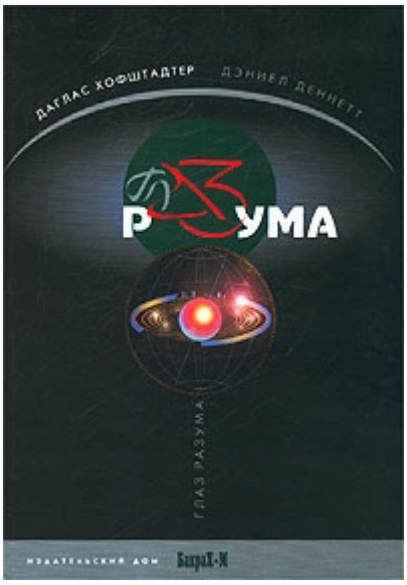
Даглас Хофштадтер, Дэниел Деннетт. Глаз разума

Даглас Хофштадтер, возможно, знаком вам по бестселлеру Гедель, Эшер, Бах. Эта бесконечная гирлянда (книга переведена на 17 языков). Глаз разума – сборник работ таких выдающихся мыслителей XX века, как Хорхе Луис Борхес, Ричард Докинз, Джон Сирл, Роберт Нозик, Станислав Лем, Алан Тьюринг и многих других. Каждая оригинальная статья сопровождается размышлениями составителей сборника – Хофштадтера и Деннетта. Читателя приглашают в удивительный и парадоксальный мир человеческого духа и «думающих» машин. Здесь представлены различные взгляды на природу человеческого мышления и природу искусственного разума, исследуются, сопоставляются, сталкиваются такие понятия, как «сознание», «душа», «личность»…

Даглас Хофштадтер, Дэниел Деннетт. Глаз разума. – Самара: Бахрах-М, 2003. – 432 с. Первое издание на английском языке вышло в 1981 г.



Купить книгу в Ozon или Лабиринте

Часть I. САМООЩУЩЕНИЕ
Глава 1. ХОРХЕ ЛУИС БОРХЕС. Борхес и я

Размышления. Великий аргентинский писатель Хорхе Луис Борхес имеет заслуженную международную репутацию, которая создает любопытный эффект. Борхесу кажется, что он — это два человека, общественная персона и частное лицо. Его слава увеличивает этот эффект, но в той или иной мере подобное чувство знакомо каждому из нас. Вы читаете ваше имя в списке, видите собственную беспристрастную фотографию или случайно слышите, как знакомые говорят о ком-то… и внезапно вы понимаете, что это вы.

Хотя Роберт Бёрнс, может быть, и прав, говоря, что уметь видеть себя глазами других — это дар, мы не можем и не должны желать находиться в подобном положении все время. Например, Пит стоит в очереди в кассу в супермаркете и вдруг видит притаившуюся под потолком телекамеру — оружие против магазинных воришек. Глядя на толпу народа на мониторе, он замечает, что у мужчины в плаще, стоящем в левом углу монитора, обхватив двумя руками пакет с покупками, в этот момент вытаскивают деньги из кармана. Пит в удивлении подносит руку ко рту и видит, что жертва на мониторе повторяет точно такой же жест. Внезапно до него доходит: это же меня обкрадывают! Этот драматический сдвиг в восприятии является для него открытием; Пит узнал нечто важное, что еще минуту назад было для него неизвестным. Если бы он был неспособен на подобные мысли, которые сейчас гальванизируют его и заставляют действовать, он не смог бы себя защитить. Разумеется, перед тем, как это открытие произошло, он не был в полном неведении. Он думал о «человеке в плаще» и видел, что его обкрадывают — и, поскольку человек в плаще — он сам, Пит думал о себе. Однако он не думал о себе как о себе; иными словами, он думал о себе «неправильно».

Глава 2. Д. Е. ХАРДИНГ. Как я был без головы

Размышления. Способность выделять классы помимо тех, что запрограммированы в нас от рождения — это довольно выдающаяся способность интеллекта. К примеру, пчелы отлично усвоили класс «цветы», но мало вероятно, что они смогли бы сформировать понятия «камин» или «человеческое существо». Собаки и кошки уже могут сформировать некоторые классы, такие как «плошка для еды», «дверь», «игрушка» и т.п. Люди, однако, намного опередили других живых существ в умении создавать новые классы один за другим.

Ребенком я сформулировал абстрактное понятие «человеческое существо», наблюдая внешние по отношению ко мне объекты, имеющие между собой нечто общее — внешний вид, поведение и так далее. На более поздних ступенях когнитивного развития ребенок делает открытие, что этот абстрактный класс может «заглотать» его самого. Это, наверное, потрясающее событие, хотя, скорее всего, мало кто из нас его помнит.

Способность отождествлять себя с другими, по-видимому, принадлежит лишь высшим животным. Видя вашу голову, я прихожу к заключению, что и у меня должна быть голова, хотя я ее и не вижу. Однако этот шаг «из себя» — гигантский и в какой-то мере самоотрицающий шаг. Он противоречит большинству моих непосредственных знаний о себе самом. Это напоминает два разных значения глагола «видеть» у Хардинга: в приложении ко мне это совсем не то же самое, что в приложении к вам. Все мы постоянно колеблемся между субъективным и объективным взглядом на мир, и это

Глава 3. ХАРОЛЬД ДЖ. МОРОВИЦ. Новое открытие разума

В последнее время изучение жизни на всех уровнях, от социального до молекулярного, опиралось для объяснения явлений на понятие редукционизма. Этот подход к знаниям пытается понять явления одного уровня научных феноменов в терминах другого, низшего, предположительно более фундаментального уровня. В химии сложные реакции объясняются с точки зрения поведения молекул. Таким же образом физиологи изучают деятельность живой клетки в терминах процессов, происходящих на уровне органелл и других субклеточных элементов. В геологии происхождение и свойства минералов описываются в терминах составляющих их кристаллов. Основным во всех этих случаях является поиск объяснения на низших уровнях.

Примером редукционизма на психологическом уровне является бестселлер Карла Сагана «Драконы Эдема» (см. также Карл Эдвард Саган. Мир, полный демонов. Наука — как свеча во тьме). Он пишет: «Моя основная посылка касательно мозга заключается в том, что его работа — которую мы иногда называем «разумом» — это следствие его анатомии и физиологии, и ничего более.» Дальнейшим подтверждением этой мысли является отсутствие в списке терминов Сагана таких слов как разум, сознание, восприятие, осознание и мысль. Вместо них в книге используются синапс, лоботомия, белки и электроды.

До девятнадцатого века дуализм «тело-разум», центральное понятие в философии Декарта, помещал человеческий разум вне сферы биологии. Эволюционисты с их вниманием к нашей «обезьянности» сделали нас объектом биологических исследований при помощи методов, применявшихся к человекообразным обезьянам и, по аналогии, к другим животным. Павловская школа развила эту тенденцию, и она легла в основу многих бихевиористских теорий.

Сложная система может быть описана только в терминах вероятности того или иного результата эксперимента. Чтобы узнать, каков именно результат данного эксперимента, необходимо произвести измерения. Именно эти измерения и являются физическим событием, в отличие от вероятности, являющейся математической абстракцией. Единственное простое и последовательное описание измерения включает наблюдателя, осознающего результат. Таким образом физическое событие и человеческий разум становятся неразделимы. Эта связь заставила физиков рассматривать сознание как существенную часть структуры физики.

Информатика в ее приложении к термодинамике утверждает, что основное понятие этой науки, энтропия, не что иное как мера незнания наблюдателем деталей атомной структуры системы. Измеряя давление, объем и температуру объекта, мы не знаем многого о точных позициях и скоростях атомов и молекул, составляющих этот объект. Числовое значение количества недостающей информации пропорционально энтропии (см. Введение в теорию информации).

Мы можем свести воедино перспективы трех крупных областей науки: психологии, биологии и физики. Во-первых, человеческий разум, включая сознание и мысли о самом себе, может быть объяснен в терминах деятельности центральной нервной системы, которая, в свою очередь, может быть сведена к уровню биологической структуры и функций данной физиологической системы. Во-вторых, биологические явления могут быть полностью поняты в терминах атомной физики, то есть в терминах действия и взаимодействия составляющих систему атомов углерода, азота, кислорода и т.д. И наконец, атомная физика, наиболее полно понимаемая в терминах квантовой механики, должна содержать разум в качестве основного компонента системы. Таким образом, мы, шаг за шагом, описали эпистемологический круг — от разума назад к разуму.

Если мы не согласны с этой эпистемологической кругообразностью, нам остаются две противостоящих области: физика, утверждающая, что она полна, поскольку описывает всю природу, и психология, считающая себя всеобъемлющей, поскольку она имеет дело с разумом, единственным источником нашего знания о мире. Однако оба эти взгляда не свободны от проблем, так что нам, может быть, не мешает вернуться к кругу и взглянуть на него с большей симпатией. Возможно, что этот круг представляет из себя наилучший подход к теоретической психологии.

Кодирование информации в генетических молекулах внесло возможность серьезных нарушений равновесия в законах, управляющих вселенной. Например, перед приходом генетической жизни колебания температуры и шума были уравновешены, из чего следовали точные законы планетарного развития. Однако после этого единственное молекулярное событие на уровне термального шума могло привести к макроскопическим последствиям. Если этим событием оказывалась мутация в самовоспроизводящейся системе, оно могло изменить весь ход эволюции. Единичное молекулярное событие может убить кита, вызвав у него рак, или разрушить экосистему, произведя на свет сильнейший вирус, атакующий основные виды системы. Появление жизни не отменяет законы физики, но добавляет к ним новую особенность: глобальные последствия молекулярных событий. Это изменение в правилах делает историю эволюции неопределенной и, таким образом, представляет из себя яркий случай дискретности.

Некоторые современные биологи считают, что появление разума в процессе эволюции приматов — еще один пример подобной дискретности, изменяющей правила. Так же, как прежде, новая ситуация не меняет биологических законов, но требует новых подходов к проблеме. Эволюционный биолог Лоренс Б. Слободкин определил новую черту системы как интроспективное представление о себе. Эта особенность, утверждает он, изменяет ответ на эволюционные проблемы и делает невозможным объяснение исторических событий как прямых следствий законов биологического развития. Слободкин предполагает, что правила изменились и что человека нельзя понять согласно тем же законам, какие приложимы к остальным млекопитающим, чей мозг имеет похожую физиологию (аналогичную позицию высказывает Вилейанур Рамачандран в книге Мозг рассказывает. Что делает нас людьми).

Эта возникающая в процессе эволюции черта не может быть «убрана на полку» только лишь для того, чтобы сохранить чистоту редукционистского подхода. Эта дискретность должна быть полностью изучена и оценена, но вначале ее необходимо признать. Приматы сильно отличаются от остальных животных, а человеческие существа сильно отличаются от приматов. Стопроцентный редукционизм не является ответом на загадку разума.

Размышления.

> Действительное кажется плавающим в море возможного, откуда оно было выбрано; нондетерминизм утверждает, что где-то эти возможности существуют и являются частью истины.
> *Виллиам Джеймс*

Идея о том, что загадки квантовой механики и загадки разума едины, весьма привлекательна. В частности, квантовая механика придает привилегированный каузальный статус неким системам, называемым «наблюдателями», не давая при этом точного определения, что такое «наблюдатель». Чтобы объяснить ситуацию, прибегнем к помощи метафоры «квантового водопроводного крана».

Представьте себе водопроводный кран с двумя ручками, для холодной и для горячей воды, каждую из которых вы можете поворачивать постепенно. Из крана вытекает вода — но эта система имеет

странную особенность: вода либо совершенно холодная, либо совсем горячая, без промежуточных градаций. Эти состояния называются «собственными температурными состояниями» воды. Единственный способ определить, в котором из двух состояний находится вода — сунуть руку под струю. Правда, в ортодоксальной квантовой механике дело обстоит посложнее. Именно ваш акт засовывания руки под струю приводит воду в то или иное температурное состояние. Вплоть до этого момента говорится, что вода находится в ситуации наложения собственных температурных состояний. Вплоть до момента измерения система ведет себя так, словно не находится ни в одном из собственных состояний.

Вы можете вообразить множество экспериментов с водой из квантового крана, позволяющих вам узнать ее температуру, не суя руку под струю (мы предполагаем, что видимые признаки, такие, как пар, при этом отсутствуют). Например, вы можете включить стиральную машину, наполнив ее водой из-под крана. Но и в этом случае вы не узнаете, сел ли ваш шерстяной свитер, пока не откроете дверцу машины (измерение, произведенное сознательным наблюдателем). Заварите чай водой из-под крана. Вы не узнаете, холодный или горячий он получился, пока не попробуете его (снова взаимодействие с сознательным наблюдателем). Прикрепите термометр под струей воды. Пока вы не увидите ртутный столбик, вы не можете знать температуру воды. Вы не можете быть более уверены в положении столбика термометра, чем в том, что вода имеет определенную температуру. Критический момент здесь состоит в том, что свитер, чай и термометр, не имея статуса сознательных наблюдателей, ведут себя так же, как и вода, и испытывают наложение собственных состояний — севший-не севший, холодный-горячий, столбик высоко-столбик низко.

Может показаться, что эта проблема напоминает древние философские загадки типа «Шумит ли падающее в лесу дерево, если рядом нет никого, кто услышал бы этот шум»? Интересующимся читателям рекомендую книгу Ричарда Фейнмана Характер физического закона.

Проблема усугубляется, когда вы осознаете, какие последствия имеет эта теория лично для вас, здесь и теперь. Ведь в каждой точке каждой из квантово-механических ветвей вашей жизни (а этих точек были уже биллионы биллионов) вы разделялись на двух или больше вас, и каждый из вас развивался на одной из параллельных, но разъединенных ветвей одной гигантской «универсальной волновой функции».

И все же достаточно спросить себя: «Чувствую ли я, что нахожусь всего в одном мире?» Согласно Хью Эверетту, вы этого не чувствуете — вы одновременно воспринимаете все альтернативы, и лишь тот из вас, который находится на данной ветви, этих альтернатив не воспринимает. Все это шокирует… (подробнее см. Множественность миров Хью Эверетта).

И все же это не столь удивительно. Между воображаемыми мирами у нас в голове и альтернативными мирами, развивающимися параллельно с нашим, есть явная связь. Пресловутый молодой человек, обрывающий лепестки ромашки и бормочущий: «Любит — не любит…», представляет себе по меньшей мере два возможных мира, основанных на двух различных представлениях о его возлюбленной. Или, может быть, вернее было бы сказать, что у него в голове лишь одна модель возлюбленной, и эта модель является мысленным аналогом квантово-механического наложения собственных состояний?

Часть II. В ПОИСКАХ ДУШИ

Глава 4. А. М. ТЬЮРИНГ. Вычислительные машины и разум

(подробнее о Тьюринге см. Эндрю Ходжес. Игра в имитацию и Чарльз Петцольд. Читаем Тьюринга)

Игра-имитация. Я предлагаю рассмотреть вопрос «Могут ли машины думать?». Проблема может быть описана в терминах игры, которую мы будем называть «игра-имитация». В ней принимают участие три человека: мужчина (А), женщина (Б) и экзаменатор (В). Экзаменатор сидит в отдельной комнате. Цель игры заключается в том, что экзаменатор должен определить, кто из остальных двоих — женщина и, кто — мужчина. Теперь мы спрашиваем: а что произойдет, если вместо А в игре примет участие машина? Будет ли экзаменатор ошибаться так же часто, как и в игре с женщиной и мужчиной? Эти вопросы заменяют наш первоначальный вопрос о том, могут ли машины думать.

Сначала я изложу собственное мнение. Я считаю, что лет через пятьдесят (статью Тьюринг написал в 1950 г.) компьютеры смогут играть в игру-имитацию настолько успешно, что обычный экзаменатор

после пяти минут опроса будет иметь не более чем 70% шансов на верную идентификацию. Далее я рассмотрю мнения, противоположные моему.

1. Теологический аргумент. Мышление есть функция бессмертной души человека. Бог дал бессмертную душу всем мужчинам и женщинам, но не животным и не машинам, и посему ни одно животное и ни одна машина не способна думать.
2. Аргумент «голова в песке». «Последствия машинного мышления были бы слишком ужасны. Давайте же надеяться, что машины ни на что подобное не способны.»
3. Математическое возражение. Некоторые результаты математической логики могут быть использованы для доказательства того, что мощь дискретных машин ограничена. Самый известный из этих результатов — теорема Гёделя. Она показывает, что в любой достаточно сложной непротиворечивой логической системе можно сформулировать высказывания, относительно которых нельзя средствами этой системы доказать, истинны они или ложны. Краткий ответ на этот аргумент таков: действительно, возможности каждой отдельной машины ограничены — но никто пока не пытался строго доказать, что подобных ограничений не существует для человеческого интеллекта. Я думаю, некоторые люди умнее, чем определенные машины, но некоторые машины умнее этих людей… и так далее. Мы не можем одержать победу одновременно над всеми машинами.
4. Аргумент от сознания. Этот аргумент прекрасно выражен профессором Джефферсоном в 1949 году: «Пока машина не напишет сонета или концерта, вдохновленного чувствами, а не полученного в результате случайного сочетания символов, мы не сможем согласиться с тем, что машина равна мозгу. Никакой механизм не может почувствовать (а не просто показать при помощи какого-либо несложного ухищрения) удовлетворения от удачи, печали от того, что его контакты перегорели, испытать удовольствие от похвалы, расстроиться из-за своих ошибок, быть очарован сексом, сердиться или впадать в депрессию, когда он не может получить желаемого.» В действительности, это солипсистская[1] точка зрения. Возможно, это наиболее логичная точка зрения, но она делает обмен идеями затруднительным. А расположен считать: «А думает, а Б — нет», в то время, как Б считает: «Б думает, а А — нет». Чтобы избежать вечных споров по этому поводу, принято считать, что каждый человек способен думать.
5. Аргумент леди Лавлейс. Об Аналитической Машине Баббиджа нам лучше всего известно из мемуаров леди Лавлейс. Она пишет: «Аналитическая машина не претендует на создание чего-либо нового. Он может делать лишь то, что мы умеем ей приказать».
6. Аргумент от непрерывности нервной системы. Разумеется, нервная система не является машиной дискретных состояний. Малейшая ошибка в интенсивности нервного импульса в одном из нейронов может вызвать значительную разницу в выходящем импульсе. Имея это в виду, можно утверждать, что невозможно имитировать поведение нервной системы с помощью какой-либо дискретной машины. Безусловно, дискретная машина отличается от непрерывной. Но если мы будем соблюдать условия игры-имитации, экзаменатору не удастся получить от этой разницы никакой выгоды. Ситуация станет понятнее, если мы рассмотрим некоторые непрерывные машины попроще. Для этой цели отлично подойдет дифференциальный анализатор. (Дифференциальный анализатор — это недискретная машина, используемая для некоторых вычислений.) Цифровой компьютер не сможет предсказать, какие ответы на определенные вопросы будет выдавать дифференциальный анализатор, но сам сможет решить задачу правильно. Например, если попросить его определить значение числа пи (около 3,1416), он может сделать выбор между 3,12, 3,13, 3,14, 3,15 и 3,16 с вероятностью, скажем, 0,05, 0,15, 0,55, 0,19 и 0,06. В этих условиях экзаменатору было бы очень трудно отличить дифференциальный анализатор от цифрового компьютера.
7. Аргумент от неформального поведения. Невозможно выработать свод правил, предписывающих человеку, что тот должен делать во всех возможных обстоятельствах. Можно, например, придумать правило, что человек должен останавливаться, когда он видит красный свет светофора, и идти на зеленый свет. Но что ему делать, если из-за технической

[1] Солипсизм (от лат. solus — «единственный» и лат. ipse — «сам») — философская, характеризующаяся признанием собственного индивидуального сознания в качестве единственной и несомненной реальности, и отрицанием объективной реальности окружающего мира.

неполадки оба зажигаются одновременно? Возможно, надежнее будет остановиться. Но из этого решения могут последовать затруднения в дальнейшем. Попытаться составить свод правил поведения, учитывающих любую возможность — даже сломанный светофор, — не представляется возможным. Со всем этим я согласен.

Глава 5. ДАГЛАС Р. ХОФШТАДТЕР. Тест Тьюринга: Беседа в кафетерии

Участники: Крис, студент-физик; Пат, студентка-биолог; Сэнди, студентка-философ.

СЭНДИ: Тест Тьюринга — это гениальный «зонд», подобный ускорителю частиц в физике. Думаю, что тебе понравится эта аналогия. В физике, чтобы узнать, что происходит на атомном или субатомном уровне, поскольку вы не можете это увидеть, вам приходится бомбардировать ваш объект ускоренными частицами и наблюдать, как они себя ведут после столкновения. По результатам наблюдения вы делаете заключения о внутренней природе объекта. Тест Тьюринга применяет ту же идею к интеллекту. Интеллект в нем рассматривается как «объект», невидимый глазом; структура этого объекта может быть вычислена абстрактно. «Бомбардируя» вопросами разум-объект, мы узнаем нечто о том, что происходит у него внутри, так же, как в физике.

ПАТ: Мне понадобится большее, чем простое физическое сходство с человеком, прежде чем я решу, что некое существо способно думать.

СЭНДИ: Главное здесь — не телесное, органическое и химическое сходство, а сходство внутренней, организационной структуры, программы. Вопрос о том, может ли некое существо думать, сводится для меня к вопросу о том, может ли его внутренняя организация быть описана определенным образом, и я считаю, что тест Тьюринга позволяет определить наличие или отсутствие этой организации.

КРИС: Почему разум должен обязательно идти рука об руку с эмоциями? Почему вы не можете представить себе разум, который только вычисляет, но при этом ничего не чувствует?

СЭНДИ: Любой разум должен иметь мотивацию. Когда система становится достаточно сложной, нам приходится сменить наш уровень описания. Деннетт называет подобное изменение уровней у наблюдателя «принятие интенциональной позиции». Думаю, что по-настоящему интересные вещи в Искусственном Интеллекте начнут происходить тогда, когда программа примет интенциональную позицию в отношении самой себя.

Размышления. Тест Тьюринга довольно категорично определяет, что в интеллекте важнее всего. Тьюринг утверждает, что важно вовсе не то, какой тип серого вещества имеется у кандидата в голове и не то, как он выглядит и пахнет, но то, способен ли он на разумное поведение. Подобно хорошему пари, тест Тьюринга очень соблазнителен; он выглядит справедливым, трудным, но возможным, и объективным в принятии окончательного решения. Тьюринг говорит нам: «Вместо того, чтобы спорить о конечной природе и сущности разума, почему бы нам не договориться, что любой, кто сможет пройти этот тест, наверняка разумен, и затем задаться вопросом, как сконструировать машину, которая сможет его пройти?» По иронии судьбы, Тьюрингу не удалось остановить этот вечный спор — он лишь направил его по иному руслу.

Часть III. ОТ «ЖЕЛЕЗОК» К ПРОГРАММАМ
Глава 9. АЛЛЕН УИЛЛИС. Дух

Жак Моно пишет: «Мы знаем или полагаем, что река и камень возникли в результате свободной игры физических сил, которым мы не можем приписать никакой цели, никакого проекта или намерения (разумеется, если мы принимаем основное положение научного метода, утверждающее, что природа объективна, а не проективна)».

Подобный подход выглядит очень привлекательно. Мы еще помним, как всего несколько поколений назад царило обратное мнение: камни желали падать, река хотела петь или разливаться. Своенравные духи наполняли мир и делали с природой все, что хотели. Мы знаем, насколько возросло наше понимание и контроль, когда мы решили рассматривать природные объекты и явления, как лишенные намерений. Камень ничего не хочет, у вулкана нет цели, река не жаждет влиться в море, ветер не стремится попасть куда-либо. Но существует и другая точка зрения.

Предположение, что природа «просто существует» и не имеет никакой цели, служит нам, когда речь идет о днях, годах и человеческих жизнях, но может послужить плохим путеводителем по равнинам вечности.

Глава 10. РИЧАРД ДОКИНЗ. Эгоистические гены и эгоистические мемы

(подробнее см. Ричард Докинз. Слепой часовщик. Как эволюция доказывает отсутствие замысла во Вселенной)

Дарвинское «выживание наиболее приспособленного» на самом деле является частным случаем более общего закона выживания наиболее устойчивого. Вселенную населяют устойчивые вещи. Устойчивая вещь — это набор атомов, достаточно долговременный или обычный для того, чтобы заслуживать названия.

Гемоглобин — хороший пример того принципа, что атомы образуют устойчивые сочетания. Еще до возникновения на земле жизни могла начаться некая рудиментарная эволюция молекул, движимая законами физики и химии. Здесь нет нужды думать о замысле, цели или направленности. Если группа атомов под воздействием некой энергии принимает устойчивую конфигурацию, скорее всего, она такой и останется. Самая ранняя форма натурального отбора была просто выбором устойчивых форм и отклонением неустойчивых форм.

Из этого, разумеется, не вытекает, что мы можем объяснить возникновение таких сложных существ, как человек, действием только этого принципа. Бесполезно брать нужное количество атомов и трясти их в присутствии некой внешней энергии до тех пор, пока они случайно не улягутся в нужном порядке и на свет не появится Адам. И здесь теория Дарвина, в наиболее общей ее форме, приходит нам на выручку. Теория Дарвина стартует там, где кончается медленное случайное формирование молекул.

Когда-то в незапамятные времена естественный отбор представлял собой дифференцированное выживание репликаторов, плававших в первичном бульоне. Сейчас естественный отбор предпочитает репликаторов, преуспевших в создании механизмов для выживания, генов, отлично умеющих контролировать эмбриональное развитие. При этом репликаторы так же бессознательны и лишены целенаправленности, как всегда. Как и раньше, продолжают происходить те слепые и неотвратимые процессы автоматического отбора среди конкурентов, согласно критериям долгожительства, плодовитости и аккуратности воспроизводства. Гены не обладают даром предвидения. Они не строят планов на будущее. Они просто существуют, и некоторые из них делают это более удачно. В этом и есть весь смысл эволюции.

Естественный отбор в самой общей форме означает дифференцированное выживание особей. Некоторые особи живут, другие умирают; однако, чтобы эта смерть оказала какое-то влияние на мир, должно быть соблюдено одно дополнительное условие. Каждая особь должна быть представлена в виде множества копий, и, хотя бы некоторые из особей должны быть потенциально способны к выживанию — в форме копий — в течение долгого периода эволюционного времени. Маленькие генетические единицы имеют эту особенность; индивиды, группы и классы — нет. Великое достижение Грегора Менделя состояло в том, что он показал, что на практике наследственные единицы можно рассматривать как неделимые и независимые частицы.

Еще один аспект корпускулярности гена состоит в том, что он никогда не стареет. Достигнув миллиона лет, он остается таким же, каким был в сто. Он перескакивает из одного тела в другое, путешествуя по поколениям, манипулирует телами для достижения собственных целей и оставляет смертные тела, одно за другим, до того, как они состарятся и умрут. Гены бессмертны.

Одна из самых удивительных особенностей поведения механизмов для выживания — это их кажущаяся целеустремленность. Каждый из нас знает путем собственного интроспективного анализа, что по крайней мере в одном из современных механизмов для выживания эта целеустремленность породила свойство, которое мы называем «сознанием».

Причина, по которой они не могут прямо управлять нами, как кукловод — марионеткой — запаздывание во времени. Гены работают, контролируя синтез белков. Это могучее орудие для управления миром, но действует оно достаточно медленно. Чтобы построить эмбрион, требуются

месяцы терпеливого дергания за белковые «струны». Поведение, с другой стороны, — вещь быстрая и действует на временных отрезках протяженностью не в месяцы, но в секунды и в доли секунд. Гены могут только постараться предусмотреть все заранее, построив себе быстрый компьютер и вложив в программу правила и «советы» на все случаи жизни, которые они способны «предусмотреть». Но, как и в игре в шахматы, в жизни слишком много различных гипотетических возможностей, и предусмотреть их все невозможно. Как программист, работающий над шахматной программой, гены должны «проинструктировать» собственные механизмы для выживания не по поводу специфических деталей, но дать им общие советы о стратегии и тактике выживания.

Одним из интересных способов предсказания будущего является имитация. Когда вам приходится принимать трудное решение в ситуации, включающей неизвестные будущие величины, вы прибегаете к определенного типа имитации. Вы представляете себе, что произойдет, если вы примете ту или иную возможную альтернативу. Вы создаете в голове модель мира — той его части, которая, как вам кажется, важна в данном случае. Вы ясно видите эту модель своим внутренним зрением и можете манипулировать ее составляющими.

Вершиной эволюции умения предсказывать стало, по-видимому, появление субъективного сознания. То, почему это произошло, кажется мне глубочайшей загадкой, стоящей перед современной биологией. Каковы бы ни были философские проблемы, возникающие в связи с самосознанием, для целей данной статьи мы можем считать его кульминацией эволюционной тенденции, ведущей к освобождению механизмов для выживания как субъектов, принимающих решения, от их абсолютных хозяев, генов. Мозг не только осуществляет повседневный контроль за делами механизма для выживания; он также научился предсказывать будущее и действовать в соответствии со своими предсказаниями. Он достаточно силен, чтобы восстать против диктата генов, например, отказавшись иметь столько детей, сколько физически возможно.

Диктуя то, как строятся механизмы для выживания и их нервная система, гены осуществляют контроль над поведением. Однако повседневные, сиюминутные решения — дело нервной системы. Гены создают общую политику, мозг воплощает ее в жизнь. По мере того как мозг развивается, он начинает перехватывать инициативу в создании общей политики, используя такие приемы как обучение и имитация. Логическим завершением этой тенденции была бы единственная общая инструкция, даваемая генами: делай то, что ты считаешь наилучшим, чтобы сохранить нас в живых.

Я думаю, что новый тип репликатора появился недавно на нашей планете. Новый бульон — это бульон человеческой культуры. Новому репликатору нужно имя, такое имя, которое бы несло идею культурной передачи, или единицы имитации – мем. Примерами мемов являются мелодии, идеи, крылатые фразы, фасоны одежды, приемы изготовления горшков или построения арок. Подобно тому, как гены распространяются в генофонде, переходя от тела к телу при помощи яйцеклеток и сперматозоидов, мемы распространяются в мемофонде, переходя от мозга к мозгу путем процесса, который в широком смысле можно назвать имитацией.

Когда мы умрем, мы сможем оставить после себя две вещи: гены и мемы. Наши гены могут обладать бессмертием, но определенный набор генов, представляющий каждого конкретного человека, недолговечен. С каждым поколением вклад ваших генов уменьшается вдвое. Однако, если вы внесли свой вклад в мировую культуру, ваше детище может продолжать жить через много лет после того, как ваши гены растворятся в общем генофонде.

Размышления. Докинз излагает редукционистскую гипотезу о том, что жизнь и разум возникли из кипящего молекулярного хаоса, когда крохотные, случайно возникшие единицы снова и снова проходили через безжалостный фильтр жестокой борьбы за ресурсы для самовоспроизводства. Редукционизм считает, что весь мир можно свести к физическим законам, и не оставляет места для так называемых «возникающих» (эмерджентных) свойств — структур высшего уровня, предположительно необъяснимых с точки зрения управляющих ими законов.

Глава 11. ДАГЛАС Р. ХОФШТАДТЕР. Прелюдия… и Муравьиная фуга

Ахилл и Черепаха пришли в гости к Крабу, чтобы познакомиться с его другом Муравьедом. После того, как новые знакомые представлены друг другу, вся четверка садится за чай (этот фрагмент присутствует так же в книге Даглас Хофштадтер. Гедель, Эшер, Бах. Эта бесконечная гирлянда).

АХИЛЛ: Знакомы ли вы с гравюрой М. К. Эшера под названием «Куб с магическими лентами»?

ЧЕРЕПАХА: На которой изображены изогнутые ленты с искривлениями в виде пузырей, которые кажутся попеременно то выпуклыми, то вогнутыми?

КРАБ: Я помню эту картину (рис. 1). Кажется, что пузыри на ней все время перескакивают из одного состояния в другое: они, то выпуклые, то вогнутые, в зависимости от того, с какого угла на них посмотреть. Невозможно одновременно увидеть их и выпуклыми, и вогнутыми — почему-то мозг этого просто не позволяет. У нас просто есть два разных способа воспринять эти пузыри.

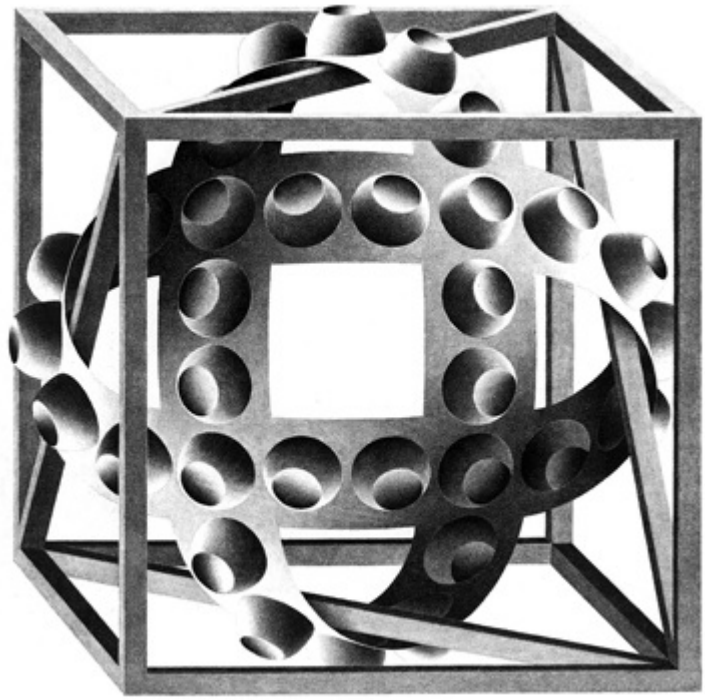


Рис. 1. М. К. Эшер. Куб с магическими лентами

КРАБ: ХОЛИЗМ означает, что целое больше, чем сумма его частей.

МУРАВЬЕД: РЕДУКЦИОНИЗМ означает, что целое может быть полностью понято, если вы понимаете и его части, и природу их «суммы».

КРАБ: Я отрицаю редукционизм. К примеру, можете ли вы объяснить мне, как понять мозг с помощью редукционизма? Любое редукционистское описание мозга неизбежно столкнется с трудностями, пытаясь объяснить, откуда в мозгу берется сознание.

МУРАВЬЕД: Я отрицаю холизм. К примеру, можете ли вы объяснить мне, как холистское описание муравьиной колонии может помочь понять ее лучше, чем описание отдельных муравьев, их взаимоотношений и ролей внутри колонии.

АХИЛЛ: «МУ» — это старинный ответ дзен-буддизма, «развопросивающий» вопрос. Нашим вопросом было: «Должны ли мы понимать мир холистским или редукционистским способом?» Ответ «МУ» отрицает самую постановку этого вопроса предположение, что необходимо выбрать лишь один из двух способов. Развопросивая этот вопрос, «МУ» открывает нам истину высшего порядка: существует более широкий контекст, куда вписываются и холистский и редукционистский подходы (рис. 2).

Рис. 2. Картина МУ; если присмотреться, то можно увидеть несколько уровней написания слов МУ, ХОЛИЗМ и РЕДУКЦИОНИЗМ

ЧЕРЕПАХА: Мне кажется, что муравейник напоминает человеческий мозг, состоящий из нейронов. Чтобы объяснить человеческую способность к разумной беседе, никто не стал бы утверждать, что отдельные нервные клетки — разумные существа.

АХИЛЛ: Хотя в моем мозгу происходит бурная деятельность, я способен воспринять ее только на одном уровне — уровне символов; при этом я полностью нечувствителен к низшим уровням.

КРАБ: Именно это и произошло, когда вы прочитали «МУ», не замечая низших уровней, «ХОЛИЗМА» и «РЕДУКЦИОНИЗМА».

Размышления. Некоторым читателям этот диалог напоминает о поведении стран — оно кажется целеустремленным, эгоистическим и направленным на выживание; оно каким-то образом возникает из привычек и обычаев граждан данной страны, из их системы образования, юридических установлений, религии, ресурсов, стиля потребления и уровня запросов и так далее. Когда из различных особей формируется жесткая организация, несводимая к вкладам отдельных особей на низших уровнях, мы рассматриваем подобную структуру как индивида высшего уровня, которого часто очеловечиваем. В одной газетной статье о террористической группе было написано, что та «играет, держа свои карты слишком близко к груди». О России часто говорят, что она «желает» добиться мирового признания своей мощи, поскольку «страдает» от «застарелого комплекса неполноценности» по отношению к Западной Европе. Хотя это не более, чем метафоры, они показывают, насколько сильно в нас стремление очеловечивать организации.

У индивидов, составляющих организацию — рабочих, секретарш, шоферов автобусов, управляющих и т.д. — есть свои цели в жизни, цели, которые, как кажется, должны быть в конфликте со структурой высшего уровня, частями которой они являются. Однако существует эффект, могущий показаться предательским и зловещим многим, изучающим политологию. Он состоит в том, что организация ассимилирует и эксплуатирует именно эти цели, используя гордость, необходимость самоуважения и так далее, для получения собственной выгоды. Таким образом из многих целей низших уровней возникает единый импульс, включающий их в себя, и именно этот импульс самосохраняется.

Есть ли какой-нибудь смысл «сравнивать» себя с целой страной? Есть ли у страны мысли или убеждения? Все сводится к тому, есть ли у страны некий символический уровень. Вместо того, чтобы говорить, что система «имеет символический уровень», мы могли бы сказать: «это репрезентативная система». Концепция «репрезентативной» системы — основная мысль этой книги. Под «репрезентативной системой» мы подразумеваем активный, самоактуализирующийся набор структур, организованный так, что он «отражает» мир в его эволюции. Репрезентативная система построена на категориях; она сортирует входящие данные по различным категориям, с необходимостью уточняя или расширяя информационную сеть внутренних категорий.

Компьютерная программа, которая может «посмотреть» на некую сцену и сказать, какие объекты она «видит», находится ближе к репрезентативной системе. Программа, которая может посмотреть на некую сцену и не только сказать, какие объекты перед ней находятся, но и определить, что вызвало эту сцену к жизни и как события могут развиваться дальше, была бы, согласно нашему определению, настоящей репрезентативной системой.

Синонимом редукционизма является «каузальность, направленная снизу вверх», а синонимом холизма — «каузальность, направленная сверху вниз» Существуют и соответствующие понятия для временно̀й шкалы: редукционизму соответствует идея предсказания будущего на основании прошлых данных, не учитывая «целей» организмов. Холизму соответствует идея о том, что так можно предсказать будущее лишь для неодушевленных объектов; в случае же одушевленных организмов их действия определяются в основном их целями, намерениями и желаниями.

Если взять целизм и холизм вместе, вы получите соулизм (от англ. «soul» — душа). Предсказанионизм и редукционизм, взятые вместе, дают механицизм.

Мозг – тонкая, сложная и самоуглубленная система, которая реагирует на внешние стимулы и на состояние разных уровней своей внутренней конфигурации. Почти невозможно отличить ответы такой системы на окружающий мир от ее ответов на состояние собственных частей, поскольку малейшая внешняя пертурбация порождает мириады крохотных взаимосвязанных событий, нарастающих лавинообразно. Если думать об этом как о «восприятии» системой входных данных, ясно, что ее собственное состояние воспринимается системой так же. Самовосприятие здесь неотличимо от восприятия.

Когда подобные системы эволюционируют путем естественного отбора, обязательно будет какая-то причина того, что некоторые выживают, а некоторые — нет: это осмысленная внутренняя организация, позволяющая системе использовать окружающий мир и, по крайней мере частично, его контролировать.

Часть IV. РАЗУМ КАК ПРОГРАММА
Глава 17. КРИСТОФЕР ЧЕРНЯК. Загадка Вселенной и ее решение

Размышления. Парадокс лжеца и ему подобные — следствия формализации наблюдения, старо, как само человечество: любой объект находится в совершенно особых отношениях сам с собой, что ограничивает его способность воздействовать на самого себя так, как он может воздействовать на все другие объекты. Карандаш не может писать сам на себе; мухобойка не может прихлопнуть муху, сидящую на ее ручке; змея не может съесть саму себя, и так далее.

В последнее время наибольший прогресс в работе над искусственным интеллектом был достигнут при попытках снабдить программу набором понятий о ее собственных внутренних структурах и способом реагировать на определенные замеченные в них изменения. Эта идея возникла как одно из ключевых требований к настоящей гибкости, синониму интеллекта.

Часть V. СОЗДАННОЕ САМОСОЗНАНИЕ И СВОБОДНАЯ ВОЛЯ
Глава 20. РЭЙМОНД М. СМОЛЛЯН. Даоист ли Бог?

Размышления. Внутренний конфликт — это одна из наиболее знакомых и тем не менее еще не понятых особенностей человеческой природы. Известный рекламный слоган картофельных чипсов, «Спорим, ты не сможешь съесть только один!», эффективно напоминает нам о нашем внутреннем расколе. Существует распространенный миф, что каждый человек является единством, единой организацией, обладающей своей волей. Как раз наоборот, человек — это сочетание разных суб-людей, каждый из которых обладает своей волей. Эта иерархическая организация индивидуальности не слишком льстит нашему достоинству, но есть много указаний на то, что дело обстоит именно так.

Смоллян затрагивает темы, к которым мы обращались на протяжении всей книги — попытка примирить детерминизм и направленную «снизу вверх» каузальность законов природы со свободой воли и каузальностью, направленной «сверху вниз», которую мы все осуществляем. Смоллян говорит устами Бога (или наоборот?), что детерминизм и свобода воли ближе, чем кажутся. Смоллян разработал элегантную возможность примирения этих противоположных взглядов, которая зависит от нашей готовности сменить точку зрения — прекратить думать «дуалистично» (то есть разбивать

мир на «Я» и «Не-я») и увидеть всю вселенную как безграничное пространство, с объектами, перетекающими друг в друга, накладывающимися друг на друга. В такой вселенной нет точно определенных категорий, нет углов.

Сначала кажется, что это очень странное убеждение для логика — но кто сказал, что логики всегда должны быть строгими и несгибаемыми? Одна из любимых фраз Марвина Мински — «Логика неприложима к реальному миру.» В каком-то смысле это так и есть. Это одна из трудностей, с которыми сталкиваются разработчики искусственного интеллекта. Они понимают, что разум не может быть основан только на логических рассуждениях — точнее, что изолированные рассуждения невозможны, поскольку они зависят от предварительно установленной системы понятий, классов, категорий — как бы вы их ни называли — в терминах которых понимаются затем все ситуации. Именно здесь рождаются склонности и предпочтения. Рассуждающая способность должна быть готова принять первую характеристику ситуации, выданную воспринимающей способностью. Однако, если она усомнится в этой характеристике, воспринимающая способность должна, в свою очередь, быть готова согласиться с этими сомнениями и пересмотреть первоначальную интерпретацию ситуации. Так создается постоянная замкнутая петля между уровнями. Именно взаимодействие воспринимающего «я» и рассуждающего «я» и порождает полную индивидуальность.

Глава 22. ДЖОН Р. СИРЛ. Разум, мозг и программы

Я понимаю рассказы по-английски, в меньшей степени, по-французски и в еще меньшей степени, по-немецки. По-китайски я не понимаю ничего. С другой стороны, моя машина и мой калькулятор не понимают ничего — это не их дело. Мы часто метафорически приписываем «понимание» и другие когнитивные предикаты машинам, калькуляторам и другим приспособлениям, но это ничего не доказывает. Мы говорим: «Дверь знает, когда нужно открыться, благодаря своим фотоэлектрическим элементам», «Калькулятор может складывать и вычитать, но не способен умножать» и «термостат воспринимает изменения температуры». Причина, по которой мы приписываем все это машинам, довольно интересна — мы наделяем их собственной интенциональностью.[2] Наши инструменты — продолжение наших целей, и мы находим естественным метафорически приписывать им интенциональность.

Запрограммированный компьютер понимает столько же, сколько автомобиль или калькулятор, то есть совершенно ничего. Понимание компьютера не частично или неполно (как мое понимание французского или немецкого) — оно равно нулю.

Может ли нечто думать, чувствовать, всего лишь будучи компьютером с соответствующей программой? Может ли работа программы быть достаточным условием возникновения понимания? Ответ отрицательный. Почему? Потому что сама по себе манипуляция формальными символами не обладает интенциональностью и не имеет смысла. Нельзя даже сказать, что это манипуляция символами, поскольку символы ничего в данном случае не символизируют. Говоря языком лингвистов, они обладают только синтаксисом, но лишены семантического наполнения. Кажущаяся интенциональность компьютеров находится в головах тех, кто их запрограммировал и использует, кто вводит некие данные и интерпретирует ответ.

Свойство Искусственного Интеллекта, которое казалось таким привлекательным — различение между программой и ее реализацией — оказывается роковым для гипотезы о том, что симуляция может быть идентичной копией. Различие между программой и ее реализацией на определенной аппаратуре повторяет различие между уровнем мышления и уровнем мозговой деятельности. Если бы нам удалось описать мыслительные процессы в виде формальной программы, возможно, удалось бы описать основные характеристики разума, не прибегая к интроспективной психологии или нейрофизиологии мозга. Однако уравнение «разум для мозга = программа для аппаратуры» имеет слабые стороны. Программа чисто формальна, в то время как интенциональные состояния этим типом формальности не обладают. Они определяются своим содержанием, а не формой.

[2] Интенциональность – черта некоторых мысленных состояний, направляющих их на объекты и события в мире. Убеждения, желания и намерения – примеры интенциональных состояний, а ненаправленная тревожность и депрессия – нет.

Существует несколько причин того, что ИИ должен был казаться — и кажется многим до сих пор — воспроизведением и объяснением феномена мышления. Нам не удастся развеять это заблуждение, пока мы не выясним причин, по которым оно возникло. Первая и, возможно, наиболее важная причина — это путаница с понятием «обработка информации». Запрограммированный компьютер не обрабатывает информацию в том смысле, в каком это делает человек, когда он решает арифметические задачи, читает или пытается ответить на вопрос о прочитанном рассказе. Компьютер занимается только тем, что манипулирует формальными символами.

ИИ изобилует остаточным бихевиоризмом или операционализмом. Тест Тьюринга, бессовестно бихевиористский и операционалистский, типичен для этой традиции. Мне кажется, что, если бы специалисты по ИИ полностью отказались бы от бихевиоризма и операционализма, это положило бы конец путанице между подражанием и идентичной копией. Остаточный операционализм усугубляется остаточным дуализмом. Сильная версия ИИ имеет смысл только тогда, если принимается на вооружение дуалистическая гипотеза: там, где дело касается разума, мозг не важен. Программы не зависят от их реализации в машинах: одна и та же программа может быть реализована в электронном компьютере, картезианской мыслительной субстанции или гегельянском мировом духе. Однако, феномен человеческого разума может находиться в зависимости от физико-химических свойств человеческого мозга.

Сторонники ИИ пытаются воспроизвести и объяснить ментальные явления при помощи программ. Однако, поскольку программа полностью независима от ее реализации, этот проект может быть приведен в исполнение лишь в том случае, если разум не только концептуально, но и эмпирически независим от мозга. Если вы не верите в то, что разум можно отделить от мозга как концептуально, так и эмпирически — дуализм в сильной форме, — то вы не можете надеяться имитировать ментальные процессы путем написания и пуска в действие программ, поскольку программы должны быть независимы как от мозгов, так и от любой другой формы воплощения.

Чем бы еще ни была интенциональность, это прежде всего биологический феномен, который, скорее всего, находится в такой же каузальной зависимости от специфической биохимии своего происхождения, как лактация, фотосинтез и любой другой биологический феномен.

Размышления. Статья Сирля основана на изобретательном примере «мысленного эксперимента китайской комнаты». В этом эксперименте читателю предлагается представить себя в роли человека, вручную проделывающего те же шаги, которые предположительно проделывает хитроумная программа ИИ, когда она читает рассказы по-китайски и отвечает на поставленные вопросы настолько похоже на то, как это сделал бы человек, что может пройти тест Тьюринга. Нам кажется, что Сирл допускает серьезнейшую ошибку, предполагая, что человек на это способен. Принимая идею Сирля на веру, читатель оказывается незаметно втянутым в совершенно нереалистичную концепцию отношений между разумом и манипуляцией символами.

Нашим ответом на это является, в основном, «ответ системы». Мы считаем, что ошибочно приписывать понимание человеческому компоненту системы. Скорее, понимание — свойство всей системы в целом. Когда человеческая «система» изучает иностранный язык. Второй язык не действует «поверх» первого, как некая программа-паразит; вместо этого он, как и первый язык, в большей или меньшей степени входит в саму структуру аппаратуры. Усвоение второго языка влечет за собой глубокие изменения в поддерживающей «аппаратуре» — разнообразные изменения в том, как возбуждаются нейроны. Как следствие, создаются новые способы для феноменов высшего уровня — символов — активировать друг друга.

Подобным образом в компьютерной системе программа высшего уровня должна вносить изменения в «демона», выполняющего свою программу. Это совершенно чуждо современному типу имплементации, при котором уровни располагаются строго один над другим по вертикали и каждый из них изолирован от других. Способность высшего уровня «загибаться назад» и затрагивать низшие уровни — свой собственный фундамент — это магический трюк, который, как нам кажется, близок самому сердцу сознания. Может быть однажды станет ясно, что это ключ к достижению все большей гибкости компьютерного дизайна и, в конце концов, разработке искусственного интеллекта.

Глава 23. РЭЙМОНД М. СМОЛЛЯН. Несчастный дуалист

Размышления. Дуалисты считают, что разум и тело — раздельные субстанции. Это означает, что мы имеем дело по крайней мере с двумя видами субстанций — физической и ментальной. Наш разум сделан из чего-то, не обладающего ни материальной массой, ни физической энергией; возможно, он даже не занимает места в пространстве. Этот взгляд настолько загадочен, так систематически избегает попыток разъяснения, что непонятно, как он может кого-либо привлекать. Одна из дорог, ведущих к дуализму, проходит через следующий (никуда не годный) аргумент:

- Некоторые факты не касаются свойств, обстоятельств и отношений физических объектов.
- Следовательно, некоторые факты касаются свойств, обстоятельств и отношений нефизических объектов.

Этот аргумент ошибочен.

Часть VI. ВНУТРЕННЕЕ ОКО
Глава 24. ТОМАС НАГЕЛЬ. Каково быть летучей мышью?

Самосознание — вот то, что делает проблему тело-разум практически неразрешимой. Недавняя волна редукционистской эйфории породила несколько работ, анализирующих феномен мышления и связанные с ним понятия. Однако все эти проблемы являются общими как для данного типа редукции, так и для других ее типов. То, что делает проблему тело-разум уникальной и непохожей на проблемы вода — H_2O, или машина Тьюринга — компьютер IBM, или молния — электрический разряд, или ген — ДНК, или дуб — углеводород, в этих работах игнорируется. В настоящее время мы не представляем себе, каким могло бы быть объяснение физической природы мыслительных явлений. Если бы самосознания не было, загадка тело-разум была бы гораздо менее интересной. С самосознанием эта проблема кажется безнадежной.

Размышления о том, на что похоже быть летучей мышью, привели нас к следующему заключению: существуют некие факты, которые не могут быть выражены средствами человеческого языка. Нам приходится признавать существование этих фактов, хотя мы не можем их понять и выразить.

Не будем обсуждать эту тему более подробно. Ее значение для нашего вопроса, проблемы тело-разум, заключается в том, что она позволяет нам сделать общий вывод о субъективном характере опыта. Каков бы ни был статус переживаний человека, марсианина или летучей мышью, все эти переживания включают в себя определенную точку зрения.

Говоря о психофизической редукции, мы сталкиваемся с общей трудностью. В других областях движение в сторону редукции — это движение в сторону большей объективности, к более аккуратной картине действительной природы вещей. Однако сам личный опыт, по-видимому, не укладывается в эту схему. В этом случае движение от видимого к действительному теряет смысл. Кажется маловероятным, что мы приблизимся к действительной природе человеческого опыта, оставив в стороне особенности человеческой точки зрения и составляя описание этого опыта таким образом, чтобы оно могло быть понято существами, отличными от нас, которые никогда не смогут вообразить, на что похоже быть человеком. Если субъективный характер опыта возможно понять с единственной точки зрения, тогда любой сдвиг в сторону большей объективности — то есть меньшей зависимости от определенной точки зрения — не приближает нас к действительной природе данного феномена, а, напротив, удаляет от нее.

Большинство нео-бихевиористских течений современной философской психологии основаны на стремлении выработать объективную концепцию разума, которая может быть полностью редуцирована. Если мы согласимся с тем, что физическая теория разума должна иметь в виду субъективный характер опыта, мы должны согласиться и с тем, что никакая из существующих на данный момент теорий не объясняет нам, как это можно сделать. Эта проблема уникальна. Если мыслительные процессы являются в действительности процессами физическими, следовательно, испытывать эти процессы по определению на что-то похоже. На что — остается загадкой.

Глава 26. ДАГЛАС ХОФСТАДТЕР. Беседа с мозгом Эйнштейна

Размышления. Мысленные эксперименты могут быть систематическими, и вытекающие из них следствия могут являться результатом строгих умозаключений. Вспомните, как Галилей

неопровержимо доказал методом reductio ad absurdum несостоятельность гипотезы о том, что тяжелые предметы падают быстрее легких. Он предлагает нам мысленно взять тяжелый предмет A и легкий предмет B и перед тем, как сбросить их с башни, связать их вместе веревкой или цепью. Согласно гипотезе, B будет падать медленнее, а следовательно, задерживать A в его падении. Таким образом, предмет A, привязанный к B, будет падать медленнее, чем A сам по себе. Однако A, привязанный к B, уже является новым объектом C, который тяжелее, чем A — а следовательно, должен бы падать быстрее, чем A сам по себе. A и B не могут в одно и то же время падать и быстрее, и медленнее, чем A (противоречие!), что доказывает ошибочность исходной гипотезы.