенно случайно этот след таким образом изгибается и перекрещивается, что образует узнаваемую карикатуру на Уинстона Черчилля. Нарисовал ли муравей картинку Уинстона Черчилля – картинку, которая изображает Черчилля?

Большинство людей, не слишком вдаваясь в размышления, ответит отрицательно. В конце концов, муравей никогда не видел Черчилля или хотя бы изображения Черчилля; у него не было намерения изображать Черчилля. Он просто провел линию (и даже это непреднамеренно), которую мы можем “рассматривать как” изображение Черчилля.

Мы можем выразить это, сказав, что линия “сама по себе” не репрезентирует ¹ что бы то ни было в большей степени, чем нечто иное. Сходства (некоторого очень сложного вида) с чертами Уинстона Черчилля еще недостаточно для того, чтобы позволить чему-то репрезентировать Уинстона Черчилля или осуществлять к нему референцию. Но сходство и не является для этого необходимым: в нашем сообществе напечатанные значки “Уинстон Черчилль”, произнесенные слова “Уинстон Черчилль” и многие другие вещи используются для репрезентации Черчилля (хотя и не изобразительной), не обладая при этом сходством с Черчиллем того же вида, что и изображение, даже абрисный рисунок. Если сходство не является необходимым или достаточным для того, чтобы сделать нечто репрезентацией чего-то еще, то как вообще что бы то ни было может быть необходимым или достаточным для этой цели? Как вообще одна вещь может репрезентировать (“представлять” и т. п.) другую, отличную от нее?

Ответ может показаться легким. Предположим, что муравей видел Уинстона Черчилля, и предположим, что у него достаточно ума и способностей, чтобы его нарисовать. Предположим, что он создал эту карикатуру намеренно. Тогда эта линия действительно представляла бы

¹В этой книге термины “репрезентация” и “референция” всегда указывают на отношение между словом (или другими видами знаков, символов или представлений) и чем-то, что существует в действительности (т. е. не является лишь “мысленным объектом”). Есть некоторый смысл, в котором я могу “указывать” (осуществлять референцию) на то, что не существует, но это не тот смысл, в котором я здесь употребляю слово “референция”. Раньше то, что я называю “репрезентацией” и “референцией”, называли денотацией.

Кроме того, я следую образцу современных логиков и использую слово “существует”, подразумевая “существует в прошлом, настоящем или будущем”. Таким образом, Уинстон Черчилль “существует”, и мы можем “репрезентировать” Уинстона Черчилля или “осуществлять референцию” к нему даже несмотря на то, что его нет в живых.

Черчилля.

С другой стороны, предположим, что линия имела бы форму УИН-СТОН ЧЕРЧИЛЛЬ. И предположим, что это произошло бы совершенно случайно (не станем обращать здесь внимание на невероятность такого события). Тогда “печатные значки” УИНСТОН ЧЕРЧИЛЛЬ не репрезентировали бы Черчилля, хотя эти же печатные значки репрезентируют Черчилля практически в каждой книге, где они сегодня встречаются.

Поэтому может показаться, что для репрезентации необходимо (или главным образом необходимо) намерение.

Но для того, чтобы иметь намерение, что что бы то ни было, даже индивидуальный язык (и даже слова “Уинстон Черчилль”, произнесенные мною про себя, а не вслух), должно репрезентировать Черчилля, я должен прежде всего быть способным подумать о Черчилле. Если линии на песке, звуки и т. п. не могут ничего репрезентировать “сами по себе”, то как могут что-то репрезентировать “сами по себе” мысленные формы? Или все-таки могут? Как мышление может обратиться к чему-то внешнему и ухватить его?

В прошлом некоторые философы перескакивали от такого рассмотрения к тому, что они считали доказательством сущностно не-физической природы сознания. Аргумент здесь прост: сказанное нами о муравьиной кривой применимо к любому физическому объекту. Ни один физический объект не может сам по себе указывать на одну вещь скорее, чем на другую; тем не менее, мысли в сознании очевидно преуспевают в указании на одну вещь скорее, чем на другую. Поэтому мысли (и, соответственно, сознание) по природе своей сущностно отличаются от физических объектов. Мысли обладают характеристикой интенциональности — они могут указывать на нечто иное; ничто физическое не обладает “интенциональностью”, не считая того обстоятельства, что интенциональность производна от некоторого употребления этой физической вещи сознанием (или, по крайней мере, так признается). Однако это слишком торопливое решение; простое постулирование таинственных сил сознания само по себе ничего не дает. Но проблема вполне реальна. Как возможна интенциональность, референция?

2 Магические теории референции

Мы видели, что созданная муравьем “картинка” не имеет никакой необходимой связи с Уинстоном Черчиллем. Тот факт, что “картинка” имеет определенное “сходство” с Черчиллем, сам по себе не делает ее настоящей картинкой или не делает ее репрезентацией Черчилля. Если только

муравей не является разумным муравьем (каковым он не является) и не знает Черчилля (а он не знает), начерченная им кривая — не картинка и даже не репрезентация чего бы то ни было. Некоторые примитивные народы верят, что репрезентации (в особенности имена) связаны со своими носителями с необходимостью; что знание “подлинного имени” кого-либо или чего-либо дает подлинную власть над этим человеком или вещью. Эта власть происходит из волшебной связи между именем и его носителем; когда же мы понимаем, что имена связаны со своими носителями лишь контекстуально, случайно, конвенционально, нам уже трудно увидеть, каким образом знание имени может иметь какое бы то ни было мистическое значение.

Важно понять, что для физических изображений действуют те же правила, что и для ментальных образов и вообще для ментальных репрезентаций; ментальные репрезентации связаны с тем, что они представляют, не с большей необходимостью, чем физические. Обратное предположение — пережиток магического мышления.

Возможно, легче понять это замечание в случае с ментальными образами. (Возможно, первым философом, ухватившим огромное значение этого замечания, даже если он не первым его обнаружил, был Витгенштейн). Предположим, что где-то есть планета, населенная человеческими существами — появившимися там в результате эволюции, или высадки инопланетян, или как бы то ни было. Предположим, что эти люди — во всем остальном полностью похожие на нас — никогда не видели деревьев. Предположим, что они никогда не представляли деревьев (возможно, растительность существует на их планете лишь в виде плесени). Предположим, что однажды на их планету случайно попала картинка, изображающая дерево; ее обронил пролетающий мимо космический корабль, никак не вступивший с ними в контакт. Представьте себе, как они ломают голову над этой картинкой. Что бы это могло быть? Им приходят на ум всевозможные предположения: может быть, это какое-то строение? Балдахин? Или какое-то животное? Но они, как мы продолжаем предполагать, не приближаются к истине.

Для нас картинка является репрезентацией дерева. Для этих же людей картинка представляет лишь некоторый странный объект, чьи природа и назначение неизвестны. Предположим, что в результате взгляда на картинку у одного из них возник ментальный образ, в точности подобный одному из моих ментальных образов дерева. Однако его ментальный образ не является репрезентацией дерева. Это лишь репрезентация странного предмета (чем бы он ни был), который представляет таинственная картинка.

Здесь можно возразить, что этот ментальный образ в действитель-

ности является репрезентацией дерева при условии, что картинка, вызвавшая этот ментальный образ, сама изначально была репрезентацией дерева. Существует каузальная цепь, ведущая от действительных деревьев к ментальному образу, даже если эта цепь весьма необычна.

Но можно представить себе ситуацию, где даже и эта каузальная цепь отсутствует. Предположим, что “картинка дерева”, которую обронил космический корабль, была на самом деле не картинкой, изображающей дерево, а случайным результатом разлития красок. Даже если бы этот результат выглядел в точности как картина дерева, он поистине не в большей степени изображал бы дерево, чем муравьиная “карикатура” на Черчилля изображала бы Черчилля. Мы даже можем представить себе, что космический корабль, обронивший картинку, прилетел с планеты, где тоже ничего не знают о деревьях. Тогда у рассматривающих картинку людей возникли бы ментальные образы, количественно идентичные моему образу дерева, но эти образы не представляли бы дерево в большей степени, нежели что-то другое.

То же самое относится и к словам. Ряд значков на бумаге может быть прекрасным описанием деревьев, но если он был произведен обезьянами, наугад нажимающими клавиши на машинке в течение миллионов лет, то эти слова ни на что не указывают. И если бы кто-то запомнил эти слова и произнес их про себя, не понимая, то они ни на что не указывали бы, будучи помысленными.

Представим себе, что человек, произносящий эти слова про себя, оказался загипнотизирован. Предположим, что эти слова — на японском, а нашему герою внушили, что он понимает японский. Предположим, что по мере того, как он мыслит эти слова, у него появляется “чувство понимания”. (Хотя если кто-нибудь вмешается в ход его мысли и спросит его, что значат те слова, которые он мыслит, то он обнаружит, что не может ответить.) Возможно, иллюзия нашего героя будет настолько полной, что он сможет обмануть японца-телепата. Но если он не мог бы использовать слова в правильных контекстах, отвечать на вопросы о том, что он “думает” и т. д., то он не понимал бы их.

Комбинируя рассказанные мною научно-фантастические истории, мы можем выдумать такой случай, когда некто произносит в уме слова, которые фактически являются описаниями слов в некотором языке, и одновременно у него возникают соответствующие ментальные образы, но при этом он не понимает эти слова и не знает, что такое дерево. Мы даже можем представить себе, что ментальные образы были вызваны разлитием красок (хотя наш герой под воздействием гипноза думает, что это — образы чего-то соответствующего его мысли, но не может ответить на вопрос о том, что же это такое). И мы можем представить, что язык, на

котором думает наш герой — это такой язык, о котором ни гипнотизер, ни гипнотизируемый никогда и не слышали. Возможно, эти “бесмысленные предложения”, каковыми их полагает гипнотизер, являются описаниями деревьев на японском лишь в силу простого совпадения. Короче говоря, все происходящее в сознании нашего героя может быть количественно идентично тому, что происходит в сознании носителя японского языка, который действительно думает о деревьях — но ничто из этого не будет указывать на деревья.

Конечно же, все это в действительности невозможно, в том смысле, что в действительности невозможно, чтобы обезьяна случайно напечатала “Гамлета”. Иными словами, обратная вероятность столь велика, что это никогда не произойдет в действительности (как мы полагаем). Но это не является логически невозможным или даже физически невозможным. Это могло бы произойти (в соответствии с физическими законами и, возможно, с действительным состоянием вселенной, если в ней много планет с разумными существами). И если бы это произошло, это было бы поразительное проявление важной концептуальной истины: что даже обширная и сложная система репрезентаций, как словесных, так и визуальных, не имеет внутренней, встроенной, волшебной связи с тем, что она представляет — связи, независимой от того, как она причинно обусловлена, и от того, что каковы диспозиции говорящего или думающего. И это справедливо и в том случае, когда система репрезентаций (слов и образов, в случае наших примеров) воплощена физически — слова написаны или произнесены, а картинки являются физическими изображениями, — и в том, когда она воплощена лишь в сознании. Помысленные слова и ментальные картины не представляют внутреннего, на что они указывают.

3 Дело о мозгах в бочке

Одна из научно-фантастических возможностей, обсуждаемых философами, такова. Представим себе, что человеческое существо (допустим, вы сами) подверглось операции, сделанной злодеем-ученым. Ваш мозг был отделен от тела и помещен в бочку с питательным раствором, поддерживающим этот мозг живым. Нервные окончания были присоединены к некоему супернаучному компьютеру, поддерживающему у той личности, которой принадлежит этот мозг, полную иллюзию, что все в порядке. Как будто вокруг люди, предметы, небо и т. д.; но в действительности все, что ощущает наш герой, то есть вы, является результатом электронных импульсов, передающихся от компьютера к нервным окон-

чаниям. Компьютер такой умный, что если наш герой пытается поднять руку, то реакция компьютера послужит причиной того, что он “увидит” и “почувствует”, как рука поднимается. Более того, меняя программу, злодей-ученый может заставить жертву “почувствовать” (или прогаллюцинировать) любую ситуацию или окружение, какие хочет. Он может также стереть память об операции на мозге, в результате чего жертва будет считать, что всегда была в таком окружении. Жертве даже может казаться, что она сейчас сидит и читает вот эти слова о забавном, но совершенно абсурдном предположении, что есть такой злодей-ученый, который удаляет людям мозги и помещает их в бочки с питательным раствором, поддерживающим эти мозги живыми. При этом нервные окончания присоединены к некоему супернаучному компьютеру, поддерживающему у той личности, которой принадлежит этот мозг, полную иллюзию, что...

Когда такая возможность упоминается на лекциях по теории познания, то цель, конечно, состоит в том, чтобы современными средствами поставить классическую проблему скептицизма по отношению к внешнему миру. (Откуда вы знаете, что вы сами не находитесь в этом затруднении?) Но это затруднение также полезно в качестве инструмента для постановки вопросов, касающихся отношения между сознанием и миром.

Вместо одного-единственного мозга в бочке теперь представим себе, что все человеческие существа (возможно, все существа, наделенные ощущениями) представляют собой мозги в бочке (или нервные системы в бочке для тех существ с минимальными нервными системами, которые уже считаются “чувствующими”). Конечно же, злодей-ученый должен находиться снаружи. Или не должен? А может быть, нет никакого злодея-ученого, а просто вселенная (хотя это и абсурдно) состоит из автоматических механизмов, обслуживающих бочку с мозгами и нервными системами?

На этот раз предположим, что автоматические механизмы запрограммированы так, чтобы дать нам массовую галлюцинацию, а не множество отдельных, не связанных галлюцинаций. Таким образом, когда мне кажется, что я разговариваю с вами, вам кажется, что вы слышите мои слова. Конечно, это не значит, что мои слова действительно достигают ваших ушей — поскольку у вас (в действительности) нет ушей, а у меня рта и языка. Скорее, когда я порождаю слова, происходит следующее:

центробежные импульсы проходят от моего мозга к компьютеру, что заставляет меня “слышать” мой собственный голос, произносящий эти слова, “чувствовать”, как мой язык двигается, и т. д., а вас заставляет “слышать” мои слова, “видеть”, как я говорю и т. д. В этом случае мы, в некотором смысле, действительно общаемся. Я не заблуждаюсь отно-

сительно вашего действительного существования (только относительно существования вашего тела и “внешнего мира”, отдельно от мозгов). С определенной точки зрения, даже неважно, что “целый мир” — это лишь коллективная галлюцинация; поскольку вы, в конце концов, действительно слышите мои слова, когда я говорю с вами, даже если механизм этого процесса не таков, как мы предполагаем. (Конечно, если бы мы были влюбленными, занимающимися любовью, а не просто двумя собеседниками, то предположение о том, что мы всего лишь два мозга в бочке, разочаровало бы нас).

Теперь я хочу задать вопрос, который покажется очень глупым и очевидным (по крайней мере некоторым людям, включая некоторых весьма изощренных философов), но который довольно быстро заведет нас в настоящие философские бездны. Предположим, что вся эта история — действительно правда. Могли бы мы, если бы мы были такими мозгами в бочке, сказать или подумать, что мы таковы?

Я буду доказывать, что ответ на этот вопрос отрицательный. Фактически я буду утверждать, что предположение о том, что мы на самом деле мозги в бочке, не может быть истинным, хотя оно не нарушает физических законов и вполне совместимо со всем нашим опытом. Оно не может быть истинно, потому что в определенном смысле оно само себя отрицает.

Аргумент, который я здесь хочу представить, необычен, и мне понадобилось несколько лет, чтобы убедить себя в его правильности. Но это корректный аргумент. Станным его заставляет выглядеть то обстоятельство, что он оказывается связанным с некоторыми самыми глубокими философскими вопросами. (Он впервые пришел мне в голову, когда я размышлял над одной из теорем современной логики — теоремой Сколема-Левенгейма, и внезапно увидел связь между этой теоремой и некоторыми аргументами “Философских исследований” Витгенштейна.)

“Самоотрицающее предположение” — это такое утверждение, истинность которого предполагает его же ложность. Например, рассмотрим тезис “Все общие утверждения ложны”. Это общее утверждение. Тогда, если оно истинно, то оно должно быть ложно. Следовательно, оно ложно. Иногда тезис называют самоотрицающим, если его ложность вызывается предположением, что тезис вообще поддержан или провозглашен. Например, тезис “Я не существую” является самоотрицающим (для любого “я”). Поэтому если кто-то думает о том, что он существует, то можно быть уверенным, что он существует (как утверждал Декарт).

Я покажу, что предположение о том, что мы представляем собой мозги в бочке, обладает именно этим свойством. Если мы вообще можем оценивать его с точки зрения истинности или ложности, то оно не истинно

(как я покажу). Следовательно, оно не истинно.

Прежде, чем я приведу аргумент, рассмотрим, почему нам кажется таким странным то обстоятельство, что такой аргумент вообще может быть дан (по крайней мере некоторым философам — приверженцам концепции истины как подобия). Мы признали, что то положение, что существует мир, где все наделенные ощущениями существа представляют собой мозги в бочке, совместимо с законами физики. Как выражаются философы, есть “возможный мир”, где все наделенные ощущениями существа представляют собой мозги в бочке. (Рассуждение о “возможном мире” таково, что оно как бы предполагает существование места, где всякое абсурдное предположение истинно; поэтому в философии оно может вводить в заблуждение.) Человеческие существа в этом возможном мире обладают теми же знаниями, что и мы. У них те же мысли (по крайней мере, через их сознание проходят те же слова, образы, формы мысли и т. д.). Однако я утверждаю, что мы можем дать аргумент, показывающий, что мы — не мозги в бочке. Как возможен такой аргумент? И почему его не могли бы точно так же выдвинуть люди в возможном мире, действительно являющиеся мозгами в бочке?

Ответ будет (вкратце) таков: хотя люди в этом возможном мире могут мыслить и “произносить” все слова, которые мы мыслим и произносим, они не могут (как я утверждаю) указывать на то же, на что и мы. В частности, они не могут помыслить или сказать, что они — мозги в бочке (даже думая “мы -мозги в бочке”).

4 Тест тьюринга

Предположим, что кому-то удалось изобрести компьютер, который может поддерживать осмысленную беседу с человеком (на столько тем, на сколько мог бы разумный человек). Как можно решить, “разумен” ли компьютер?

Английский логик Алан Тьюринг предложил следующий тест²: пусть некто поддерживает разговор с компьютером и с неизвестным человеком. Если испытатель не сможет различить, где компьютер, а где человек (предположим, что тест повторяется достаточное количество раз с разными собеседниками), то компьютер разумен. Вкратце, вычислительная машина разумна, если она может пройти “тест Тьюринга”. (Разговор ведется, разумеется, не лично, поскольку испытатель не должен знать

²Turing A. M. Computing Machinery and Intelligence // Mind, 1950. Перепечатано в: A. R. Andfrson (ed.). Minds and Machines.

облик своих собеседников. Нельзя использовать и голос, так как механический голос может просто звучать иначе, чем человеческий. Скорее представим себе, что беседа ведется с помощью электрической пишущей машинки. Испытатель печатает на ней утверждения, вопросы и т. п., а его партнеры — машина и человек — отвечают с помощью электрической клавиатуры. Кроме того, машина может лгать — на вопрос “Ты машина?” она может ответить: “Нет, я тут лаборант”.)

Предложение такого теста в качестве решающего теста на разумность было подвергнуто критике многими авторами (в принципе никоим образом не враждебными идее о том, что машина может быть разумной). Но не это сейчас занимает нас. Я хочу использовать общую идею теста Тьюринга, общую идею диалогического теста на компетентность, для другой цели — цели исследования понятия референции.

Представим ситуацию, в которой задача состоит не в том, чтобы определить, является ли собеседник человеком или машиной, а в том, чтобы определить, использует ли он слова с той же референцией, что и мы. Очевидно, тест состоял бы в том, чтобы так же вести разговор и, если при этом не возникает проблем, если собеседник оказывается неотличимым от кого-то, о ком нам заранее достоверно известно, что он говорит на том же языке, то мы заключаем, что собеседник указывает на предметы тем же способом, что и мы. В тех случаях, когда тест Тьюринга направлен описанным образом на определение существования (общей, разделяемой) референции, я буду называть его Тестом Тьюринга на Референцию. Подобно тому, как философы обсуждали вопрос, является ли изначальный тест Тьюринга решающим тестом на разумность — т. е. вопрос о том, является ли машина, проходящая тест не один раз, а регулярно, разумной с необходимостью, — я хочу обсудить вопрос о том, является ли предложенный Тест Тьюринга на Референцию решающим тестом на общую референцию.

Ответ окажется отрицательным. Тест Тьюринга на Референцию не является решающим. На практике это, конечно, превосходный способ проверки; однако логически возможно (хотя, конечно, крайне маловероятно), чтобы кто-то прошел Тест Тьюринга на Референцию, ни на что не указывая. Отсюда следует, как мы увидим, что мы можем расширить наше наблюдение о том, что слова (и целые тексты и дискурсы) не связаны со своими референтами с необходимостью. Даже если мы рассматриваем не слова сами по себе, а правила, определяющие, какие слова могут быть надлежащим образом произведены в данном контексте — даже если мы рассматриваем, выражаясь компьютерным жаргоном, программы для использования слов, — здесь все еще не появится определенная референция, если только эти программы сами не указывают

на нечто внеязыковое. Это будет решающим шагом в процессе достижения заключения о том, что Обитатели Мира Мозгов-в-Бочке вообще не могут указывать на что-либо внешнее (и тем самым не могут сказать, что они — Обитатели Мира Мозгов-в-Бочке).

Предположим, например, что я нахожусь в Тьюринг-ситуации (веду “имитационную игру”, в терминологии Тьюринга) и мой собеседник в действительности машина. Предположим, что машина запрограммирована так, что способна выиграть (“пройти” тест). Представим, что машина запрограммирована давать прекрасные ответы по-английски на утверждения, вопросы, замечания и т. д. на английском языке, но у машины нет ни органов чувств (кроме подсоединения к электрической пишущей машинке), ни органов движения (кроме электрической пишущей машинки). (Насколько я себе представляю, Тьюринг не утверждает, что обладание органами чувств или движения необходимо для наличия сознания или разума). Допустим, что не только у машины нет электронных ушей, глаз и т. д., но также в ее программе для ведения имитационной игры отсутствуют соответствующие модули для введения данных с таких органов чувств или для контроля над телом. Что мы скажем о такой машине?

Для меня очевидно, что мы не можем и не должны приписывать референции такому устройству. В самом деле, машина может прекрасно рассуждать, скажем, о видах Новой Англии, но она не может распознать яблоню или яблоко, гору или корову, поле или колокольню, если бы они находились перед ней.

Мы получили машину для производства предложений в ответ на предложения. Но ни одно из этих предложений никак не связано с реальным миром. Если соединить две машины и заставить их играть в имитационную игру друг с другом, то они будут “обманывать” друг друга бесконечно, даже если весь остальной мир исчезнет! К тому, чтобы рассматривать речь машины о яблоках как указывающую на яблоки реального мира, не больше оснований, чем рассматривать “рисунок” муравья” как указывающий на Уинстона Черчилля.

Иллюзию референции, значения, разумности и т. д. здесь производит тот факт, что речь машины указывает на яблоки, колокольни, Новую Англию и т. д. в пределах нашей конвенции по поводу репрезентации. Точно так же может существовать иллюзия, что муравей изобразил Черчилля — по той же самой причине. Но мы можем воспринимать яблоки и поля, воздействовать на них и взаимодействовать с ними. Наша речь о яблоках и полях тесно связана с невербальными отношениями между нами и этими предметами. Есть “правила языкового входа”, которые ведут нас от знаний о яблоках к таким произнесениям, как “Я вижу яб-

локо”, и “правила языкового выхода”, которые ведут нас от выраженных в языковой форме решений (“Куплю яблок”) к Действиям — другим действиям, нежели говорение. В отсутствие либо правил языкового входа, либо правил языкового выхода нет причин рассматривать речь машины (или двух машин, если мы берем случай с двумя машинами, играющими в имитационную игру друг с другом) как нечто большее, чем синтаксическая игра. Синтаксическая игра, которая походит на осмысленную речь, — пожалуй, но лишь в той степени (и не в большей), в какой муравьиная кривая походит на язвительную карикатуру. В случае с муравьем мы могли бы привести такой аргумент: муравей провел бы ту же кривую, если бы Черчилль никогда не существовал. В случае с машиной мы не можем рассуждать подобным образом; если бы яблоки, деревья, колокольни и поля не существовали, то, по всей вероятности, программисты не создали бы именно эту программу. Хотя машина не воспринимает яблоки, поля или колокольни, ее создатели-проектировщики воспринимали все это. Существует некоторая каузальная связь между машиной и яблоками и т. д. реального мира — связь через перцептуальный опыт и знания создателей-проектировщиков. Но такая слабая связь едва ли достаточна для референции. Прежде всего, логически возможно, хотя и фантастически невероятно, что та же самая машина могла бы существовать и в том случае, если бы яблоки, поля и колокольни не существовали. Но и кроме того, что более важно, машина в высшей степени нечувствительна к продолжающемуся существованию яблок, полей, колоколен и т. д. Даже если бы все эти вещи прекратили существовать, машина все же успешно вела бы речь прежним образом. Вот почему вообще нельзя рассматривать машину как осуществляющую референцию.

Для нашей дискуссии важно заметить, что в тесте Тьюринга нет ничего, что отсеяло бы машину, запрограммированную лишь на имитационную игру и не умеющую делать ничего другого, а машина, не умеющая делать ничего, кроме как вести имитационную игру, явно указывает на что-либо не в большей степени, чем магнитофон.

5 Мозги в бочке (еще раз)

Сравним гипотетические “мозги в бочке” с описанными машинами. Очевидны важные различия. У мозгов в бочке нет органов чувств, но есть входы для данных органов чувств; т. е. есть центростремительные нервные окончания, есть ввод данных с этих центростремительных нервных окончаний, и эти входные данные играют такую же роль в “программе” мозгов в бочке, что и в программе нашего мозга. В силу этих причин

может показаться абсурдным отрицать их наделенность сознанием или разумом. Но тот факт, что они сознательны и разумны, не означает, что их слова указывают на то же, на что и наши. Интересующий нас вопрос таков: действительно ли их вербализации, содержащие, скажем, слово “дерево”, указывают на деревья? В более общем виде: могут ли они вообще указывать на внешние объекты? (В противоположность, например, объектам, составляющим образ, порожденный автоматическим механизмом).

Для того, чтобы сформулировать наши идеи, уточним, что автоматические механизмы, как предполагается, появились в результате какой-то космической случайности или совпадения (или, может быть, всегда существовали). В этом гипотетическом мире у автоматических механизмов нет разумных создателей-проектировщиков. Фактически, как мы заметили в начале главы, мы можем представить, что все наделенные ощущениями существа (как минимальна бы ни была их чувствительность) находятся внутри бочки.

Это допущение не способствует достижению успеха. Нет связи между словом “дерево”, употребляемым этими мозгами, и настоящими деревьями. Они будут употреблять слово “дерево” тем же способом, будут иметь те же мысли и образы, даже если деревьев не будет существовать. Их образы, слова и т. д. количественно идентичны с образами, словами и т. д., которые представляют деревья в нашем мире; но мы уже видели (вспомним муравья!), что качественное подобие чему-либо, что представляет объект (Уинстона Черчилля или дерево), само по себе не дает репрезентации. Коротко говоря, мозги в бочке не думают о реальных деревьях, когда они думают “передо мной дерево”, потому что нет ничего такого, в силу чего их мысль “дерево” представляла бы реальное дерево.

Если такой вывод покажется поспешным, подумайте над следующим: мы видели, что слова не обязательно указывают на деревья, даже если они организованы в последовательность, идентичную такой речи, которая (если бы она возникла в одном из наших сознаний) безусловно относилась бы к деревьям в действительном мире. Так же и “программа” — в смысле правил, практик, Диспозиций мозгов к вербальному поведению — не указывает с необходимостью на деревья и не вызывает референцию к деревьям через те связи, которые она устанавливает между словами и Другими словами, между языковыми репликами и языковыми ответами. Если эти мозги думают о деревьях, имеют референцию к деревьям, репрезентируют деревья (реальные деревья, находящиеся за пределами бочки), то это возможно в силу того, что “программа” соединяет языковую систему с невербальными входами и выходами. В мире Мозгов-в-Бочке существуют такие невербальные входы и выходы (опять

эти центробежные и центростремительные нервные окончания!), но мы уже видели, что “чувственные данные”, произведенные автоматически механизмами, не репрезентируют деревья (или что бы то ни было внешнее), даже если они в точности схожи с нашими образами деревьев. Подобно тому, как пятна краски могут походить на изображение дерева, не являясь изображением дерева, также, мы видели, “чувственные данные” могут быть количественно идентичны “образу дерева”, не будучи образом дерева. Как может тот факт, что, в случае мозгов в бочке, язык связан программой с сенсорными входами, которые не репрезентируют, ни внутренне, ни внешне, деревья (или что бы то ни было внешнее) — как может это привести к тому, что полная система репрезентаций, используемый язык, указывает на нечто внешнее или репрезентирует его?

Ответ состоит в том, что это невозможно. Полная система чувственных данных, двигательные сигналы на центробежных окончаниях и вербально или концептуально опосредованная мысль, связанная “правилами языкового входа” с чувственными данными (или чем бы то ни было) в качестве входных данных и “правилами языкового выхода” с двигательными сигналами в качестве выходных данных — эта система не более связана с деревьями, чем муравьиная кривая с Уинстоном Черчиллем. Как только мы увидели, что качественное подобие (приводящее, если хотите, к качественной идентичности) между мыслями мозгов в бочке и чьими-то мыслями в реальном мире никоим образом не ведет к подобию референции, то становится ясно, что совершенно нет оснований рассматривать мозг в бочке как указывающий на внешние объекты.

6 Посылки аргумента

Итак, я привел обещанный аргумент, который должен был показать, что мозги в бочке не могут подумать или сказать, что они мозги в бочке. Остается только дать его окончательную, явную формулировку и исследовать его структуру.

Исходя из того, что было сказано выше, мы принимаем, что когда мозг в бочке (в мире, где каждое наделенное ощущениями существо является и всегда было мозгом в бочке) думает: “Вот передо мной дерево”, то его мысль не указывает на настоящие деревья. Согласно некоторым теориям, которые мы будем обсуждать, она может указывать на образы деревьев или на электронные импульсы, которые вызывают представления о деревьях. или на детали программы, управляющие этими импульсами. Тем, что было сказано выше, эти теории не исключаются, поскольку существует тесная каузальная связь между употреблением слова

“дерево” в бочковом английском и наличием образов деревьев, наличием электронных импульсов и наличием определенных деталей машинной программы. Согласно этим теориям, мозг прав, а не ошибается, когда думает: “Вот передо мной дерево”. Принимая во внимание то, на что в бочковом английском указывает слово “дерево”, и то, на что указывает слово “перед”; принимая, далее, что одна из этих теорий правильна, мы получим, что условия истинности предложения “Вот передо мной дерево” в бочковом английском языке состоят в том, что образ дерева состоит с обсуждаемым “я” (принадлежащим той же образной системе) в отношении “быть перед”. Или же, возможно, эти условия истинности состоят в том, что электронные импульсы того вида, который обычно производит такой опыт, исходят из автоматических механизмов; или же — в том, что задействованы те особенности или детали этих механизмов, которые, как предполагается, производят знание “Вот передо мной дерево”. И эти условия истинности явно выполняются.

В силу того же аргумента “бочка” в бочковом английском указывает на образ бочки или что-либо связанное с ним (электронные импульсы или детали программы), но явно не на реальную бочку, поскольку употребление слова “бочка” в бочковом английском каузально не связано с реальными бочками (за исключением связи, состоящей в том, что мозги в бочке не могли бы употреблять слово “бочка”, если бы не наличие конкретной бочки — той, в которой они сидят; но эта связь возникает и между употреблением каждого слова в бочковом английском и этой конкретной бочкой; это не особенная связь между употреблением конкретного слова, “бочка” и бочками). Точно так же “питательный раствор” указывает в бочковом английском на образ жидкости или что-либо связанное с ним (электронные импульсы или детали программы). Следовательно, если их “возможный мир” в действительности является реальным миром, а мы в действительности представляем собой мозги в бочке, тогда то, что мы теперь подразумеваем под выражением “Мы — мозги в бочке” — это что мы образно представляем собой мозги в бочке или что-то в этом роде (если мы вообще подразумеваем этим что бы то ни было). Но в гипотезу о том, что мы являемся мозгами в бочке, входит представление о том, что мы не являемся образом мозгов в бочке (т. е. то, что мы “галлюцинируем” — это не что мы мозги в бочке). Итак, если мы — мозги в бочке, то предложение “Мы — мозги в бочке” сообщает нечто ложное (если оно вообще сообщает что бы то ни было). Короче говоря, если мы — мозги в бочке, то “Мы — мозги в бочке” ложно. Следовательно, это (с необходимостью) ложно.

Предположение о том, что такая возможность имеет смысл, возникает из сочетания двух ошибок: (1) слишком серьезного отношения к

физической возможности и (2) бессознательного обращения к магической теории референции — теории, согласно которой определенные ментальные репрезентации с необходимостью указывают на определенные внешние предметы и виды предметов.

Имеется “физически возможный мир”, в котором мы являемся мозгами в бочке — что это значит, кроме того, что имеется описание такого положения дел, которое совместимо с законами физики? Поскольку в нашей культуре (начиная еще с XVII столетия) существует тенденция воспринимать физику как метафизику, т. е. рассматривать точные науки как долгожданное описание “полного и окончательного содержания вселенной”, постольку существует, как ее неосредственное следствие, тенденция принимать “физическую возможность” за истинный критерий того, что может иметь место в действительности. С такой точки зрения, истина есть физическая истина; возможность — физическая возможность, а необходимость — физическая необходимость. Но мы только что видели, даже если пока лишь на весьма искусственном примере, что такие взгляды ошибочны. Существование “физически возможного мира”, в котором мы являемся мозгами в бочке (всегда были ими и всегда будем) не означает, что мы можем в действительности, по-настоящему, с определенной вероятностью быть мозгами в бочке. И устраняет эту возможность не физика, а философия.

Некоторые философы, стремясь предъявить и одновременно свести к минимуму требования своей профессии (типичное состояние умов в англо-американской философии XX века), сказали бы: “Ну да, вы показали, что некоторые из тех вещей, которые кажутся физически возможными, в действительности являются концептуально невозможными. Что же в этом удивительного?”

Разумеется, мой аргумент может быть охарактеризован как “концептуальный”. Но характеристика философской деятельности как поиска “концептуальных” истин превращает ее в исследование значений слов. А это вовсе не то, чем мы собирались заняться.

Наша задача — рассмотрение предварительных условий мышления. о чем-либо, репрезентации чего-либо, указания на что-либо и т. п. Мы исследовали эти предпосылки не путем исследования значений этих слов и фраз (как это мог бы сделать, например, лингвист), а путем априорного рассуждения — не в старом “абсолютном” смысле (поскольку мы не заявляем, что магические теории референции априорно ошибочны), но в смысле исследования того, как возможно разумно допустить определенные общие посылки или в смысле принятия определенных, достаточно широких теоретических допущений. Такая процедура не является ни “эмпирической”, ни вполне “априорной”, но включает в себя элементы обоих

способов исследования. Несмотря на погрешность моей процедуры и на ее зависимость от допущений, которые могут быть охарактеризованы как “эмпирические” (например, допущение о том, что сознание не имеет другого доступа к внешним вещам или свойствам, кроме предоставляемого чувствами), она тесно связана с тем, что Кант называл “трансцендентальным” исследованием, поскольку это — повторяю еще раз — исследование предпосылок референции и тем самым мышления предпосылок, встроенных в саму природу нашего сознания, хотя и не полностью независимых (как надеялся Кант) от эмпирических допущений.

Одна из посылок аргумента очевидна и состоит в том, что магические теории референции ошибочны — ошибочны не только для физических, но и для ментальных репрезентаций. Другая посылка состоит в том, что невозможно указывать на определенные виды вещей без всякого каузального взаимодействия с ними³ или с теми вещами, в терминах которых они могут быть описаны. Но почему мы должны принимать эти посылки? Поскольку они представляют широкий фон, на котором я рассуждаю, пора рассмотреть их подробнее.

7 Причины отрицания необходимых связей между репрезентациями и их референтами

Как я уже упоминал, некоторые философы (наиболее известный из которых Brentano) приписывали сознанию некоторую силу- “интенциональность”, – в точности позволяющую ему осуществлять референцию. Очевидно, я отбросил эту идею как не дающую решения. Но что дает мне право так поступить? Может быть, я поспешил?

Эти философы не утверждали, что мы можем мыслить о внешних вещах или свойствах, вообще не используя репрезентаций. Приведенный мной при сравнении чувственных данных с муравьиной “картиной” аргумент (аргумент от научно-фантастической истории о “картине” дерева, возникшей из пролитой краски и породившей чувственные данные, качественно подобные нашим “визуальным образам деревьев”, но не сопровождавшейся никаким понятием дерева) был бы принят ими как показывающий, что образы не осуществляют референцию с необходимостью.

³Если бы Мозги в Бочке имели каузальную связь, скажем, с деревьями в будущем, то, возможно, они могли бы указывать на деревья сейчас при помощи дескрипции “вещи, на которые я буду указывать как на “деревья” в такой-то и , такой-то момент в будущем”. Но мы должны представить себе такой случай, в котором Мозги в Бочке никогда не выберутся из бочки, а следовательно, никогда не вступят в каузальную связь с деревьями и т. д.

Если ментальные репрезентации осуществляют референцию (к внешним вещам) с необходимостью, то они должны отражать природу понятий, а не природу образов. Но что такое понятия?

При интроспекции мы не воспринимаем как таковые “понятия”, проплывающие по сознанию. Остановив поток мысли во времени или в пространстве, мы обнаружим слова, образы, ощущения, чувства. Когда я высказываю вслух свои мысли, я не мыслю их дважды. Я слышу свои слова по мере того, как говорю. Разумеется, я чувствую разницу между тем, когда я произношу слова, которым верю, и когда я произношу слова, которым не верю (но иногда, когда я нервничаю или перед враждебной аудиторией, я могу чувствовать себя как если бы я лгал, тогда, когда я говорю правду); и я чувствую разницу между тем, когда я произношу слова, которые понимаю, и когда я произношу слова, которые не понимаю. Но я без труда могу представить себе кого-либо, кто мыслит именно эти слова (т. е. проговаривает их в уме) и переживает то же самое понимание, утверждение и т. п., что и я, но через минуту (или будучи разбужен гипнотизером) осознает, что вообще не понимал того, что только что прошло через его сознание: что он даже не понимает того языка, которому принадлежат эти слова. Я не утверждаю, что это весьма вероятно: я просто хочу сказать, что в этом нет ничего невообразимого. И это показывает нам не то, что понятия и есть слова (или образы, ощущения и т. д.), а что приписывать кому-либо “понятие” или “мысль” — это совсем не то, что приписывать ему ментальные “представления”, какие бы то ни было сущности или события, доступные внутреннему самонаблюдению. Понятия не являются ментальными представлениями, в силу присущих им свойств указывающими на внешние объекты — по той решающей причине, что они вообще не являются ментальными представлениями. Понятия — это употребленные определенным способом знаки; они могут быть публичными или частными, ментальными или физическими сущностями, но даже когда знаки являются “ментальными” и “частными”, сам по себе знак, в отрыве от своего использования, еще не есть понятие. И знаки не обладают референцией в силу своих свойств.

Мы можем увидеть это, проделав очень простой мысленный эксперимент. Предположим, что вы, как я, не можете отличить вяза от бука. Мы все же скажем, что референция слова “вяз” в моей речи — та же, что и референция слова “вяз” в речи других людей, а именно дерева-вяза, и что множество всех букв является экстенсионалом термина “бук” (т. е. множеством всех вещей, которым может быть истинно предсказано слово “бук”) — как в моей, так и в вашей речи. Можно ли согласиться, что различие между тем, на что указывает “вяз” и тем, на что указывает “бук”, вызвано различием в наших понятиях? Со стыдом признаю, что

мое понятие вяза точно таково, как мое понятие бука. (Кстати, это показывает, что референция определена социально, а не индивидуально: и вы, и я считаемся с мнением экспертов, умеющих отличить бук от вяза.) Если же кто-то отважно возьмется утверждать, что различие между референцией “вяза” и референцией “бука” в моей речи объясняется различием моих психологических состояний, то пусть представит себе Землю-Двойник, где эти слова поменялись местами. Земля-Двойник очень похожа на Землю; по сути, за исключением того факта, что “вяз” и “бук” заменены друг на друга, читатель может предположить, что Земля-Двойник похожа на Землю в точности. Предположим, что на Земле-Двойнике у меня есть личный двойник, идентичный со мной до молекулы (в том смысле, в котором могут быть .-идентичными” два галстука). Если вы — дуалист, предположите, что мой двойник мыслит те же вербализованные мысли, что и я, обладает теми же чувственными данными, теми же диспозициями и т. д. Абсурдно было бы думать, что его психологическое состояние отлично от моего, однако же его слово “вяз” репрезентирует буки, а мое слово “бук” репрезентирует вязы. (Точно так же, если “вода” на Земле-Двойнике — это какая-то другая жидкость, скажем, XYZ, а не H₂O, то “вода” будет репрезентировать разные жидкости на Земле-Двойнике и на Земле и т. д.) Вопреки доктрине, сопутствующей нам с XVII века, значения не находятся в голове.

Мы видели, что обладание понятием — это не вопрос обладания образами (например, деревьев — или даже образами, “визуальными” или “акустическими”, предложений или целых дискурсов), так как можно обладать какой угодно системой образов и не иметь способности употреблять предложения ситуационно подходящими способами (учитывая в определении “ситуационной уместности” как лингвистические факторы — о чем говорилось ранее, — так и внелингвистические). Человек может обладать всеми образами, какими угодно, но тем не менее испытывать сильное затруднение, когда его попросят “указать на дерево” даже в присутствии множества деревьев. Он может даже обладать образом того, что от него ожидается, и все же не знать, чего от него ожидают. Ведь если образ не сопровождается способностью действовать определенным образом, то он — всего лишь картинка, а действовать в соответствии с картинкой — это уже способность, которой можно владеть или не владеть. (Человек может изобразить себя указывающим на дерево лишь ради созерцания чего-то логически возможного; указывающим на дерево после того, как некто произнес (для него бессмысленную) последовательность звуков “укажите на дерево”). Он все еще не будет знать, что от него ожидалось указание на дерево, и все еще не будет понимать выражение “укажите на дерево”.

Я рассмотрел способность употреблять определенные выражения как критерий обладания полноценным понятием, но этот критерий легко может быть ослаблен. Мы можем допустить, например, такую символику, которая состоит из элементов, не являющихся словами естественного языка, и такие ментальные феномены как образы и другие типы внутренних событий. Здесь существенно то, что они должны обладать такой же сложностью, такой же способностью сочетаться друг с другом и т. д., что и предложения естественного языка. Хотя для некоторого математика определенное представление — например, синяя вспышка — может служить внутренним выражением целого доказательства теоремы простых чисел, но вряд ли нам придет в голову утверждать (и было бы ложным утверждать) это, если бы этот математик не мог развернуть свою “синюю вспышку” в отдельные шаги и логические связи доказательства. Но, вне зависимости от того, какой вид внутренних феноменов мы допускаем в качестве возможных выражений мысли, аргументы, в точности подобные приведенному выше, показывают, что понимание конституируется вовсе не феноменами самими по себе, но способностью мыслителя применять эти феномены, производить правильные феномены в правильных обстоятельствах.

Вышеизложенное представляет собой весьма сокращенный вариант аргумента, данного Витгенштейном в “Философских исследованиях”. Если он справедлив, то попытка понять мышление с помощью того, что называют “феноменологическим” исследованием, исходит из неверных оснований, поскольку феноменологи не видят, что описывают внутреннее выражение мысли, тогда как понимание этого выражения — чье-либо понимание собственных мыслей — представляет собой не явление или событие, но способность. Наш пример с человеком, якобы мыслящим по-японски (и обманувшим японского телепата), уже показывает тщетность феноменологического подхода к проблеме понимания. Даже если есть какое-то доступное интроспекции свойство, присутствующее тогда и только тогда, когда некто действительно понимает нечто (хотя интроспективно это кажется ложным), все же это свойство лишь коррелирует с пониманием и возможно, что человек, обманываемый японского телепата также обладает этим качеством, но в то же время не понимает ни слова по-японски.

С другой стороны, рассмотрим вполне возможного человека без какого бы то ни было “внутреннего монолога”. Он говорит на прекрасном английском и на просьбу высказать свое мнение по тому или иному вопросу даст пространственный ответ. Но он никогда не мыслит (в словах, образах и т. д.), когда не высказывается вслух; ничто не “проходит через его голову” — за исключением, конечно, того, что он слышит свой соб-

ственный голос и располагает обычными чувственными впечатлениями о своем окружении, а также общим “чувством понимания”. (Может быть, у него есть привычка разговаривать с самим собой.) Когда он печатает письмо, идет в магазин и т. д., у него нет никакого внутреннего “потока мысли”; однако его поступки разумны и целесообразны, и если кто-то подойдет и спросит его “Что вы делаете?”, он даст вполне связный ответ.

Такого человека вполне можно себе представить. Мы не колеблясь, станем утверждать, что он обладает сознанием, что он, — если он неоднократно выражал стойкое отвращение к рок-н-роллу, — не любит рок-н-ролл и т. д., лишь на том основании, что у него нет осознанных мыслей, за исключением тех случаев, когда он высказывается вслух.

Из этого следует, что (а) никакое множество ментальных событий — образов или более “абстрактных” ментальных происшествий и свойств — не конституирует понимание; и что (б) никакое множество ментальных событий не является необходимым для понимания. В частности, понятия не могут быть идентичны ментальным объектам любого вида. Мы видели, что — в предположении, что под ментальным объектом мы подразумеваем нечто доступное интроспекции — чем бы ни являлся такой объект, его может не быть у человека, который понимает соответствующее слово (и тем самым обладает полноценным понятием), и он может наличествовать у человека, который вообще не обладает этим понятием.

Возвращаясь к нашей критике магических теорий референции (тема, также связанная с Витгенштейном), мы видим, что, с одной стороны, те “ментальные объекты”, которые мы можем обнаружить интроспективно — слова, образы, чувства и т. д. — не указывают в силу своих свойств на что бы то ни было в большей степени, нежели это делает муравьиная картинка (и это происходит по тем же причинам). В то же время попытки постулировать особые ментальные объекты, “понятия”, которые обладают необходимой связью со своими референтами и которые может обнаружить лишь подготовленный феноменолог, представляют собой логическую ошибку, поскольку понятия являются (по крайней мере частично) способностями или умениями, а не явлениями. Доктрина, согласно которой ментальные репрезентации могут с необходимостью указывать на внешние вещи, плоха не только с естественнонаучной точки зрения — она плоха также феноменологически и ведет к путанице в понятиях.

8 Послесловие. М. В. Лебедев. “he matrix never has you”. Хилари Патнэм и традиционная философия

Каковы отношения между нашим восприятием и действительностью? Каковы отношения между сознанием и телом? Каким образом мы придаем значение выражениям нашего языка? На основании чего мы заключаем о мире? Наконец, в чем состоит истинность, и является ли истина свойством или отношением?

Эти традиционные вопросы всегда были и остаются темой философских дебатов. Сразу же бросающаяся в глаза — и в самом деле поразительная — особенность книги Хилари Патнэма “Разум, истина и история” состоит в том, что он показывает единство этих вопросов. Усмотрение такого единства носит отчетливо прагматистский характер, поскольку в качестве связующего аспекта выбрана тема рациональности, понимаемой деятельностно, как снятие противоречия между спекулятивными крайностями. Подобный прагматизм характерен для Патнэма на всех этапах его карьеры, и даже регулярно декларируемая им смена позиций означает не перемену направления на противоположное, но скорее радикализацию *Aufhebung* на основе соотнесения с практикой. Это относится и к двум блокам исследований, концептуально предваряющим знаменитый аргумент о “мозгах в бочке” — по функционализму и по теории референции, — и к исследованиям по проблеме реализма, наследующим ему.

Философская позиция, которую Патнэм сформулировал и назвал функционализмом в 1960-ых годах, стала очень влиятельной в философии сознания. Функционализм предлагает интерпретацию познавательного процесса как процесса обработки информации — формальной манипуляции репрезентациями или символами согласно некоторым правилам. Согласно этой теории, сознание аналогично вычислительной машине, и поэтому для изучения познания релевантны организация и структура, а не материальный субстрат сознания. Аргумент Патнэма здесь основан на интуиции о том, что одно и то же ментальное состояние может быть реализовано на различных материальных носителях. Его позднейшие ревизии функционализма связаны с расширением требования о множественности реализации ментального — в духе свойственного прагматистам плюрализма. Это расширение основано на инверсии исходной интуиции: если различные состояния вычислительной машины или машины Тьюринга могут реализовывать одно и то же ментальное состояние, то сознание не может быть идентифицировано ни с какой конкретной

вычислительной машиной. Поддержка Патнэмом сильной формы множественной реализуемости ментального приводит его к пересмотру самого статуса психофизической проблемы как ключевой для философии сознания.

Это определяет тот способ, которым Патнэм связывает свою философию сознания со своей философией языка. В этом отношении он следует другому крупному прагматисту XX века — Куайну, давшему мощное орудие для встраивания проблемы языкового значения в очень широкий эпистемологический контекст, рассмотрев в качестве значения референцию — указание на предмет. Патнэм становится одним из родоначальников так называемых “расширенных” теорий референции, согласно которым референциальная связь между родовым термином и его референтом может быть установлена даже в том случае, если говорящий не информирован или плохо информирован о природе референта. Это возможно благодаря тому, что выбор индивида при установлении референции к нему зависит от “каузальной истории” использования данного имени, а она восходит к некоторому акту первоначального именованя. В своей типичной форме этот первоначальный акт устанавливает референцию посредством указательного местоимения или прямого указания на объект реального мира. Поэтому интенционал собственного имени не фиксирован на каком-либо одном представлении в головах говорящих, общем для всех, кто использует это имя как единицу данного языка. Вообще можно сказать, что для каждого термина, введение которого в язык хотя бы частично было основано на остенсивном акте фиксации референции, интенционал этого термина хотя бы частично устанавливается на основе свойств объекта реального мира, первоначально вовлеченных в акт введения термина, хотя выделение объекта происходит не через указание совокупности индивидуальных свойств объекта, а через описание тех обстоятельств, в которых однозначная связь термина и объекта не подвергается сомнению. Поэтому незнание говорящим экстенционала термина или “каузальной истории” термина ведет к незнанию интенционала, но не меняет референциальной ситуации. Согласно Патнэму, границы между экстенционалами терминов устанавливаются более или менее произвольно: в этом смысле содержание существующего определения любого данного термина может быть конвенционально.

Возражения Патнэма против традиционных теорий значения основываются на фундаментальной несовместимости, с его точки зрения, двух важных для них тезисов. Один из них заключается в том, что экстенционал определяет значения общего термина, т. е. то, относительно чего термин истинен. В соответствии с этим тезисом, два термина с одинаковым значением имеют одинаковый экстенционал. Согласно другому осново-

полагающему для таких теорий тезису, существуют определенные психологические состояния людей, соответствующие определенному уровню знания значения.

Таким образом, подход Патнэма противостоит фрегеанской теории значения (и подобным ей семантикам, содержащим менталистские компоненты) в следующем отношении: согласно этому подходу, экстенционал не определен ментальным состоянием говорящего. Фрегеанская теория значения предполагает, что

- (1) знание значения состоит из нахождения в соответствующем ментальном состоянии, и что
- (2) значение определяет экстенционал.

Иными словами, для менталистов знание значения определяет экстенционал термина; ментальное состояние определяет интенционал, а интенционал определяет экстенционал. Патнэм полагает, что такая картина существенно ущербна. По его мнению, само существование реальных сущностей и сама природа референции устанавливают, что значения находятся “не в голове”. Между предметами-референтами и их ментальными (или любыми иными) репрезентациями нет необходимых связей.

По сути, это очередная переформулировка центрального аргумента аналитической философии сознания и языка середины XX века, приписываемого позднему Витгенштейну. “Если он справедлив, — пишет Патнэм, — то попытка понять мышление с помощью того, что называют “феноменологическим” исследованием, исходит из неверных оснований, поскольку феноменологи не видят, что описывают внутреннее выражение мысли, тогда как понимание этого выражения — чье-либо понимание собственных мыслей — представляет собой не явление или событие, но способность”. (Аналогичным образом известный историк феноменологии Г. Шпигельберг задает вопрос: как феноменология все еще возможна, когда она больше не может отрешиться от “ангажированности” в феномены?) Однако классический аналитический аргумент переформулирован с явным постмодернистским акцентом (“значения не находятся в голове” — “значения ускользают”). Это обстоятельство — как и все, связанное с постмодернизмом — в равной степени способствовало как его успеху у философствующих гуманитариев, так и ограниченному доверию со стороны профессиональных философов; последние явно предпочли более глубокий вариант выражения сходных интуитивных представлений, который дал Кит Доннелан, и более прозрачный, данный Солом Кришке.

Таков наиболее общий замысел семантики Патнэма (я опускаю здесь эволюцию деталей), непосредственно связанный с аргументом о “мозгах в бочке” — вызовом, брошенным Патнэмом метафизическому реализму.

Метафизический реализм, не допускающий никаких эпистемических

ограничений на наши понятия действительности или истины, сталкивается с проблемой объективности или трансцендентности этих понятий — с проблемой возможного отличия того, чем действительность является, от того, чем мы ее считаем, как мы ее воспринимаем или представляем себе. Эта скептическая возможность неоднократно в истории философии была драматизирована различными мысленными экспериментами, классический пример которых — Декартов *malin genie*⁴. Патнэм предлагает вариант такого эксперимента, сегодня ставший широко известным по лженаучно-популярному фильму “Матрица”: допустим, что все мы — мозги в бочке, присоединенные к мощному компьютеру, который моделирует наш опыт о мире. Патнэм проводит косвенную атаку на метафизический реализм, выдвигая требование, что это невозможно, мы не можем быть мозгами в бочке. Это требование носит отчетливо антискептический характер: впрямую оно направлено против скептического аргумента об отсутствии (или невозможности наличия) надежных оснований нашего знания о мире — например, аргумента Чжуан-цзы: философ спит и ему снится, что он бабочка, которая спит и ей снится, что она философ, который спит и т. д. Патнэм пытается разрубить этот гордиев узел дамокловым мечом, заявляя, что невозможность для нас быть мозгами в бочке априорна, а скептический аргумент тем самым некорректен.

Аргумент Патнэма связан с его теорией референции. Итак, если мы являемся мозгами в бочке, то мы не можем действительно указывать на внешние объекты, поскольку мы не связаны каузально с реальными вещами в мире; в лучшем случае мы каузально связаны с программированием компьютера. Из этого следует, что, когда мозг в бочке говорит “Это — стол”, он делает ложное утверждение постольку, поскольку он намеревается указать на реальные столы, с которыми он каузально связан. То же самое относится ко всем его референциальным утверждениям. Теперь рассмотрим утверждение “Я — мозг в бочке”, сделанное мозгом в бочке. Подразумевается, что это референциальное утверждение о реальных мозгах и реальных бочках. Однако нашему мозгу в бочке никогда не были даны никакой реальный мозг и никакая реальная бочка; все, чем он располагает — это его машинно-генерированные галлюцинации этих вещей. Ему дано только электрическое возбуждение, но не дано никакого опыта реального мира, в котором эти референты действительны. Но если такие референциальные утверждения, сделанные мозгом в бочке, ложны, из этого следует, что и утверждение “Я — мозг в бочке” также является ложным. Патнэм называет такой аргумент “трансцендентальным” в кантианском смысле, поскольку если мы видим, каковы

⁴Злой гении (фр.). —Прим. авт.

условия референции, то отсюда априорно будет следовать, что мы не можем быть мозгами в бочке. И этого достаточно, чтобы опровергнуть метафизическое реалистическое утверждение, что не может быть никакого априорного доказательства против скептицизма.

Философское сообщество восприняло этот аргумент весьма неоднозначно; его критика может исходить, в частности, из того, что автореферентные утверждения — это в любом случае достаточно специальный класс утверждений, и их характеристики вовсе не обязательно должны распространяться на все референциальные утверждения языка. Кроме того, даже если не рассматривать этот аргумент как некоторую семантическую уловку, его позитивная суть все равно остается не до конца проясненной. Отвечая на эту критику, Патнэм отметил, что реальное значение аргумента состоит в том, чтобы проиллюстрировать “отчужденный” характер метафизического реалистического способа рассмотрения мира. “Отчуждение” здесь состоит в следующем. Такие глобальные возможности, как оказаться мозгами в бочке или жертвами обмана коварным демоном, традиционно рассматривались философами как ставящие глубокие эпистемологические и метафизические проблемы. Но допущение подобных возможностей опирается на идею о том, что мы можем выйти за пределы самих себя, то есть видеть себя и свое состояние с некоторой объективной “точки зрения Бога”; что мы способны, по выражению Томаса Нагеля, на “взгляд ниоткуда”. Но, по мнению Патнэма, мы — не Бог (этот популистский аргумент явно апеллирует к тому, что большинство из нас с ним согласится, хотя и по разным причинам), и любая точка зрения — это взгляд откуда-нибудь. Поэтому даже если аргумент о мозгах в бочке не представляет собой полное доказательство чего-либо, он тем не менее ясно иллюстрирует возможную смену угла зрения в обсуждении проблемы реализма.

Патнэм характеризует метафизический реализм как позицию, согласно которой существует одно и только одно истинное и полное описание мира, где мир — это полная и фиксированная совокупность независимых от сознания объектов, а истина — это корреспонденция (соответствие) между утверждениями и вещами (с “точки зрения Бога”). Что же создает связь между утверждениями и вещами? Согласно многим современным реалистам, причинные связи — это такие связи, когда мои мысли о столах вызваны реальными столами и т. п., то есть основанные на адекватной корреспонденции. Возражение Патнэма заключается в том, что это далеко не всегда так — например, мои мысли об электронах вызваны учебниками, но они не являются мыслями об учебниках. Реалист ответит, что мы должны иметь правильный вид каузальной цепи. И вот здесь Патнэм с помощью своей теории референции ловит реалистов на цирку-

лярности: мы не будем знать, что данный вид каузальной цепи правилен до тех пор, пока мы не знаем, на что указывают наши слова – а именно это мы и намереваемся установить.

Альтернативная точка зрения, ассоциирующаяся в разных аспектах с антиреализмом, верификационизмом, когерентизмом, прагматизмом, исходит из того, что вопрос “Какие объекты имеются в наличии?” имеет смысл только в пределах соответствующей концептуальной схемы. Здесь, вообще говоря, еще нет ничего неожиданного по сравнению с непосредственными предшественниками Патнэма. Однако из требования, согласно которому наша точка зрения всегда интегрирована в некоторую концептуальную схему, Патнэм делает вывод о том, что истина — это “идеализированная рациональная приемлемость”. Основную авторскую нагрузку здесь получает слово идеализированная. истинным является не просто то, что было бы рационально принять, исходя из нашей нынешней концептуальной схемы, но то, что мы примем “при эпистемически идеальных условиях” (например, 3000 лет назад теория о том, что Земля плоская, была рационально приемлема, но не истинна). Такая позиция в целом остается в рамках характерного для прагматизма фаллибилизма – восходящего к Пирсу взгляда, согласно которому не существует окончательных верификаторов; нет никаких утверждений, которые не могли бы быть пересмотрены в свете нового опыта. Однако Патнэм предлагает некоторые точки опоры внутри этой бесконечной цепи обоснования нашего знания, обозначает некоторый надежный горизонт – и это горизонт рациональной приемлемости.

Такая трактовка призвана прояснить само понятие рациональности. Важнейшее требование Патнэма состоит в том, что обе наиболее влиятельных концепции рациональности в философии двадцатого столетия неудовлетворительны; фактически, обе опровергают сами себя.

Первая из них – это концепция логического позитивизма, согласно которой утверждения имеют значение лишь в том случае, если они поддаются проверке, и существует установленная процедура для подтверждения (проверки) всех значимых предложений. Эта процедура (множество процедур) – чистый логический анализ для утверждений, истинных в силу содержащихся в них символов, плюс эмпирическая проверка в терминах (референциального) содержания для утверждений, заключающих о мире вокруг нас. Рациональность — это не что иное, как соответствие этой установленной процедуре; если мы утверждаем вещи, которые не являются ни аналитическими утверждениями, истинными в силу правил для символов, ни эмпирическими утверждениями, истинными в силу своего референциального содержания, то мы иррациональны. Самоопровержение этой позиции, по мысли Патнэма, состоит в том, что защита

самой процедуры не может считаться рациональной: процедура не является истинной в силу содержащихся в ней символов, и при этом она и не истинна в силу некоторого референциального содержания.

Реакцией против позитивизма Патнэм считает релятивизм (куда он относит как Куна и Фейерабенда, так и Фуко и других континентальных философов). Согласно этой позиции, нет никакой установленной формулы или установленной процедуры; что считать “рациональным”, может быть установлено каждой культурой (или субкультурой) различным способом. Релятивисты исходят из того, что культурные факторы имеют огромное значение даже в науке и что научные теории зачастую принимаются или отклоняются по причинам политики и финансирования, а не только по причинам чистой проверки в духе позитивизма. Поэтому заключение релятивизма состоит в том, что различные культуры имеют различные стандарты или “парадигмы” рациональности; мы можем оценивать то или иное требование только относительно данной культуры (так, в некоторых культурах рациональны некоторые требования об электронах, а в других культурах рациональны некоторые требования о привидениях, и так далее), и мы не всегда можем перевести наши требования из одной культурной схемы в другую — культурные схемы “несоизмеримы”. Самоопровержение релятивизма (который он называет “ментальным самоубийством”) Патнэм видит в следующем: формула о том, что является истинным во всех культурах и субкультурах, сама по себе является видом универсального и кросскультурного утверждения, которые релятивисты считают нелегитимными.

Итак, Патнэм старается уйти от эпистемологических крайностей как фундаментализма, так и релятивизма. Однако в глазах фундаменталиста он, несомненно, оказывается релятивистом, так как отрицает наличие и самое возможность предельного основания знания — скорее он выдвигает требование достаточного основания. Для релятивиста же на этом уровне будет не очень внятен предъявляемый упрек: релятивизм может вообще не апеллировать к допущению о наличии объемлющей, единой эпистемологии. Патнэм же не может — в отличие, например, от того же Фуко — не мыслить в присутствии возможности такого допущения. Мы могли бы сказать, что здесь сама попытка снять глобальное эпистемологическое противоречие, выйти из очерченной им плоскости оказывается важнее результата этой попытки с ее не вполне убедительной концепцией “идеализации”. Патнэма, очевидно, не устраивает такое положение, и он предпринимает еще более решительный шаг.

До сих пор в анализе рациональности, как и в анализе реализма, Патнэм применяет даже не постаналитическую, а скорее ортодоксальную аналитическую технику, апеллируя практически исключительно к

системам аргументации. В самом деле, до сих пор его протест против абсолютистской концепции трансцендентной истинности или окончательного, предельного основания для наших убеждений до некоторой степени напоминает тезис Куайна об онтологической относительности или критику Дональдом Дэвидсоном различения концептуальной схемы и ее содержания. Но, в отличие от Куайна или Дэвидсона, Патнэм расширяет этот подход на моральную и политическую сферу, доказывая, что крах абсолютистской перспективы в эпистемологии и метафизике делает возможным подтвердить истинностную релевантность морального, политического и других видов управляемого нормами дискурса.

Ход Патнэма состоит здесь в том, чтобы привести доводы против принятой дихотомии между фактическими утверждениями и ценностными суждениями и усмотреть такую позицию, откуда это различие снимается — сохранив, таким образом, единство корпуса человеческого знания. Фактически это возврат на новом уровне к той хорошо известной в истории философии традиции — от Платона до Гегеля, — которая проводила отождествление онтологической и аксиологической проблематики и изначально наделяла бытие ценностным измерением.

Разрыв с этой традицией происходит относительно недавно, когда неокантианцы фрайбургской школы разрабатывают учение о ценностно-нормативном компоненте чистого сознания, выявляемом трансцендентальной рефлексией, и уже у Хайдеггера онтологический анализ оборачивается отвержением аксиологической проблематики. Аргумент здесь примерно таков: если допустить, что бытие (и производные от него образования) само по себе ценностно нейтрально, то тогда требует решения вопрос, каким образом возникает видение вещей с точки зрения их аксиологической значимости, которое пронизывает культуру и которое невозможно отрицать. Отсюда берет свое начало линия поиска истоков ценностного сознания в самом человеке и его культуросозидающей деятельности, чрезвычайно характерная для европейской философии XX века.

Позиция Патнэма, насколько я понимаю, направлена не на отрицание, но на переосмысление этой тенденции. Его возврат к традиции на новом уровне проводится по иному основанию, нежели в светлые времена схоластики, поскольку исходит из критики реализма, и исходным для него оказывается тезис о том, что “каждый факт нагружен ценностью, и каждая из наших ценностей нагружает некоторый факт”. Факт (или истина) и рациональность оказываются взаимозависимыми понятиями.

Патнэм стремится доказать, что даже обычные фактические утверждения (например, “вода состоит из водорода и кислорода”) вовлекают некоторые ценностные обязательства, и что оценочные суждения имеют

некоторых из особенностей, обычно связываемых с фактическими утверждениями (в частности, оценочные суждения могут быть объективны). Если мы отвергаем метафизический реализм, то неверна картина, согласно которой мы “считываем” факты с мира и измеряем, насколько хорошо наши утверждения соответствуют этим независимым от сознания фактам. Природа процесса познания совершенно иная: мы создаем систему понятий и стандартов, которые будут наилучшим образом соответствовать миру по нашим когнитивным способностям, и оцениваем, что является истинным в соответствии с этими понятиями и стандартами. А если понятия истины или объективности не определены в терминах некоторого отношения к независимым от сознания объектам, но скорее связаны с нашими методами оценки и выбора, то нет препятствий к тому, чтобы рассматривать моральные и политические требования как самостоятельно объективные и истинные.

Для аналитических философов предшествующего периода характерным считалось отношение к этике, политической и социальной философии и т. п. как к расплывчатым и неясным областям, где изобилуют догадки и спекуляции, но нет четких аргументов, с логической необходимостью ведущих к ясно артикулированным результатам, обладающим истинностным значением. Преодолевая (хотя, разумеется, не первым) эту ограниченность, Патнэм не отрицает различия между науками и такими предметами как этика или политика, но указывает, что абсолютное различие между ними по критериям требований истины и объективности не является надежным; исторически больше не может рассматриваться как надежное. Сами точные науки также нормативны: заключения о том, какие теории принимать, или даже о том, какие вопросы ставить, делаются учеными под влиянием ценностей и систем убеждений, и принципиальным образом не отличается от того, что это происходит в моральном и политическом дискурсе. Впоследствии эта позиция привела Патнэма к теории “идеальной коммуникации” Юргена Хабермаса, акцентирующей общественный характер ценностей, как к способу показать, каким образом моральная объективность могла бы быть достигнута в современном обществе, которое является, по общему признанию, субъективистским по характеру; как к способу согласовать между собой философию языка, эпистемологию и социальную теорию — задача, вряд ли решенная Хабермасом или Патнэмом, но, безусловно, представляющая собой передний край современной философии, и их заслуги в ее постановке трудно переоценить.

Итак, “сверхзадача” Патнэма здесь состояла в том, чтобы разрушить ту власть, которую естествознание обрело над философской мыслью в XX столетии (“появление в культуре философской тенденции, которая

была загниотизирована успехом науки до такой степени, что не могла представить возможности знания и разума вне того, что нам нравится называть наукой — это такой феномен, которого следовало ожидать, учитывая чрезвычайно высокий престиж науки в культуре в целом и учитывая снижающийся престиж религии, абсолютистской этики и трансцендентальной метафизики”). Хотя Патнэм не враждебен науке, он отклоняет отождествление рационального мышления с научным мышлением и отклоняет идею о том, что наука дает единственное истинное описание действительности. Подчеркивая уникальную, ни к чему не редуцируемую ценность самостоятельного философского усмотрения, он, несомненно, остается традиционным философом в лучшем смысле обоих этих слов.