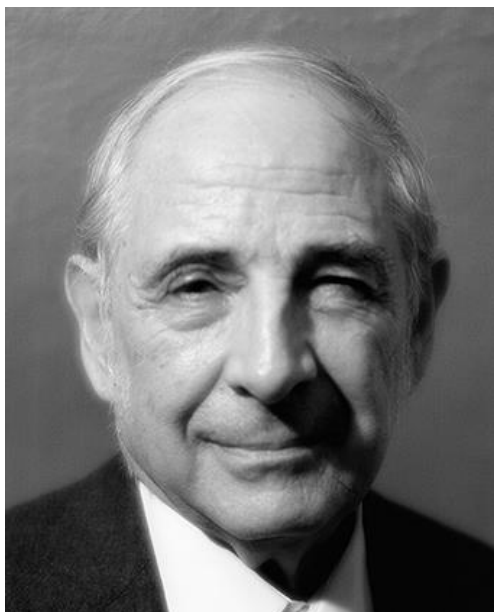


СОЗНАНИЕ, МОЗГ И НАУКА



Джон Сёрль

Searle J. Minds, Brains and Science. 1984. Сёрль Дж. Р. Сознание, мозг и наука. Перевод на русский язык: А. Ф. Грязнов. — М., 1993. // Электронная публикация: Центр гуманитарных технологий. — 10.06.2013.

Джон Роджерс Сёрль (John Rogers Searle) — американский философ, представитель аналитического направления в философии (см. Аналитическая философия). В 1960–1970-х годах занимался развитием теории речевых актов. Автор понятия косвенного речевого акта. С 1980-х годов — один из ведущих специалистов по философии искусственного интеллекта. Профессор философии Калифорнийского университета в Беркли. Автор широко известного мысленного эксперимента «Китайская комната», отвергающего возможность воспроизведения семантической составляющей человеческого интеллекта синтаксическими средствами.

ПРЕДИСЛОВИЕ РЕДАКТОРА-ПЕРЕВОДЧИКА

Вот уже три десятилетия имя Джона Сёрля известно философскому миру. Джон Сёрль родился в 1932 году в Денвере, штат Колорадо, США, высшее образование получил в университетах Висконсина и Оксфорда. С 1959 года он является профессором философии Калифорнийского университета в г. Беркли. Известность американскому философу принесла его первая книга «Речевые акты» (1969), за которой последовали «Выражение и значение» (1979) и «Интенциональность» (1983). Русскому читателю лишь недавно стали доступны небольшие публикации Сёрля, посвящённые специальным проблемам философии языка.¹ Таким образом «Сознание, мозг и наука» — первая книга американского автора, которая публикуется в России.

Джон Сёрль — один из наиболее самобытных современных представителей аналитической традиции в западной философии, представленной такими выдающимися именами, как Б. Рассел, Дж. Мур, Л. Витгенштейн, А. Айер, Р. Карнап, Г. Райл. Особенно широкое распространение метода философского анализа в последние годы, его выход за пределы англо-саксонского философского мира и популярность в иных регионах объясняются тем, что многие философы (и не только аналитики, но и, скажем, герменевтики или феноменологи) видят в анализе современный стиль мышления, характеризующийся особым вниманием к строгости аргументации, точности, избирающий и качестве своего предмета структуру философского рассуждения, его языковые средства.

Отечественному читателю прежних лет, знакомившемуся с «критикой буржуазной философии» по учебникам, в отношении аналитической философии повезло, пожалуй, меньше всего, Единственное исключение в этом плане составляют отдельные издания конца 1950-х — начала 1960-х годов, а также публикации текстов философов и методологов науки (которых трудно безоговорочно отнести к аналитической традиции) в 1970-е годы. До середины 1980-х годов в нашей литературе почти не использовался термин «аналитическая философия». Эталоном неопозитивизма считалась ранняя доктрина логического позитивизма Венского кружка с её экстремистской «антиметафизической» программой разрыва с традиционной философской проблематикой. Как известно, аналитическая философия давно размежевалась с таким «уклоном», восстановив свою связь с историей и мировым философским процессом. Более того, именно аналитическая философия из всех направлений современной западной философии в наибольшей степени сохраняет преемственность по отношению к классическому наследию,

ассимилировав его главные проблемы. Об этом свидетельствует и публикуемый текст Сёрля, в котором рассмотрены едва ли не все «вечные» философские вопросы. Конечно, терминология, в которой формулируются и решаются эти вопросы, существенно иная, чем в философии классической.

Незнание аналитической философии, думается, обеднило ситуацию, сложившуюся в отечественной философии самых последних лет. Освободившись от идеологического контроля и догматических пут, она приобрела одностороннюю экзистенциально-антропологическую направленность. Разумеется, значительная часть работ философов-аналитиков трудна для понимания, требует специальной логико-лингвистической подготовки, знания некоторых конкретных наук. Тем не менее ведущие аналитики за рубежом ищут всё более широкую аудиторию, стремясь придать своим идеям общемировоззренческое звучание. Аналитическая философия сегодня уже не делается в «башне из слоновой кости», она перестаёт быть академической «забавой». И Рейтовские лекции Джона Сёрля — хороший тому пример. Кстати, его знаменитый предшественник Бертран Рассел, открывший эту серию в 1948 году, сам был замечательным популяризатором «строгой» философии. По стилю изложения и широкому охвату фундаментальных вопросов книга Сёрля сопоставима с работой Рассела «Проблемы философии» (1912), которую до сих пор многие считают лучшим современным введением в философию. ²

Что касается собственно теоретического содержания концепции Сёрля, то оно ближе не линии Рассела, а линии позднего Витгенштейна. Прямой его предшественник и учитель — английский философ Джон Остин (1911–1960), создавший теорию «речевых актов». Данная теория явилась развитием концепции Витгенштейна о «языковых играх» как целостных системах лингвистических и нелингвистических действий человека. Теория «речевых актов» нашла ряд приложений и вне сферы философии — в лингвистической прагматике.

Основной заслугой Сёрля является сближение понятия «речевой акт» с понятием «интенциональность», которое рассматривается как фундаментальное свойство психической репрезентации. Им, в частности, разработан формальный аппарат так называемой иллокутивной логики, придающий концептуальную строгость теории Остина. Не следует путать теорию интенциональности, которая разрабатывается в рамках аналитической философии, и теорию интенциональности, принадлежащую феноменологии. Их роднит лишь самое общее понимание смысла данного термина. Для Сёрля интенциональны такие ментальные состояния, как убеждения (верования), страхи, радости, надежды, желания, фантазии и прочее. Интенциональности

присущ и своеобразный каузальный аспект. Долингвистическое, биологическое в своей основе свойство интенциональности, согласно Сёрлю, получает реализацию в «языке-действии», что позволяет исследовать её, не прибегая к таинственной «идеации» и трансценденталистской терминологии. Таким образом, Сёрль выступает против отождествления интенциональности с сознанием, и уж тем более с «чистым сознанием» Гуссерля.

Конечно, указание на биологические корни интенциональности может вызвать подозрение в односторонней физикалистской ориентации Сёрля, в его симпатиях к различным «научно-материалистическим» теориям тождества психического и телесного (нейрофизиологического), столь популярным сегодня на Западе. Однако уже первые главы Рейтовских лекций не оставляют сомнения в более глубокой, философской постановке и решении им психофизической проблемы о том, что Сёрлю не свойственны сциентистские настроения, зачарованность новыми достижениями научно-технической мысли, свидетельствуют его критические рассуждения о когнитивизме, а также мысленный эксперимент («китайская комната»), указывающий пределы возможностей концепции «сильного искусственного интеллекта». Сёрль, как и его коллеги по Университету в Беркли Хьюберт и Стюарт Дрейфус,³ не устаёт доказывать невозможность отождествления психики (в том числе сознания) и алгоритмизированной компьютерной программы, подчёркивая содержательно-семантический характер первой и формально-синтаксический характер второй.

Трезвый, реалистический подход к модным «вычислительным» моделям психики выделяет Сёрля среди многих его коллег. ⁴

Внимание читателя несомненно привлекут и изложенные в последних главах книги мысли Сёрля о специфике социального знания в сопоставлении со знанием естественнонаучным, о характере социальных законов, а также о свободе и детерминизме. Материал четвёртой главы книги относится к пока ещё малоизвестной в России дисциплине — аналитической «философии действия» и поэтому, можно надеяться, вызовет особый интерес.

Александр Феодосиевич Грязнов, 1993. ⁵

Примечания:

1. Сёрль Дж. Что такое речевой акт? — «Новое в зарубежной лингвистике». Вып. XVII. — М., 1986; Классификация иллокутивных актов (там же); Косвенные речевые акты (там же); Природа интенциональных состояний (первая глава из книги «Интенциональность». — Прим. авт.) — Философия. Логика. Язык. — М., 1987; Перевернутое слово. — «Вопросы философии», 1992, № 4.

2. Russell B, The problems of philosophy. — L, 1912.

3. См. Dreyfus H. L., Dreyfus S. E. Mind over machine. The power of human intuition and expertise in the era of the computer. — NY, 1986.

4. Справедливости ради следует упомянуть концепцию наиболее серьёзного оппонента Сёрля Дэниэла Деннета, который, фиксируя эволюционно-биологическое (генетическое) происхождение человеческой интенциональности, разрабатывает широкую «интенциональную стратегию», позволяющую рассматривать в качестве «интенциональных систем» различные объекты (например, артефакты), поведение которых в определённом смысле предсказуемо и рационально (см. Dennett D. The Intentional stance. — Cambridge, Mass., 1987).

5. Грязнов А. Ф. О Джоне Сёрле. // Путь. Международный философский журнал, 1993, № 4. С. 67–69.

ПРЕДИСЛОВИЕ АВТОРА

Приглашение прочитать Рейтовские лекции 1984 года было для меня огромной честью. Бертран Рассел начал всю эту серию в 1948 году и оставался единственным философом, участвовавшим в работе над Рейтовскими лекциями по радио.

Но это не просто честь, но и личный вызов. В идеале лекций должно быть шесть, каждая продолжительностью в полчаса, лекция должна иметь самостоятельное значение и в то же время оставаться частью единого курса, основывающегося на уже проделанной исследовательской работе лектора и в то же время содержащего новый оригинальный материал. И наконец (самое, вероятно, трудное условие), лекции должны быть понятны аудитории, пусть даже заинтересованной и живой, но в большинстве своём не знакомой ни с предметом, ни с терминологией, ни со специалистами в данной области знания. Не думаю, что все эти цели достижимы, но во всяком случае я старался выполнить поставленные условия. Одним из главных моих мотивов было также убеждение в том, что результаты и методы современной аналитической философии доступны самому широкому кругу людей.

В книжной версии этих лекций я намеревался расширить главы и ответить на всевозможные возражения со стороны моих придирчивых собратьев-философов, не говоря уже о замечаниях коллег по когнитивной науке, искусственному интеллекту и некоторым другим областям. Короче говоря, лекции ждала участь превращения в обычную книгу со сносками и так далее. В конце концов я отказался от этого плана, потому что он повредил бы тому, что является привлекательной стороной Рейтовской серии, а именно её доступности. Нижеследующие главы, в сущности, повторяют названные лекции. Некоторые из них немного расширены, однако я старался сохранить стиль, тон и неформальный характер моих бесед на радио.

Главная тема лекций — место человеческих существ во Вселенной. В частности, в них затрагивается вопрос о совместимости традиционного менталистского представления человека о самом себе с явно противоречащим ему представлением о Вселенной как о чисто физической системе, или как о множестве взаимодействующих физических систем. Каждая глава является ответом на конкретные вопросы: каково отношение между сознанием и мозгом? может ли у компьютеров быть сознание по той единственной причине, что у них правильные программы с правильными результатами на входе и выходе? насколько правдоподобна модель сознания, представляющая

его как компьютерную программу? какова структура человеческого действия? каков статус социальных наук? каким образом, если это вообще достижимо, можно примирить убеждение в существовании свободной воли с представлением о Вселенной как о физической системе, или как о множестве взаимодействующих физических систем?

В ходе работы над лекциями наметились и некоторые другие важные темы, о которых мне хотелось бы сказать. Надеюсь, что это поможет читателю лучше понять последующее.

Первая из этих тем — скудость наших знаний о деятельности человеческого мозга, — скудость, предопределяющая некоторые теоретические претензии. Как писал в 1978 году нейрофизиолог Дэвид Хьюбел, «наши знания о мозге примитивны. В то время как функции некоторых его областей нам известны, имеются другие, размером с кулак, о которых мы знаем столько же, сколько знали о сердце до того, как поняли, что оно качает кровь». И в самом деле, если взять полдюжины учебников о мозге и поискать там ответы на возникающие вопросы, то в результате мы скорее всего испытаем разочарование. Какова нейрофизиология сознания? Почему мы нуждаемся во сне? Почему алкоголь делает нас пьяными? Как именно воспоминания хранятся в мозге? Мы просто не знаем, как ответить на любой из этих фундаментальных вопросов. В основе многих утверждений лежит незнание. Многие заявления о работе мозга, которые мы встречаем в различных дисциплинах — от фрейдовской психологии до исследований по искусственному интеллекту, кажутся правдоподобными вследствие того, что наше знание просто изобилует пробелами.

С точки зрения традиционного подхода, рассматривающего нейрон в качестве фундаментальной функциональной единицы, всё разнообразие данных, которые получает мозг, — фотоны, ударяющиеся о сетчатку, звуковые волны, достигающие барабанной перепонки, воздействие на кожу, которое активирует нервные окончания, реагирующие на давление, жару, холод, боль и так далее, — все эти данные, имеющиеся на входе (inputs), преобразуются в то или иное нейронное возбуждение. Степени возбуждения нейронов в нейронных сетях и локальных состояниях мозга и порождают нашу ментальную жизнь во всём её разнообразии. Запах роз, синева неба, вкус лука, мысль о математической формуле, — все это порождается нейронным возбуждением в сетях и соответствует различным локальным состояниям мозга.

Но что это за нейронные сети и локальные состояния, которыми объясняют нашу ментальную жизнь? Никто этого точно не знает, но есть данные, что определённые области мозга специализируются на определённых видах опыта. Одна часть коры головного мозга выполняет особую роль в зрительном восприятии, другая — в слуховом и так далее. Представим себе, что слуховые стимулы поступают в зрительную часть коры, а зрительные — в слуховую. Что из этого получится? Насколько мне известно, никто не ставил такого эксперимента, но разумно было бы предположить, что слуховой стимул будет «увиден», то есть вызовет зрительные ощущения (experiences), а зрительный будет «услышан», то есть вызовет слуховые ощущения, и всё это произойдёт в силу специфических, хотя во многом и неизвестных, свойств (features) зрительной и слуховой частей коры. Эта гипотеза носит спекулятивный характер, однако давайте задумаемся, почему удар в глаз вызывает визуальную вспышку («искры из глаз»), — ведь никакого оптического стимула здесь нет.

Вторая тема — введённое в нас неприятие подхода к сознанию как к обычному биологическому феномену. Это возвращает нас в XVII век, к Декарту. Декарт делил мир на два рода субстанций: ментальные и физические. Физические субстанции находятся в ведении науки, а ментальные принадлежат сфере религии. Нечто похожее на это деление мы встречаем и сегодня. Например, сознание и субъективность часто рассматриваются как вне научные предметы. И такое нежелание исследовать сознание и субъективность отчасти обусловлено определённым пониманием науки: наука якобы должна заниматься объективно наблюдаемыми феноменами. Выступая перед биологами и нейрофизиологами, я обнаружил, что многие из них не считают психику в целом и сознание в частности за предметы, подлежащие научному исследованию.

Третья тема — неадекватность традиционной терминологии. Из трёх терминов, составляющих название «Сознание, мозг и наука», только второй имеет вполне определённый смысл. Под «сознанием» имеется в виду последовательность мыслей, чувств и впечатлений, сознательных или бессознательных, составляющих нашу ментальную жизнь. Но этот термин вызывает к жизни множество опасных призраков, целый сонм дряхлых философских теорий. Очень трудно избавиться от того, чтобы представлять психику как вещь, или как арену, или как чёрный ящик, — как нечто такое, внутри чего происходят ментальные процессы.

Со словом «наука» дело обстоит ещё хуже. Будь моя воля, я бы вообще запретил этот термин. «Наука» превратилась во что-то вроде почётного звания, и различные дисциплины, совершенно непохожие ни на физику, ни на

химию, претендуют на то, чтобы называть себя «наукой». Можно даже вывести эмпирический закон: всё, что именует себя наукой, скорее всего таковой не является: возьмём, например, «христианскую науку», или «военную науку», или даже «когнитивную науку», или «социальную науку». Слово «наука» вызывает у нас представление о людях в белых халатах, потряхивающих пробирками и бросающих многозначительные взгляды на приборы. Для многих наука — это нечто таинственное и непогрешимое. Мне кажется, что образ науки должен быть иным. В дисциплинах интеллектуального характера мы стремимся к знанию и пониманию. И существуют только знание и понимание — и в математике, и в литературной критике, и в истории, физике и философии. Некоторые дисциплины при этом более систематичны, — их-то и можно было бы обозначить словом «наука».

За помощь в подготовке Рейтовских лекций — как прочитанных по радио, так и в книжной их версии — я признателен многим моим ученикам, коллегам и друзьям. Особенно хочу поблагодарить Алана Коуда, Риджейн Кэррион, Стивена Дейвиса, Хьюберта Дрейфуса, Уолтера Фримена, Барбару Хорен, Пола Кьюба, Карла Прибрама, Гюнтера Стента и Ванессу Ванг. Я также признателен радиокорпорации BBC. Джордж Фишер, возглавляющий отдел бесед, очень помог мне, а Джефф Диэн оказался превосходным продюсером. Но всего более я обязан моей жене Дагмар Сёрль, которой я и посвящаю эту книгу.

ОГЛАВЛЕНИЕ

Джон Сёрль:

Сознание, мозг и наука. Глава 1. Проблема отношения сознания и тела

Джон Сёрль:

Сознание, мозг и наука. Глава 2. Могут ли компьютеры мыслить?

Джон Сёрль:

Сознание, мозг и наука. Глава 3. Когнитивная наука

Джон Сёрль:

Сознание, мозг и наука. Глава 4. Структура действия

Джон Сёрль:

Сознание, мозг и наука. Глава 5. Перспективы социальных наук

Джон Сёрль:

Сознание, мозг и наука. Глава 6. Свобода воли

ГЛАВА 1.

ПРОБЛЕМА ОТНОШЕНИЯ СОЗНАНИЯ И ТЕЛА

На протяжении тысячелетий люди пытались понять своё место во Вселенной. В силу ряда причин современные философы не расположены заниматься трудными проблемами. Но проблемы от этого не исчезают, и в этой книге я хотел бы взяться за некоторые из них.

Самой трудной, на мой взгляд, является следующая проблема. У человеческих существ есть определённое представление о самих себе, составленное на основании здравого смысла, и это представление очень трудно сообразовать с «научным» представлением о физическом мире. Мы считаем себя сознательными, свободными, разумными существами, действующими в мире, который, как утверждает наука, состоит из лишённых сознания, бессмысленных физических частиц. Как же согласовать эти два представления? Возможно ли, чтобы мир не содержал ничего, кроме не обладающих сознанием физических частиц, и в то же время содержал сознание? Возможно ли, чтобы механическая Вселенная содержала интенциональные человеческие существа, то есть человеческие существа, которые способны формировать представления о мире? Короче говоря, возможно ли, чтобы бессмысленный мир содержал смыслы?

Эти проблемы могут звучать и более современно. Например, можно задать такие вопросы: как интерпретировать недавние результаты компьютерной науки и исследований по искусственному интеллекту, занимающихся конструированием разумных машин? В частности, можно ли считать, что цифровой компьютер есть изображение человеческого сознания? Почему социальные науки не дали результатов в отношении человека, сопоставимых с результатами естественных наук в отношении природы? Как соотносятся обыденные, здравые объяснения человеческого поведения и научные способы его объяснения?

В данной главе мне хотелось бы взяться за проблему, которую многие философы считают наиниболее трудной: каково отношение нашего сознания ко всей остальной Вселенной? Отношение сознания и тела относится к числу традиционных проблем. В современной формулировке это — проблема отношения сознания и мозга.

Думаю, что проблема отношения сознания и тела решается достаточно просто, и её решение согласуется и с нейрофизиологией, и со здравым смыслом. Но прежде чем предложить это решение, зададимся вопросом: почему проблема отношения сознания и тела считается такой трудной? Почему на протяжении столетий она не находит решения? Ведь решена же «проблема отношения пищеварения и желудка?» Почему феномен сознания столь загадочен?

Отчасти сложности проистекают из той настойчивости, с которой мы продолжаем пользоваться словарём XVII века. Ещё в студенческие времена меня поражала неадекватность выбора, предлагавшегося философией сознания: можно быть либо монистом, либо дуалистом. Если вы монист, то вы либо материалист, либо идеалист. Если вы материалист, то вы либо бихевиорист, либо физикалист. И так далее. В дальнейшем изложении я попытаюсь освободиться от этих устаревших категорий. Никто ведь не обязан выбирать между монизмом и дуализмом, решая «проблему отношения пищеварения и желудка». Почему же нужно поступать иначе, решая проблему отношения сознания и тела?

Но дело не только в терминологии, существует и сама по себе проблема, или скорее семейство проблем. Начиная с Декарта, проблема отношения сознания и тела формулируется следующим образом: как объяснить связь двух по видимости совершенно различных родов вещей? С одной стороны, имеются ментальные вещи, наши мысли и чувства; мы считаем их субъективными, принадлежащими сознанию и нематериальными. С другой стороны, имеются физические вещи; мы считаем, что они обладают массой, занимают место в пространстве и взаимодействуют с другими физическими вещами. Большинство решений проблемы отношения сознания и тела состоит в отрицании существования или по крайней мере в снижении статуса одного или другого из этих типов вещей. Учитывая успехи физических наук, неудивительно, что снижается статус ментальных реальностей (entities). Большинство модных (вплоть до недавнего времени), материалистических концепций сознания, таких как бихевиоризм, функционализм и физикализм, приходили к отрицанию, явному или неявному, существования таких вещей, как сознание. Иначе говоря, отрицалось, что человеку присущи субъективные, сознательные, ментальные состояния и что эти состояния столь же реальны и нередуцируемы, как и все остальные состояния, существующие во Вселенной.

Но почему высказываются такие мнения? Почему многие теоретики приходят к отрицанию того, что ментальные феномены в своей сущности несут ментальный характер? Если мы ответим на этот вопрос, то понятно

будет и то, почему отношение сознания и тела в течение столь длительного времени считалось наиболее трудной проблемой.

Ментальные явления обладают свойствами, из-за которых их невозможно вписать в наше «научное» представление о мире. Именно эти свойства и делают отношение сознания и тела столь трудной проблемой. Они так удивительны, что заставили многих мыслителей — философов, психологов и специалистов по искусственному интеллекту, говорить о сознании странные и неправдоподобные вещи.

Самым важным свойством является сознание (consciousness). Ия, когда пишу эти строки, и вы, когда их читаете, — оба мы обладаем сознанием. Очевидно, что мир содержит сознательные ментальные состояния и события. Трудно понять, однако, как чисто физические системы могут обладать таким свойством. Как эта бело-серая мешанина внутри моей черепной коробки может обладать сознанием?

Кажется удивительным само существование сознания. Но сознание — это главное в специфически человеческом существовании; без него невозможны все другие специфически человеческие аспекты нашего существования — язык, любовь, юмор и так далее. И разве это не скандальный факт, что современная философия и психология почти ничего не способны сказать о сознании?

Вторым трудным для понимания свойством является то, что называют интенциональностью. Это свойство, благодаря которому наши ментальные состояния «направлены на», «говорят о», «указывают на» объекты и состояния дел в мире. «Интенциональность» — это не просто намерения (intentions), но также и убеждения, надежды, страхи, любовь, ненависть, вожеление, отвращение, стыд, гордость, раздражение, радость и все те ментальные состояния (сознательные или бессознательные), которые имеют отношение к миру, или которые говорят о мире, существующем отдельно от сознания. Вопрос об интенциональности весьма напоминает вопрос о сознании. Как вещество внутри моей головы может быть о чём-либо? Как оно может отсылать к чему-либо? В конечном счёте оно состоит из «атомов и пустоты», как и вся остальная материальная реальность. Но каким образом эти атомы и пустота могут иметь представление о чём-либо?

Третьим свойством, которое, по-видимому, трудно совместить с научным представлением о реальности, является субъективность, ментальных состояний. Например, я могу ощущать свою боль, а вы — нет. Я вижу мир со своей точки зрения; вы видите его со своей точки зрения. Я знаю, что я и мои

внутренние ментальные состояния совершенно отличаются от «я» других людей и их ментальных состояний. Начиная с XVII века мы стали думать о реальности как о том, что равнодоступно всем её наблюдателям; иначе говоря, мы считаем, что реальность объективна. Но как примирить реальность субъективных ментальных феноменов с научным представлением об объективной реальности?

Наконец, четвёртой проблемой является природа ментального причинения (causality). С точки зрения здравого смысла, наши мысли и чувства определяют наше поведение, они оказывают каузальное воздействие на физический мир. Например, я хочу поднять руку и — смотрите-ка — моя рука поднимается. Но если наши мысли и чувства ментальны, как они могут воздействовать на что-либо физическое? Как ментальное может воздействовать на физическое? Верно ли, что наши мысли и чувства могут химически воздействовать на мозг и на всю нервную систему в целом? Как это возможно? Означает ли это, что мысли способны обвиваться вокруг аксонов, трясти дендриты или проникать сквозь стенки клеток и атаковать клеточное ядро?

Но если между сознанием и мозгом такой связи нет, то не следует ли из этого, что сознание есть нечто незначительное, вроде пены, образующейся при движении волны? Если бы пена обладала сознанием, она бы наверное думала: «как же это тяжело — весь день тянуть волны на берег, а потом обратно!» Но ведь пена здесь не при чём. С какой же стати нам считать, что ментальная жизнь чем-то отличается от пены на волне физической реальности?

Эти четыре свойства — сознательность, интенциональность, субъективность и ментальное причинение — и делают проблему отношения сознания и тела такой трудной. Однако всё это — реальные черты ментальной жизни. Не всякое ментальное состояние обладает сразу всеми этими свойствами. Но любая теория о сознании и об отношении сознания и тела должна их учитывать. Если отрицается хотя бы одно из этих свойств, это означает, что где-то допущена ошибка.

Мой первый тезис состоит в следующем.

Ментальные феномены, все ментальные феномены, сознательные или бессознательные, зрительные или слуховые, боли, щеготка, чесотка, мысли, вся наша ментальная жизнь, — все это причинно обусловлено процессами, происходящими в мозге.

Опишем каузальные процессы детально, взяв в качестве примера одно ментальное состояние — боль. Следующее поколение, наверное, очень

удивится тому, что мы сейчас скажем, ведь наше знание о работе мозга будет изменяться. Однако форма объяснения может оставаться правильной, даже если изменяются детали. С сегодняшней точки зрения, болевые сигналы передаются от сенсорных нервных окончаний к спинному мозгу посредством по крайней мере двух типов нервных волокон — волокон типа Дельта А, передающих колющие ощущения, и волокон типа С, передающих ощущения жжения и ноющие ощущения. В спинном мозге они проходят через область, называемую трактом Лиссауэра, и заканчиваются в нейронах. Поднимаясь по спинному мозгу, импульсы достигают головного мозга двумя различными способами: через колющую боль и через боль от жжения. Оба пути пролегают через таламус, но в дальнейшем колющая боль локализуется в большей степени в сомато-сенсорной коре, в то время как путь, которым передаётся боль от жжения, посылает импульсы не только вверх в кору, но и латерально, в гипоталамус и другие отделы у ствола головного мозга. Поэтому нам значительно легче локализовать колющее ощущение — например, мы точно знаем, где именно в кожу воткнулась игла. В то же время боль от жжения и ноющая боль более мучительны, они затрагивают большую часть нервной системы. При ощущении боли возбуждаются не только основные области головного мозга, особенно таламус, но и сомато-сенсорная кора.

Итак, мы должны усвоить следующее: болевые ощущения причинно обусловлены рядом событий, которые начинаются в нервных окончаниях и заканчиваются в таламусе и других областях головного мозга. В самом деле, что касается ощущений, то событий исключительно внутри центральной нервной системы вполне достаточно для того, чтобы вызвать боль: например, мы знаем о болевых иллюзиях у пациентов с ампутированными конечностями и о болях, которые вызываются воздействием на соответствующие, области головного мозга. Рискну предположить: что верно в отношении боли, то верно и в отношении всех ментальных феноменов. Грубо говоря, всю центральную нервную систему можно считать частью головного мозга, и тогда всё, что имеет значение для нашей ментальной жизни, все наши мысли и чувства причинно обусловлены процессами, происходящими внутри мозга. Важно то, что происходит в голове, а внешние или периферические стимулы значения не имеют. Аргументация здесь простая. Если произошли события вне центральной нервной системы, но при этом ничего не произошло в мозге, то не произойдёт и ментальных событий. Но если что-то произойдёт в мозге, то ментальные события произойдут и в отсутствие внешнего стимула. (Между прочим, именно этот принцип положен в основу хирургической анестезии, предотвращающей воздействие внешнего стимула на центральную нервную систему.)

Но если боль и другие ментальные феномены причинно обусловлены процессами в мозге, то хотелось бы знать: что же такое боли? что это такое на самом деле? Очевидный ответ состоит в том, что это неприятные виды ощущений. Однако ответ этот всё же неудовлетворителен, остаётся неясным, каким образом боль вписывается в наше общее представление о мире. Здесь требуется определённое разъяснение. К нашему первому утверждению (что боль и другие ментальные феномены причинно обусловлены процессами, происходящими в мозге) необходимо добавить второе. Боли и другие ментальные феномены являются свойствами мозга (и вероятно, всей остальной центральной нервной системы).

Одна из наиболее важных задач данной главы — показать, истинность и того и другого утверждения. Как возможно, чтобы мозг причинно обуславливал сознание, а сознание было лишь свойством, присущим мозгу? Думаю, что именно неспособность представить себе, что оба эти утверждения могут быть истинными одновременно, мешала решению проблемы отношения сознания и тела. Здесь могут возникнуть недоразумения разного уровня. Если ментальные и физические феномены находятся в отношении причины и действия, то как одно может быть свойством другого? Не означает ли это, что сознание причинно обусловило само себя, и не возвращаемся ли мы тогда к внушающему благоговейный ужас учению о *causa sui*? Но в конечном счёте наше замешательство вызвано непониманием того, что такое причинность. Принято думать, что всегда, когда А причинно вызывает В, имеют место два дискретных события, одно — причина, другое — действие, а всякое причинение подобно соударению бильярдных шаров. Эта грубая модель причинных отношений заставляет нас занять своего рода дуалистическую позицию; мы склоняемся к тому мнению, что события в материальной, «физической» области причинно обуславливают события в другой, не субстанциальной, «ментальной» области. Мне кажется, что это — ошибка. Исправить её можно, выдвинув более тонкую концепцию каузальности.

Отвлечемся на время от отношений между сознанием и мозгом и обратимся к другим видам каузальных отношений, существующих в природе. В физике общепринято различение микро- и макросвойств систем.

Рассмотрим, к примеру, письменный стол, за которым я сейчас сижу, или стакан воды на столе передо мной. Каждый объект состоит из микрочастиц. Микрочастицы обладают свойствами как на уровне молекул и атомов, так и на более глубоком субатомном уровне. Но объекты обладают также определёнными свойствами, например твёрдостью, текучестью, прозрачностью, которые являются поверхностными, или общими свойствами физических систем. Многие из таких поверхностных, или общих свойств

могут быть объяснены причинно, указанием на поведение элементов на микроуровне. Например, твёрдость стола объясняется решетчатой структурой молекул дерева. Текучесть воды объясняется характером взаимодействий между молекулами H_2O . Таким образом, макросвойства причинно объясняются поведением элементов на микроуровне.

Полагаю, что это даёт нам модель для объяснения загадочных отношений между сознанием и мозгом. В случае с текучестью, твёрдостью и прозрачностью нам нетрудно предположить, что поверхностные свойства причинно обусловлены поведением элементов на микроуровне, и в то же время мы допускаем, что поверхностные феномены как раз и являются свойствами самих рассматриваемых систем. Думаю, что яснее всего эту мысль можно выразить, сказав, что поверхностное свойство одновременно и причинно обусловлено поведением микроэлементов, и реализуется в системе, состоящей из микроэлементов. Поверхностные свойства представляют собой свойства более высокого уровня той системы, поведение которой на микроуровне и обуславливает эти свойства.

Кто-то возразит, что текучесть, твёрдость и тому подобная суть свойства микроструктуры. Так, например, мы можем определить твёрдость как решетчатую молекулярную структуру, то есть так же, как определяется теплота: теплота — это кинетическая энергия молекулярного движения. Само по себе это правильно. Но в чем же состоит возражение? Для научного прогресса весьма характерно, что определения в терминах поверхностных свойств, свойств, доступных чувствам, в дальнейшем находят выражение в терминах микроструктуры, причинно обуславливающей эти поверхностные свойства. Так, стол передо мной твёрд в обыденном смысле слова: ему присуща жёсткость, он сопротивляется определённым воздействиям, выдерживает вес книг, он непроницаем для большинства других объектов, таких как другие столы и так далее. Таково понятие твёрдости с точки зрения здравого смысла. С научной точки зрения, твёрдость — это микроструктура, причинно обуславливающая наблюдаемые макросвойства. Так что можно было бы сказать, что твёрдость — это решетчатая структура системы молекул и что определённая таким образом твёрдость причинно обуславливает, например, сопротивление воздействию и давлению. Или же можно было бы сказать, что твёрдость — это свойства высшего уровня, такие как жесткость и сопротивление, и что она причинно обусловлена поведением элементов на микроуровне.

С учётом этого несложно объяснить отношение между сознанием и мозгом в терминах мозговой деятельности. Подобно тому как свойство текучести причинно обусловлено поведением элементов на микроуровне и в то же время

реализуется в системе микроэлементов, ментальные феномены причинно обусловлены процессами, происходящими в мозге на нейронном или модульном уровне, и одновременно они реализуются в системе, состоящей из нейронов. И точно так же, как мы нуждаемся в различении «микро» и «макро» уровней в отношении любой физической системы, мы нуждаемся в этом различении в отношении мозга. Хотя мы и можем сказать о системе частиц, что её температура 10 градусов по Цельсию или что она твёрдая или жидкая, мы не можем сказать о любой данной частице, что она твёрдая или жидкая, или что её температура 10 градусов по Цельсию. Невозможно вытащить молекулу из стакана воды и сказать; «Эта молекула мокрая».

Подобно этому, хотя мы и можем сказать о каком-то конкретном мозге: «этот мозг обладает сознанием» или «этот мозг испытывает жажду или боль», мы не можем сказать о каком-либо конкретном нейроне в мозге: «этот нейрон испытывает боль» или «этот нейрон испытывает жажду». И хотя работа мозга остаётся загадкой, не существует никаких логических, философских или метафизических препятствий, которые мешали бы чтобы объяснять отношение между сознанием и мозгом так же, как мы объясняем все остальные вещи в природе. Здесь нет ничего необычного — одновременно быть и следствием какой-то микроструктуры и её реализацией, а именно таково отношение между сознанием и мозгом.

Вернёмся теперь к тем четырём проблемам, которые возникают, когда мы пытаемся объяснить отношение между сознанием и телом.

Во-первых, как возможно сознание?

Лучший способ показать, как что-либо возможно, — показать, как это существует в действительности. Мы уже говорили о том, как нейрофизиологические процессы, происходящие в таламусе и коре, причинно обуславливают боль. Но многие люди таким ответом не удовлетворяются. Полагаю, что это чувство разочарования можно несколько рассеять. В течение длительного времени многие биологи и философы считали, что невозможно в принципе объяснить существование жизни, исходя из чисто биологических оснований.

Они полагали, что биологические процессы должны быть дополнены другим необходимым элементом: должен быть постулирован некий *elan vital*, придающий жизнь тому, что в противном случае осталось бы мёртвой и инертной материей. Предыдущее поколение ещё помнит всю серьёзность спора между витализмом и механицизмом, однако сегодня он уже не вызывает никакого интереса. Почему? Думаю, что дело не в победе, одержанной

механицизмом; просто мы пришли к более глубокому пониманию процессов, происходящих в живых организмах. Объяснение, которое мы даем свойствам живых существ, кажется, разрешает загадку жизни. Думаю, что такие же соображения применимы и к сознанию, Что загадочного в том, что это вещество, эта серо-белая, похожая на овсянку субстанция мозга обладает сознанием? Это не более загадочно, чем существование совокупности нуклеопротеиновых молекул, заключённых в кальциевую оболочку. Короче говоря, решение загадки — в познании процессов. Мы ещё не понимаем вполне эти процессы, но общий их характер ясен: в нейронах или нейронных модулях происходят определённые электрохимические процессы, и эти процессы являются причиной сознания.

Вторая проблема заключалась в том, как могут «атомы и пустота» обладать интенциональностью. Как они могут быть о чём-то?

Рассмотрим чувство жажды. Как известно, по крайней мере некоторые виды жажды причинно обусловлены рядом нервных возбуждений в гипоталамусе. Эти возбуждения, в свою очередь, причинно обусловлены действием ангиотензина, а ангиотензин синтезируется из ренина, секреторно выделяемого почками. Жажда, по крайней мере в этих своих разновидностях, причинно обусловлена рядом событий в центральной нервной системе, главным образом в гипоталамусе, и она реализуется в гипоталамусе. Жажда — это, помимо прочих вещей, желание выпить воды. Поэтому жажда является интенциональным состоянием: у жажды есть содержание, и это содержание определяет, при каких условиях она утоляется; жажда обладает также всеми прочими свойствами, которые присущи интенциональным состояниям.

Разгадать тайну интенциональности можно, описав в деталях причинную обусловленность феноменов биологическими процессами, — феноменов, которые в то же время реализуются в биологических системах. Зрительные и слуховые ощущения, тактильные ощущения, голод, жажда, сексуальное влечение, — все это причинно обусловлено мозговыми процессами и реализовано в структуре мозга, и Всё это — интенциональные феномены.

Дело не в том, что мы должны избавиться от ощущения загадочности природы. Напротив, приведённые мной примеры в каком-то смысле загадочны. Но они не более и не менее загадочны, чем все остальные свойства мира, такие как сила тяготения, процесс фотосинтеза или величина Млечного Пути.

Третья проблема; как сообразовать субъективный характер ментальных состояний с объективностью реального мира?

Мне кажется, ошибочно считать, что реальность исключает субъективность. Если «наука» есть название для совокупности объективных и систематических истин, то существование субъективности есть такой же объективный научный факт, как и все остальные научные факты. В научное описание входит и субъективность ментальных состояний. Очевидно, что биологическая эволюция породила биологические системы определённого вида, а именно — мозг человека и некоторых животных, обладающий свойством субъективности. Состояние моего сознания в настоящий момент есть свойство моего мозга, но моё сознание доступно мне и недоступно вам. А состояние вашего сознания есть свойство вашего мозга, и ваше сознание доступно вам и недоступно мне. Итак, субъективность есть объективный биологический факт. Часто «науку» определяют через существующие научные теории. Но это — провинциальный предрассудок. Любая область фактов может стать предметом систематического исследования. Так, например, если бы Бог существовал, то его существование было бы таким же фактом, как и все другие факты. Я не знаю, существует ли Бог, но не сомневаюсь в существовании субъективных ментальных состояний, потому что нахожусь сейчас в одном из них, подобно тому как вы находитесь в другом. Если факт субъективности не согласуется с определением «науки», то мы должны отбросить не факт, а это определение.

В-четвёртых, существует проблема ментального причинения. Следует объяснить, как ментальные события причинно обуславливают физические события. Каким образом, например, нечто столь «невесомое» и «эфирное», как мысль, порождает действие?

Ответ состоит в том, что мысли не являются невесомыми и эфирными. Когда происходит мышление, происходит мозговая деятельность. Деятельность мозга причинно обуславливает, через физиологические процессы, телесные движения. Далее, поскольку ментальные состояния являются свойствами мозга, они описываются на двух уровнях — высшем, в ментальных терминах, и низшем, в терминах физиологических. Одни и те же причины можно описать на любом из этих уровней.

Чтобы проиллюстрировать эти отношения, возьмём пример из области физики, а именно забивание гвоздя с помощью молотка. И молотку, и гвоздю присуща определённая твёрдость. Молотки из ваты или масла бесполезны, а молотки из воды или пара — вообще не молотки. Твёрдость есть реальное свойство молотка, его причина. Но твёрдость сама причинно обусловлена поведением частиц на микроуровне, и она реализуется в системе, состоящей из микроэлементов. Существование двух уровней описания мозга,

макроуровня ментальных процессов и микроуровня нейронных процессов, аналогично существованию двух уровней описания молотка. Сознание, например, есть свойство мозга, способное причинно обуславливать происходящие вещи. На высшем уровне моя сознательная попытка поднять руку причинно обуславливает её движение. На низшем уровне возбуждение нейронов даёт начало цепи событий, приводящих к сокращению мышц. Как и в случае с забиванием гвоздя, одна и та же последовательность событий имеет два уровня описания. Оба они реальны как причины, и причинные свойства высшего уровня и причинно обусловлены, и реализуются в структуре низшего уровня.

Подведём итоги, Сознание и тело взаимодействуют, но это не две различные вещи, поскольку ментальные феномены суть просто свойства мозга. Можно рассмотреть эту позицию с точки зрения и физикализма, и ментализма. Согласно «наивному физикализму», в мире существуют только физические частицы с их свойствами и отношениями. Влияние физической модели настолько велико, что трудно поставить наивный физикализм под сомнение. По «наивному ментализму» ментальные феномены существуют в реальности. Некоторые из ментальных состояний являются сознательными; многие из них интенциональны; все они субъективны; и многие каузальны и детерминируют физические события, происходящие в мире. Таким образом, тезис этой первой главы можно сформулировать достаточно просто. Наивный ментализм и наивный физикализм прекрасно согласуются друг с другом. И все свидетельствует не только о согласии, но и об истинности того и другого.

ГЛАВА 2.

МОГУТ ЛИ КОМПЬЮТЕРЫ МЫСЛИТЬ?

В первой главе я обрисовал в общих чертах решение проблемы отношения сознания и тела. Хотя мы и не знаем в подробностях, как функционирует мозг, у нас достаточно знаний, чтобы составить общее представление об отношениях между мозговыми и ментальными процессами. Ментальные процессы причинно обусловлены деятельностью элементов мозга. В то же время они реализуются в структуре, состоящей из этих элементов. Полагаю, что это согласуется со стандартными подходами к биологическим феноменам. Вообще говоря, это — представление, основанное на здравом смысле. Однако такую точку зрения разделяет меньшинство. В философии, психологии и исследованиях по искусственному интеллекту превалирует взгляд, подчёркивающий аналогии между функционированием человеческого мозга и цифровых компьютеров. Согласно крайней версии этого взгляда, мозг является цифровым компьютером, а сознание — компьютерной программой. Можно суммировать данный взгляд, — я называю его «сильной версией искусственного интеллекта», или «сильным ИИ», — сказав, что сознание относится к мозгу так же, как программа — к аппаратным средствам компьютера.

Следовательно, в человеческом сознании нет ничего биологического. Мозг оказывается одним из видов компьютерной аппаратуры, способной выполнять программы, которые составляют человеческий интеллект. Согласно этому взгляду, сознанием обладает любая физическая система с правильной программой и правильными устройствами ввода и считывания данных. Например, вы можете соорудить компьютер из пустых пивных банок, а в качестве источника энергии выбрать ветряные мельницы. Если в компьютер заложена правильная программа, то он обладает сознанием. И дело не в том, что он мог бы обладать мыслями и чувствами. Он должен ими обладать, потому что для этого требуется только одно: выполнение правильной программы.

Считается, что мы ещё не разработали программу, которая является сознанием. Но при этом все, кто придерживается сильной версии ИИ, соглашаются с тем, что это лишь вопрос времени. Специалисты по компьютерам и искусственному интеллекту рано или поздно построят компьютерную аппаратуру и программы, эквивалентные человеческому мозгу

и сознанию. Это будут искусственный мозг и искусственное сознание, во всём равные человеческому мозгу и человеческому сознанию.

Люди, не знакомые с областью исследований по искусственному интеллекту, очень удивляются, когда узнают, что кто-то вообще придерживается такого взгляда. Поэтому, прежде чем подвергнуть его критическому анализу, приведу высказывания учёных, которые занимались этим предметом. Герберт Саймон из университета Карнеги-Меллона утверждает, что у нас уже есть машины, которые в буквальном смысле мыслят. Речь идёт не о какой-то будущей машине — уже существующие цифровые компьютеры мыслят точно так же, как и мы с вами. Представляете? На протяжении столетий философов волновало, может машина мыслить или нет, а сейчас обнаруживается, что в Карнеги-Меллоне мыслящие машины уже имеются. Коллега Саймона Алан Ньюэлл заявляет: мы открыли (заметьте, что он говорит «открыли», а не «предположили» или «сочли вероятным»), что интеллект есть всего лишь манипуляция физическими символами и не связан существенным образом с какой-либо особой биологической или физической аппаратурой. Интеллектом обладает любая система, способная правильно манипулировать физическими символами. К чести Саймона и Ньюэлла, оба они подчёркивают, что в их заявлениях нет ничего метафорического: смысл этих заявлений буквален.

Далее, Фримен Дайсон утверждает, что если говорить о дальнейшей эволюции, то компьютеры имеют преимущество перед людьми. Поскольку сознание — это формальные процессы, в компьютерах они могут выполняться на таких субстанциях, которые имеют гораздо больше шансов выжить, чем мы — существа, состоящие из влажных и мягких материалов. Марвин Мински из Массачусетского технологического института утверждает, что следующее поколение компьютеров будет настолько разумным, что нам «повезёт, если они захотят оставить нас при себе в качестве домашних животных». Но наилучшим преувеличением являются слова Джона Маккарти, изобретателя самого термина «искусственный интеллект». С его точки зрения, даже «столь простые машины, как термостаты, придерживаются определённых убеждений». Убеждения есть практически у любой машины, способной решать проблемы. Я восхищаюсь мужеством Маккарти. Однажды я спросил его: «Какие же именно убеждения есть у вашего термостата?» Он ответил: «У моего термостата три убеждения: здесь слишком жарко, здесь слишком холодно, здесь нормальная температура». Как философу, эти тезисы нравятся мне по одной простой причине: им присуща ясность и они допускают простое и решающее опровержение. Именно такое опровержение я и собираюсь изложить в этой главе.

Опровержение это не имеет в виду ту или иную стадию развития компьютерной технологии. Мы склонны полагать, что решение всех наших проблем заключено в каком-то ещё не случившемся технологическом чуде. Наше опровержение не имеет ничего общего с такого рода аргументацией. Оно касается самого определения цифрового компьютера.

Для нашей концепции цифрового компьютера существенно важно, чтобы его операции можно было определять чисто формальным образом: иначе говоря, мы определяем шаги в терминах абстрактных символов — например, в виде последовательностей нулей и единиц. Типичное компьютерное «правило» определяет, что когда машина находится в определённом состоянии и на ленте напечатан определённый символ, она выполняет некую операцию, например стирает символ или печатает другой символ; затем она переходит к другому состоянию, например передвигает ленту на один квадрат влево. Но у символов нет семантического содержания, и они должны быть конкретизированы исключительно в терминах своей формальной или синтаксической структуры. Например, нули и единицы — всего лишь символы чисел; они даже не являются заменой для чисел. И именно это свойство цифровых компьютеров делает их столь мощными инструментами. Одна и та же компьютерная аппаратура может быть непользована для выполнения бесконечного числа различных программ. А одна и та же программа может выполняться на различных типах компьютерной аппаратуры.

Но это свойство программ — их чисто формальный или синтаксический характер — фатально для взгляда, согласно которому ментальные и программные процессы тождественны друг другу. Объяснить это можно очень просто. Сознание есть нечто большее, чем формальные или синтаксические процессы. Наши внутренние ментальные состояния, по определению, обладают содержанием. Если я думаю о Канзас-Сити, или не прочь был бы выпить холодного пива, или интересуюсь, не упадёт ли процентная ставка, — в каждом из этих случаев моё ментальное состояние обладает определённым содержанием в дополнение к любым формальным свойствам. Иначе говоря, даже если мои мысли представляются мне в виде цепочек из символов, в самом мышлении должно быть нечто большее, потому что цепочки сами по себе не обладают никаким смыслом. Если мои мысли являются мыслями о чём-либо, то и цепочки должны иметь смысл, и это делает их мыслями об этих вещах. Словом, в сознании есть не только синтаксис, но и семантика. Компьютерная программа не может стать сознанием, если это чисто синтаксическая программа. Сознание же семанлично в том смысле, что оно обладает не только формальной структурой, но и содержанием.

Проведем следующий мысленный эксперимент. Представим себе, что группа программистов составила программу, позволяющую компьютеру имитировать понимание китайского языка. Если компьютеру задаётся вопрос по-китайски, он сверяет его со своей памятью, или базой данных, а затем выдаёт подходящие ответы, тоже по-китайски. Допустим, ответы компьютера оказываются не хуже ответов настоящего китайца. Можно ли сказать, что компьютер понимает по-китайски не хуже самих китайцев? Вообразите, что вы находитесь в комнате, где расставлено несколько корзинок с китайскими иероглифами. Вы не понимаете по-китайски ни слова, но вам дали книжку на английском языке, содержащую правила манипулирования этими иероглифами. Это чисто формальные правила, касающиеся синтаксиса, а не семантики. Например: «Возьми знак фигли-мигли из корзины № 1 и положи его рядом со знаком фогли-могли из корзины № 2». Предположим теперь, что в комнату поступили другие китайские иероглифы, а вас снабдили новыми правилами, предписывающими, как возвращать эти иероглифы. Предположим также, что находящиеся снаружи люди называют неизвестные вам иероглифы вопросами, а те, что вы возвращаете назад, — ответами. Допустим также, что программисты поднаторели в составлении программ, а вы — в манипулировании символами. Ваши ответы практически неотличимы от ответов настоящих китайцев. Итак, вы заперты в комнате, перекладываете иероглифы с места на место и посылаете одни в ответ на другие. В этой ситуации нет никакой возможности изучить китайский язык, дело сводится к простому манипулированию этими чисто формальными символами.

Смысл примера состоит в следующем: выполняя формальную компьютерную программу, вы, с точки зрения внешнего наблюдателя, ведёте себя так, как будто понимаете по-китайски, однако на самом деле вы не понимаете по-китайски ни единого слова. Но если выполнения компьютерной программы недостаточно для того, чтобы вы стали понимать по-китайски, то этого недостаточно и для того, чтобы его стал понимать любой другой цифровой компьютер. Объясняется это очень просто. Если вы не понимаете по-китайски, то и никакой другой компьютер не сможет понять этот язык, потому что у цифрового компьютера нет ничего такого, чего бы не было у вас. Всё, что у него имеется, — это формальная программа для манипулирования неинтерпретированными иероглифами. Повторим, у компьютера есть синтаксис, но нет семантики. Суть притчи проста. Понимать язык, или иметь ментальные состояния — значит иметь нечто большее, чем просто совокупность формальных символов. Предполагается, что у вас есть интерпретация, или смысл, которым наделены символы, А у цифрового компьютера не может быть ничего, кроме формальных символов, потому что работа компьютера, как я уже говорил, состоит в выполнении программ. Эти программы чисто формальны и не имеют семантического содержания.

Силу этого аргумента можно оценить, если сопоставить вопросы и ответы на английском языке с вопросами и ответами на языке, слова которого для нас не имеют никакого смысла. Вообразите, что в китайской комнате вам также задают вопросы по-английски: сколько вам лет, или о вашей жизни, — а вы на них отвечаете. В чём же отличие «китайского» и «английского» случаев? Если вы не понимаете по-китайски, но понимаете по-английски, тогда отличие очевидно. Вы понимаете вопросы, заданные по-английски, потому что они выражены в символах, смыслы которых вам известны. Когда вы на них отвечаете, вы производите осмысленные символы. Ничего подобного нет в «китайском» случае. Тут вы просто манипулируете формальными символами в соответствии с компьютерной программой и не наделяете их никаким смыслом.

Ответы на это возражение были предложены как исследователями в области искусственного интеллекта и психологами, так и философами. Но в них есть нечто общее: все они неадекватны. Причина неадекватности очевидна: одного синтаксиса недостаточно, а цифровые компьютеры имеют, по определению, один только синтаксис.

Рассмотрим аргументы, выдвигаемые против моей позиции 1.

Первое возражение заключается в том, что в примере с китайской комнатой китайский язык понимает вся система в целом. И хотя я, находясь в комнате и манипулируя символами, не понимаю китайского языка, однако представляю собой центральную часть компьютерной системы. Вся система в целом, включая комнату, корзинки с символами, папки с программами и вероятно с чем-то ещё, — все это вместе взятое понимает китайский язык. Но против этого аргумента можно выдвинуть такое же возражение, как и раньше: не существует способа, который позволил бы системе перейти от синтаксиса к семантике. Даже я, центральный процессор, не способен выяснить, что означает каждый из этих символов; но тогда это не способна сделать и вся система в целом.

Другое характерное возражение, выдвигаемое против моей позиции, заключается в следующем. Вообразите, что в робота вложена программа понимания китайского языка. Если робот движется и взаимодействует с миром, то не свидетельствует ли это о том, что он понимает по-китайски? Однако неустранимость различий между семантикой и синтаксисом сводит на нет и этот способ аргументации. До тех пор, пока у робота вместо мозга компьютер, он не может перейти от синтаксиса китайского языка к его семантике. Вообразите, что я — компьютер. Находясь в комнате и

располагаясь в голове робота, я передвигаю с места на место символы и не знаю, что одни символы поступают ко мне от телевизионных камер, прикрепленных к голове робота, а другие — выходят и управляют движением его рук и ног. И до тех пор, пока все, с чем я имею дело, есть формальная компьютерная программа, я не способен наделять смыслом какой-либо из символов. Тот факт, что робот взаимодействует с внешним миром, не поможет мне в наделении символов смыслом, пока у меня не будет способа получить информацию об этом факте. Представим, что робот выбирает гамбургер, и вследствие этого символ, обозначающий гамбургеры, появляется в комнате. Поскольку всё, что у меня есть, — это символ без знания его причин или того, как он сюда попал, установить его значение невозможно. Причинные взаимодействия между роботом и остальным миром не играют роли, если только они не представлены в том или ином сознании. Но их и нельзя представить, если всё, что составляет так называемое сознание, есть множество чисто формальных, синтаксических операций.

Важно точно понять, что утверждает и чего не утверждает мой аргумент. Зададимся старым вопросом: «Может ли машина мыслить?» В некотором смысле все мы машины. Мы можем считать, что вещество в наших головах есть машина. И разумеется, все мы способны мыслить. Так что в одном из смыслов слова «машина», а именно в том, что машина — всего лишь физическая система, способная выполнять определённые операции, мы все машины и мы можем мыслить. Итак, существуют машины, которые могут мыслить. Но ведь нас занимал вовсе не этот вопрос. Сформулируем его иначе. Может ли мыслить артефакт? Может ли мыслить машина, сотворённая человеком? Конечно, это зависит от того, о каком артефакте идёт речь. Допустим, мы сконструировали машину, которая ни на одну молекулу не отличается от человеческого существа. Тогда, если вы можете дублировать причины, вы, вероятно, сможете дублировать и следствия. Ответ на поставленный вопрос, по крайней мере в принципе, тривиален: «да». Если бы вы построили машину, имеющую ту же структуру, что и человеческое существо, то эта машина была бы способна мыслить. Конечно, это был бы суррогат. Но сделаем ещё одну попытку.

Вопрос состоит не в том, может ли машина мыслить, или может ли артефакт мыслить, а в том, может ли мыслить цифровой компьютер. Но мы должны быть очень осторожны с интерпретацией. С математической точки зрения, цифровым компьютером может быть названо всё что угодно. Всё что угодно может быть описано как пример или выполнение компьютерной программы. Даже ручка, лежащая передо мной на столе, может быть названа цифровым компьютером. Хотя компьютерная программа у неё весьма и весьма скучная. Программа говорит: «Оставайся на месте». Поскольку в этом

смысле всё что угодно является цифровым компьютером, мы вновь получаем тривиальный ответ на наш вопрос. Конечно, наш мозг является цифровым компьютером, поскольку он выполняет любое количество компьютерных программ. И конечно, наш мозг способен мыслить. Итак, мы опять получили тривиальный ответ. Но на самом деле вопрос был задан неверно. Его следовало задать так: «Может ли цифровой компьютер мыслить?» Или иначе; «Достаточно ли для мышления выполнения правильной компьютерной программы с правильными данными на входе и выходе, конституирует ли это мышление?» На этот вопрос следует ответить: «нет». «Нет», потому что компьютерная программа определяется чисто синтаксически. Но мышление есть нечто большее, чем манипулирование не имеющими смысла символами, оно предполагает осмысленные семантические содержания. Эти семантические содержания я и имею в виду, когда говорю о «смысле».

Ещё раз следует подчеркнуть, что речь идёт не о той или иной конкретной стадии развития компьютерной технологии. Будущие открытия ничего не изменят. Дело не в грядущих удивительных достижениях компьютерной науки, не в различении последовательных и параллельных процессов, не в объёме программ, не в скорости компьютерных операций и не в компьютерах, способных взаимодействовать со средой, и даже не в изобретении роботов. Конечно, развитие компьютеров является весьма значительным шагом вперёд. И в будущем мы вправе ожидать дальнейшего прогресса в этой области. В будущем мы сможем имитировать человеческое поведение на компьютерах гораздо лучше, чем сегодня, и уж наверное лучше, чем вчера. Но с моей точки зрения, все эти имитации просто не имеют отношения к делу. Неважно, насколько хороша технология или какова скорость вычислений, производимых компьютером. В компьютере операции определяются синтаксически, в то время как сознание, мысли, чувства, эмоции и тому подобное предполагают нечто большее, чем синтаксис. Эти свойства компьютер не способен воспроизвести по определению, как бы хорошо он их ни имитировал. Существует большое различие между воспроизведением и имитацией. Никакая имитация сама по себе не может быть воспроизведением.

До сих пор я просто подводил фундамент под то ощущение абсурдности, которое вызвали цитаты, приведённые в начале главы. Остаётся один только вопрос: как кому-то вообще пришла в голову мысль, что компьютеры могут мыслить, чувствовать, обладать эмоциями и так далее? В конце концов, мы можем осуществить компьютерную имитацию любого процесса, поддающегося формальному описанию. Так, мы можем имитировать движение денег в британской экономике, или разделение власти в лейбористской партии. Мы способны имитировать грозы в прилегающих к Лондону графствах или пожары на складах восточного района Лондона. В

любом из этих случаев ни у кого не возникает и мысли, что компьютерная имитация — это что-то, что происходит на самом деле; никто не ожидает, что компьютерная имитация грозы оставит всех нас насквозь промокшими, а имитация пожара превратит в пепелище наш дом. Никто, будучи в здравом уме, не примет компьютерную имитацию ментальных процессов за действительные ментальные процессы. Откровенно говоря, сама идея кажется мне совершенно безумной. Но стоит всё же немного поразмышлять на эту тему.

Во-первых, в отношении сознания многие люди до сих пор придерживаются позиции бихевиоризма. Они полагают, что если система ведёт себя так, как если бы она понимала китайский язык, то она и в самом деле его понимает. Но мы уже опровергли эту форму бихевиоризма, прибегнув к помощи аргумента «китайской комнаты». Согласно другому предположению, сознание не является частью биологического мира, частью мира природы. Точка зрения «сильной версии ИИ» состоит в том, что сознание — это чисто формальный процесс и не может считаться конкретным продуктом биологических процессов. Короче говоря, это нечто вроде останков дуализма. Приверженцы ИИ убеждены, что сознание — не просто часть природного биологического мира; они считают, что оно полностью формализуемо. Парадокс заключается в том, что литература по ИИ полна инвектив по поводу дуализма, но фактически сильная версия ИИ сама основывается на своего рода дуализме: за сознанием отрицается его принадлежность к обычным биологическим феноменам.

Сведём вместе тезисы первой и второй главы. Если это сделать, получается весьма впечатляющая картина. Аргументация имеет очень простую логическую структуру, так что вы сами сможете увидеть, правильна она или нет. Первая посылка такова:

1. Мозг причинно обуславливает сознание.

Разумеется, это несколько грубая формулировка. Имеется в виду, что ментальные процессы, составляющие сознание, причинно обусловлены, полностью причинно обусловлены мозговыми процессами. Но мы будем говорить кратко: мозг причинно обуславливает психику. Это просто факт, говорящий о том, как устроен мир.

Теперь запишем второе суждение:

2. Синтаксиса недостаточно для семантики.

Это суждение является чисто понятийной истиной. В нём различено формальное и обладающее содержанием. К этим двум суждениям — что мозг причинно обуславливает психику и что синтаксиса недостаточно для семантики — добавим третье и четвёртое:

Компьютерные программы полностью определяются своей формальной или синтаксической структурой.

Это суждение истинно по определению, будучи частью того, что мы имеем в виду под компьютерной программой.

Сознание обладает ментальными содержаниями, в частности — содержаниями семантическими.

По-моему, это просто очевидный факт. Мои мысли, убеждения и желания всегда говорят о чём-то, они указывают на что-то, они касаются положения дел в мире; и это происходит потому, что их содержание направлено на это положение дел. Из указанных посылок можно вывести наше первое заключение; оно с очевидностью следует из посылок 2, 3 и 4.

Заключение 1:

Компьютерной программы самой по себе недостаточно, чтобы у системы появилось сознание. Короче говоря, программы — это не сознание, и самих по себе программ недостаточно для того, чтобы появилось сознание.

Это очень сильное заключение, оно означает, что проект создания сознания исключительно с помощью программ обречён с самого начала. Важно ещё раз подчеркнуть, что дело не в конкретном состоянии технологии или степени сложности программы. Это заключение — чисто формальный, или логический результат, полученный из множества аксиом, которые принимают все (или почти все) участники соответствующей дискуссии. Даже большинство энтузиастов искусственного интеллекта согласны, что мозговые процессы причинно обуславливают ментальные состояния, а программы определяются чисто формальным образом. Но если сложить вместе эти заключения и некоторые другие известные нам вещи, мы увидим, что проект сильной версии ИИ неосуществим. Посмотрим, что ещё можно вывести из этих аксиом.

Заключение 2:

Причинная обусловленность сознания деятельностью мозга не может определяться выполнением компьютерной программы.

Это второе заключение следует из соединения первой посылки с нашим первым заключением. Иначе говоря, из того, что мозг причинно обуславливает сознание и что компьютерных программ для этого

недостаточно, следует, что мозг порождает сознание не благодаря выполнению некоей компьютерной программы. Это тоже важный результат, из него следует, что мозг не является цифровым компьютером, или по крайней мере не является только цифровым компьютером. В тривиальном смысле слова цифровым компьютером может быть названо всё что угодно, и мозг не составляет здесь исключения. Но значимость данного заключения в том, что вычислительных свойств мозга просто недостаточно для того, чтобы можно было объяснить появление ментальных состояний. И это в любом случае должно казаться нам здравым научным заключением, оно напоминает о том, что мозг является биологической реальностью; биология мозга здесь имеет значение. И неправы те специалисты по искусственному интеллекту, которые считают несущественным то, что сознание реализуется в человеческом мозге.

Из нашей первой посылки мы можем вывести ещё одно заключение.

Заключение 3:

Чем бы ни было обусловлено сознание, оно должно обладать каузальной силой, по крайней мере эквивалентной силе мозга.

Это третье заключение есть тривиальное следствие из первой посылки. Вывод аналогичен следующему рассуждению: поскольку карбюраторный двигатель обеспечивает скорость в семьдесят пять миль в час, то и любой дизельный двигатель, способный на это, обладает мощностью, по крайней мере эквивалентной мощности карбюраторного двигателя. Наверное, могла бы существовать и какая-то другая система, которая бы обуславливала ментальные процессы, обладая химическими или биохимическими свойствами, полностью отличными от тех, что имеются в мозге. Может оказаться, что на других планетах или в других солнечных системах есть существа с совершенно другой биохимией, но обладающие ментальными состояниями. Предположим, на Землю высадились марсиане, и мы пришли к выводу, что они обладают ментальными состояниями. Но предположим также, что при вскрытии их голов там обнаружилась одна только Зелёная слизь. Эта Зелёная слизь, порождающая сознание и всю в целом ментальную жизнь, должна иметь причинную силу, равную силе человеческого мозга. Итак, из нашего первого заключения, что одних программ недостаточно, и нашего третьего заключения, что любая другая система должна была бы обладать причинной силой, равной силе мозга, следует заключение 4.

Заключение 4:

Для любого артефакта с ментальными состояниями, эквивалентными человеческим, недостаточно одной лишь компьютерной программы. Артефакт должен обладать силой, равной силе человеческого мозга.

Это обсуждение, я полагаю, напоминает нам о том, что мы всегда знали, а именно что ментальные состояния суть биологические феномены. Сознание, интенциональность, субъективность и ментальная причинность — все это части биологической истории нашей жизни, наряду с ростом, воспроизводством, выделением желчи и пищеварением.

Примечания:

1. Основные комментарии учёных по поводу мысленного эксперимента с «китайской комнатой» систематизированы в статье: Searle J. Minds, brains and programs. // Artificial Intelligence. — L., 1987, pp. 18–40.

ГЛАВА 3.

КОГНИТИВНАЯ НАУКА

Мы не колеблясь говорим такие, например, вещи: «Бэзил голосовал за консерваторов, потому что Тэтчер повела себя наилучшим образом в ситуации с Фолклендами». Однако мы не знаем, как относиться к таким, например, заявлениям: «Бэзил голосовал за консерваторов, потому что таким было состояние его гипоталамуса», Иначе говоря, мы привыкли объяснять поведение людей с точки зрения их ментальности, их желаний, страхов, надежд и так далее. Существует и нейрофизиологическое объяснение человеческого поведения, которое формулируется в терминах мозговых процессов. Сложность в том, что объяснение первого типа работает на практике, но не является научным. Объяснение второго типа является научным, но у нас нет ни малейшей идеи, как сделать так, чтобы оно работало на практике.

Между мозгом и сознанием продолжает зиять брешь. Значительные интеллектуальные усилия XX века были направлены на заполнение этой бреши, на разработку науки о человеческом поведении, которая не была бы ветхозаветной психологией и не основывалась на одном только здравом смысле, но в то же время не сводилась бы к нейрофизиологии. Вплоть до настоящего времени все попытки в этом направлении не приносили успеха. Самой яркой неудачей был, конечно, бихевиоризм, но за свою жизнь я наслушался претенциозных заявлений и от лица теории игр, кибернетики, теории информации, структурализма, социобиологии и так далее; все эти дисциплины также принесли одно только разочарование. Предваряя дальнейшее изложение, я хотел бы сказать, что усилия «заполнить брешь» безуспешны потому, что нет самой этой бреши.

Последние попытки опираются на аналогию между человеческими существами и цифровыми компьютерами. В соответствии с крайней точкой зрения, которую я называю сильной версией искусственного интеллекта, или просто сильной версией ИИ, мозг является цифровым компьютером, а сознание — компьютерной программой. Именно этот взгляд я и опровергал в предыдущей главе. Его часто называют «когнитивизмом», поскольку попытки «заполнить брешь» исходят из работ в области когнитивной психологии и искусственного интеллекта и образуют основное направление новой дисциплины — «когнитивной науки». Подобно «сильной версии ИИ», когнитивизм считает компьютер адекватным изображением сознания, а не

просто метафорой. Но, в отличие от «сильной версии ИИ», когнитивизм не заявляет или по крайней мере не должен заявлять, что компьютеры мыслят и чувствуют в буквальном смысле слова.

Суммируем исследовательскую программу когнитивизма. Мышление есть обработка информации, но обработка информации — не что иное, как манипулирование символами. Компьютеры осуществляют манипулирование символами. Поэтому наилучшим путём изучения мышления (или, как они предпочитают его называть, «познавания», «cognition») является изучение вычислительных символа-манипуляционных программ, неважно где — в компьютерах или в мозге. Задачей когнитивной науки является изучение мозга не на уровне нервных клеток и не на уровне сознательных ментальных состояний, а на уровне его функционирования как системы, обрабатывающей информацию. Таким именно образом брешь и заполняется.

Вначале этот исследовательский проект казался новым словом в науке о сознании. Его сторонники полагали, что он наконец поставит психологию на надёжный научный фундамент и освободит её от заблуждений бихевиоризма.

В этой лекции я собираюсь выступить против когнитивизма, однако начну всё же с его привлекательных черт. Мы знаем, что есть уровень наивной, здоровой, старой психологии, а также уровень нейрофизиологии — нейронов и нейронных модулей, синапсов и медиаторов, нервных узлов и так далее. Какие же у нас имеются основания полагать, что между двумя этими уровнями есть также уровень ментальных процессов, которые являются вычислительными процессами? И в самом деле, почему мы считаем, что именно на этом уровне мозг выполняет функции, существенно важные для выживания организма, а именно функции обработки информации?

Причин тому несколько. Скажу об одной сомнительной, однако весьма значимой причине. Скудные знания о мозге всё время заставляют нас считать новейшую технологию его моделью. В детстве, например, нам говорили, что мозг — это телефонный коммутатор. Удивительно, но Шеррингтон, великий английский учёный, полагал, что мозг работает наподобие телеграфа. Фрейд сравнивал мозг с гидравлической и электромагнитной системой! Лейбниц — с мельницей, а кто-то из греков считал, что мозг подобен катапульте. В настоящее время такой метафорой служит цифровой компьютер.

Сколько всякой чепухи можно услышать сегодня по поводу компьютеров и роботов. Пресса постоянно заверяет, что у нас в домах вотвот появятся роботы, которые будут делать всю работу по хозяйству, сидеть с детьми, развлекать нас беседой, заботиться о нас в старости. Разумеется, все это просто болтовня.

Мы и близко не подошли к созданию таких роботов. Наиболее удачные модели способны решать очень узкие задачи в очень ограниченном контексте, например, на линиях сборки автомобилей.

Вернёмся, однако, к более серьёзным вещам. Во-первых, в пользу истинности когнитивизма приводят психологические свидетельства. Эксперименты на время реакции показывают, что решение различных интеллектуальных задач требует и различного времени. И если различия во времени реакций людей идут параллельно различиям во времени реакции компьютера, то это якобы свидетельствует в пользу того, что человеческая система работает на тех же принципах, что и компьютер. Другие свидетельства содержатся в лингвистике, в частности в работах Хомски и его коллег по теории порождающей грамматики. Формальные правила грамматики, которым люди следуют в речи, подобны тем формальным правилам, которые выполняет компьютер.

Я не стану много распространяться по поводу свидетельства, основанного на времени реакции; все, полагаю, согласятся, что оно неубедительно и может быть интерпретировано как угодно. К «лингвистическому» же свидетельству я ещё вернусь.

Однако за «вычислительной» интерпретацией этих свидетельств лежит гораздо более глубокое и, полагаю, более веское основание: поскольку компьютеры следуют правилам, когда обрабатывают информацию, а человеческие существа также следуют правилам, когда мыслят, то имеется некое единство, мозг и компьютер функционируют сходным, а может быть и одинаковым образом.

Третье допущение, лежащее за исследовательской программой когнитивизма, имеет весьма древнее происхождение. Оно возвращает нас к Лейбницу, а может быть даже к Платону. Ментальное событие должно иметь теоретические причины. Если информация на выходе осмысленна, например, если осмысленна наша когнитивная способность изучать языки или узнавать лица, то должна существовать теория, интериоризированная нами, которая стоит за этой способностью.

Наконец, имеется причина, в силу которой люди, в особенности люди, склонные к философии, делают приверженцами когнитивистской исследовательской программы. Они не видят другого способа понять отношение между сознанием и мозгом. Отношение компьютерной программы к компьютерной аппаратуре даёт прекрасную модель, может быть даже единственную модель, которая позволяет объяснить отношение между сознанием и мозгом.

Как же относиться к этим аргументам в пользу когнитивизма? Вряд ли существует опровержение когнитивизма — наподобие моего опровержения сильной версии ИИ. Но я убеждён в слабости этих аргументов. Выявление их слабости поможет нам понять важные различия, существующие между поведением человеческих существ и функционированием компьютеров.

Начнём с понятия следования правилу. Человеческие существа следуют правилам, и компьютеры также следуют правилам. Но в случае с человеческими существами их всегда направляет содержание или смысл правила, смыслы причинно обуславливают поведение. Разумеется, его обуславливают не только они, но они несомненно играют роль причин в выработке поведения. Рассмотрим такое правило; в Великобритании следует вести машину по левой стороне дороги. Когда я приезжаю в Великобританию, я должен об этом помнить. Как действует это правило? Сказать, что я подчиняюсь правилу, — значит сказать, что смысл этого правила, то есть его семантическое содержание, выполняет роль причины того, что я делаю. Заметьте, имеется множество других правил, описывающих происходящее. Но я следую не этим правилам. Так, например, если я нахожусь на дороге с двусторонним движением и рулевое колесо моей машины расположено справа, вы можете сказать, что моё поведение согласуется с правилом: веди машину так, чтобы руль был как можно ближе к разделительной линии. Это — верное описание моего поведения. Но это не то правило, которому я следую в Великобритании. Правило, которому я следую, гласит: веди машину по левой стороне дороги.

Приведу другой пример. Когда мои дети ходили в автошколу, их там обучали правилу парковки. Правило заключалось в следующем: направляйте машину к краю тротуара, выворачивая руль в крайнее правое положение, пока передние колеса вашей машины не сравняются с задними колёсами машины, стоящей перед вами. Затем выворачивайте руль в крайнее левое положение. Если следовать этому правилу, его смысл должен играть причинную роль в поведении. Мне было интересно узнать о существовании этого правила, поскольку сам я ему не следую. Я вообще не следую никакому правилу, когда паркую машину. Я просто смотрю на край тротуара, пытаюсь подъехать к нему как можно ближе и в то же время не врезаться в машину впереди и в машину сзади. Может, правда, окажется, что моё поведение, если посмотреть на него со стороны, ничем не отличается от поведения того, кто следует правилу парковки. Тем не менее, неправильно было бы говорить обо мне, что я следую правилу. Чтобы утверждать, что правилу следуют, формальных свойств поведения недостаточно. Чтобы правилу следовали, смысл правила должен играть в поведении роль причины.

Мораль проста: в том смысле, в котором человеческие существа следуют правилам (между Прочим, человеческие существа следуют правилам в значительно меньшей степени, чем считают когнитивисты), компьютеры правилам вообще не следуют. Они действуют в соответствии с определёнными формальными процедурами. Программа компьютера определяет различные шаги машины; она определяет трансформацию одного состояния в другое. Это следование правилам в метафорическом смысле. Но в буквальном смысле, в том смысле, в котором человеческие существа следуют правилам, компьютеры не следуют правилам, они только действуют так, как если бы они следовали правилам. Подобные метафоры совершенно безвредны, более того, в науке они широко распространены и выполняют полезные функции. Мы можем высказываться метафорически о любой системе, как если бы она следовала правилам, например о солнечной системе. Метафора становится вредной, когда её понимают буквально. Вполне допустимо использовать психологическую метафору при объяснении работы компьютера. Путаница возникает тогда, когда метафора понимается буквально и метафорический смысл следования правилу используется для объяснения его психологического смысла, на котором данная метафора в первую очередь и основывалась.

Вернёмся теперь к «лингвистическому» свидетельству в пользу когнитивизма. Если люди следуют правилам синтаксиса, это ещё не доказывает, что они ведут себя подобно цифровым компьютерам, ибо в том смысле, в каком люди следуют правилам синтаксиса, компьютеры этим правилам не следуют. Они выполняют формальные процедуры.

Итак, существуют два смысла следования правилу: буквальный и метафорический. Их легко спутать. Обратимся теперь к понятию обработки информации. В нём воплощено столь же распространённое заблуждение. Когда я мыслю, я обрабатываю информацию; вычислительная машина тоже обрабатывает информацию, когда преобразует данные на входе и выдаёт информацию на выходе. Поэтому в процессах обработки информации должно быть нечто общее. Но это заблуждение. Я осуществляю обработку информации в том смысле, что сознательно или неосознанно произвожу определённые ментальные процессы. Но в этом смысле калькулятор не осуществляет обработки информации, поскольку в нём не происходит никаких ментальных процессов. Он просто копирует или имитирует формальные свойства моих ментальных процессов. Иначе говоря, даже если шаги, осуществляемые калькулятором, формально такие же, как у меня, это не доказывает, что машина делает похожие вещи, и причина здесь очень простая — калькулятор вообще не обладает ментальными состояниями. Складывая 6

и 3, он не знает, что цифра «6» обозначает число шесть, цифра «3» — число три, а знак «плюс» — операцию сложения. Калькулятор вообще ничего не знает. Поэтому нам и нужны калькуляторы. Они способны производить вычисления быстрее и аккуратнее, чем мы, не предпринимая никаких ментальных усилий. В том смысле, в каком мы осуществляем обработку информации, они этого не делают.

Необходимо, следовательно, различить два смысла понятия обработки информации или, по крайней мере, два радикально отличающихся друг от друга вида обработки информации. Первый из них, «психологическая обработка информации», предполагает ментальные состояния. Сформулируем это как можно проще: когда люди выполняют ментальные операции, они мыслят, а в мышлении происходит обработка информации того или иного рода. Но имеется и другой смысл обработки информации, который не предполагает ментальных состояний. В этих случаях происходят процессы, как будто бы предполагающие ментальную обработку информации. Назовём эту вторую разновидность «обработкой информации „как если бы“». Нет вреда в том, чтобы пользоваться и той, и другой разновидностью. Однако в когнитивизме их всё время путают.

В свете этого различия видна одна из главных слабостей когнитивистской аргументации. Из того, что я, мысля, осуществляю обработку информации и из того, что компьютер также осуществляет обработку информации — даже такую, которая имитирует формальные свойства моего мышления, — не следует, что компьютерная программа имеет отношение к психологии. Чтобы показать её психологическую релевантность, следовало бы иметь какой-то аргумент в пользу того, что «обработка информации „как если бы“» психологически релевантна. Понятие обработки информации используется для обозначения двух совершенно различных феноменов. Короче говоря, путаница, обнаруженная нами в отношении следования правилу, находит точную параллель в понятии обработки информации.

В понятии обработки информации заключено, однако, и более глубокое недоразумение. Заметьте, что в смысле обработки информации «как если бы» любая система может быть описана так, как если бы она осуществляла обработку информации. Мы даже могли бы описать её так, как если бы она занималась получением информации. Дело не просто в использовании калькуляторов и компьютеров. Возьмём, к примеру, текущую вниз по холму воду. Мы можем описать её таким образом, как если бы она занималась обработкой информации или даже получением информации. Мы, например, могли бы использовать воду для получения информации о линии наименьшего сопротивления в рельефе холма. Но из этого не следует, что бегущая вниз вода

обладает психологическим смыслом. В действии силы тяготения на воду никакой психологии нет.

Из всего вышеизложенного следует извлечь определённые уроки. Очевидно, что в мозге имеется уровень реальных психических информационных процессов. Люди мыслят, и мышление происходит в мозге. Разнообразные вещи, происходящие в мозге на нейрофизиологическом уровне, причинно обуславливают наши мыслительные процессы. Но многие люди полагают, что, кроме этих двух уровней — психологии и нейрофизиологии, — должен также существовать уровень вычислительной обработки информации. Почему же они так считают? Это происходит, по-моему, отчасти из-за смешения психологического уровня обработки информации с возможностью описания процессов, «как если бы» происходящих в мозге. Если говорить о стекающей вниз по холму воде, каждому ясно, что психология не имеет к этому никакого отношения. Однако в отношении мозга все это не столь очевидно.

В мозге происходят психологические процессы, и он имеет нейрофизиологию, которая причинно обуславливает и реализует эти процессы. Но тот факт, что мы можем описывать некоторые другие процессы в мозге с точки зрения «обработки информации „как если бы“», ещё не свидетельствует, что эти процессы психологически реальны или вообще имеют отношение к психологии. Поскольку мы говорим о мозге, увидеть здесь путаницу сложнее, чем в примере с водой.

Следующее предположение заключается в том, что за всяким осмысленным поведением должна скрываться некая внутренняя теория. Его высказывает не только когнитивная психология. Так, например, теория универсальной грамматики Хомски основывается на предположении, что если имеются свойства, общие для всех языков, и если эти свойства определяются общими свойствами человеческого мозга, то в мозге должны содержаться правила универсальной грамматики. Гипотеза могла бы быть и более простой: физиологическая структура мозга определяет возможные грамматики непосредственно, без промежуточного уровня правил или теорий. Дело не только в большей простоте этой гипотезы: само существование универсальных свойств языка, определяемое внутренними свойствами мозга, свидетельствует, что одного нейрофизиологического уровня описания вполне достаточно. Нет нужды предполагать сверх нейрофизиологических структур ещё и некие правила.

Пара аналогий, надеюсь, прояснит вопрос. Возьмём простой факт: человеческое зрение не воспринимает инфракрасных и ультрафиолетовых волн. Почему? Может быть, существует универсальное правило визуальной

грамматики, гласящее: «Не воспринимай инфракрасного или ультрафиолетового?» Разумеется, нет, Это происходит потому, что наш зрительный аппарат нечувствителен к этим двум крайним полосам спектра. Конечно, можно было бы описать себя так, как если бы мы следовали правилу визуальной грамматики, но ведь мы не следовали никакому правилу. Возьмём другой пример: предположим, что мы провели теоретический анализ человеческой способности сохранять равновесие при ходьбе. При этом нам показалось, что в нас как будто происходят более или менее сложные ментальные процессы, что мы как будто принимаем различные сигналы, с помощью которых решаем некое множество квадратных уравнений (бессознательно, конечно), и это позволяет нам ходить не падая. Но мы хорошо знаем, что такого рода теория не нужна для объяснения способности ходить, сохраняя равновесие. Способность ходить в значительной мере обусловлена функционированием во внутреннем ухе определённых жидкостей, которые ничего не вычисляют. Если вертеться на месте достаточно долго, положение жидкостей изменится и вы вероятно упадёте. Возможно, значительная часть наших когнитивных достижений также обусловлена подобными вещами. Мозг их просто реализует. У нас нет оснований предполагать, что в дополнение к уровням ментальных состояний и нейрофизиологии имеется ещё какое-то бессознательное вычисление.

Как мы узнаем лица? Лица наших друзей, родственников и знакомых мы распознаем, не прилагая никаких усилий, и в настоящее время у нас имеются свидетельства в пользу того, что за это отвечают определённые участки мозга. Как это происходит? Предположим, мы сконструировали компьютер, способный узнавать лица. В его задачу входила бы масса вычислений геометрических и топологических свойств. Но свидетельствует ли это о том, что распознавание лиц обязательно предполагает вычисление? Когда мы ступаем по мокрому песку и оставляем на нём след, ни ноги, ни песок не делают никаких вычислений. Что касается вычисления топологии следа ноги по информации о различных давлениях, оказываемых на песок, то это довольно сложная задача. Цифровая имитация природного феномена предполагает сложную обработку информации, но это не доказывает, что сам феномен предполагает такую обработку. Узнавание лица столь же простое и автоматическое дело, как оставление следов на песке.

Если последовательно развивать аналогию с компьютером, можно обнаружить в нём массу вещей, которые точно так же не являются вычислительными процессами. Например, если вы спросите; «Как калькулятор умножает трижды на семь?», ответом будет; «Он прибавляет семь раз по три». Но если вы затем спросите; «А как он прибавляет по три?», то не услышите никакого «вычислительного» ответа. Вам ответят: «Он просто

делает это — и все». За подавляющим большинством базовых способностей, таких как наша способность видеть или обучаться языку, нет никакого теоретического ментального уровня; мозг просто способен на все это. Наша нейрофизиология такова, что попадание фотонов на клетки фоторецепторы позволяет нам видеть, а стимулы в виде речи других людей и взаимодействия с ними позволяют нам обучиться языку.

Я не утверждаю, что правила не играют никакой роли в нашем поведении. Напротив, правила языка или правила игр, по-видимому, выполняют решающую роль в поведении. Но очень трудно установить, какие части поведения выполняются по правилам, а какие нет. И у нас нет оснований предполагать, что за осмысленным поведением стоит некая система правил.

Вероятно, следует сказать, что хотя я и не настроен оптимистически по поводу исследовательского проекта когнитивизма, затраченные на него усилия наверняка принесут немало результатов. Тем более мне не хочется мешать никому, кто захочет доказать, что я не прав. Но даже если я прав, много хорошего можно извлечь и из неудавшихся исследовательских проектов; примерами служат бихевиоризм и фрейдская психология. В случае с когнитивизмом особенно ценными мне представляются исследования Дэвида Марра по зрительному восприятию и многих других учёных в области «естественного языка», то есть компьютерной имитации процессов продуцирования и интерпретации обычной человеческой речи.

Мне бы хотелось закончить эту главу на более оптимистической ноте. Рассмотрим следствия когнитивистского подхода для изучения сознания. Позвольте мне изложить альтернативный подход к решению проблем, стоящих перед социальными науками. Отбросим идеи о том, будто между сознанием и мозгом стоит компьютерная программа. Рассмотрим сознание и ментальные процессы как биологические феномены, такие как рост, пищеварение или выделение желчи. Наш визуальный опыт, например, можно рассмотреть как конечный продукт событийного ряда, начинающегося с попадания фотонов на сетчатку и заканчивающегося где-то в мозге. Причинное объяснение зрения у животных предполагает два основных уровня описания. Во-первых, уровень нейрофизиологии, на котором мы можем обсуждать нейроны, синапсы и потенциалы. Внутри этого нейрофизиологического уровня будут свои нижний и верхний уровни описания. Мы не должны ограничиваться только нейронами и синапсами. Мы можем говорить о поведении групп или модулей нейронов, таких как различные уровни типов нейронов на сетчатке или строение коры головного мозга. И мы можем говорить о поведении нейрофизиологических систем на значительно более высоких уровнях сложности, например, мы можем

говорить о роли бороздчатой коры для зрения, о роли зон 18 и 19 для визуальной области коры или же о взаимодействии визуальной области коры и всего остального мозга в обработке зрительных стимулов. Таким образом, в рамках нейрофизиологического уровня будет своя группа уровней описания.

Кроме того, имеется ментальный уровень описания. Мы знаем, например, что восприятие есть функция ожидания. Если вы ожидаете увидеть нечто, то вероятность того, что вы это увидите, будет гораздо более высокой. На восприятие могут воздействовать различные ментальные феномены. Настроение и эмоции способны влиять на то, каким образом и что именно человек воспринимает. В рамках этого ментального уровня также имеются различные уровни описания. Мы можем говорить не только о воздействии индивидуальных убеждений и желаний, но также о воздействии таких ментальных феноменов, как личностные способности, обусловленные происхождением и средой, или мировоззрение. Но нет никакой необходимости в допущении ещё одного уровня — уровня цифровых вычислительных процессов. И нет никакого вреда в том, чтобы думать как об уровне ментальных состояний, так и об уровне нейрофизиологии как о процессах обработки информации, если не смешивать реальную психологическую форму обработки информации с «обработкой информации „как если бы».

Итак, как же можно оценить когнитивистскую исследовательскую программу? Конечно, её ложность не доказана. Программа может оказаться и истинной. Полагаю, что шансы её не хуже, чем у бихевиоризма пятьдесят лет назад. Иначе говоря, шансов на успех у неё нет. Моя аргументация в пользу этого вывода сводится к трём моментам. Во-первых, как только мы выявляем главные предпосылки, лежащие в основе когнитивизма, их неправдоподобие сразу же становится очевидным. Эти предпосылки глубоко спрятаны в нашей интеллектуальной культуре, некоторые из них очень трудно выкорчевать, трудно даже вполне осознать, что они существуют. Итак, моё первое утверждение состоит в том, что как только мы осознаем природу этих предпосылок, их неправдоподобие станет очевидным. Во-вторых, у нас не хватает эмпирических свидетельств для предположения о том, что эти предпосылки истинны, поскольку неоднозначными остаются некоторые важные механизмы, такие как обработка информации и следование правилам. И в-третьих, как в данной, так и в первой главе я изложил альтернативный взгляд на отношение между мозгом и сознанием; нет нужды в постулировании какого-то промежуточного уровня алгоритмических вычислительных процессов, опосредующих нейрофизиологию мозга и интенциональность сознания.

Итак, в дополнение к уровню ментальных состояний, таких как убеждения и желания, а также к уровню нейрофизиологии не нужен никакой другой уровень, не требуется ничего, что бы заполняло брешь между сознанием и мозгом. Ибо не существует самой бреши. Компьютер, вероятно, не лучшая и не худшая метафора, если сравнивать её с механическими метафорами. Мы узнаем о мозге, говоря, что он компьютер, ровно столько же, сколько мы о нём узнаем, сравнивая его с телефонным коммутатором, телеграфной системой, водяным насосом или паровым двигателем.

Предположим, никто не знает, как работают часы. Страшно трудно выяснить, как они работают, потому что никто не знает способа их изготовления, а попытки выяснить это приводят к их поломке. Находится группа исследователей, которые говорят: «Мы поймём, как работают часы, если спроектируем машину, функционально эквивалентную часам, то есть такую машину, которая станет показывать время». Они делают циферблат и заявляют: «Теперь мы знаем, как работают часы», или они говорят: «Если бы мы могли сделать циферблат так же аккуратно, как и часы, мы бы наконец поняли, как работают часы». Поставьте в этой притче «мозг» на место «часов», «цифровую компьютерную программу» на место «циферблата», а понятие интеллекта на место понятия определения времени, и вы получите современное состояние когнитивизма и многих (хотя и не всех) исследований по искусственному интеллекту.

Моей главной задачей в данной книге является выяснение некоторых наиболее трудных вопросов, касающихся места человеческих существ, во Вселенной. В первой главе я пытался дать решение проблемы отношения сознания и тела. Во второй — отверг некоторые экстремистские заявления, отождествляющие человеческие существа с цифровыми компьютерами. В третьей главе я выразил определённые сомнения по поводу когнитивистской исследовательской программы. Далее я хотел бы обратить внимание на структуру человеческого действия, обсудить вопрос о природе социальных наук и заняться проблемой свободы воли.

ГЛАВА 4.

СТРУКТУРА ДЕЙСТВИЯ

В этой главе мы займёмся структурой человеческого действия, которое необходимо объяснить в силу ряда причин. Следует показать, что наша концепция действия совместима с объяснением отношения сознания и тела, что она не противоречит отрицанию программы искусственного интеллекта. Для меня важно показать ментальный компонент в действии и его отношение к физическому компоненту, а также отношение структуры действия к объяснению действия. И я должен иметь основания для обсуждения природы социальных наук и возможности свободы воли.

При рассмотрении человеческих действий бросаются в глаза различия между ними и другими событиями в мире природы. Считать, будто типы действий или поведения тождественны типам телесных движений, ошибочно. Один и тот же набор человеческих движений может быть танцем, сигнализацией, упражнением, проверкой мускулов или ещё чем-нибудь. Более того, подобно тому как одни и те же типы физических движений могут образовывать совершенно различные виды действий, один и тот же вид действия может быть выполнен совершенно разными типами физических движений. Например, мы посылаем весточку другу. Можно написать её на листе бумаги. Можно напечатать. Можно послать её с кем-нибудь или дать телеграмму. Можно зачитать её по телефону. Каждый из этих способов посылки одной и той же вести может быть выполнен с помощью разнообразных физических движений.

Можно написать записку левой или правой рукой, пальцами ноги или даже держа карандаш в зубах. Ещё одной необычной чертой действий, отличающей их от других событий, является то, что действия предполагают описания особого рода. Допустим, я гуляю в Гайд-парке. Здесь происходит множество самых разных вещей, но их описания не касаются моих интенциональных действий, потому что то, что я делаю, в значительной степени зависит от того, что я при этом думаю. Так, например, я могу двигаться в направлении Патагонии, трясти шевелюрой, изнашивать ботинки и задевать при ходьбе молекулы воздуха. Однако ни одно из этих описаний не схватывает существа моего действия.

Третье характерное свойство действий состоит в том, что личность находится в особой позиции, она знает то, что она делает. Человеку нет необходимости наблюдать себя со стороны или проводить исследование, чтобы понять, какое

действие он выполняет или пытается выполнить. Так что, если вы меня спросите: «Что вы делаете: прогуливаетесь по Гайд-парку или же идёте в направлении Патагонии?», я не стану колебаться с ответом, хотя физические движения, которые я совершаю, могут подходить для любого из предлагаемых мне ответов.

Примечательно, что мы способны идентифицировать и объяснить поведение самих себя и других людей, не прилагая никаких усилий. Полагаю, что эта способность основывается на бессознательном владении определённым набором принципов, так же как способность распознавать предложения, написанные по-английски, основывается на бессознательном владении принципами английской грамматики. По-видимому, имеется набор принципов, которые предполагаются, когда мы говорим такие обыденные вещи, как например то, что Бэзил голосовал за тори, потому что думал, что они решат проблему инфляции, или что Сэлли переехала из Бирмингема в Лондон, полагая, что там больше возможностей получить работу, или даже что человек, совершающий некие странные движения, на самом деле точит топор или чистит ботинки.

Люди, признающие существование этих принципов, насмехаются над ними, они говорят, что это просто «народная наука» и что её следует заменить более научным объяснением человеческого поведения. Яс недоверием отношусь к подобным заявлениям; столь же подозрительно я бы стал относиться и к заявлению, что мы должны заменить нашу неосознанную (implicit) теорию английской грамматики — ту, что мы усваиваем, обучаясь языку, — какой-то другой теорией. Причина моей подозрительности в любом из этих случаев одна и та же: с моей точки зрения, имплицитная теория входит в выполнение действия так же, как правила грамматики входят в речевой процесс. Поэтому, хотя мы и можем что-то добавить или открыть разного рода интересные вещи о языке или поведении, маловероятно, чтобы мы могли заменить теорию, которая имплицитна и отчасти конститутивна по отношению к феномену, некоторым внешним «научным» объяснением этого феномена.

Аристотелю и Декарту наши объяснения человеческого поведения показались бы весьма знакомыми. Обычно говорят, что в отличие от биологии и физики наука о человеческом поведении осталась на прежнем уровне. Думаю, дело не в этом. Аристотель и Декарт уже обладали точной и сложной теорией человеческого поведения. Думаю также, что многие с виду «научные» объяснения человеческого поведения, такие как фрейдовская теория, фактически применяют принципы неявной теории человеческого поведения.

Подведу итог: типы действий не сводятся к типам физических движений; действия требуют описания особого рода; люди знают, что они делают, не осознавая этого; принципы, с помощью которых мы идентифицируем и объясняем действия, сами являются частью действий, иначе говоря — они конститутивны в отношении действий.

А сейчас мне хотелось бы заняться вопросом о структуре поведения.

В предварительном порядке необходимо ввести одно или два понятия. Ключевым является понятие интенциональности. Сказать, что ментальное состояние интенционально, значит сказать, что оно — «о чем-то». Убеждение, например, всегда является убеждением в том, что то-то и то-то имеет место, а желание всегда есть желание того, чтобы то-то и то-то произошло или имело место. Намерение (*intending*) в обычном смысле слова не играет особой роли в теории интенциональности. Намерение сделать что-то есть лишь одна из разновидностей интенциональности наряду с верованием, желанием, надеждой, боязнью и так далее.

Интенциональное состояние, скажем убеждение, желание или намерение, включает два компонента. В нём есть то, что можно было бы назвать «содержанием», это делает его говорящим о чём-то; в нём также есть так называемый «психологический модус», или «тип». Такое различие необходимо, ибо в различных типах вы можете иметь одно и то же содержание. К примеру, я могу желать покинуть комнату; я могу быть убеждён, что покину комнату; и я могу намереваться покинуть комнату. В каждом из этих случаев у нас имеется одно и то же содержание — что я покину комнату, но существует оно в различных психологических модусах, или типах: в модусах убеждения, желания и намерения.

Кроме того, различие содержания и типа служит для соотнесения ментального состояния и мира. В конце концов сознание представляет мир в различных ментальных состояниях, представляет, каков он, каким бы мы хотели его видеть, какого мира мы боимся, что мы намереваемся сделать в отношении него и так далее. Наши убеждения истинны, если они соответствуют миру, и ложны, если они ему не соответствуют; точно так же наши желания осуществляются или не осуществляются, а намерения реализуются или не реализуются. Поэтому, вообще говоря, интенциональные состояния имеют свои «условия удовлетворения». Каждое состояние само определяет, при каких условиях оно истинно (если это убеждение), или при каких условиях оно осуществляется (если это желание), или при каких условиях оно реализуется (если это намерение). В каждом случае ментальное состояние диктует свои собственные условия удовлетворения.

Третьей отличительной чертой таких состояний является то, что иногда они причинно обуславливают происходящие вещи. Например, если я хочу пойти в кино и в самом деле иду в кино, моё желание причинно обуславливает само событие, то есть мой поход в кино. В таких случаях между причиной и действием имеется внутренняя связь, или причина является репрезентацией положения дел, обуславливаемого причинно. Причина одновременно и репрезентирует действие, и вызывает его. Я называю такого рода отношения между причиной и действием «интенциональной причинностью». Как мы увидим, интенциональная причинность играет существенно важную роль в структуре и объяснении человеческого действия. В ряде аспектов это объяснение полностью отличается от стандартных объяснений причинности, описывающих, как один бильярдный шар ударяет по другому и это и является причиной движения. Существенно важным моментом в интенциональной причинности является то, что в случаях, которые мы будем рассматривать, сознание вызывает именно то положение дел, которое оно себе представляет.

Итак, в нашем анализе человеческого поведения необходимо иметь в виду три момента. Во-первых, интенциональные состояния включают содержания, которые существуют в определённом ментальном типе. Во-вторых, они сами определяют условия своего удовлетворения, то есть они будут или не будут удовлетворяться в зависимости от того, насколько мир будет соответствовать содержанию состояния. И в-третьих, иногда они причинно обуславливают происходящие вещи и с помощью интенциональной причинности вызывают вышеупомянутое соответствие; иначе говоря, они вызывают и то положение дел, которое репрезентируют, и свои собственные условия удовлетворения.

Используя эти идеи, перейду к главной задаче настоящей главы: описанию того, что можно было бы назвать структурой действия или структурой поведения. Под поведением я имею в виду свободное интенциональное человеческое действие. Я имею в виду такие вещи, как прогулка, бег, еда, любовь, голосование на выборах, женитьба, покупка и продажа, уход в отпуск, работа на службе. И не имею в виду такие вещи, как пищеварение, старение или храпение. Но даже если мы ограничимся интенциональным поведением, нам предстанет удивительное разнообразие типов. Мы должны будем различить индивидуальное поведение и социальное поведение, коллективное социальное поведение и индивидуальное поведение в рамках социального коллектива, а также делание чего-то ради чего-то ещё и делание чего-то ради самого себя. Пожалуй, труднее всего объяснить поведенческую «мелодию», слышную на протяжении длительного времени. Человеческие действия — это не ряд неподвижных моментальных снимков, наша жизнь скорее похожа на фильм.

На все вопросы мне не ответить, но мои слова, полагаю, будут звучать здраво. Впрочем, здравые объяснения никогда не казались очевидными. Начать с того, что бихевиористская традиция в философии и психологии привела к отрицанию ментального компонента в действиях. Бихевиористы стремились определить действия, да и всю нашу ментальную жизнь, в терминах чисто физических движений. Кто-то однажды охарактеризовал бихевиористский подход, и на мой взгляд справедливо, как не удавшуюся анестезию. Другой крайностью был ментализм, утверждавший, что единственные действия, которые мы выполняем, — это внутренние ментальные акты воления. Согласно этому взгляду, мы, строго говоря, никогда не поднимаем рук. Мы только «желаем», чтобы наши руки поднялись вверх. И если они действительно поднимаются, то это следует отнести скорее к удаче, чем к нашему действию.

До недавнего времени философия действия в известной мере игнорировалась. Западная традиция настойчиво проводила мысль, что познание важнее делания. Теории знания и смысла представлялись ей более важными, чем теория действия.

Попытаюсь соединить ментальные и физические аспекты действия. Структуру поведения лучше всего описать некоторым множеством принципов, объясняющих как ментальные, так и физические аспекты действия. Я не стану обсуждать, откуда берутся убеждения, желания и так далее, в мою задачу будет входить анализ той роли, которую они играют в нашем поведении.

Проще всего будет сформулировать эти принципы, а затем попытаться их обосновать.

Принцип 1:

Действие состоит из двух компонентов: ментального и физического.

Представим себе, например, что мы толкаем машину. С одной стороны, здесь имеется определённый сознательный опыт усилия. В случае его удачи такое усилие даст в результате движение вашего тела и соответствующее движение машины. В случае неудачи будет иметь место ментальный компонент, то есть опыт толкания машины, а также некоторые физические компоненты, такие как напряжение мускулов, ощущение давления на машину и прочее. Это позволяет нам сформулировать следующий принцип.

Принцип 2:

Ментальный компонент интенционален.

Он обладает интенциональностью, он — «о чем-то»; и он определяет, что считать успешным или неуспешным действием; если действие успешно, оно причинно обуславливает телесное движение, которое в свою очередь причинно обуславливает другие движения, такие как движение машины, что и образует все действие в целом. С точки зрения теории интенциональности, действие состоит из двух компонентов: ментального и физического. Если действие успешно, ментальный компонент причинно обуславливает и репрезентирует физический компонент. Эту форму причинности я называю интенциональной причинностью.

Лучший способ выявить природу различных компонентов действия — выделить эти компоненты и исследовать каждый из них по отдельности. В лаборатории это сделать нетрудно, и нейрофизиологами уже проведены некоторые эксперименты. Стимулируя при помощи электрического тока определённую часть моторной области коры головного мозга, Уайлдер Пенфилд из Монреаля вызывал движение конечностей испытуемого. При этом пациенты неизменно характеризовали происходившее следующим образом: «Я этого не делал — это сделали вы». В таком случае телесное движение лишено интенции. Следует отметить, что телесные движения могут быть теми же самыми, что и в интенциональном действии, однако ясно, что между ними имеются различия. Какие же? Вспомним эксперименты, проведённые ещё Уильямом Джемсом, где выделялся только ментальный компонент действия. Рука испытуемого анестезировалась, затем его помещали в тёмную комнату и он получал приказ поднять руку. Он делал то, что, как ему представлялось, было подчинением приказу, однако в дальнейшем очень удивлялся, когда ему сообщали, что его рука так и не поднялась. В таких случаях ментальный компонент, то есть интенция, отделяется от телесного движения. Ибо интенция у испытуемого имелась. Иначе говоря, можно с полным правом сказать, что он действительно пытался поднять руку.

Обычно в действии присутствуют оба компонента. Как правило, у нас имеется и интенция, и телесное движение. Но они не независимы друг от друга, В первых двух принципах выражен способ их связи. Ментальный компонент, являясь одним из условий удовлетворения, должен как репрезентировать, так и причинно обуславливать физический компонент. Следует отметить, что в формулировке этих принципов мы используем обширный словарь: «пытаться», «достигать успеха», «терпеть неудачу», «интенциональное» и «неинтенциональное», «действие» и «движение».

Принцип 3:

Интенциональная причинность существенно важна и в структуре действия, и в его объяснении.

Телесные движения в действии причинно обусловлены интенциями, Интенции носят причинный характер потому, что благодаря им происходят различные вещи; но они также обладают содержанием и играют роль в процессе логического рассуждения. Они могут служить причинами и обладать логическими свойствами, потому что причинность, о которой мы говорим, есть ментальная, или интенциональная причинность. А в случае интенциональной причинности ментальные содержания воздействуют на мир. Весь этот аппарат работает потому, что он реализуется в мозге (тем способом, о котором говорилось в первой главе).

Причинность, которую мы здесь обсуждаем, весьма отличается от стандартной формы причинности, описываемой в учебниках по философии. Это не регулярности и не охватывающие законы или постоянные связи. Думаю, что она ближе к нашему обыденному понятию причинности — как тому, благодаря чему происходит нечто другое. Для интенциональной причинности характерно, что это такое ментальное состояние, которое заставляет происходить нечто другое; а нечто другое — это то самое положение дел, которое репрезентировано в ментальном состоянии, обуславливающим его причинно.

Принцип 4:

В теории действия проводится фундаментальное различие тех действий, которые заранее продуманы и являются результатом некоего предварительного планирования, и тех действий, которые спонтанны и совершаются без предварительной рефлексии.

Соответственно этому различию можно различить предварительные интенции, то есть интенции, образуемые до выполнения действия, и интенции в действии, то есть интенции, имеющие место при актуальном выполнении действия.

Распространённое мнение, будто все интенциональные действия являются преднамеренными, что это порождение цепи практических рассуждений, ошибочно. Очевидно, что во многом это не так. Зачастую мы просто что-то делаем, и здесь нет никакой предварительной рефлексии. Например, беседуя, мы обычно не рефлектируем над дальнейшими словами, а просто говорим те или иные вещи. В таких случаях интенция, конечно же, имеется, но она не образуется до выполнения самого действия. Это — интенция в действии. В

каких-то случаях мы формируем предварительные интенции. Мы рефлектируем по поводу того, что мы желаем и каков наилучший путь для достижения желаемого. Этот процесс рефлексии (Аристотель называл его «практическим рассуждением») либо приводит к формированию предварительной интенции, либо (как отмечал тот же Аристотель) проявляется в самом действии.

Принцип 5:

Формирование предварительных интенций есть (обычно) результат практического рассуждения, то есть рассуждения о наилучшем способе выбора между конфликтующими желаниями.

Мотивом большинства действий человека (и животных) является желание. Убеждения лишь позволяют нам выяснить, как лучше удовлетворить наши желания. Например, я хотел бы поехать в Париж, и я убеждён, что, учитывая все моменты, лучше всего лететь туда самолётом; поэтому у меня формируется интенция лететь в Париж самолётом. Таков типичный и обыденный пример практического рассуждения. Практическое рассуждение существенно отличается от рассуждения теоретического, говорящего о том, что имеет место. Цель практического рассуждения — сделать наилучший выбор из тех разнообразных конфликтующих желаний, которые в нас борются. Представим, что я собрался в Париж, и при этом выясняется, что лучше всего лететь туда самолётом. Однако осуществить это желание я могу, только победив множество других желаний. Мне вовсе не хочется тратить деньги, стоять в очередях в аэропортах, долго сидеть в самолёте, есть пищу» подаваемую в полёте; мне не хочется, чтобы сосед клал свой локоть на то место, на которое мне самому хотелось бы облокотиться и так далее до бесконечности. Тем не менее, несмотря на все подавляемые желания, я всё же путём рассуждения, учитывая все обстоятельства, прихожу к тому, что лучше всего лететь в Париж самолётом. И это не просто типичный случай. Полагаю, что практическое рассуждение постоянно разрешает споры между конфликтующими желаниями.

Сформулированные пять принципов рисуют нам картину, из которой следует, что ментальная энергия, питающая действие, есть энергия интенциональной причинности. Это — форма энергии, благодаря которой в причине, в виде желаний или намерений, репрезентируется то самое положение дел, которое эта причина обуславливает.

Вернёмся к тем компонентам действия, о которых мы говорили вначале. Мы отметили, что у действий есть свои особые описания и что здравый смысл позволяет их идентифицировать. Теперь мы понимаем, что это особое

описание действия состоит в описании интенции в действии. То, что человек делает или пытается делать, сводится к интенции, с которой он это делает. Например, я знаю, что мне нужно попасть в Гайд-парк, а не в Патагонию, потому что именно с такой интенцией я пошел прогуляться. И я знаю об этом, не производя наблюдений, поскольку такое знание есть знание не о внешнем поведении, а о внутренних ментальных состояниях.

Кроме того, это показывает некоторые логические особенности тех способов, с помощью которых мы объясняем человеческое действие. Объяснить действие — значит выяснить его причины. Его причины — это психологические состояния. Эти состояния имеют отношение к действию, будучи либо шагами в практическом рассуждении, приводящем к интенциям, либо самими интенциями.

Наиболее важная особенность объяснения действия, однако, заслуживает того, чтобы её сформулировали в качестве отдельного принципа.

Принцип 6:

Действие должно объясняться через то содержание, какое имелось в голове человека, выполнявшего действие или рассуждавшего относительно интенции к выполнению действия.

Если объяснение желает действительно что-то объяснить, его содержание должно быть тождественно содержанию, обуславливающему поведение через интенциональную причинность.

В этом отношении действия отличаются от других природных событий. Когда мы объясняем землетрясение или ураган, содержание в нашем объяснении только репрезентирует, что произошло и почему это произошло. Оно не обуславливает само событие. Что касается человеческого поведения, то и причина и объяснение обладают содержанием, и объяснение только потому является объяснением, что у него имеется то же содержание, что и у причины. До сих пор разговор шёл так, как если бы интенции возникали у людей сами собой. Но это, конечно же, совершенно неверно. Поэтому сейчас мы коснёмся некоторых трудностей, которые приблизят наш анализ к повседневной жизни. Интенция никогда не возникает просто так. Например, я намереваюсь поехать на машине в Оксфорд. Такая интенция могла возникнуть у меня спонтанно, но тем не менее у меня должен быть ряд других интенциональных состояний. Я должен быть убеждён, что имей машину и что Оксфорд находится на доступном для неё расстоянии. Кроме того, естественно желание, чтобы дорога была незагруженной, а погода не слишком скверной. Кроме того (и здесь мы подходим к понятию объяснения действия), я не просто хочу поехать

в Оксфорд, у меня имеется вполне определённая цель. Я буду заниматься практическим рассуждением — той формой рассуждения, которая ведёт не к убеждениям или заключениям, основанным на аргументах, но к интенциям и последующему действию. Осознав, что это за рассуждение, мы делаем значительный шаг к объяснению действий. Назовём другие интенциональные состояния, придающие моему интенциональному состоянию вполне конкретный смысл, «системой интенциональности». В заключение можно сказать следующее.

Принцип 7:

Любое интенциональное состояние функционирует только в качестве элемента системы интенциональных состояний.

Под «функционирует» я понимаю определение условий удовлетворения относительно множества других интенциональных состояний.

Когда мы переходим к системе, то открываем ещё один интересный феномен. Деятельность нашего сознания не сводится к ментальным состояниям. Сами ментальные состояния функционируют соответствующим образом только потому, что это происходит на фоне возможностей, способностей, умений, привычек, образов действия и общего мироотношения, которые не сводятся к интенциональным состояниям. Для того чтобы у меня сформировалось намерение поехать в Оксфорд, я по крайней мере должен уметь водить машину. Но умение водить машину не сводится к множеству других интенциональных состояний. Чтобы уметь водить машину, требуется нечто большее, чем совокупность убеждений и желаний. Я должен обладать навыком. Это тот случай, когда моё «знание как» не сводится только к «знанию что». Назовём совокупность навыков, привычек, способностей и так далее, на фоне которых функционируют интенциональные состояния, «фоном интенциональности». И к тезису о системе интенциональности, то есть о том, что любое интенциональное состояние функционирует только в качестве части системы, мы добавим тезис о фоне интенциональности.

Принцип 8:

Система интенциональности функционирует на фоне человеческих возможностей, которые сами не являются ментальными состояниями.

Я уже говорил, что многие претендующие на научность объяснения поведения пытаются избежать этой модели или предложить какую-либо иную. Думаю, что им это не удастся: изложенные выше принципы не просто описывают феномены, они сами — часть этих феноменов. Рассмотрим, например, фрейдистские объяснения. Во фрейдовской метапсихологии, то есть в теории

того, что делает сам Фрейд, часто используются научные сравнения. Проводится немало аналогий между психологией и электромагнитной теорией или гидравликой; и сознание, например, функционирует по аналогии с принципами гидравлики и так далее. Но когда Фрейд исследует пациента, описывает природу его неврозов, многие его объяснения оказываются объяснениями с точки зрения здравого смысла. Дора ведёт себя соответствующим образом, потому что она любит г-на К., или потому что она подражает своей кузине, которая уехала в Марицелл.

К точке зрения здравого смысла Фрейд добавляет только то наблюдение, что ментальные состояния, причинно обуславливающие наше поведение, часто являются бессознательными. По сути дела, они подавлены. Мы часто не признаем определённых интенциональных состояний, потому что нам за них стыдно или по какой-то другой причине. Кроме того, Фрейд добавляет теорию преобразований ментальных состояний, описывающую, как интенциональное состояние одного рода трансформируется в интенциональное состояние другого рода. Но с этими и другими добавлениями фрейдистское объяснение становится подобным объяснению с точки зрения здравого смысла. Я думаю, что здравый смысл сохранится даже тогда, когда мы предложим какие-то более научные объяснения поведения. Поскольку структура объяснения должна соответствовать структуре объясняемого феномена, усовершенствования в объяснении вряд ли выявят какие-то новые и неслыханные доселе структуры.

В данной главе я попытался объяснить, как и в каком смысле поведение одновременно и содержит внутренние ментальные состояния, и причинно обуславливается ими. Удивительно, но психология и когнитивная наука пытались отрицать наличие таких отношений. В следующей главе будут анализироваться некоторые следствия, которые вытекают из этого взгляда на человеческое поведение и которые имеют значение для социальных наук. В чём причина неудач и успехов этих наук, и чему мы можем у них научиться?

ГЛАВА 5.

ПЕРСПЕКТИВЫ СОЦИАЛЬНЫХ НАУК

В этой главе мне бы хотелось обсудить одну из наиболее болезненных интеллектуальных проблем нашего времени: почему методы естествознания не принесли успеха в изучении человеческого поведения? И на какие «социальные» или «бихевиоральные» науки мы вообще можем положиться? На мой взгляд, между человеческим поведением и феноменами, которые изучает естествознание, имеются существенные различия. Эти различия объясняют как неудачи, так и успехи наук о человеке.

Вначале я хотел бы подчеркнуть различие объяснений с точки зрения здравого смысла и научных объяснений. Стандартная теория гласит, что научное объяснение некоторого феномена заключается в выведении его из научных законов. Эти законы являются обобщениями. Например, если у вас имеется закон падения тела и вам известно, где это тело начало падать, вы сможете вывести и всё дальнейшее. Подобно этому, если вам нужно объяснить некоторый закон, вы можете вывести его из закона более высокого уровня. Так что объяснение и предсказание симметричны. Вы предсказываете, дедуцируя то, что уже произошло. Однако какими бы достоинствами ни обладал этот тип объяснения в естественных науках, от него нет никакой пользы в сфере человеческого поведения. И бесполезен он не потому, что нет законов для объяснения индивидуальных случаев человеческого поведения. Даже если бы такие законы были, от них всё равно не было бы никакой пользы. Вообразите, что бы вышло, если бы у нас был «закон», то есть универсальное обобщение, касающееся какого-то аспекта поведения.

Допустим, на прошедших выборах вы голосовали за консерваторов, и делали вы это в надежде, что они лучше других будут способствовать решению проблемы инфляции. Допустим также, что причина вашего предпочтения является таким же обыкновенным фактом, как и ваш выбор. Представьте далее, что некоторые социологи делают универсальное обобщение в отношении людей, которые в точности вам подобны — и по социоэкономическому статусу, и по уровню доходов, образованию, интересам и так далее. И пусть из этого обобщения с необходимостью следует, что люди, похожие на вас, всегда голосуют за консерваторов. Как же объяснить ваш выбор? По-моему, обобщение здесь не годится. Обобщение устанавливает регулярность. Знание такой регулярности может быть полезным для предсказания, но оно не способно ничего объяснить в индивидуальных

случаях человеческого поведения. Оно лишь приглашает к дальнейшему объяснению. Например, почему все люди этой группы голосуют за консерваторов? Ответ очевиден. Они голосовали за тори, потому что беспокоились по поводу инфляции — вероятно, на людей этой группы инфляция действует в наибольшей степени и поэтому они делают один и тот же выбор.

Короче говоря, обобщение не объясняет ни нашего собственного, ни чьего-либо ещё поведения. Для самого обобщения потребовалось бы объяснение. Там, где речь идёт о человеческом поведении, мы обычно ищем объяснения с точки зрения ментальных состояний — убеждений, страхов, надежд, желаний и так далее, которые являются причинами, продуцирующими поведение (см. предыдущую главу).

Вернёмся, однако, к первому нашему вопросу: почему у нас нет законов социальных наук, подобных законам естественных наук? На это имеется несколько стандартных ответов. Некоторые философы считают, что у нас нет науки о поведении по той же причине, по какой у нас нет науки о мебели. У нас не может быть такой науки, потому что не существует общих физических свойств, которыми бы обладали предметы мебели — стулья, столы и прочее и которые бы всех их подводили под некие «мебельные» законы. Кроме того, мы не нуждаемся в такой науке, потому что всё, что мы желали бы объяснить — например, почему деревянные столы тверды или почему железная мебель ржавеет, может быть объяснено с помощью уже имеющихся наук. Подобно этому, нет свойств, которые были бы присущи всем видам человеческого поведения. А кроме того, конкретные вещи, которые мы желали бы объяснить, объяснимы с помощью физики, физиологии и других наук.

Некоторые философы указывают, что понятия, служащие для описания нас самих и других человеческих существ, не согласуются с понятиями таких фундаментальных наук, как физика и химия. Вероятно, говорят они, наука о человеке сродни науке о погоде. Метеорология — не строгая наука, потому что вещи, интересующие нас в отношении погоды, не согласуются с категориями, которыми оперирует физика. Такие понятия, как «светлые области над центральными графствами» или «в Лондоне кое-где туман», не связаны каким-либо систематическим образом с понятиями физики. Например, Джерри Фодор¹ считает, что частные науки, такие как геология или метеорология, занимаются свойствами мира, которые могут быть представлены и в физике. Подобная нежёсткая связь между частной наукой и наукой базисной характерна и для социальных наук. Подобно тому, как горы и бури реализуются в различных микрофизических структурах, деньги могут быть физически реализованы в виде золота, серебра или бумажек. И такие

дизъюнктивные связи между высшим и низшим порядком явлений, позволяя иметь содержательно богатые науки, не позволяют иметь строгих законов, ибо форма нежёстких связей предполагает законы, которые допускают исключения.

Ещё одним аргументом служит утверждение Дональда Дэвидсона 2 о том, что понятия о ментальных феноменах не связаны систематическим образом с понятиями физики. По выражению Дэвидсона, они не имеют никакого «отзвука» в физике. Однако множество пауков содержит фундаментальные понятия, которые также не имеют своего «отзвука» в физике, и тем не менее как науки они весьма основательны. Биология, например, пользуется понятием организма, и хотя это понятие не имеет своего «отзвука» в физике, биология не перестаёт быть из-за этого наукой.

Ещё один распространённый взгляд заключается в том, что сложные взаимоотношения наших ментальных состояний не позволяют нам сформулировать систематическое множество законов, которые бы соединяли ментальные и нейрофизиологические состояния. Ментальные состояния — это сложные, взаимосвязанные системы, которые невозможно спроецировать на типы состояний мозга. Однако этот аргумент не совсем ясен. Допустим, Ноам Хомски прав, считая, что мы обладаем сложным набором правил универсальной грамматики, запрограммированных в мозге. В сложности и взаимозависимости правил универсальной грамматики нет ничего, что бы препятствовало их систематической реализации в нейрофизиологии мозга. Взаимозависимость и сложность сами по себе ещё не служат аргументом против возможности строгих психофизических законов.

Все указанные объяснения, на мой взгляд, полезны, но я не считаю, что они адекватно схватывают действительно радикальные различия, существующие между ментальными и физическими науками. Отношение социологии и экономики, с одной стороны, и физики — с другой, отнюдь не похоже на отношение метеорологии, геологии, биологии и других частных естественных наук и физики. Но в чем же тогда состоит это различие? Думаю, мы должны раз и навсегда отбросить идею, что социальные науки подобны доньютоновской физике и что мы дожидаемся момента, когда они откроют свои собственные «ньютоновские» законы сознания и общества.

В чём же собственно заключается проблема? Можно было бы сказать: «Конечно, социальные и психические феномены не менее реальны, чем все остальные. Так почему же не может быть законов, описывающих их поведение?» Почему есть законы поведения молекул, но нет законов поведения обществ? Один из способов опровергнуть какой-либо тезис —

предположить, что он истинен, а затем показать, что предположение о его истинности абсурдно. Предположим, что у нас есть законы общества и истории, которые позволяют нам предсказывать войны и революции и делать это с такой же точностью, с какой мы предсказываем ускорение падения в вакууме, происходящее на уровне моря.

Проблема заключается в следующем: что бы собой ни представляли войны и революции, это прежде всего множество молекулярных движений. Поэтому любой строгий закон, касающийся войн и революций, должен полностью согласовываться с законами движения молекул. Чтобы революция началась в некий определённый день, все соответствующие молекулы должны устремиться в одном направлении. Если так, то законы на уровне революций и их участников должны быть теми же законами, что и законы движения молекул на уровне физических частиц. Переформулируем теперь наш вопрос. Почему законы на более высоком уровне, уровне революций, не могут полностью соответствовать законам на низшем уровне, уровне частиц? Чтобы доказать невозможность этого, рассмотрим некоторые случаи, в которых имеется полное соответствие законов высшего и низшего порядка; мы увидим все их отличие от того, что происходит в обществе.

Примером успешного сведения законов высшего уровня к законам низшего уровня является сведение законов для газов, законов Бойля и Шарля, к законам статистической механики. Насколько полезным оказалось такое сведение? Законы для газов касаются отношения между давлением, температурой и объёмом. Например, они предсказывают, что повышение температуры газа в цилиндре увеличивает давление на его стенки. Законы статистической механики касаются поведения большого количества малых частиц. Например, они предсказывают, что увеличение скорости движения частиц приведёт к тому, что всё больше частиц будет ударяться о стенки цилиндра и они будут это делать со всё большей силой. Эти две группы законов удаётся согласовать потому, что объяснение температуры, давления и объёма может быть дано в терминах поведения частиц. Повышение температуры увеличивает скорость частиц, а увеличение числа и скорости частиц, ударяющихся о стенки цилиндра, повышает давление. Отсюда следует, что увеличение температуры производит повышение давления. Предположим теперь, что объяснения давления и температуры в терминах поведения частиц не существует. Тогда любые законы на уровне давления и температуры будут казаться чудом. Совпадение того, что происходит с давлением и температурой, с тем, что происходит с частицами, следует отнести к чудесам, ибо отсутствует систематическая связь между поведением системы на уровне давления и температуры и поведением системы на уровне частиц.

Возьмём пример посложнее. Согласно закону «науки о питании», потребление калорий равно их затратам плюс или минус отложение жира. Не слишком впечатляюще, зато вполне реалистично. Последствия этого закона хорошо известны: если много есть и мало двигаться, становишься толстым. Этот закон, в отличие от законов для газов, основан не на поведении частиц, а на достаточно сложном ряде процессов переработки пищи в жировые отложения. Тем не менее обоснование данного закона через поведение более фундаментальных частиц всё же имеется. При прочих равных условиях, если вы много едите, молекулы вашего тела устремятся в том направлении, которое приведёт к вашему потолстению.

Продолжим нашу аргументацию. «Война», «революция», «свадьба», «деньги» и «собственность» — это понятия, которые не основываются систематическим образом на поведении более фундаментальных элементов, как это происходит с явлениями, описываемыми с помощью таких понятий, как «отложение жира» и «давление». Следует отметить, что именно такие обоснования позволяют нам достигать существенных результатов в науке. Открытие ДНК оказало чрезвычайно важное влияние на биологию, а микробная теория — на медицину, потому что в этих случаях свойства высшего уровня, такие как наследственные черты и симптомы заболевания, объясняются систематически и в терминах более фундаментальных элементов.

Но возникает вопрос: почему же подобным образом не обосновываются социальные и психологические феномены? Почему не соотнести такие социальные явления, как войны и революции, с молекулярными движениями?

Чтобы понять, почему это невозможно, посмотрим, какие свойства социальных явлений позволяют нам объединять их в категории. Каковы те фундаментальные принципы, на основе которых мы категоризируем психологические и социальные феномены? Один из них состоит в следующем: для огромного числа социальных и психологических феноменов понятие, обозначающее феномен, является составной частью самого этого феномена. Чтобы нечто можно было считать свадебной церемонией или профсоюзом, собственностью или деньгами, даже войной или революцией, люди, включённые в соответствующие деятельности, должны иметь соответствующие мысли. Вообще говоря, по их мнению, все так и должно быть. Например, для того чтобы жениться или приобрести собственность, вы и другие люди должны думать, что вы делаете именно то, что вы делаете. Эта особенность социальных феноменов является существенно важной. Ничего подобного нет в биологических или физических науках. Что-то может быть деревом или растением, а кто-то может болеть туберкулёзом, даже если никто не думает при этом; «вот дерево», или «вот растение», или «вот случай

туберкулёза», и даже если вообще никто ничего не думает. В случае же социальных феноменов многие термины, которые их описывают, сами должны входить в состав этих феноменов. Таким терминам присуща особого рода самореферентность. Термин «деньги» указывает на то, что используется и рассматривается в качестве денег. Термин «обещание» указывает на то, что люди намереваются сделать и рассматривают в качестве обещаний. Дело не в том, что для того, чтобы иметь институт денег, люди обязательно должны иметь в своём словаре слово «деньги». Но у них должны быть определённые мысли и установки в отношении чего-то, чтобы это что-то стало деньгами; и эти мысли и установки входят в качестве составных частей в само определение денег.

Ещё одно важное следствие заключается в том, что никаких физических пределов тому, что может рассматриваться в качестве физической реализации таких социальных феноменов, не устанавливается. А это означает, что невозможны никакие систематические связи между физическими свойствами феномена и его социальными или ментальными свойствами. Социальные свойства отчасти детерминируются нашими собственными установками в отношении этих свойств. Эти установки никак не ограничиваются физическими свойствами феноменов. Поэтому никакого соответствия ментального уровня и уровня физического, необходимого для того, чтобы сделать возможными строгие законы социальных наук, не существует.

Главным аргументом в пользу радикального разрыва между социальными и естественными науками является ментальный характер социальных феноменов. Но именно это-то их свойство и игнорируется.

Рассмотрим утверждение Фодора о том, что социальные законы будут иметь исключения, поскольку феномены на социальном уровне свободно или дизъюнктивно проецируются на физические феномены. Но даже если такого рода дизъюнкция и существует, обычно имеется возможность что-то к ней добавить. Деньги, скажем, всегда имели ограниченное число физических форм — они существовали в виде золота, серебра и бумажных чеков. Но возможно ведь, чтобы какой-то человек или какое-то общество стали считать деньгами нечто совсем иное. Физическая реализация не имеет значения для свойств денег, пока не препятствует им выступать в качестве средства обмена.

Кто-то может возразить: «Для того чтобы иметь строгие законы социальных наук, не требуется строгого соответствия свойств всего на свете. Требуется лишь строгое соответствие между психологическими свойствами и свойствами мозга. Обосновать экономику и социологию необходимо не через обращение к свойствам объектов вокруг нас, а через обращение к физическим

свойствам мозга. Так что даже если мысль о том, что нечто является деньгами, существенно важна для того, чтобы это нечто действительно было деньгами, всё же такая мысль может быть и с вашей точки зрения действительно является мозговым процессом. Поэтому, чтобы показать, что строгие законы социальных наук невозможны, вы должны показать невозможность строгих корреляций типов ментальных состояний и типов состояний мозга. Однако вы этого пока не показали».

Чтобы понять, почему такие законы невозможны, рассмотрим области, в которых ожидается появление некой строгой нейропсихологии, строгих законов, коррелирующих ментальные и нейрофизиологические феномены. Рассмотрим боль. Разумно было бы предположить, что нейрофизиологические причины боли, по крайней мере у человеческих существ, достаточно специфичны. Мы уже обсуждали некоторые из них в предыдущей главе. И, по-видимому, нет никаких принципиальных препятствий, чтобы появилась некая точная нейрофизиология боли. То же самое и со зрением. Трудно увидеть какие-либо принципиальные препятствия на пути к созданию адекватной нейрофизиологии зрения. Мы могли бы даже дать точное описание нейрофизиологических условий появления определённых видов зрительного опыта. Например, опыта видения чего-то красного. В моём подходе нет ничего, что препятствовало бы появлению такой нейрофизиологической психологии.

Но здесь нас ожидает главная трудность: хотя мы и можем установить систематические корреляции нейрофизиологии и боли, или нейрофизиологии и зрительного восприятия, мы не можем дать такого же объяснения деньгам. Почему? Всякий раз, как вы видите перед собой деньги, происходит некий нейрофизиологический процесс. Почему же не предположить, что это один и тот же процесс? Из того факта, что деньги способны выступать в неограниченном количестве физических форм, следует, что они могут оказывать неограниченное количество стимульных воздействий на нашу нервную систему. Но если деньги вызывают в нашей зрительной системе неограниченное количество изображений, то о каком же одинаковом нейрофизиологическом воздействии на мозг может идти речь?

То, что присуще видению чего-либо в качестве денег, в ещё большей степени присуще убеждению в том, что это деньги. Невозможно, чтобы всякий раз, когда кто-либо убеждён, что ему не хватает денег, это производило один и тот же тип нейрофизиологической реализации. Количество возможных нейрофизиологических стимулов, способных породить убеждение, бесконечно. Парадокс заключается в том, что сам способ, каким ментальное

воздействует на физическое, препятствует появлению строгой науки о сфере ментального.

Следует отметить, что в тех случаях, когда нет подобного взаимодействия социальных и физических феноменов, отсутствуют и препятствия к появлению строгих социальных наук. Рассмотрим гипотезу Хомски об универсальной грамматике. Допустим, каждый из нас обладает врождёнными правилами универсальной грамматики. Поскольку эти правила запрограммированы в мозге изначально и не зависят от каких-либо отношений организма со средой, ничто не запрещает, чтобы существовали строгие психофизические законы, связывающие эти правила и свойства мозга. Далее, многим животным присущи сознательные ментальные состояния, но у них отсутствует самореференциальность, сопровождающая человеческие языки и социальные институты. Ничто не запрещает, чтобы существовала наука о поведении животных. Например, могут существовать строгие законы, устанавливающие корреляцию состояний мозга птиц и их поведения, связанного со строительством гнёзд.

Я обещал рассказать вкратце о своих аргументах, Посмотрим, сдержал ли я своё обещание. Представим процесс аргументации как ряд шагов.

Чтобы появились законы социальных наук, подобные законам физики, должна иметь место систематическая корреляция феноменов, идентифицируемых в социальных и психологических терминах, и феноменов, идентифицируемых в физических терминах. Она может быть такой же сложной, как связь погодных явлений с явлениями физическими, Говоря языком сегодняшнего дня, должны быть некие «принципы», соединяющие высшие и низшие уровни.

Социальные феномены в значительной степени определяются психологическими установками. То, что считается деньгами, обещанием или женитьбой, во многом зависит от мыслей об этих вещах как о деньгах, обещании или женитьбе.

Из этого следует, что данные категории физически не определены. Строго говоря, отсутствует какой-либо физический предел тому, что мы могли бы рассматривать или обозначать в качестве денег, обещания или церемонии бракосочетания. А это означает, что не может быть никаких принципов, «соединяющих» социальные и физические свойства мира, то есть соединяющих феномены, описываемые в социальных терминах, и те же феномены, но описываемые в физических терминах. У нас даже не может быть тех нестрогих дизъюнктивных принципов, которые управляют погодой или пищеварением.

Далее, невозможно получить принципы, «соединяющие» феномены, описываемые в ментальных терминах, и феномены, описываемые в нейрофизиологических терминах, то есть мозг и сознание. Это невозможно потому, что для любого данного социального понятия имеется бесконечный спектр стимульных условий. И этот спектр препятствует такой реализации невстроенных в нас понятий, которая устанавливала бы систематическую корреляцию ментальных и физических свойств.

В заключение данной главы изложу свой взгляд на социальные науки. Социальные науки имеют дело с различными аспектами интенциональности. Экономика занимается производством и распределением товаров и услуг. Экономист может счесть интенциональность за нечто само собой разумеющееся. Он исходит из того, что предприниматели стремятся делать деньги, а потребители предпочитают быть богатыми, а не бедными. «Законы» экономики в таком случае устанавливают систематические результаты, или следствия этих предпосылок. Из этих предпосылок можно вывести, что предприниматели станут продавать там, где предельная цена будет давать предельный доход. Закон не предсказывает вопросов, которые бизнесмен задаёт сам себе: «Продаю ли я там, где предельная цена равна предельному доходу?» Закон не устанавливает содержания интенциональности. Скорее, он говорит о следствиях индивидуальной интенциональности. Теория фирм в микроэкономике разрабатывает следствия, вытекающие из определённых допущений, касающихся желаний и возможностей потребителей и предприятий, которые участвуют в закупках, производстве и сбыте. Макроэкономика выводит такие следствия в отношении целых наций и обществ. Но экономист не занимается вопросами типа: «чем в действительности являются деньги?» или «чем в действительности является желание?» Если он занимается экономикой благосостояния, он может изучать в деталях желания предпринимателей и потребителей, но даже в этом случае систематическая часть экономической дисциплины заключается в выведении следствий из тех фактов, которые касаются интенциональности.

Поскольку экономика базируется не на системе фактов о физических свойствах, таких как молекулярная структура, но скорее на фактах, касающихся человеческой интенциональности, желаний, деятельности, состояния технологии и знания, из этого следует, что она не свободна от истории и контекста. Экономика как наука предполагает определённые факты, касающиеся людей и обществ, — это факты истории, которые сами по себе не являются составной частью экономики. Когда меняются эти факты, меняется и экономика. Например, до недавнего времени казалось, что кривая Филипса, формула, связывающая ряд факторов, которые действуют в индустриальных обществах, даёт правильное описание экономических реалий этих обществ. Но теперь оказалось, что она неадекватна. Большинство экономистов

полагают, что кривая недостаточно точно описывала реальность. Правда, можно было бы предположить, что она точно описывала ту реальность, какая имела место в своё время. После нефтяного кризиса и Других событий семидесятых годов реальность изменилась. Экономика — это систематическая формализованная наука, но она зависит от контекста и от истории. Её предметом являются человеческие практики, но эти практики сами не являются вневременными, вечными или неизбежными. Если по какой-то причине деньги будут делаться из льда, то строгим законом экономики окажется то, что деньги тают при температуре выше нуля по Цельсию. Этот закон будет действовать до тех пор, пока деньги будут делать из льда, хотя он ничего не говорит о самом интересном, что есть в деньгах.

Теперь вернёмся к лингвистике. Общепризнанной целью лингвистики является установление различных правил — фонологических, синтаксических и семантических, — соотносящих звуки и смыслы в естественных языках. В идеале лингвистика — это полный набор правил естественного языка. Правильна ли эта цель, достижима ли она — не уверен, для нас важно, что лингвистика — это прикладная наука, имеющая дело с интенциональностью и ничем не напоминающая химию или геологию. Лингвистика занимается теми исторически детерминированными интенциональными содержаниями в сознании людей, говорящих на различных языках, которые обуславливают лингвистическую компетенцию людей. Как и в экономике, ключевым моментом в лингвистике является человеческая интенциональность.

Сформулируем теперь основной вывод данной главы. Радикальный разрыв между социальными и естественными науками не обусловлен существованием дизъюнктивной связи социальных и физических феноменов. Не обусловлено это и тем, что социальные дисциплины содержат конститутивные понятия, не имеющие «отзвука» в физике; не является причиной и чрезвычайная сложность социальной жизни. Многие дисциплины, такие как геология, биология и метеорология, обладают этими свойствами, но не перестают быть систематическими естественными науками. Радикальный разрыв есть следствие существенно ментального характера социальных и психологических феноменов.

Тот факт, что социальные науки имеют своим предметом сознание, является источником их слабости. Но в этом также и источник их силы. То, что мы стремимся получить от социальных наук и что в лучшем случае от них получаем, — это теории чистой и прикладной интенциональности.

ГЛАВА 6.

СВОБОДА ВОЛИ

На страницах этой книги я пытался ответить на некоторые грудные вопросы, касающиеся места человеческих существ во Вселенной. Наше представление о себе как о свободных деятелях лежит в основе нашего представления о самих себе. В идеале хотелось бы сохранить и точку зрения здравого смысла, и точку зрения науки. Думается, в отношении сознания и тела мне это удалось. Что касается примирения свободы и детерминизма, то здесь я, как и многие другие философы, потерпел поражение.

Все говорит за то, что после 2000 лет обсуждения проблема свободы воли к настоящему времени должна быть решена. Многие философы так и считают. Проблема была решена, считают они, Томасом Гоббсом и Давидом Юмом, а также другими эмпирически ориентированными философами, в XX же столетии эти решения были усовершенствованы. Лично я думаю, что эта проблема так и осталась нерешённой. В данной лекции я постараюсь разъяснить её суть и показать, почему современное решение является неадекватным. В заключение я попытаюсь объяснить, почему она скорее всего так и не найдёт своего решения.

Природа состоит из частиц и их отношений, и всё может быть объяснено в терминах этих частиц и их отношений, поэтому для свободы воли просто не остаётся места. Не имеет никакого значения, является ли физика, подобно ньютоновской физике, детерминистической или же допускает индетерминизм на уровне физики элементарных частиц, как это имеет место в современной квантовой механике.

Индетерминизм вовсе не свидетельствует в пользу учения о свободе воли; во-первых, статистический индетерминизм на уровне частиц ещё не говорит об индетерминизме на уровне объектов, например человеческих тел; во-вторых, даже если в поведении физических частиц и присутствует элемент индетерминизма — даже если они предсказуемы лишь статистически, из этого ещё не следует человеческая свобода воли — из того факта, что частицы детерминированы статистически, не следует, что человеческое сознание может заставить статистически детерминированные частицы отклониться со своего пути. Индетерминизм ничего не говорит о существовании ментальной энергии человеческой свободы, способной двигать молекулами в тех

направлениях, по которым бы они в противном случае не двигались. Так что всё, что мы знаем о физике, склоняет нас к отрицанию человеческой свободы.

Ярче всего это выражено Лапласом: если бы идеальный наблюдатель знал положения всех частиц в некоторый данный момент, а также все законы, управляющие их движением, он мог бы предсказать и рассказать всю историю Вселенной. Некоторые из предсказаний современного квантово-механического Лапласа носили бы статистический характер, но это по-прежнему не оставляло бы места для свободы воли.

Но довольно о детерминизме. Обратимся к аргументу, который выдвигается в пользу свободы воли. Как отмечали многие философы, все мы знакомы с одним простым фактом: наши выборы, решения, рассуждения и познания небезразличны для нашего поведения. Например, мы сделали какую-то вещь; тем не менее мы прекрасно знаем, что могли сделать нечто совсем другое. Мы знаем это, потому что поступаем так, а не иначе в силу определённых причин. Но мы знаем, что были основания и для того, чтобы выбрать нечто другое. Можно сформулировать это и иначе. Самым обычным эмпирическим фактом является то, что наше поведение невозможно предсказать наподобие того, как мы предсказываем поведение объектов, скатывающихся по наклонной плоскости. Причина заключается в том, что мы можем действовать так, как нам того захочется. Человеческая свобода есть факт нашего опыта. Если необходимо какое-то эмпирическое доказательство, мы просто укажем, что мы в состоянии опровергнуть любые предсказания, которые кто-то захочет сделать в отношении нашего поведения. Если кто-то предсказывает какое-то моё действие, то я могу сделать нечто совершенно иное. Такого рода выбор не могут сделать ледники, спускающиеся по горным склонам, или шары, катящиеся по наклонным плоскостям, или планеты, движущиеся по своим эллиптическим орбитам.

Перед нами философская головоломка. С одной стороны, ряд очень веских аргументов приводит нас к заключению, что свободе воли нет места во Вселенной. С другой стороны, имеются аргументы, которые склоняют нас к заключению, что свобода воли существует, — об этом говорят факты нашего опыта.

Стандартное решение головоломки — тезис о совместимости свободы воли и детерминизма. Конечно, всё в мире детерминировано, но некоторые человеческие действия свободны. Утверждать это — не значит отрицать их детерминированность; просто эти действия не являются вынужденными. Никто не заставляет их совершать. Например, если человека заставляют что-то делать под дулом пистолета, или он страдает от какого-то невроза, то его

поведение несвободно. Но если он действует свободно, если он действует, как говорится, по своей собственной свободной воле, тогда его поведение свободно. Конечно, оно тоже детерминированно, по-скольку любой аспект поведения обусловлен действием физических сил. Так что свободное поведение существует лишь в маленьком уголке детерминированного мира, в котором отсутствуют определённые виды обстоятельств и вынужденности.

Этот взгляд обычно называется позицией совместимости (compatibilism). Думаю, что он неадекватен. Проблема свободы воли заключается не в том, имеются или нет внутренние психологические основания, которые причинно обуславливают наши действия подобно внешним физическим причинам и внутреннему принуждению. Проблема в том, являются ли эти причины достаточными для того, чтобы вещи происходили так, как они происходят. Можно поставить эту проблему и по-другому. Имеем ли мы право вообще говорить, что человек имел возможность поступить иначе в тех же самых условиях? Например, если известно, что человек предпочёл голосовать за консерваторов, мог ли он сделать иной выбор? Точка зрения «совместимости» не оставляет места для обыденного понятия о свободе воли. С этой точки зрения, поведение детерминировано таким образом, что оно не могло быть иным при тех же самых обстоятельствах. Всё, что произошло, было детерминировано. Просто дело заключается в том, что некоторые вещи были детерминированы внутренними психологическими причинами («причинами действия»), а не внешними силами или психотической вынужденностью. Так что проблема остаётся. Оправданно ли говорить о человеческом существе, что оно могло поступить как-то иначе?

Концепция «совместимости» не даёт ответа на вопрос: «Можно ли было поступить иначе в тех же условиях?», который не противоречил бы нашей вере в нашу свободную волю. Короче говоря, концепция «совместимости» отрицает субстанцию свободной воли, но сохраняет её словесную оболочку.

Поэтому начнём с начала. Мы уверены, что наша собственная свободная воля основывается на фактах человеческого опыта. Перед нами открывается несколько возможностей выбора, и мы размышляем, что лучше сделать, а затем приходим к решению и делаем то, что решили. Но можно ли положиться на этот опыт? Быть может, убеждение в том, что наш опыт свидетельствует в пользу учения о человеческой свободе, иллюзорно?

Рассмотрим такой пример. Под гипнозом испытуемый получает постгипнотическое внушение. Вы можете сказать испытуемому, чтобы он сделал какие-то самые тривиальные и безвредные вещи, например прополз по полу. После сеанса испытуемый принимает участие в беседе, сидит и пьёт

кофе; и вдруг он говорит нечто вроде: «Какой восхитительный пол в этой комнате!», или «Я хочу проверить, что это за ковер», или «Я предполагаю вложить деньги в ковровые покрытия и потому хотел бы изучить этот пол». И начинает ползать по полу. Такие случаи интересны тем, что испытуемый всегда приводит какую-то более или менее правдоподобную причину, объясняющую, что он собирается делать. Иначе говоря, с его собственной точки зрения, он действует свободно. У нас же имеются достаточные основания полагать, что его поведение отнюдь не свободно, что его объяснения неправильны, что его поведение было детерминировано заранее, что на самом деле он находился во власти постгипнотического внушения. Его поведение можно было предсказать заранее. Итак, поставить проблему детерминизма или, по крайней мере, один из аспектов этой проблемы можно было бы следующим образом: «Можно ли сказать, что всё человеческое поведение подобно поведению человека, совершающего поступки под действием постгипнотического внушения?»

С эмпирической точки зрения это маловероятно. Иногда на людей действует гипноз, иногда они находятся во власти бессознательных неконтролируемых побуждений. Но все ли люди так поступают? Все ли наше поведение детерминировано подобной психологической вынужденностью? Если трактовать психологический детерминизм как утверждение о фактах нашего поведения, оно явно ложно. Тезис психологического детерминизма состоит в том, что первоначальные психологические причины детерминируют всё наше поведение так же, как они детерминируют поведение находящегося под гипнозом субъекта или поведение наркомана, употребляющего героин. Согласно этому взгляду, всё наше поведение так или иначе носит психологически вынужденный характер. Однако имеющиеся у нас данные свидетельствуют о том, что этот тезис ложен. Мы обычно действуем, исходя из наших интенциональных состояний — убеждений, надежд, страхов, желаний и так далее, — и в этом смысле наши ментальные состояния действуют как причины. Но это отношение причины и следствия не является формой детерминизма. При тех же самых ментальных состояниях мы могли бы и не сделать того, что мы сделали. С другой стороны, гипноз и психологически вынужденное поведение обычно носят патологический характер, и их легко отличить от нормального свободного действия. С психологической точки зрения, возможность для существования человеческой свободы остаётся.

Но является ли это решение шагом вперёд по сравнению с позицией «совместимости?» Мы как будто снова говорим: «Да, все поведение детерминировано, но то, что мы называем свободным поведением, есть поведение, детерминированное рациональными мыслительными

процессами». Иногда сознательные, рациональные мысленные процессы не играют никакой роли, как в случае с гипнозом, по обычно всё происходит наоборот. Обычно мы говорим, что агент действительно свободен. Но ведь и нормальные рациональные мыслительные процессы являются детерминированными, как и всё остальное в мире. Итак, не пришли ли мы вновь к тому выводу, что всё, что делается, уже было записано в книге истории за миллиарды лет до того, как мы родились, и потому никакое наше действие не является свободным в философском смысле слова? Называть или не называть наше поведение свободным — вопрос чисто терминологический. Продолжаем же мы говорить о заходе солнца, хотя и знаем, что в буквальном смысле солнце никуда не заходит. Точно так же мы продолжаем говорить о свободной воле, хотя такого явления не существует.

Проверить философский или какой-то иной тезис можно, спросив: «Что от этого меняется? Чем будет отличаться мир в том случае, если этот тезис истинен, а не ложен?» Отчасти детерминизм привлекает нас тем, что соответствует реальному положению дел в мире, по крайней мере в мире физическом. Иначе говоря, окажись детерминизм истинным, мир развивался бы точно так же, как и сейчас, и только некоторые из наших представлений оказались бы ложными. Однако эти представления для нас значимы. Они связаны с убеждением, что мы могли бы поступить иначе, чем поступили на самом деле, а это, в свою очередь, связано с убеждениями, касающимися моральной ответственности и нашей природы как личностей. Но если бы тезис о свободной воле был истинным, то в наши представления о мире пришлось бы внести некоторые весьма радикальные изменения. Необходимо было бы постулировать существование внутри каждого из нас некоторой самости (self), способной влиять на причинный порядок природы. Иначе говоря, в нас должна была бы существовать некая внутренняя реальность (entity), способная изменять направление движения молекул. Не знаю, насколько разумен такой взгляд, но он явно не согласуется с тем, что мы знаем об устройстве мира из физики. И нет ни малейшего основания для того, чтобы отказаться от физической теории в пользу такого взгляда.

Итак, до сих пор мы, по-видимому, так и не пришли ни к какому результату в разрешении конфликта между детерминизмом и верой в свободу воли. Наука не оставляет места для свободы воли, а индетерминизм в физике не может служить свидетельством в её пользу. С другой стороны, мы не можем отказаться от веры в свободу воли. Так давайте ещё немного поразмыслим об этом.

Почему, с точки зрения современной науки, для свободы воли не остаётся места? Основные объяснительные механизмы в физике работают снизу вверх.

Иначе говоря, мы объясняем поведение такого феномена, как прозрачность стекла или текучесть воды, в терминах поведения микрочастиц, например молекул. Объяснение отношения сознания и мозга как раз и является примером такого механизма. Ментальные свойства и причинно обуславливаются, и реализуются в нейрофизиологических феноменах. Но имеет место и причинное обуславливание, которое идёт от сознания к телу, то есть своего рода перевёрнутая причинность; она имеет место, потому что феномены на высшем и низшем уровне происходят одновременно. Представим себе, что мне хочется вызвать выделение ацетилхолина вставочным нейроном в аксонных рецепторах моих моторных нейронов. Можно сделать это, решив поднять руку и затем действительно её подняв. В данном случае ментальное событие, — намерение поднять руку, — причинно обуславливает физическое событие, выделение ацетилхолина. Но перевёрнутая причинность работает только потому, что ментальные события базируются на нейрофизиологии. Так что, соответственно описанию причинных отношений, действующих сверху вниз, есть и другое описание тех же событий, где причинные отношения имеют место только внизу, то есть в сфере нейронов и нейронного возбуждения в синапсах и так далее. В этой концепции устройства природы вряд ли остаётся место для свободы воли, ибо в соответствии с такой концепцией сознание может воздействовать на природу только потому, что само является частью природы. В таком случае, подобно всей остальной природе, свойства сознания детерминированы базисными физическими микроуровнями.

Это — абсолютно фундаментальный момент, так что позвольте мне повторить сказанное ещё раз. Форма детерминизма, которая вызывает трудности, не является психологическим детерминизмом. Тезис о том, что наших состояний сознания достаточно для детерминации всего, что мы делаем, вероятно, просто ложен. Форма детерминизма, которая вызывает трудности, более важна и фундаментальна. Поскольку все поверхностные свойства мира полностью обусловлены системами микроэлементов и реализованы в них, их достаточно для детерминации всего, что происходит. Такая «вверх ногами картина мира допускает «вверх ногами» причинность (например, воздействие нашего сознания на наше тело). Но эта причинность действует только потому, что высший уровень уже обусловлен низшими уровнями и реализован в них.

Ну, а сейчас зададимся вопросом: что в нашем опыте не позволяет нам отказаться от веры в свободу воли? Если свобода — иллюзия, почему мы не можем от неё отказаться? Первое, на что надо обратить внимание, — это то, что свобода существенным образом связана с сознанием. Мы говорим о свободе только в отношении со-знательных существ. Например, робота, полностью лишённого сознания, мы никогда не назовём свободным. Даже

если мы обнаружим, что его поведение произвольно и непредсказуемо, мы всё же не скажем, что он действовал свободно. Если же кто-то сконструирует робота, обладающего сознанием в том же смысле, в каком мы сами им обладаем, вопрос о его свободе останется по крайней мере открытым. Во-вторых, следует отметить, что в существовании свободы убеждает не просто любое состояние сознания. Если бы жизнь состояла только в пассивном получении восприятий, то мы бы никогда не смогли сформулировать идей человеческой свободы. Вообразите себя совершенно неподвижным, неспособным определять даже направление собственных мыслей, но всё же воспринимающим стимулы, например периодические лёгкие болевые ощущения. У вас не будет никакого основания для заключения, что вы обладаете свободой воли.

Большинство философов полагают, что убеждение в существовании человеческой свободы существенным образом связано с процессом рационального принятия решений. Но это верно только отчасти. Фактически, взвешивание оснований — лишь частный случай того опыта, который убеждает нас в существовании свободы. Опыт, который нас действительно убеждает, — это опыт участия в произвольных, интенциональных человеческих действиях. Обсуждая интенциональность, мы сосредоточили внимание на сознательных интенциях в действии, на той форме интенциональности, которая является каузальной в вышеуказанном смысле и условия удовлетворения которой заключаются в том, что происходят некоторые телесные движения, и происходят они как причинно обусловленные самой интенцией в действии. Именно этот опыт и служит краеугольным камнем для нашей веры в свободу воли. Почему? Поразмыслим о характере того опыта, который у нас есть, когда мы принимаем участие в нормальных, повседневных человеческих действиях. Мы почувствуем возможность альтернатив, встроенных в этот опыт. Поднимите руку, или пройдитесь по комнате, или возьмите стакан воды, и вы почувствуете в этом опыте наличие альтернатив.

Различие опыта восприятия и опыта действия заключается в том, что в восприятии человек имеет ощущение: «Это происходит со мной», а в действии человек чувствует: «Я действую таким образом, чтобы это произошло». Но в ощущении «я действую таким образом, чтобы это произошло» содержится также ощущение «я мог бы сделать и что-то другое». В случае нормального поведения всё, что мы делаем, содержит в себе убеждение, верное или неверное, что здесь и сейчас, в тех же самых условиях, мы могли бы делать и что-то другое. В этом — источник нашей непоколебимой веры в наличие у нас свободной воли. Следует подчеркнуть, что речь идёт о нормальном человеческом действии. Если кто-то находится под влиянием сильной страсти,

например сильного гнева, он теряет это ощущение свободы и иногда сам удивляется тому, что делает.

Как только мы это замечаем, легко объяснимыми становятся и многие из тех загадочных явлений, о которых упоминалось ранее. Например, почему мы чувствуем, что человек, находящийся в состоянии постгипнотического внушения, не свободен, несмотря на его полнейшую уверенность в собственной свободе? Дело в том, что в одном весьма существенном смысле он не знает, что делает. Его интенция-в-действии является бессознательной. Выбор, который, как ему представляется, для него доступен, не имеет отношения к действительной мотивации его поступков. Следует также отметить, что примеры «вынужденного» поведения во многих случаях всё же предполагают опыт свободы. Если мне под дулом пистолета приказывают что-то сделать, то даже в этом случае в моём опыте присутствует ощущение альтернативы. Если, например, под дулом пистолета мне приказывают пройти по комнате, в буквальном смысле слова на каждом шагу передо мной открыта возможность совершить какое-то другое действие. Таким образом, опыт свободы является существенным компонентом любого случая интенциональной деятельности.

Можно убедиться в этом, сопоставив нормальное действие с экспериментами Пенфилда, в которых воздействие на моторную кору вызывает произвольное движение руки или ноги. В этом случае испытуемый воспринимает движение пассивно, подобно тому как он воспринимал бы звук или ощущение боли. В отличие от интенциональных действий, здесь возможности выбора в опыт не встроены. Вообразите, что ваша жизнь представляет собой как бы увеличенный в масштабах эксперимент Пенфилда. Вы обнаруживаете, что ваше тело движется по комнате; вы слышите слова, вами произносимые. В этом случае ваш опыт окажется опытом совершенно пассивной, но сознательной куклы. Вообразив это, вы отсечете в этом своём воображении опыт свободы. Но в типичном интенциональном действии опыт свободы неустраним. Это — существенный момент опыта деятельности, Это объясняет и то, почему мы не можем отказаться от нашей веры в существование свободы. Мы легко откажемся от веры в то, что Земля плоская, если поймём аргументы в пользу гелиоцентрической теории.

Подобно этому заход Солнца есть просто иллюзия, создаваемая вращением Земли. В каждом из этих случаев от точки зрения здравого смысла можно отказаться, потому что заменяющая её гипотеза объясняет и тот опыт, который привёл к этой точке зрения, и великое множество других фактов, объяснить которые здравый смысл неспособен. Вот почему мы отказались от предположения о том, что Земля плоская, и от «закатов» Солнца в пользу

коперниканской концепции солнечной системы. Но мы не можем отказаться от веры в свободу, потому что свобода включена во всякое нормальное, сознательное интенциональное действие. Это убеждение служит нам для идентификации и объяснения действий. Чувство свободы не есть просто выбор. Это часть любого действия, задуманного заранее или осуществляемого спонтанно. В сущности, это не обдумывание; обдумывание — всего лишь частный случай чувства свободы.

Мы не пускаемся в мореплавание, исходя из предположения о том, что Земля плоская; но мы действуем, исходя из предпосылки о существовании свободы. Мы вообще не можем действовать иначе, и это никак не зависит от наших знаний об устройстве мира как детерминированной физической системы.

Теперь можно сформулировать заключения, всё время неявно присутствовавшие в нашем обсуждении. Во-первых, если трудности с детерминизмом сводятся к тому, что в этом случае всё наше повеление оказывается психологически вынужденным, то эти трудности неоправданны. В той мере, в какой психологический детерминизм есть эмпирическая гипотеза, подобная любой другой гипотезе, имеющиеся у нас сегодня свидетельства говорят о том, что он ложен. Таким образом, мы имеем дело с модификацией концепции «совместимости», согласно которой психологический либертарианизм и физический детерминизм не противоречат друг другу.

Во-вторых, это даёт нам ощущение «мог бы»; поведение человека детерминировано, однако он мог бы поступить и иначе; речь здесь идёт о психологических факторах. Понятие способности — того, что мы можем делать и что мы могли бы делать, подчиняется определённым критериям. Например, я мог бы голосовать за Картера на американских выборах 1980 года. Но я не мог бы голосовать за Джорджа Вашингтона, Так что в соответствии со смыслом «мог бы» мне доступен самый широкий выбор возможностей, и в этом смысле я мог бы сделать множество вещей, которых, однако, я так и не сделал. Подобно этому, поскольку воздействующие на меня психологические факторы не всегда и не во всём вынуждают меня к определённому поведению, я часто мог бы делать совсем не то, что я фактически делал.

В-третьих, эта форма концепции «совместимости» так и не разрешает конфликт между свободой и детерминизмом. Пока мы придерживаемся концепции физического объяснения «снизу — вверх», — а именно на этой концепции последние триста лет основывалась наука, — психологические факты, как и любые другие факты высшего уровня, полностью объяснимы в

терминах (и реализуемы в системах элементов) фундаментального микрофизического уровня. Наша концепция физической реальности просто не допускает существования радикальной свободы. Наконец, в-четвёртых, по причинам, которые мне не очень понятны, эволюция предоставила нам такой опыт произвольного действия, в котором опыт свободы, так сказать опыт ощущения альтернативных возможностей, встроен в саму структуру сознательного, произвольного, интенционального человеческого поведения. Поэтому никакие аргументы никогда не убедят нас, что наше поведение несвободно.

В данной книге я попытался описать отношения между нашим представлением о себе как о рациональных, свободных, сознательных, разумных деятелях и нашим представлением о мире как совокупности лишённых сознания и смысла физических частиц. Можно предположить, что подобно тому, как здравый смысл в значительной мере неадекватен реальному устройству мира, так же и наше представление о самих себе и собственном поведении является совершенно ложным. Однако различие реальности и кажимости неприменимо к самому существованию сознания. Ибо если мне кажется, что я сознательное существо, то я и на самом деле сознательное существо. Мы способны открыть самые поразительные вещи о самих себе и о своём поведении; но мы не способны открыть отсутствия у себя сознания, мы не способны открыть, что сознание не содержит сознательных, субъективных, интенциональных ментальных состояний. И мы никогда не обнаружим отсутствия у самих себя попыток произвольного свободного интенционального действия. Проблема состоит не в том, чтобы доказать существование всех этих вещей, но в том, чтобы определить их статус и последствия, которые они имеют для наших представлений о природе. Мой общий вывод состоит в том, что, за некоторыми существенно важными исключениями, наше обыденное менталистское представление о самих себе прекрасно согласуется с нашим представлением о природе как физической системе.

