

Б. П. ПОНУКОВ

ИИИ  
11111

ПОСМОТРЕТЬ

УМ

Хороля

11111

# Р. Пенроуз "Новый ум короля"

## Глава 9

### Реальный мозг и модели мозга

#### Как же устроен мозг?

У нас в голове находится великолепное устройство, которое управляет нашими действиями и каким-то образом дает нам представление об окружающем мире. Правда, как однажды отметил Алан Тьюринг [1], внешне оно больше всего напоминает миску холодной овсянки! Трудно представить, как столь заурядного вида объект умудряется совершать чудеса, на которые, как мы знаем, он способен. Однако при ближайшем рассмотрении оказывается, что он имеет гораздо более сложное строение и замысловатую организацию.

Большая покрытая извилинами (и более всего похожая на овсянку) часть мозга, расположенная сверху, называется собственно *головным (или большим) мозгом*. Он четко делится посередине на правое и левое *полушария* и, более условно, в передне-заднем направлении — на лобную долю и три остальные: височную, теменную и затылочную. Еще дальше и несколько книзу расположен небольшой и округлый *мозжечок*, чем-то похожий на пару клубков шерсти. Глубоко внутри мозга, как бы укрытый им, находится целый ряд любопытных и сложных на вид структур: варолиев мост и продолговатый мозг, которые вместе с ретикулярной формацией — областью, к которой мы обратимся позднее — составляют ствол мозга, а также таламус, гипоталамус, гиппо-камп, мозолистое тело и еще много других, странных как по виду, так и по названиям, частей.

Большой мозг — предмет особой гордости человека и не только потому, что он является самой большой частью человеческого мозга, но и потому, что пропорция между этой частью и мозгом в целом у *человека* больше, чем у животных. (Мозжечок человека тоже превосходит размерами таковой у большинства других животных.) Головной мозг и мозжечок имеют сравнительно тонкий наружный слой *серого вещества*, под которым расположено значительно большее по массе *белое вещество*. Эти области серого вещества называют, соответственно, корой головного мозга и корой мозжечка. Считается, что в сером веществе происходят различные вычислительные действия, а белое вещество, состоящее из длинных нервных волокон, отвечает за передачу сигналов из одной части мозга в другую.

Каждой из различных областей коры головного мозга присущи свои специфические функции. Зрительная кора расположена в затылочной доле, прямо в задней части мозга, и занимается восприятием и распознаванием зрительных образов. Забавно, что природа именно там решила разместить интерпретатор визуальной информации, получаемой зрительными органами, которые (по крайней мере, у человека) находятся прямо спереди! Но природа вытворяет и куда более странные вещи. Так, за *левую половину* человеческого тела практически полностью отвечает *правое* полушарие, тогда как за правую — почти исключительно левое, поэтому чуть ли не все нервы, идущие в головной мозг или выходящие из него, по необходимости должны перекрещиваться! При этом в случае зрительной коры правая ее часть связана не с левым глазом, а с *левой частью поля зрения обоих глаз*. Аналогично, левая часть зрительной коры связана с правой частью поля зрения обоих глаз.

Это означает, что нервы от правой части сетчатки каждого из глаз должны идти к правой половине зрительной коры (вспомните, что изображение на сетчатке перевернуто по

отношению к источнику), а нервы от левой части сетчатки — к левой половине коры. Таким образом в левой и правой частях зрительной коры формируется четкое отображение правой и левой областей поля зрения, соответственно.

Сигналы от ушей приходят на противоположные части мозга столь же замысловатым образом. Правая слуховая кора (часть правой височной доли) обрабатывает в основном звуки, поступающие слева, а левая слуховая кора — звуки, поступающие справа. Обоняние кажется здесь исключением из общего правила. Правая часть обонятельной коры, которая расположена в передней части большого мозга (в передней доле — что уже само по себе является исключением для сенсорной области), отвечает в основном за правую ноздрю, а левая часть — за левую ноздрю.

*Осязание* связано с областью затылочной доли мозга, которая носит название *соматосенсорной коры*. Эта область находится как раз за условной границей, разделяющей лобную и теменную доли. Между различными частями поверхности тела и отдельными участками соматосенсорной коры существует довольно своеобразное соответствие. Иногда оно изображается графически в виде так называемого «соматосенсорного гомункулуса» — искаженной человеческой фигуры, изображаемой лежащей вдоль соматосенсорной коры.

Правая часть соматосенсорной коры принимает осязательные сигналы, идущие от левой стороны тела, а левая — с правой. В лобной доле непосредственно *перед* границей с теменной долей находится участок коры, известный как *двигательная кора*. Он приводит в движение различные части нашего тела. И опять мы встречаемся с точно определенным соответствием между мышцами нашего тела и зонами этого участка мозга. Как и в случае с осязанием, эти связи можно графически изобразить в виде «двигательного гомункулуса». И снова правая часть двигательной коры отвечает за движение левой стороны тела, а левая — правой.

Все упомянутые выше зоны коры головного мозга (зрительная, слуховая, обонятельная, осязательная и двигательная) называются *первичными*, поскольку именно они непосредственно осуществляют прием поступающих в мозг и передачу исходящих из него сигналов. Рядом с ними расположены *вторичные* зоны, предназначенные для более тонкой и сложной обработки сенсорной информации.

Сенсорная информация, полученная зрительной, слуховой или соматосенсорной зоной коры головного мозга, обрабатывается соответствующими вторичными областями, после чего вторичная двигательная область вырабатывает план движения, который переводится первичной двигательной областью на язык прямых команд, непосредственно адресованных мышцам. (Мы не будем касаться обонятельного участка коры, поскольку он функционирует иным и малоизученным пока образом.) Остальные участки коры головного мозга относятся к разряду *третичных* (или *ассоциативных*). В этих областях в основном и выполняется наиболее сложная и характеризующаяся высокой степенью абстрагирования часть умственной деятельности. Именно здесь при определенном участии периферической нервной системы собирается воедино и подвергается всестороннему анализу информация, поступающая от различных сенсорных участков; здесь происходит запоминание, складываются картины внешнего мира, намечаются и оцениваются планы действий, распознается и генерируется речь.

Речь представляет для нас особый интерес, поскольку ее обычно относят к разряду способностей, присущих исключительно человеческому интеллекту. Интересно, что (по крайней мере у подавляющего большинства правшей и большей части левшей) *речевые* центры находятся в основном в *левой* половине мозга. К важным участкам относятся *зона Брока*, расположенная в задней нижней части лобной доли, и *зона Вернике*, которая располагается внутри и вокруг верхней задней части височной доли. Зона Брока отвечает за

построение предложений, а зона Вернике — за понимание языка. Повреждение зоны Брока приводит к нарушению речи, но не ее пониманию, тогда как при повреждении зоны Вернике речь остается беглой, но, в основном, бессмысленной. Пучок нервных волокон, который связывает между собой две эти области, называется *дуговидным пучком*. При его повреждении ни речь, ни ее понимание не нарушаются, но мысль не может быть выражена словами.

Мы теперь можем составить очень приблизительную картину того, что делает головной мозг. *Входные данные* для мозга представляют собой зрительные, слуховые, осязательные и прочие сигналы, которые сначала регистрируются в первичных областях (главным образом) *задних долей* (теменной, височной и затылочной). *Выходные сигналы* мозга, приводящие к различным движениям тела, вырабатываются в основном *лобными* долями мозга. А где-то между ними происходит обработка информации и принятие решений. В общем, можно сказать, что активность мозга, начавшись в первичных областях задних долей, перемещается затем во вторичные области, где входные данные анализируются, и, далее, в третичные области задних долей, где информация становится полностью осмысленной (как, например, в случае с пониманием речи в зоне Вернике).

*Дуговидный пучок* — упомянутый выше пучок нервных волокон, но теперь уже с обеих сторон мозга, — переносит эту информацию в лобную долю, где ее третичными областями вырабатывается общий план действий (например, как это происходит при генерации речи в зоне Брока). Эти общие планы действий преобразуются в более конкретные представления о движениях тела во вторичных двигательных областях, откуда активность мозга перемещается в первичную двигательную кору, которая, в конце концов, посылает соответствующие сигналы различным группам мышц тела (и часто несколькими одновременно).

Создается впечатление, что перед нами предстает картина превосходного вычислительного устройства. Сторонники сильного ИИ (см. главу 1 и далее) рассматривают мозг как великолепный образец алгоритмического компьютера - по сути, машины Тьюринга — в котором есть входные данные (как на ленте слева от машины Тьюринга) и выходные данные (как на ленте справа от машины Тьюринга) и который способен выполнять всевозможные нетривиальные вычисления на промежуточных этапах. Конечно, активность мозга может не прекращаться и в отсутствие внешних раздражителей. Это происходит в тех случаях, когда человек думает, занимается вычислениями или предаётся воспоминаниям. Приверженцы сильного ИИ отнесли бы это на счет продолжающейся алгоритмической деятельности и предположили бы, что явление «осознания» возникает как раз в те моменты, когда подобная деятельность достигает определенного уровня сложности.

Но, хотя такая логика и напрашивается сама собой, мы не будем торопиться с выводами. Общая картина работы мозга, приведенная выше, довольно груба. Прежде всего, даже зрительное восприятие не происходит по такой простой схеме, как это было мной представлено ранее. В коре, по-видимому, существует несколько различных (хотя и менее значимых) областей, на которые отображаются поля зрения, очевидно, с какими-то другими целями. (Похоже, именно они отвечают за различия в том, как мы *осознаем* увиденное). Скорее всего, по коре разбросаны также и другие дополнительные сенсорные и двигательные области (например, движение глаз может быть вызвано сигналами из определенных точек *задних долей*).

В своем описании мозга я затронул только его кору и ни разу не коснулся вопроса о назначении прочих частей. Какую роль выполняет, например, *мозжечок*? Ясно, что он отвечает за координацию и контроль движений тела, его равновесие, своевременность и точность действий. Представьте себе артистичность танцора, отточенность движений профессионального игрока в теннис, мгновенную реакцию гонщика, уверенные движения рук

музыканта или художника; подумайте о грациозных прыжках газели или крадущейся кошке. Без мозжечка подобная точность движений была бы невозможна, они стали бы неуверенными и неуклюжими. По-видимому, в процессе приобретения новых навыков, будь то ходьба или вождение машины, сначала человеку приходится детально обдумывать каждое свое действие, и за это отвечает кора головного мозга, но когда достигнут определенный уровень мастерства и действия начинают выполняться «автоматически», управление ими передается мозжечку. Более того, хорошо известно, что как только профессионал *задумывается* о своих действиях, он на время теряет легкость их координации. *Думание*, по-видимому, сопровождается переходом контроля к коре головного мозга и, хотя при этом, как следствие, появляется гибкость действий, «мозжечковая» плавность и точность движений на время утрачивается. Такое описание, без сомнения, является чересчур упрощенным, но тем не менее позволяет нам в общих чертах понять функцию мозжечка).

При описании функций головного мозга до сих пор вообще не упоминалось о других частях мозга. Например, *гиппокамп* играет важнейшую роль в формировании долговременной (постоянной) памяти, хотя сама память располагается где-то в коре головного мозга, возможно, во многих местах одновременно. Мозг способен также сохранять образы различными способами с помощью кратковременной памяти в течение нескольких минут или даже часов (просто, что называется, «держит их в голове»). Но для того, чтобы человек мог вспомнить эти образы после того, как его внимание с них переключилось, необходимо сохранить их в долговременной памяти, и здесь уже не обойтись без гиппокампа. (Повреждение этого участка мозга приводит к ужасному состоянию, когда человек не способен запомнить ничего нового и все сразу забывается, как только его внимание переключается на другой объект.) *Мозолистое тело* — это область, ответственная за связь между двумя полушариями мозга. (Далее мы увидим, к каким поразительным явлениям приводит рассечение мозолистого тела.) *Гипоталамус* представляет собой эмоциональный центр, в котором гнездятся удовольствие, ненависть, страх, отчаяние, голод, и который служит посредником между эмоциями и их ментальными и физическими проявлениями. Между гипоталамусом и различными частями мозга идет постоянный обмен сигналами. *Таламус* функционирует как важный обрабатывающий центр и переключающий узел, который передает значительную часть импульсов, поступающих извне, в кору головного мозга. *Ретикулярная формация* отвечает за общее состояние готовности мозга и его отдельных частей к осознанному восприятию. Все эти и многие другие жизненно важные части мозга соединены многочисленными нервами.

Вышеприведенное описание дает только общее представление о некоторых наиболее значимых частях мозга. Мне кажется целесообразным в завершение этого раздела привести некоторые сведения о строении мозга в целом. Его различные части группируются в три отдела, которые, если двигаться от позвоночника, называются по порядку *задним* (*rhombencephalon*), *средним* (*mesencephalon*) и *передним* (*prosencephalon*) мозгом. На ранних стадиях развития эмбриона эти отделы, в том же порядке, видны как три вздутия на конце позвоночного столба. Самое дальнее — развивающееся в передний мозг — имеет два выроста в виде пузырей, по одному с каждой стороны, которые становятся большими полушариями головного мозга. Полностью развитый передний мозг включает в себя многие важные части всего мозга — не только большой головной, но и мозолистое тело, таламус, гипоталамус, гиппокамп и многие другие. Мозжечок является частью заднего мозга. Ретикулярная формация расположена частью в среднем мозге, а частью в заднем. Передний мозг является «новейшим» отделом с точки зрения эволюционного развития, а задний — наиболее «древним».

Я надеюсь, что это краткое описание, во многом неточное, даст читателю некоторое представление о том, на что похож мозг человека и как он функционирует. До сих пор я лишь

вскользь упомянул то, что служит центральной темой нашей дискуссии — *сознание*. Теперь перейдем к этому вопросу вплотную.

## Где обитает сознание?

Существует множество различных точек зрения на соотношение между состоянием мозга и феноменом сознания. Насколько очевидна важность этого явления, настолько же велико и расхождение во взглядах на него. Однако ясно, что не все части мозга в равной степени участвуют в формировании сознания. Например, как следует из вышесказанного, мозжечок по роду своей деятельности гораздо ближе к «автоматическому устройству», чем кора головного мозга. Действия, контролируемые мозжечком, происходят как будто сами собой и не требуют «обдумывания». Когда мы сознательно решаем пройти от одного места до другого, то вряд ли имеем перед собой тщательно разработанный план мышечных сокращений, который был бы необходим для управляемого движения. То же самое можно сказать и о бессознательных рефлекторных действиях, как, например, отдергивание руки от горячей печи, которое может быть опосредовано не головным мозгом, а верхней частью спинного мозга. Таким образом, напрашивается вывод о том, что феномен сознания, вероятнее всего, связан с активностью головного мозга, а не мозжечка или спинного мозга.

С другой стороны, совершенно не очевидно, что активность коры головного мозга всегда определяет осознанность наших действий. Например, как я уже указывал, в норме при ходьбе человек *не* контролирует детальные движения конечностей и работу мышц — управление этими действиями осуществляет, в основном, мозжечок (с помощью других частей головного мозга и спинного мозга), — однако первичные двигательные области головного мозга *тоже* вовлекаются в этот процесс. Более того, то же можно сказать и о первичных сенсорных областях: мы можем совершенно не осознавать меняющееся при ходьбе давление на подошвы ног, тем не менее соответствующие участки соматосенсорной коры постоянно активируются.

Уайлдер Пенфилд, выдающийся американско-канадский нейрохирург (среди заслуг которого — составление в 1940-х и 1950-х годах детальных карт двигательных и сенсорных областей мозга человека), считал, что сознание не связано просто с активностью коры головного мозга. На основании опыта проведения многочисленных операций на мозге пациентов, находившихся в сознании, он предположил, что область, которую он называл *верхней частью ствола мозга*, включающая, в основном, таламус и средний мозг (см. Пенфилд, Джаспер [1947]), хотя он имел в виду главным образом ретикулярную формацию, в некотором смысле может быть названа «центром сознания». Верхняя часть ствола мозга связана с корой головного мозга, и, согласно Пенфилду, «акт осознания» или «осознанное действие» происходит каждый раз, когда эта область ствола мозга непосредственно обменивается сигналами с определенным участком коры, отвечающим именно за те чувства, мысли, воспоминания или действия, которые в данный момент осознанно воспринимаются или совершаются. Он указывал, что можно, например, стимулировать определенный участок двигательной коры мозга, который отвечает за движение правой руки (и правая рука на самом деле будет двигаться), но это не вызовет у подопытного *желания* двигать правой рукой. (Более того: он может даже постараться остановить ее движение левой рукой - совсем как доктор Стрэнджлав из популярного фильма!) Пенфилд предполагал, что *желание* совершить действие связано скорее с таламусом, нежели с корой головного мозга. Согласно его представлениям сознание — это проявление активности верхней части ствола мозга, однако, поскольку должно еще быть что-то, *что* осознается, то эта активность не ограничивается стволом мозга, но включает в себя еще и те участки коры, с которыми у

верхней части ствола мозга в этот момент существует активная связь и которые представляют собой субъект (чувственное восприятие или воспоминание) или объект (волевое действие) сознания.

Другие нейрофизиологи тоже высказывали предположение о том, что ретикулярную формацию можно было бы назвать «местонахождением» сознания, если таковое на самом деле существует. Ведь, как бы там ни было, эта область отвечает за пребывание мозга в активном состоянии. Ее повреждение приводит к потере сознания. Всегда, когда мозг находится в бодрствующем сознательном состоянии, активна и ретикулярная формация, и наоборот. На самом деле существует явная связь между активностью ретикулярной формации и тем состоянием человека, которое мы традиционно называем «сознательным». Однако ситуация осложняется тем, что во сне, когда мы на самом деле «сознаем», что мы спим, активные в норме участки ретикулярной формации активности не проявляют. И еще один факт мешает ученым признать за ретикулярной формацией столь почетный статус: с точки зрения эволюции, эта часть мозга является очень *древней*. Если все, что нужно для обладания сознанием — активность ретикулярной формации, то им должны быть наделены лягушки, ящерицы и даже треска!

Лично я не расцениваю последний довод как достаточно весомый. Разве у нас есть неоспоримые свидетельства того, что ящерицы и треска *не* обладают неким зачаточным сознанием? Какое право мы имеем утверждать, как это некоторые делают, что человеческие существа — единственные обитатели нашей планеты, наделенные свыше настоящим «сознанием»? Неужели на Земле мы единственные, кому дозволено «осознавать»? Позвольте усомниться в этом. Конечно, лягушки, ящерицы и уж тем более треска не вызывают у меня ощущения, что «кто-то в них» взирает на меня, когда я рассматриваю эти создания, но я очень явственно ощущаю присутствие «сознания», когда смотрю в глаза кошке, собаке или, особенно, когда на меня смотрят обезьяны или мартышки в зоопарке. Я не требую, ни чтобы они чувствовали то же, что и я, ни даже какой-либо сложности испытываемых ими чувств. Им совершенно не обязательно «сознавать себя» в каком-то строгом смысле этого слова (хотя наличие некоторого элемента самосознания у них я не исключаю). Достаточно будет, чтобы они просто *чувствовали!* Что касается состояния сна, то я бы признал, что определенная форма сознания при этом присутствует, хотя, по всей видимости, на довольно низком уровне. Если за функционирование сознания каким-то образом отвечают только участки ретикулярной формации, то они должны сохранять активность (хотя бы невысокую) и во время сна.

Другая точка зрения (О'Кифи [1985]) состоит в том, что сознание в большей мере связано с функционированием *гиппокампа*. Как я уже отмечал, гиппокамп определяет способность к долговременному запоминанию. Принимая в качестве гипотезы, что постоянная память связана с сознанием, мы должны рассматривать гиппокамп как главное действующее лицо в феномене осознанного восприятия.

Есть и другое мнение, согласно которому сознание является результатом деятельности самой коры головного мозга. Раз уж большой головной мозг служит предметом особой гордости человека (хотя у дельфинов он никак не меньше!), и умственная деятельность, понимаемая как интеллект, связана как раз с этой частью мозга, то именно в ней и должна обитать душа человека! Таким, по-видимому, мог бы быть вывод, например, сторонников сильного ИИ. Если «осознание» - не более, чем следствие *сложности* алгоритма или, возможно, его «глубины» или некой «степени изошренности», — тогда, в соответствии с представлениями сильного ИИ, сложные алгоритмы, выполняемые *корой головного* мозга, дали бы ей *преимущественное право* претендовать на способность к проявлению сознания.

Многие философы и психологи склонны считать *язык* непременным атрибутом

человеческого сознания. Соответственно, именно способность изъясняться при помощи слов позволяет достичь той тонкости мышления, которая служит отличительной чертой человека и выражением самой его сути. Именно язык, в соответствии с этой точкой зрения, отличает нас от других животных и дает нам возможность *лишить их* свободы и вести на бойню, как только в этом возникает потребность. Именно язык позволяет нам философствовать и описывать наши ощущения, так что мы можем убедить остальных, что *мы* осознаем окружающий мир и самих себя. С этой точки зрения владение языком является необходимым и достаточным условием наличия сознания.

А теперь мы должны вспомнить о том, что языковые центры находятся (у большинства людей) в *левой* половине мозга (зоны Брока и Вернике). Из вышеизложенной точки зрения должно было бы следовать, что сознание - это что-то, что связано только с левой половиной коры головного мозга! И таково, на самом деле, мнение целого ряда нейрофизиологов (в частности, Джон Эклз [1973]), которое я, как сторонний наблюдатель, считаю весьма странным по причинам, изложенным ниже.

### **Эксперименты при разделенных больших полушариях мозга**

В связи со сказанным выше я должен упомянуть целый ряд замечательных наблюдений над людьми (и животными) при полном рассечении у них мозолистого тела, которое делало взаимодействие левого и правого полушарий головного мозга невозможным. У людей операция по рассечению мозолистого тела применялась как эффективное средство лечения в случаях особо тяжелых форм эпилепсии. Роджер Сперри с сотрудниками подвергал таких пациентов, спустя некоторое *время после операции, многочисленным психологическим тестам. При этом в левом и правом полях зрения испытуемых помещались никак не связанные друг с другом предметы, так что левое полушарие получало информацию только о том, что располагалось с правой стороны, а правое полушарие — с левой. Если справа предъявлялось изображение карандаша, а слева — чашки, то тестируемый произносил: «Это карандаш», поскольку именно карандаш, а не чашку, воспринимала та половина мозга, которая явно отвечает за речевые способности. Однако левой рукой испытуемый выбирал блюдце, а не лист бумаги, считая его ассоциативно более подходящим к чашке. Левая рука находилась «в подчинении» у правого полушария, которое, хотя и не могло оперировать словами, все же было способно производить определенные, довольно сложные и типичные для человека действия. Было высказано предположение о том, что за «геометрическое мышление» (особенно пространственное воображение) и музыкальное восприятие ответственно, в основном, *правое* полушарие, а за речевые и аналитические способности — левое. Правое полушарие мозга может понимать общеупотребительные существительные и элементарные предложения, а также выполнять простейшие арифметические действия.*

Самое поразительное, что при разделении полушарий они ведут себя как две практически независимые индивидуальности, с каждой из которых экспериментатор может общаться по отдельности, хотя общение с правым полушарием носит более примитивный характер и значительно затруднено по сравнению с левым из-за отсутствия речевых способностей. Каждая половина головного мозга может поддерживать связь с другой половиной косвенным путем, например, наблюдая за движениями руки, контролируемые другой стороной, или слыша звуковые «подсказки» (такие, как стук блюдца). Но в хорошо контролируемых лабораторных условиях даже эта примитивная связь может быть устранена. Однако от одной половины к другой все же могут передаваться неясные эмоциональные ощущения, предположительно потому, что нерассеченные структуры мозга, такие как гипоталамус, по-прежнему связаны с обоими полушариями.

Возникает искушение задать вопрос: неужели перед нами — два различных индивидуума, обладающих сознанием и пребывающих в одном теле? Этот вопрос вызвал бурную полемику. Одни без сомнений отвечали на этот вопрос утвердительно, другие считали, что ни одна из сторон не должна рассматриваться как полноценная личность. Некоторые утверждали, что общность эмоциональных ощущений может служить доказательством существования только одной личности. Еще одна точка зрения состоит в том, что сознательного индивидуума представляет только левое полушарие, а правое — просто автомат. Этой точки зрения придерживаются те, кто считает речевые способности обязательной составляющей сознания. Само собой, только левое полушарие может убедительно заявить «Да!» в ответ на вопрос: «Обладаешь ли ты сознанием?». Правому полушарию, подобно кошке, собаке или шимпанзе, может быть трудно даже понять отдельные слова этого вопроса, не говоря уже о том, чтобы правильно ответить на него.

И все же пока вопрос остается открытым. В недавних экспериментах, проведенных Дональдом Вильсоном и его коллегами (Вильсон и др. [1977], Газзанига и др. [1977]), при наблюдениях за пациентом с разделенным мозгом (назовем его "P. S."), были получены весьма интересные результаты. После операции по разделению полушарий только левое полушарие обладало речью, но понимали речь *оба* полушария, а позднее правое полушарие научилось и воспроизводить речь! Несомненно, что оба полушария были наделены сознанием. Более того, это были два *отдельных* сознания, поскольку их желания и пристрастия были совершенно различны. Например, левое полушарие выражало желание стать чертежником, а правое — гонщиком!

Лично я не верю в справедливость широко распространенного убеждения в том, что обычный человеческий язык необходим для мышления или сознания. (В следующей главе я приведу некоторые доводы в пользу своей точки зрения.) Поэтому я отношусь к тем, кто верит, что, в принципе, обе половины мозга после разделения обладают сознанием независимо друг от друга. Пример с P. S. может служить весомым подтверждением тому, что, по крайней мере в этом частном случае, это так и есть. По-моему мнению, единственное действительное различие между P. S. и всеми другими случаями заключается в том, что сознание его правого полушария на самом деле смогло убедить окружающих в своем существовании!

Если мы допускаем, что P. S. действительно имеет два независимых разума, то возникает довольно пикантная ситуация. Есть все основания полагать, что *до* операции разделения полушарий у каждого пациента было только одно сознание. Однако после операции их уже два! В некотором смысле, изначально единственное сознание *раздвоилось*. Мы можем в связи с этим вспомнить гипотетического путешественника из главы 1, который воспользовался телепортационной машиной и в какой-то момент (неумышленно) был поставлен перед фактом, что будто бы «настоящее» его «я» благополучно прибыло на Венеру. В этом случае раздвоение сознания приводит к кажущемуся парадоксу. Ведь мы можем задать резонный вопрос: «А какой, собственно, маршрут выбрал поток его сознания „на самом деле“?» Если бы *вы* были этим путешественником, то какое бы из двух сознаний *вы*, в конце концов, назвали бы «собой»? Устройство для телепортации относится к области научной фантастики, однако в случае с P. S. мы имеем в чем-то аналогичную ситуацию и притом *совершенно реальную*. Какое из сознаний P. S. было бы правомерно «отождествить» с P. S. до операции? Нет сомнений, что многие философы сочли бы этот вопрос бессмысленным, ибо его решение при помощи операционалистских методов кажется невозможным. Каждое полушарие сохраняет память о «дооперационных» временах, и, естественно, каждое будет идентифицировать себя с той — еще целостной — личностью. Но все же подобная ситуация, примечательная в качестве своего рода головоломки и способная поставить в тупик, сама по себе еще не является парадоксальной.

Эта головоломка еще усложнится, если предположить, что в дальнейшем оба сознания можно было бы каким-то образом опять свести воедино. Повторное соединение разрезанных нервных волокон мозолистого тела на сегодняшнем этапе развития медицины исключается, но можно представить себе некий способ разделения полушарий, более мягкий, чем реальное разрезание нервных волокон. Например, нервные волокна могли бы быть временно заморожены или парализованы при помощи лекарственных средств. Пока я не слышал о подобных опытах, но думаю, что появление технических возможностей для их осуществления — это вопрос обозримого будущего. Тогда можно допустить, что после приведения мозолистого тела в работоспособное состояние, мы вновь получим *одно* сознание! Представьте, что это сознание ваше. Как бы вы себя чувствовали после того, как в течение какого-то времени были двумя независимыми личностями с отдельными «я»?

### «Зрение вслепую»

Эксперименты по разделению полушарий мозга, помимо прочего, ясно показали, что наличие *единственного* «места для сознания» вовсе не обязательно. Но были проведены и другие опыты, результаты которых дают основание полагать, что некоторые участки коры головного мозга в большей степени связаны с сознанием, нежели прочие. Среди подобных опытов - изучение явления *слепоты*. Повреждение тех или иных участков зрительной коры может привести к слепоте в соответствующем секторе поля зрения. Человек не видит предмет, помещенный в этот сектор — у него появляется частичная слепота, связанная с этой конкретной зоной его поля зрения.

Однако, кое-какие любопытные изыскания (см. Вайскранц [1987]) позволяют говорить о том, что дела здесь обстоят совсем не так просто, как кажется. У пациента, называемого здесь и далее "D. В.", необходимо было удалить часть зрительной коры головного мозга, и после операции у него наступила частичная слепота в описанном выше смысле. Однако, когда что-либо (как правило, изображение крестика, кружочка или наклонного отрезка прямой) помещали в «слепую зону» и просили D. В. *угадать*, что это такое, он обнаружил, что может делать это с практически стопроцентной точностью! Эта способность к «угадыванию» оказалась неожиданной и для самого D. В., который при этом продолжал утверждать, что в этой зоне он вообще ничего не видит.

Изображения, формируемые на сетчатке, в свою очередь тоже обрабатываются не только зрительной корой, но и *другими* участками мозга, при этом один из наиболее загадочных из них находится в нижней части височной доли. Вполне возможно, что D. В. строил свои «догадки» на основе информации, полученной как раз этим участком нижней части височной доли. При активации этих областей не возникало никаких *осознанных* ощущений, однако информация в них, бесспорно, содержалась, проявляя себя только в точности «догадок» D. В. На самом деле, после соответствующей тренировки D. В. научился до некоторой степени осознавать информацию, относящуюся к этим областям мозга.

Все это, по-видимому, указывает на то, что отдельные зоны коры головного мозга (как, например, зрительная кора) имеют большее отношение к сознательному восприятию, чем другие, но некоторые из этих менее важных зон, очевидно, могут быть путем тренировок открыты для непосредственного доступа сознания.

### Обработка информации в зрительной коре

Именно в *зрительной коре* процессы обработки информации изучены гораздо лучше, чем в других частях мозга. Для их описания был предложен целый ряд разнообразных моделей. На самом деле, *до того*, как визуальная информация попадает в зрительную кору, ее частичная обработка проходит еще в сетчатке. (Вообще говоря, сетчатка считается частью мозга!) Одни из первых экспериментов по исследованию процессов обработки информации в зрительной коре были проведены Давидом Хью-белом и Торстеном Визелем и в 1981 году принесли им Нобелевскую премию. В ходе этих экспериментов удалось показать, что определенные клетки зрительной коры кошки воспринимают в поле зрения линии, имеющие вполне *определенный угол наклона*. При этом соседние клетки были восприимчивы к линиям, расположенным под несколько иным углом. Часто не имело значения, что именно характеризуется таким углом наклона. Это могла быть граница между темной и светлой областью или просто темная черта на светлом фоне. Изучаемые клетки оказались способны абстрагироваться от конкретной природы объекта, имеющего свойство «угол наклона». Другие клетки были чувствительны к определенным цветам или к различиям между изображениями, регистрируемыми каждым глазом, что позволяет воспринимать объемные изображения. Продвигаясь далее от первичных областей восприятия, мы обнаруживаем клетки, которые чувствительны ко все более тонким аспектам восприятия того, что мы видим. Например, при взгляде на *рис.* мы различаем очертания белого треугольника, однако линии, образующие сам треугольник, большей частью не изображены, но *домыслены*. Клетки, способные фиксировать эти «подразумеваемые» линии, действительно были обнаружены в зрительной коре (той, что называется вторичной зрительной корой)!

В начале 1970-х годов в литературе появились заявления об открытии в зрительной коре мозга *мартышек* клеток, которые активируются только тогда, когда на сетчатку проецируется изображение *лица*. На основании этой информации была сформулирована «гипотеза бабушкиной клетки», согласно которой в мозге человека должны существовать определенные клетки, реагирующие только в тех случаях, когда в комнату входит его/ее бабушка! Недавние исследования показали, что есть клетки, реагирующие на определенные слова. Может быть, это шаг на пути к доказательству справедливости гипотезы бабушкиной клетки?

Ясно, что нам предстоит еще очень много узнать о деталях процессов обработки информации в мозге. До сих пор очень мало известно о функционировании высших отделов мозга. Мы пока оставим эти вопросы и обратимся к самим клеткам мозга, которые позволяют ему осуществлять эту удивительную деятельность.

## Как работают нервные импульсы?

Обработка информации в головном мозге (равно как и в спинном мозге и сетчатке) осуществляется уникальными по своему разнообразию клетками, которые называются *нейронами*. Попробуем разобраться, как же устроен нейрон. Его утолщенная центральная часть, немного похожая на звезду и часто имеющая форму редиски, называется *телом (сомой) нейрона* и содержит в себе клеточное ядро. С одной стороны от тела нейрона отходит сильно вытянутое нервное волокно, называемое *аксоном*. Аксон иногда достигает действительно огромной длины (у человека — часто до нескольких сантиметров), если учесть, что речь идет всего лишь об одной микроскопической клетке.

Аксон служит «проводом», по которому передается *исходящий* из клетки нервный сигнал. От аксона в стороны могут отходить более мелкие ветви и, кроме того, аксон может несколько раз разветвляться. На концах каждого из этих нервных волокон находятся *нервные*

окончания (*терминали*). По другую сторону сомы, а часто и отходя от нее во всех направлениях, располагаются короткие сильно ветвящиеся отростки — *дендриты*, по которым в клетку поступают *входные* данные. (Иногда и на концах дендритов встречаются терминали, образующие так называемые *дендро-дендритные* синапсы между дендритами. В дальнейшем я не буду их учитывать, поскольку связанное с ними усложнение общей картины несущественно).

Клетка как целое отделена от окружения клеточной мембраной, которая охватывает сому, аксон, нервные окончания, дендриты и все остальное. Для того чтобы сигналы передавались от одного нейрона к другому, надо каким-то образом обеспечить им возможность «перехода через барьер» между нейронами. Это достигается с помощью межклеточного соединения, называемого *синапсом*, в котором терминаль одного нейрона соединена с какой-либо точкой на соме или на одном из дендритов другого нейрона. На самом деле, между терминалью одного нейрона и сомой или дендритом другого остается очень узкий зазор, который называется *синаптической щелью*. При передаче от одного нейрона к другому сигнал должен преодолеть этот зазор.

В какой форме сигналы передаются по нервным волокнам и через синаптические щели? Что заставляет следующий нейрон передавать сигнал дальше? Для непосвященного, вроде меня, механизмы, которые используются здесь природой, кажутся удивительными и совершенно зачаровывающими!

Можно было бы думать, что эти сигналы распространяются точно так же, как электрический ток по проводам, но в действительности все гораздо сложнее. Нервное волокно представляет собой цилиндрическую трубку, заполненную раствором обычной соли (хлорида натрия), смешанной с хлоридом калия (с преобладанием последнего), так что внутри трубки находится смесь из ионов натрия, калия и хлора. Снаружи волокна находятся те же ионы, но в других соотношениях: ионов натрия больше, чем ионов калия. В состоянии покоя содержимое трубки имеет суммарный отрицательный заряд (т.е. ионов хлора там больше, чем ионов калия и натрия вместе; напомним, что ионы калия и натрия заряжены положительно, тогда как ионы хлора — отрицательно). Клеточная мембрана, образующая поверхность цилиндра, имеет «утечки», поэтому ионы перемещаются через мембрану таким образом, чтобы нейтрализовать избыточный заряд. Компенсацию утечек и поддержание избыточного отрицательного заряда внутри трубки осуществляет «ионный насос», который очень медленно откачивает ионы натрия через мембрану наружу. Отчасти это же помогает поддерживать избыток ионов калия по сравнению с ионами натрия во внутреннем растворе. Существует также ионный насос, который (более медленно) переносит ионы калия из наружной среды внутрь трубки (что, правда, не способствует поддержанию разности зарядов).

*Сигнал*, распространяющийся по нервному волокну, представляет собой область с *обратным* распределением зарядов (т.е. положительный заряд внутри и отрицательный снаружи), которая перемещается вдоль волокна. Вообразите, что вы находитесь на нервном волокне как раз перед такой областью с обратным распределением зарядов. По мере того, как эта область приближается, электрическое поле открывает в мембране маленькие «дверцы», называемые *натриевыми каналами*. Это позволяет ионам натрия перемещаться с наружной стороны мембраны обратно внутрь трубки (в результате совместного действия электрических сил и давления, обусловленного разностью концентраций, т.е. «осмоса»). Это приводит к тому, что заряд снаружи становится отрицательным, а внутри — положительным. Когда это происходит, мы знаем, что область обратного распределения заряда, которая и является сигналом, достигла нас. При этом позади нее открываются крошечные «дверцы» другого типа (*калиевые каналы*), которые выпускают ионы калия наружу, тем самым восстанавливая избыточный отрицательный заряд внутри. Теперь сигнал прошел! Наконец, когда сигнал уже

достаточно удалился, медленно, но верно работающие ионные насосы постепенно выкачивают ионы натрия из трубки наружу, закачивая внутрь ионы калия. Таким образом волокно возвращается в состояние покоя и готово к передаче очередного сигнала.

Обратите внимание, что сигнал представляет собой просто область обратного распределения заряда, движущуюся вдоль волокна. *Вещество* как таковое (т. е. ионы) перемещается при этом совсем немного - только внутрь и наружу через клеточную мембрану!

Этот странный, экзотический механизм действует на поверку очень эффективно. Он универсален и используется как у позвоночных, так и у беспозвоночных. Но у позвоночных он был усовершенствован за счет изоляции нервных волокон при помощи беловатого жироподобного вещества, называемого *миелином*. (Именно миелиновым покрытием объясняется цвет «белого вещества» мозга.) Такая изоляция позволяет нервным импульсам распространяться без потерь (от одной «ретрансляционной станции» к другой) и с очень приличной скоростью — до 120 метров в секунду.

Когда сигнал достигает терминали, из нее выделяется химическое соединение, называемое нейромедиатором. Это соединение пересекает синаптическую щель и достигает другого нейрона - поверхности дендрита или сомы. При этом у одних нейронов терминаль выделяет нейромедиатор, *облегчающий* возбуждение следующего нейрона, т.е. посылку нового сигнала вдоль своего аксона. Эти синапсы называются *возбуждающими*. У других нейронов терминали выделяют нейромедиатор, *затрудняющий* другому нейрону генерацию собственного импульса, и поэтому называются *тормозящими*. На каждом нейроне действие активных в данный момент возбуждающих синапсов суммируется, из результата вычитается суммарное действие тормозящих синапсов, и если полученная разность превышает определенное критическое значение, то нейрон действительно возбуждается. (Возбуждающие синапсы создают положительную разность потенциалов между внутренней и наружной сторонами мембраны следующего нейрона, а тормозящие — отрицательную. Эти разности потенциалов складываются. Нейрон возбуждится только в том случае, если результирующая разность потенциалов на мембране в начале его аксона достигнет определенной критической величины, при которой ионы калия не успевают выходить наружу достаточно быстро, чтобы восстановить равновесие.)

## Компьютерные модели

Важным свойством нервной системы является то, что сигналы, используемые для передачи информации, относятся (большой частью) к классу явлений «все или ничего». Сила сигнала не изменяется: он или есть, или его нет. Это придает деятельности нервной системы некоторое сходство с работой цифрового компьютера. На самом деле, между работой огромного количества взаимосвязанных нейронов и процессами внутри компьютера со всеми его проводниками и логическими элементами (подробнее об этом чуть позже) есть много общего. В принципе, было бы не так уж трудно создать компьютерную модель подобной системы нейронов. Но возникает вполне естественный вопрос: не означает ли это, что какой бы ни была детальная схема соединений нейронов в мозге, всегда можно построить его компьютерную модель?

Чтобы сделать это сравнение более наглядным, я должен объяснить, что такое *логический элемент*. В компьютере мы также сталкиваемся с ситуацией типа «все или ничего»: либо в проводнике есть импульс тока, либо его нет, причем когда импульс *есть*, его величина всегда одна и та же. Поскольку все в компьютере строго синхронизовано, то

*отсутствие* импульса было бы определенным сигналом, который может быть «замечен» компьютером. Вообще говоря, когда мы пользуемся термином «логический элемент», мы неявно подразумеваем, что наличие или отсутствие импульса обозначает «истину» или «ложь», соответственно. Конечно же, к реальной истине или лжи это никакого отношения не имеет и используется только как общепринятая терминология. Мы будем также обозначать «истину» (наличие импульса) цифрой «1» и «ложь» (отсутствие импульса) цифрой «0». Помимо этого, как и в главе 4, мы будем обозначать знаком «&» логическое «и» (которое является «утверждением» об «истинности» обоих аргументов, т. е. принимает значение 1 тогда и только тогда, когда оба они равны 1); «V» — логическое «или» (которое «означает», что либо один из аргументов, либо оба они «истинны», т. е. выражение становится равным 0 тогда и только тогда, когда оба аргумента имеют значение 0); знаком «=>» — «следует» (т. е.  $A \Rightarrow B$  означает утверждение «если истинно A, то истинно B», что эквивалентно утверждению «либо A ложно, либо B истинно»); « $\Leftrightarrow$ » — «тогда и только тогда» (выражение истинно, если оба аргумента «истинны» или же оба «ложны» одновременно); и использовать знак « $\sim$ » для логического «не» (выражение «истинно», если аргумент «ложно», и «ложно», если аргумент «истинен»). Результаты применения различных логических операций можно описать при помощи так называемых «таблиц истинности»: в каждой из которых A обозначает строки (т. е.  $A = 0$  дает первую строку, а  $A = 1$  - вторую), а B — столбцы. Например, если  $A = 0$  и  $B = 1$ , что во всех таблицах отвечает правому верхнему углу, то выражение  $A \Rightarrow B$  согласно *третьей* таблице примет значение 1. (Соответствующий словесный пример из области *традиционной* логики: утверждение «если я сплю, то я счастлив», очевидно, остается истинным в частном случае, когда я бодрствую и счастлив.)

Это — основные типы логических элементов. Есть еще ряд других, но все они могут быть построены из только что описанных. Итак, можем ли мы, в принципе, построить компьютер, используя соединенные между собой *нейроны*? Я собираюсь показать, что это возможно даже при самых примитивных представлениях о функциях нейрона. Посмотрим, как можно было бы, в принципе, построить логические элементы на основе соединенных между собой нейронов. Нам потребуется новый способ записи цифр, поскольку в *отсутствие* сигнала ничего не происходит. Будем считать (совершенно произвольно), что двойной импульс обозначает 1 (или «истину»), а одиночный — 0 (или «ложь»). Примем также упрощенную схему, в которой нейрон возбуждается только при получении *двух* возбуждающих импульсов (т. е. двойного импульса) одновременно. Тогда нетрудно сконструировать элемент «и» (т.е. «&»). Для этого достаточно, чтобы с выходным нейроном образовывали входные синапсы два нервных окончания. (Тогда, если по обоим окончаниям приходят двойные импульсы, то и первый, и второй импульс превысят заданный двухимпульсный порог срабатывания; а если хотя бы на один входной синапс приходит одиночный импульс, то превысит порог лишь одна конечного нейрона соответствует сигналу элемента «не» (т. е. пара импульсов, если на входе был одиночный, и наоборот). (Такая конструкция кажется абсурдно усложненной, но это наилучшее из того, что пришло мне в голову!) В качестве развлечения читатель может составить подобные «нейронные» схемы и для остальных описанных выше логических элементов.

Естественно, эти конкретные примеры не могут служить серьезными моделями того, что происходит в мозге на самом деле. С их помощью я только старался показать, что описанная выше модель возбуждения нейрона по сути логически эквивалентна конструкции электронного компьютера. Легко видеть, что с помощью компьютера можно воспроизвести любую модель соединения нейронов между собой. В то же время, подробно рассмотренные выше конструкции указывают на то, что и, наоборот, системы нейронов могут быть моделями компьютера и, следовательно, *могут* действовать как (универсальная) машина Тьюринга. Хотя при обсуждении машин Тьюринга во второй главе мы не использовали понятие логических элементов и, в действительности, для построения модели машины Тьюринга в

общем случае помимо логических элементов нам понадобилось бы еще многое другое, в этом нет ничего принципиально нового, если только мы допускаем возможность *аппроксимации* используемой в машине Тьюринга *бесконечной ленты* огромным, но конечным множеством нейронов. А это уже, как кажется, подводит нас к выводу о том, что мозг по своей сути эквивалентен компьютеру!

Но прежде, чем делать такие поспешные выводы, нам следует рассмотреть некоторые различия между деятельностью мозга и работой современных компьютеров, которые могут оказаться достаточно важными. Во-первых, я слишком упростил описание возбуждения нейрона, отнеся его к явлениям типа «все или ничего». Это справедливо для одиночного импульса, распространяющегося по аксону. На самом деле, когда нейрон возбуждается, он генерирует целую последовательность импульсов, быстро следующих друг за другом. Даже в состоянии покоя нейрон генерирует импульсы, пара возбуждающих импульсов. Я предполагаю, что все импульсы хорошо согласованы по времени, и что в случае двойного импульса, для определенности, синхронизация осуществляется по *первой паре* импульсов.) Конструкция элемента «не» (т.е. «~») значительно сложнее. Входной сигнал поступает по аксону, разделяющемуся на две ветви. Одна из ветвей имеет увеличенную длину, такую, чтобы сигнал при движении по ней запаздывал ровно на время, равное промежутку между импульсами в паре. Затем обе ветви снова разделяются, и одно из ответвлений каждой ветви отходит к тормозящему нейрону, причем аксон от ветви с задержкой предварительно разделяется снова, образуя прямую ветвь и ветвь с задержкой. На выходе тормозящего нейрона не будет *ничего* при одиночном импульсе на его входе, и *двойной импульс* (с задержкой), если на его входе также был двойной импульс. Аксон тормозящего нейрона разделяется на три ветви, каждая из которых образует тормозящий синапс на окончном нейроне. Оставшиеся два ответвления исходного аксона снова разделяются, так что к конечному нейрону подходят уже четыре терминалы, образующие возбуждающие синапсы. При желании читатель может проверить, что выходной сигнал этогогоно с гораздо меньшей частотой. Именно многократное увеличение *частоты* импульсов характеризует переход нейрона в возбужденное состояние. Кроме того, есть еще и вероятностный аспект срабатывания нейрона. Один и тот же стимул может приводит к различным результатам. Более того, в мозге нет точной синхронизации с помощью постоянной тактовой частоты, которая необходима для работы современных компьютеров. Кроме того, следует отметить, что максимальная частота срабатывания нейрона, составляющая около 1000 импульсов в секунду, гораздо меньше, чем у современных электронных устройств, у которых она более чем в 1 млн. раз выше. К тому же, по сравнению с очень высокой точностью соединений в электронном компьютере, действительные соединения между нейронами кажутся в большой степени случайными и избыточными — правда, сегодня мы знаем, что в мозге (при рождении) эти соединения установлены с гораздо большей точностью, чем считалось полвека назад.

Может показаться, что большая часть из сказанного выше характеризует мозг с невыгодной стороны по сравнению с компьютером. Но есть и другие факторы, говорящие в пользу мозга. У логических элементов может быть лишь очень ограниченное количество входов и выходов (скажем, три-четыре, не больше), тогда как нейроны могут иметь гигантское число синапсов. (Предельным случаем можно считать нейроны мозжечка, известные как клетки Пуркинье, у которых количество возбуждающих синапсов достигает 80 000.) Помимо этого, общее число нейронов в мозге также превышает максимальное количество транзисторов, входящих в состав самой большой в мире вычислительной машины — примерно  $10^{11}$  в мозге и «всего лишь»  $10^9$  у компьютера. Однако последнее число в будущем, скорее всего, возрастет. Более того, большое число клеток мозга в значительной степени обусловлено огромным количеством мелких *клеток-зерен* в мозжечке, которых насчитывается около тридцати миллиардов ( $3 \times 10^{10}$ ). Если считать, что осознанным восприятием, в отличие от современных компьютеров, мы обладаем просто благодаря

большому числу нейронов, то нам придется найти какое-то дополнительное объяснение тому, что деятельность *мозжечка* полностью *бессознательна* и в то же время сознание может быть связано с *головным мозгом*, в котором нейронов всего в два раза больше (около  $7 \times 10^{10}$ ) при значительно меньшей плотности.

## Пластичность мозга

Между деятельностью мозга и работой компьютера существуют и другие различия, на мой взгляд даже более важные, чем до сих пор упоминавшиеся, и связанные с явлением, которое называется *пластичностью мозга*. В действительности, неправомерно рассматривать мозг как *фиксированную* совокупность связанных друг с другом нейронов. Взаимосвязи нейронов на самом деле не постоянны, как это было бы в рассмотренной выше компьютерной модели, но все время меняются. Это не значит, что изменяются положения аксонов или дендритов. Многие из их сложных взаимосвязей в общих чертах формируются еще при рождении. Я имею в виду синаптические контакты, которые в действительности и обеспечивают связь между нейронами. На дендритах они часто формируются на небольших выростах, называемых *шипами*, к которым подходят терминалы других нейронов. Здесь «контакт» означает *не* соприкосновение, а узкий зазор (синаптическую щель) заданной ширины — около одной сорокатысячной доли миллиметра. При определенных условиях шипики дендритов могут исчезать, тем самым нарушая контакт, или вырастать (могут образовываться и новые) и формировать новую связь. Таким образом, если мы представим себе, что совокупность соединенных друг с другом нейронов в мозгу действительно образует компьютер, то это компьютер, способный непрерывно изменяться!

Согласно одной из ведущих теорий долговременная память обусловлена именно такими изменениями синаптических контактов. Именно они обеспечивают возможность сохранения необходимой информации. Если это так, то пластичность предстает перед нами уже не просто как несущественное усложнение деятельности мозга, но как ее важнейшее свойство.

Каков механизм этих непрекращающихся изменений? Как быстро они могут происходить? Однозначный ответ на второй вопрос вряд ли существует, хотя представители по крайней мере одной из научных школ утверждают, что такие изменения могут происходить за несколько секунд. Этого можно было ожидать, если такие изменения ответственны за долговременное запоминание, поскольку оно происходит за характерное время около одной секунды (Кан-дел [1976]). Это имело бы для нас весьма существенное значение в дальнейшем. Я вернусь к этому важному вопросу в следующей главе.

А что же можно сказать о механизмах пластичности мозга? Согласно оригинальной теории, предложенной в 1954 году Дональдом Хеббом, существуют определенные синапсы (впоследствии получившие название «синапсов Хебба»), обладающие тем свойством, что связь между нейронами А и В, обусловленная синапсом Хебба, усиливается каждый раз, когда за возбуждением А следует возбуждение В, и ослабляется, если В не возбуждается. Изменение эффективности связи между нейронами не зависит от степени участия самого синапса Хебба в возбуждении нейрона В. Это делает возможной некоторую форму «обучения». На основе этой теории был предложен целый ряд математических моделей обучения и решения задач. Они получили название *нейронных сетей*. По-видимому, нейронные сети действительно способны к какому-то элементарному обучению, но им пока еще далеко до реальных моделей мозга. В любом случае, механизмы, управляющие изменениями синаптических контактов, скорее всего более сложны, чем рассмотренные выше. Очевидно, что необходимы дальнейшие исследования.

С пластичностью связан и другой аспект выделения нейромедиаторов терминалями. Иногда нейромедиаторы выделяются вовсе не в синаптические щели, а в окружающую межклеточную жидкость, возможно, для воздействия на другие, расположенные на большом удалении нейроны. По-видимому, многие нейрохимические вещества выделяются подобным образом. Существуют различные теории памяти, в которых используются разнообразные сочетания таких веществ, участвующих в процессе запоминания. Конечно, состояние мозга зависит от наличия в нем химических соединений (например, гормонов), выделяемых различными его частями. Проблемы нейрохимии в целом весьма сложны, и пока непонятно, как можно подойти к созданию правдоподобной и полной компьютерной модели мозга.

## Параллельные компьютеры и «единственность» сознания

Многие считают, что развитие *параллельных* компьютеров содержит в себе ключ к построению машин, обладающих возможностями человеческого мозга. Далее мы кратко рассмотрим эту популярную сегодня идею.

Параллельный компьютер, в противоположность последовательному, может выполнять одновременно и независимо друг от друга огромное число отдельных операций, и результаты этих автономных операций время от времени объединяются, давая вклад в общий вычислительный процесс. Толчком к созданию такого типа компьютерной архитектуры послужили попытки моделирования нервной системы, поскольку, согласно современным представлениям, разные части мозга выполняют практически автономные вычислительные функции (например, при обработке визуальной информации в зрительной коре).

Здесь необходимо сделать два замечания. Во-первых, между параллельным и последовательным компьютерами не существует *принципиальной* разницы. По сути, оба являются *машинами Тьюринга*. Отличие может проявляться только лишь в эффективности, или скорости, вычислений в целом. Для некоторых типов вычислительных процессов параллельная организация, действительно, более эффективна, но это далеко не всегда так. Во-вторых, по крайней мере с моей точки зрения, крайне маловероятно, что классические параллельные вычисления дают ключ к тому, что происходит при *сознательном* мышлении. Характерным свойством сознательной мысли (по крайней мере в нормальном психологическом состоянии и не после операции по разделению полушарий мозга!) является ее «единственность» — в противоположность множественности выполняемых одновременно и независимо друг от друга операций.

Фразы типа: «Я же не могу думать обо всем сразу?!» можно услышать на каждом шагу. Можно ли *вообще* думать о нескольких вещах одновременно? Вероятно, кто-то *может* удерживать в голове несколько мыслей в одно и то же время, но это, скорее всего, будет похоже на постоянное перескакивание от одной мысли к другой и обратно, нежели на действительно одновременное, сознательное и независимое их обдумывание. Если бы кто-то мог думать о двух вещах совершенно независимо, то это было бы более похоже на обладание двумя *раздельными сознаниями*, пусть даже и на короткий промежуток времени, тогда как повседневный опыт (по крайней мере нормальных людей) свидетельствует о наличии *единственного* сознания, которое может иметь смутное представление о ряде вещей, но которое сконцентрировано в каждый момент времени только на *одной* из них.

Конечно, то, что мы подразумеваем здесь под «одной вещью», не совсем ясно. В следующей главе мы познакомимся с совершенно удивительными примерами «отдельных мыслей», появлявшихся в минуты вдохновения у Пуанкаре и Моцарта. Но нам вовсе не обязательно забираться так далеко, чтобы понять, что мысль человека в каждый конкретный

момент времени может неявно быть очень сложной. Представьте себе, например, процесс обдумывания обеденного меню. Одна такая мысль может включать в себя такое количество разнообразной информации, что ее полное словесное описание было бы очень долгим.

«Единственность» осознанного восприятия представляется мне идущей вразрез с концепцией параллельного компьютера. С другой стороны, эта концепция может оказаться более подходящей в качестве модели *бессознательной* деятельности мозга. Различные независимые действия (ходьба, застегивание пуговиц, дыхание и даже разговор) могут выполняться человеком одновременно и более менее автономно, причем он может не осознавать *ни одно* из них!

С другой стороны, мне кажется, что эта «единственностью» сознания может иметь что-то общее с *квантовым параллелизмом*. Вспомним, что, согласно квантовой теории, на квантовом уровне различные альтернативы могут сосуществовать в линейной суперпозиции! Отсюда следует, что *одиночное квантовое состояние* могло бы, в принципе, состоять из большого числа различных событий, происходящих одновременно. Именно это и подразумевается под квантовым параллелизмом. Мы скоро рассмотрим теоретическую концепцию «квантового компьютера», в котором, в принципе, квантовый параллелизм мог бы быть использован для выполнения большого числа одновременных операций. Если «состояние ума», соответствующее рассудочной деятельности, имеет какое-то сходство с квантовым состоянием, то некая форма «единственности», или глобальности, мысли соответствует ему в большей степени, чем в случае обычного параллельного компьютера. У этой идеи есть несколько привлекательных аспектов, к которым я вернусь в следующей главе. Но прежде, чем рассматривать эту идею всерьез, мы должны ответить на вопрос, могут ли квантовые эффекты иметь какое-либо отношение к деятельности мозга.

### **Имеет ли квантовая механика отношение к работе мозга?**

Все предыдущее обсуждение нервной деятельности проводилось целиком в рамках классических представлений, за исключением тех случаев, когда мы затрагивали физические явления, неявные причины которых отчасти обусловлены квантово-механическими эффектами (например, ионы, несущие единичные электрические заряды; натриевые и калиевые каналы; определенные химические потенциалы, определяющие триггерный характер генерации нервного импульса; химия нейромедиаторов). Но нет ли таких ключевых процессов в мозге, которые бы непосредственно определялись квантово-механическими эффектами? Для того чтобы рассуждения, описанные в конце предыдущей главы, имели какой-то смысл, такие процессы, по-видимому, должны существовать.

В действительности, можно указать, по крайней мере, одно место, где чисто квантовые явления имеют принципиальное значение для нервной деятельности, — это *сетчатая оболочка* глаза. (Вспомним, что сетчатка фактически входит в состав мозга!) Эксперименты с жабами показали, что в подходящих условиях адаптированная к темноте сетчатка вырабатывает макроскопический нервный импульс при попадании на нее *единичного фотона* (Бэйлор и др. [1979]). То же, как выясняется, справедливо и для человека (Хехт и др. [1941]), хотя в этом случае существует дополнительный механизм, который подавляет подобные слабые сигналы, тем самым очищая воспринимаемое изображение от лишнего визуального «шума». Необходимо суммарное воздействие примерно *семи* фотонов, чтобы адаптировавшийся к темноте испытуемый мог его ощутить. Тем не менее, в нашей сетчатке, по-видимому, все-таки есть клетки, чувствительные к попаданию только одного фотона.

Поскольку в теле человека *существуют* нейроны, способные срабатывать под воздействием единичного квантового события, то вполне обоснован вопрос о наличии таких клеток где-нибудь в основных отделах мозга. Насколько мне известно, это предположение не

подтвердилось. У клеток всех изученных типов есть определенный порог срабатывания и требуется очень большое число квантов, чтобы перевести клетку в возбужденное состояние. Однако можно было бы допустить, что где-то глубоко внутри мозга должны быть клетки, чувствительные к одиночным квантам. Если это окажется верным, то квантовая механика должна играть существенную роль в деятельности мозга.

Но даже при таком положении вещей роль квантовой механики оказалась бы чисто номинальной, поскольку квант используется просто как возбудитель сигнала. Никаких интерференционных эффектов, характерных для квантовых явлений, пока обнаружить не удалось. Похоже, что в лучшем случае все, что мы можем получить от квантовой механики, это неопределенность момента срабатывания нейрона. Трудно представить, как это может пригодиться нам на практике.

Однако некоторые вопросы, имеющие к этому отношение, не так тривиальны. Для их рассмотрения обратимся вновь к сетчатой оболочке глаза. Предположим, что фотон попадает на сетчатку, предварительно отразившись от полупрозрачного зеркала. Состояние фотона тогда будет представлять собой сложную линейную суперпозицию состояний, когда он попадает в клетку сетчатки и когда он проходит мимо клетки и вместо этого, скажем, улетает через окно в космос. В тот момент, когда он *мог бы* попасть в клетку сетчатки, до тех пор, пока выполняется линейная процедура  $U$  (т. е. детерминированная эволюция вектора состояния по уравнению Шредингера), мы получим сложную линейную суперпозицию наличия и отсутствия нервного сигнала. Когда это доходит до сознания наблюдателя, воспринимается только *одна* из этих двух альтернатив, и должна использоваться другая квантовая  $R$ -процедура (редукция вектора состояния). (Говоря так, я сознательно обхожу стороной теорию множественности миров, которая имеет множество своих собственных проблем!) В соответствии с рассуждениями, приведенными в конце предыдущей главы, нам следует задать вопрос, достаточное ли количество материи вовлекается в прохождение сигнала, чтобы удовлетворялся *одногравитонный критерий* (см. главу 8)? Хотя при преобразовании энергии фотона в энергию движения массы при выработке сигнала в сетчатке достигается действительно гигантское усиление, возможно, до  $10^{20}$  раз, эта масса все же значительно меньше величины планковской массы  $m_{Pl}$  (примерно в  $10^8$  раз). Однако нервный сигнал создает регистрируемое изменяющееся *электрическое поле* в окружающей среде (тороидальное поле с осью, совпадающей с нервным волокном, по которому оно перемещается). Это поле может вносить в *окружающую среду* значительное возмущение, за счет чего одногравитонный критерий будет легко удовлетворен. Таким образом, в соответствии с изложенной мной точкой зрения,  $R$ -процедура могла бы выполняться задолго до того, как мы увидим или, может случиться, не увидим вспышку света. К тому же, для редукции вектора состояния наше сознание не требуется!

## Квантовые компьютеры

Если мы *все-таки* предположим, что чувствительные к одиночным квантам нейроны играют важную роль где-то в глубине нашего мозга, то возникает вопрос, какие следствия это могло бы иметь. Для начала я изложу концепцию *квантового компьютера*, предложенную Дойчем, а затем мы выясним, можно ли ее рассматривать как имеющую отношение к теме нашей дискуссии.

Как было указано выше, главная идея состоит в использовании квантового параллелизма, в соответствии с которым два совершенно различных процесса должны рассматриваться как происходящие одновременно в виде квантовой линейной суперпозиции,

например, фотон одновременно отражается от полупрозрачного зеркала и проходит через него или один и тот же фотон проходит через каждую из двух щелей. В случае квантового компьютера этими двумя различными наложенными друг на друга процессами будут два различных *вычисления*. При этом предполагается, что нас интересуют результаты не *обоих* вычислений, а некий результат, основанный на частичной информации, полученной из суперпозиции этих процессов. Наконец, когда оба вычисления завершены, над этой парой процессов должно быть проведено соответствующее «наблюдение», позволяющее получить искомый ответ. Таким образом, это устройство могло бы сэкономить время за счет выполнения двух вычислений одновременно! До сих пор не видно никакого значительного преимущества от использования такого подхода, поскольку было бы гораздо проще непосредственно использовать два классических компьютера параллельно (или один классический параллельный компьютер), чем один квантовый. Однако реальные преимущества квантового компьютера могли бы проявиться при необходимости выполнить *очень большое*, возможно, неограниченно большое, количество параллельных вычислений, когда нас интересуют не их результаты сами по себе, а только подходящая комбинация результатов всех вычислений.

Принципиальное устройство квантового компьютера предполагает использование квантовой разновидности логических элементов, у которых выходной сигнал является результатом «унитарной операции» над входным сигналом — операции типа  $U$ , — и вся работа компьютера состояла бы в выполнении операции  $U$  до самого конца вычислений, пока конечный «акт наблюдения» не приведет к выполнению операции  $R$ .

Согласно выводам Дойча квантовые компьютеры не предназначены для выполнения неалгоритмических операций (т. е. действий, выходящих за пределы возможностей машины Тьюринга), но способны в некоторых, очень специфических случаях, достигать более высокого быстродействия, чем обычная машина Тьюринга. Для такой блестящей идеи эти выводы представляются довольно неутешительными, но будем помнить о том, что пока мы стоим у самых истоков.

Какое отношение все это может иметь к работе мозга, содержащего значительное число нейронов, чувствительных к единичным квантам? Провести аналогию здесь мешает в первую очередь то, что квантовые эффекты быстро теряются в «шуме» — мозг слишком «горяч», чтобы квантовая когерентность (поведение, которое удобно описывать как непрерывное действие  $U$ ) сохранялась в нем сколько-нибудь продолжительное время. В моей терминологии это означало бы, что постоянно удовлетворяется одногравитонный критерий, так что операция  $R$  выполняется все время, изредка прерываясь операцией  $U$ .

Таким образом, пока у нас нет повода слишком надеяться на то, что квантовая механика откроет нам нечто новое о мозге. Возможно, мы все обречены быть просто компьютерами! Лично я в это не верю, но для окончательного выяснения вопроса нам необходимо идти дальше в наших исследованиях.

## За пределами квантовой теории?

Я хочу вновь обратиться к вопросу, который проходит красной нитью через большую часть этой книги: действительно ли наши представления об окружающем мире, управляемом законами классической и квантовой физики в их современном понимании, адекватны для описания мозга и разума? «Обычное» квантовое описание нашего мозга определенно заходит в тупик, поскольку акт «наблюдения» считается важной составляющей правильной интерпретации общепринятой квантовой теории. Следует ли считать, что мозг «наблюдает

сам себя» каждый раз при осознанном восприятии или возникновении мысли? Общепринятая теория не дает нам никаких указаний на то, каким образом квантовая механика могла бы принять это в расчет и, тем самым, как применить ее к мозгу в целом. Я попытался сформулировать вполне независимый от сознания критерий включения операции R («одногравитонный критерий»), и если нечто подобное удалось бы развить до полностью согласованной теории, то появилась бы возможность построения более ясного квантового описания мозга, чем существующее ныне.

Однако я считаю, что эти фундаментальные проблемы возникают не только при наших попытках описать деятельность мозга. Работа самих цифровых компьютеров существенно зависит от квантовых эффектов, пониманию которых, по-моему мнению, мешают трудности, внутренне присущие квантовой теории. Что это за «существенная» квантовая зависимость? Чтобы понять роль квантовой механики в цифровых вычислительных машинах, мы, прежде всего, должны выяснить, как можно заставить полностью *классический* объект вести себя подобно цифровому компьютеру. В главе 5 мы рассматривали классический «компьютер из бильярдных шаров» Фредкина-Тоффоли; но, как мы видели, в этом теоретическом «устройстве» были использованы идеализации, позволяющие обойти проблему существенной нестабильности, внутренне присущей классическим системам. Эта проблема нестабильности, как указано выше, проявляется в эффективном увеличении фазового объема эволюционирующей системы, которое почти неизбежно приводит к непрерывной потере точности операций, выполняемых классическим устройством. Именно квантовая механика позволяет в конце концов остановить это снижение точности. В современных электронных компьютерах необходимо существование *дискретных состояний* (скажем, для записи цифр 0 и 1), всегда позволяющих однозначно установить, когда компьютер находится в одном, а когда в другом состоянии. Это выражает саму суть «цифровой» природы компьютерных операций. Эта дискретность, в конечном счете, достигается за счет квантовой механики. (Мы можем вспомнить здесь квантовую дискретность энергетических состояний, спектральных частот, значений спина и т. д., см. главу 6.) Даже старые механические вычислительные машины зависели от *прочности* различных своих частей, каковая, в свою очередь, непосредственно вытекает из дискретности квантовой теории.

Но квантовая дискретность не является только следствием операции U. Пожалуй, уравнение Шредингера в *еще меньшей степени* способно предотвратить нежелательное расплывание фазового объема и «потерю точности», чем уравнения классической физики! Согласно U, волновая функция изолированной частицы, изначально локализованная в пространстве, будет все больше и больше расплываться с течением времени. Если бы не действие R время от времени, более сложные системы тоже были бы подвержены такой беспричинной делокализации (вспомним кошку Шредингера). (*Дискретные* состояния атома, например, характеризуются определенными значениями энергии, импульса и полного момента импульса. Общее состояние, которое как раз «расплывается», представляет собой суперпозицию таких дискретных состояний. Именно процедура R на некотором этапе заставляет атом на самом деле «быть» в одном из этих дискретных состояний.)

Мне представляется, что ни классическая, ни квантовая механика — если только в последнюю не будут внесены дальнейшие фундаментальные изменения, которые превратили бы R в «реальный» процесс, -никогда не смогут объяснить механизм *мышления*. Возможно, что даже работа цифровых компьютеров требует более глубокого понимания взаимосвязи действий U и R. В случае с компьютерами, мы, по-крайней мере, знаем, что цифровые вычисления являются *алгоритмическими* (по самой конструкции!), и мы не пытаемся «обуздать» предполагаемую *неалгоритмичность* физических законов. Но я утверждаю, что в случае с мозгом и разумом ситуация совершенно иная. Вполне допустимо, что в процессе (сознательного) мышления участвует некая существенная неалгоритмическая составляющая.

В следующей главе я попытаюсь подробно изложить причины, заставляющие меня верить в существование этой составляющей, а также выскажу предположения о том, какими удивительными реальными физическими эффектами обусловлено «сознание», влияющее на работу мозга.

### Примечания

1. Из радиовещания BBC (см. Ходжис [1983], с. 419).
2. Первые эксперименты такого рода были проведены на кошках (см. Мире, Сперри [1953]). За дальнейшими сведениями из области экспериментов с разделением полушарий мозга я отсылаю читателя к работам Сперри [1966], Газзаниги [1970] и Мак-Кей [1987].
3. Доступное изложение принципов действия зрительной коры можно найти у Хьюбела [1988].
4. См. Хьюбел [1988], с. 221. Ранние эксперименты позволили обнаружить клетки, чувствительные только к образу руки.
5. Общепринятая сегодня теория, согласно которой нервная система состоит из отдельных клеток — нейронов — была впервые предложена и убедительно обоснована великим испанским нейрофизиологом Рамоном-и-Кахалом около 1900 года.
6. На самом деле, *любые* логические элементы могут быть построены с помощью одних только операций " $\sim$ " и " $\&$ " (или даже только одной-единственной операции  $\sim (A\&B)$ ).
7. Фактически, использование логических элементов в большей степени отвечает конструкции электронного компьютера, чем изложенные в главе 2 особенности конструкции машины Тьюринга. В главе 2 особое внимание подходу Тьюринга было уделено по теоретическим соображениям. Начало действительному развитию компьютерных технологий положили в равной степени работы Алана Тьюринга и выдающегося американского математика венгерского происхождения Джона фон Неймана.
8. Эти сравнения во многом обманчивы. Подавляющее большинство транзисторов в современных компьютерах используется в устройствах «памяти» и не участвует в логических операциях; а память можно наращивать за счет внешних устройств практически бесконечно. При более интенсивном использовании параллельных вычислений количество транзисторов, непосредственно участвующих в выполнении логических операций, могло бы быть значительно больше, чем это принято в настоящее время.
9. Дойч в своих описаниях предпочитает использовать подход «множественности миров» относительно квантовой теории. Однако важно понимать, что это совершенно не существенно, поскольку концепция квантового компьютера принципиально не зависит от точки зрения на традиционную квантовую механику.
10. Этот комментарий перестает быть правомерным, если мы рассматриваем в качестве «классических» компонентов системы шестеренки, оси и т. п. Я предполагаю, что система состоит из обычных (скажем, точечных или сферических) частиц.

## Глава 10

### Где находится физика ума?

#### Для чего нужны умы?

В дискуссиях по проблеме «ум—тело» имеются два отдельных пункта, на которых обычно сосредоточивается внимание: «Каким образом материальный объект (мозг) может в действительности *пробуждать* сознание?»; и, наоборот: «Каким образом сознание усилием воли может реально воздействовать на (явно физически обусловленное) движение материальных объектов?» Это — пассивный и активный аспекты проблемы «ум—тело». Дело выглядит так, как если бы у нас в «уме» (или, вернее, в «сознании»), существовала некая нематериальная «вещь», которая, с одной стороны, активизируется материальным миром, а с другой — может оказывать на него воздействие. Я, однако, предпочитаю в своих предварительных замечаниях к этой последней главе обсудить до некоторой степени иной и, возможно, более научный вопрос, который относится к обеим проблемам, как пассивной, так и активной, — в надежде, что наши попытки найти ответ на него смогут приблизить нас к более глубокому пониманию этих извечных фундаментальных загадок философии. Мой вопрос звучит так: «Какое *преимущество естественного отбора* дает сознание тем, кто действительно им обладает?»

Такой формулировке вопроса присущи некоторые неявные допущения. Прежде всего — это уверенность в том, что сознание — это, на самом деле, «вещь», которую можно *научно описать*) что эта «вещь» действительно «что-то делает»; и, более того, что это

«что-то» приносит пользу существам, которые им обладают, в то время как другие создания, подобные первым во всем, кроме наличия сознания, демонстрируют менее эффективное поведение. С другой стороны, можно полагать, что сознание — это лишь пассивный спутник достаточно совершенной системы управления, и само по себе, в действительности, *не «делает» ничего*. (Это последнее утверждение является, вероятно, точкой зрения сторонников «сильного» ИИ.) В качестве альтернативы можно рассмотреть иную концепцию, согласно которой существует некое божественное или таинственное предназначение сознания — быть может, носящее телеологический характер и нам пока не ведомое — так что любое обсуждение этого феномена только лишь в терминах естественного отбора неизбежно уведет нас в сторону от истины. С моей точки зрения, из всех доводов подобного толка, наиболее убедительно и научно обоснованно здесь выглядел бы так называемый *антропный принцип*, согласно которому природа нашей вселенной такова, потому что в ней в обязательном порядке требуется присутствие разумных существ-«наблюдателей» наподобие нас с вами. (Этот принцип был вкратце упомянут в главе 8, с. 288, и я еще вернусь к нему позже.)

Я собираюсь обсудить эти вопросы в должное время, но вначале мы должны заметить, что термин «ум», пожалуй, несколько уводит нас в сторону, когда мы говорим о проблеме «ум—тело». Ведь очень часто говорят о «бессознательном уме» и это указывает на тот факт, что мы не рассматриваем термины «ум» и «сознание» как синонимы. Возможно, когда мы упоминаем о *бессознательном* уме, перед нами возникает неясный образ «суфлера», который незримо присутствует в каждой сцене, но кто обычно (кроме как, возможно, в снах, галлюцинациях, навязчивых состояниях или фрейдистских обмолвках) не посягает напрямую на контроль над нашим восприятием. Возможно, бессознательный ум в действительности *имеет* собственную способность осознавать, но в обычном состоянии это осознание пребывает совершенно отдельно от той части ума, которую мы традиционно называем «я».

Это, вообще говоря, не так уж и странно, как это может показаться на первый взгляд. Существуют эксперименты, которые, по-видимому, свидетельствуют о наличии определенного рода «сознания», присутствующего даже у пациента под общим наркозом на

операционном столе — в том смысле, что разговоры, которые ведутся во время операции, могут быть впоследствии «неосознанно» восприняты пациентом или же быть «проявлены» позже под гипнозом как действительно «воспринимавшиеся» в прошлом. Более того, ощущения, которые, казалось бы, были вытеснены из сознания гипнотическим внушением, могут позднее быть выявлены во время другого сеанса гипноза как «уже пережитые», но каким-то образом оказавшиеся записанными «на другую дорожку» (см. Окли, Имз [1985]). Эти результаты мне не вполне ясны, хотя я и не думаю, что было бы правильно приписывать обычную способность осознания бессознательному уму, но у меня нет особого желания пускаться здесь в рассуждения по этим вопросам. Тем не менее, проведение различий между бессознательным и осознающим себя умом — это действительно сложная и тонкая тема, к которой нам еще придется вернуться.

Попытаемся достичь возможно большей ясности в описании того, что мы подразумеваем под «сознанием» и что считаем признаками его проявления. Я не думаю, что было бы умно на данной стадии понимания пытаться предлагать точное *определение* сознания, но мы можем в достаточной степени полагаться на наши субъективные впечатления и интуитивный здравый смысл относительно того, что этот термин означает, и когда описываемый им феномен проявляет себя. Мне более или менее понятно, когда я нахожусь в сознании, и склонен считать, что и другие люди испытывают при этом нечто подобное. Чтобы находиться в сознании, я должен, как мне кажется, осознавать *что-то*, может быть, такие ощущения, как боль или тепло, или красочный пейзаж, или звуки музыки; или, возможно, я осознаю такое чувство как изумление, отчаяние или счастье; или я могу осознавать воспоминание о некотором событии в прошлом, или начинаю понимать то, что говорит кто-то другой; или осознавать собственную новую идею; или я могу осознанно намереваться заговорить или предпринять какое-то другое действие, например, встать со стула. Я могу также «отстраниться» и осознавать подобные намерения, или иное ощущение боли, или опыт, запечатленный в памяти, или акт понимания; или я могу даже просто осознавать свое собственное сознание. Я могу находиться в состоянии сна и все равно быть до некоторой степени осознающим происходящее, если мне снится сон; или, возможно, когда я начинаю просыпаться, я сознательно воздействую на развитие этого сна. Я готов считать, что сознание — это нечто, имеющее некоторую градацию, а не просто что-то, что есть или чего нет. Я считаю слово «сознание» в сущности синонимичным слову «осознание» (хотя, возможно, «осознание» немного пассивнее, чем то, что я понимаю под «сознанием»), в то время как «ум» и «душа» имеют дополнительные оттенки смысла, которые в значительной мере *менее* отчетливо определимы в настоящее время. У нас будет много хлопот с пониманием того, что такое «сознание» само по себе, поэтому я надеюсь, что читатель меня простит, если я оставлю в покое дальнейшие проблемы, связанные с терминами «ум» и «душа»!

Существует также вопрос о том, что подразумевать под словом *интеллект*. В конце концов, именно об этом объекте — а не о более расплывчатом понятии «сознания» — предпочитают говорить люди, связанные с ИИ. Алан Тьюринг в своей знаменитой работе (Тьюринг [1950]) (см. главу 1, с. 21) рассматривал непосредственно не столько «сознание», сколько «мышление», а слово «интеллект» даже было вынесено им в заглавие. На мой взгляд, вопрос об интеллекте является вторичным по отношению к вопросу о феномене сознания. Едва ли я поверю в то, что настоящий интеллект мог бы действительно существовать, когда бы его не сопровождало сознание. С другой стороны, если в итоге и вправду окажется, что приверженцы ИИ *способны* моделировать интеллект без присутствия сознания, *тогда было бы совершенно неудовлетворительным* определять интеллект, не включая в это понятие такой моделированный интеллект. Но в этом случае «интеллект» как предмет обсуждения оказался бы вне поля моего внимания, поскольку мой интерес связан, в первую очередь, с «сознанием».

Когда я высказываю свое убеждение, что истинный интеллект требует присутствия сознания, я при этом неявно предполагаю (поскольку я не разделяю точку зрения сторонников теории «сильного» ИИ, согласно которой простое применение алгоритма способно пробуждать сознание), что интеллект не может надлежащим образом моделироваться алгоритмическими средствами, то есть путем использования компьютера так, как это делается сегодня. (См. обсуждение «теста Тьюринга» в главе 1.) Очень скоро (см., в частности, обсуждение математического мышления, приведенное тремя разделами ниже, на с. 336) я постараюсь привести самые убедительные доводы в пользу необходимости присутствия существенно *неалгоритмической* составляющей в работе сознания.

Теперь обратимся к вопросу о том, *существует ли* четкое различие между одним объектом, который обладает сознанием, — и другим, «эквивалентным» первому во всем, кроме способности сознавать. Всегда ли сознание, присущее некоторому объекту, проявляет свое присутствие? Я предпочитаю думать, что ответить на этот вопрос следует однозначно «да». Однако, эта моя вера едва ли найдет поддержку в научных кругах, если там до сих пор нет согласия даже в вопросе о том, где можно найти сознание в царстве животных. Некоторые вообще не допускают мысли, что им могут обладать какие бы то ни было животные, отличные от людей (а некоторые придерживаются того же мнения и в отношении человеческих существ, живших за 1000 или более лет до н. э.; см. Джейнс [1980]); и в то же время кто-то готов допустить наличие сознания у насекомых, у червей и даже — почему бы нет? — у камней! Что касается меня, то я склонен сомневаться в том, что червь или насекомое — не говоря уже о камнях — в значительной степени (если вообще) обладают этим качеством; но млекопитающие, в общем и целом, подчас производят на меня впечатление существ, способных на подлинное осознание. Имея столь диаметрально противоположные точки зрения, приходится констатировать, что на сегодняшний день общепринятый критерий проявления сознания отсутствует. Правда, вполне возможно, что *есть все же* критерий сознательного поведения, хотя он и не заслужил всеобщего признания. Но не вызывает сомнений, что в любом случае только *активная* роль сознания могла бы иметь принципиальное значение, поскольку невозможно представить себе, чтобы простое наличие способности осознать, без активного дополнения к ней, может быть непосредственно зафиксировано. Подтверждением этому факту послужили ужасные случаи применения в 40-е годы лекарства на основе яда кураре в качестве «анестезирующего» средства при операциях, проводимых на маленьких детях, — тогда как действительный эффект этого средства заключается в парализации воздействия двигательных нервов на мускулы, из-за чего агония, которую в буквальном смысле *испытывали* несчастные дети, оставалась на протяжении операции незаметной для хирурга (см. Деннетт [1978], с. 209).

Вернемся к той гипотетической активной роли, которую *может* иметь сознание. Верно ли, что сознание может играть — а часто и играет — активную операционально различимую роль? Я полагаю, что это должно быть так и постараюсь сейчас обосновать свою убежденность несколькими независимыми доводами. Во-первых, благодаря нашему «здравому смыслу» мы часто ощущаем, что мы непосредственно воспринимаем, что другой человек *находится* в сознании. *Такое* впечатление вряд ли может быть ошибочным<sup>1</sup>). В то время как человек, который *находится* в сознании, может (подобно детям под действием кураре) и не подавать соответствующих признаков — находящийся в *бессознательном* состоянии едва ли будет выглядеть как человек, обладающий сознанием! Следовательно, должен существовать некий тип поведения, который можно было бы назвать характерным для человека, пребывающего в сознании (хотя даже и *не всегда* подтверждаемый самим сознанием), который мы бы воспринимали именно так благодаря нашим «интуитивным представлениям».

Во-вторых, примем во внимание безжалостный процесс естественного отбора. Будем рассматривать этот процесс в свете того факта, что, как мы видели в предыдущей главе, не вся активность мозга непосредственно доступна сознанию. И действительно, более «древний» мозжечок — обладающий значительным (по сравнению с остальными частями головного мозга) превосходством в плотности нейронов — производит, по-видимому, весьма сложные действия безо всякого вмешательства со стороны сознания. Однако, природа избрала для эволюционного развития таких сознающих себя и окружающий мир существ, как мы, вместо того, чтобы удовлетвориться созданиями, которые вполне могли бы существовать при помощи абсолютно бессознательных механизмов управления. Если сознание не служит целям селекции, то зачем природа занялась созданием «*сознательных*» *разновидностей* мозга, тогда как не наделенные сознанием «мозги-автоматы», наподобие мозжечка, могли бы функционировать не менее успешно?

Более того, существует простая «основополагающая» причина для предположения о том, что сознание должно иметь *какое-то* активное влияние, даже если его результат *не* является преимуществом при естественном отборе. Ибо почему еще мы (или существа, нам подобные) можем иной раз мучиться при попытке ответить на вопрос — особенно, если изучается эта тема — «о самих себе». (Мне так и хочется сказать: «Почему *вы* читаете эту главу?») или «Почему у *меня* было сильное желание написать книгу именно на эту тему?») Трудно себе представить, чтобы полностью лишенный сознания автомат стал бы тратить время на подобные вещи. А поскольку обладающие сознанием существа, с другой стороны, время от времени поступают *как раз таким* вот смешным образом, то их поведение *отличается* от поведения остальных — откуда следует, что сознание все-таки производит *определенное* активное воздействие! Разумеется, не составит труда специально запрограммировать компьютер так, чтобы он вел себя столь же нелепым образом (например, он мог бы в согласии с заложенным в него алгоритмом постоянно повторять на ходу: «О Господи, ну в чем же смысл жизни? Почему я здесь нахожусь? Что такое, черт побери, это „Я“, которым я себя ощущаю?»). Но почему же естественный отбор позаботился о создании благоприятных условий именно для такой расы индивидов, когда жестокий закон джунглей наверняка давно бы выщрал с корнем подобную бесполезную ерунду!

Мне кажется очевидным, что все эти размышления и бормотание, которым мы (времененно становясь философами) предаемся, не могли быть самоценными для процесса естественного отбора, а являются просто необходимым с его (естественного отбора) точки зрения «багажом», который должны нести существа, *обладающие* подлинным сознанием — при том, что само оно возникло в ходе естественного отбора по совершенно другой и, вероятно, очень серьезной причине. Этот багаж не слишком обременителен, и легко (хотя, быть может, и неохотно) переносится, скорее всего, именно непреклонными силами естественного отбора. В тех случаях, когда на земле царят мир и процветание, которыми человеческий род время от времени имеет счастье наслаждаться (ибо нам не всегда приходится бороться со стихиями (или нашими соседями) за выживание) — тогда, возможно, сокровища, содержащиеся в нашем багаже, становятся предметом удивления и любопытства. Именно в такие моменты, глядя на окружающих тебя философствующих людей, всерьез *убеждаешься* в том, что они, как и ты, тоже обладают умом.

### **Что в действительности делает сознание?**

Давайте согласимся с тем, что наличие у данного существа сознания является его реальным преимуществом в ходе естественного отбора. В чем конкретно может заключаться это

преимущество? Я знаком с одной точкой зрения, согласно которой способность осознавать происходящее может оказаться полезной хищнику в его попытке предугадать, что его жертва будет делать в следующий момент, с помощью мысленной «постановки себя на место» этой жертвы. *Воображая* себя своей жертвой, он мог бы увеличить шансы удачного исхода охоты.

Вполне может быть, что эта идея отчасти верна, но кое-что в ней меня смущает. Во-первых, здесь предполагается наличие некоторого сознания у самой жертвы, так как вряд ли имело бы смысл представлять себя автоматом, поскольку автомат — по определению, *не обладающий сознанием* — ни в коем случае не есть то, чем *можно* «быть»! В любом случае, я мог бы с таким же успехом представить себе, что совершенно не обладающий сознанием автомат-хищник имеет внутри в качестве подпрограммы последовательность действий, которая в точности соответствовала бы действиям автомата-жертвы. Мне вообще не кажется, что необходимость вводить сознание в отношения типа «хищник—жертва» может быть *логически* обоснована.

Разумеется, трудно понять, как случайные процессы естественного отбора могли быть достаточно умными, чтобы дать *автомату*-хищнику полную копию программы жертвы. Это бы выглядело скорее как *итионаж*, а не как естественный отбор! А *частичная* программа (в смысле отрезка «ленты» машины Тьюринга или чего-то подобного такой ленте) вряд ли была бы достаточно полезна хищнику с точки зрения естественного отбора. Ему был бы необходим какой-нибудь *завершенный* фрагмент этой ленты, а лучше (хотя и менее вероятно) — вся лента целиком. Следовательно (и как некая альтернатива этому), *частичная* правда могла бы заключаться в том, что модель поведения «хищник—жертва» подразумевает все же наличие не компьютерной программы, но определенного элемента сознания. Однако, это, по-моему, уже не относится к *настоящему* вопросу о различии между сознательным и «запрограммированным» действиями.

Суть идеи, о которой шла речь выше, по-видимому, напрямую связана с широко распространенной точкой зрения, согласно которой считается, что система может «осознавать» нечто, если в ней уже имеется модель этого объекта, и, соответственно, что она становится «*самосознающей*», когда она в *самой себе* содержит модель *самой себя*. Но если компьютерная программа содержит (скажем, в качестве подпрограммы) текст другой компьютерной программы, то первая при этом не может осознавать вторую, также как и не может приводить к самосознанию обращение программы к себе самой. Несмотря на подобные часто встречающиеся заявления, фундаментальные понятия самосознания и способности к осознанию в этих рассуждениях едва ли затрагиваются. Видеокамера не осознает сцены, которые она снимает; как не обладает самосознанием и видеокамера, направленная на зеркало .

Я хочу пойти в другом направлении. Как мы уже знаем, не вся деятельность нашего мозга сопровождается работой сознания (в особенности, это относится к действиям, которые управляются мозжечком). А какая часть из того, что мы можем делать осознанно, не может быть сделана бессознательно? Проблема становится еще менее ясной из-за того, что все, для чего нам изначально требуется сознание, похоже, может быть со временем заучено и впоследствии выполняться уже автоматически (возможно, именно мозжечком). Кажется, что сознание требуется, чтобы справляться с ситуациями, где нам приходится высказывать новые суждения, и где правила не были заданы заранее. Трудно достичь большой точности при проведении различий между теми видами умственной деятельности, которые, по-видимому, требуют подключения сознания, и теми, которые нет. Сторонники «сильного» ИИ (да и многие другие) будут настаивать на том, что «*формирование новых суждений*» — это не более, чем повторное применение ряда точно сформулированных алгоритмических правил — только теперь на некоем полумистическом «высоком уровне», так, что их действия нами при этом не осознаются. Однако, как мне кажется, даже в нашем обиходном лексиконе есть такие термины — используемые нами в повседневной жизни для разделения умственной деятельности на

осознанную и бессознательную — которые уже сами по себе могут по меньшей мере *навести на мысль* о различиях между действиями неалгоритмической и алгоритмической природы.

Возможно, эти различия не всегда достаточно четко очерчены хотя бы потому, что в наши сознательные суждения входит немало *неосознанных* факторов: опыт, интуиция, предрассудки, даже наше привычное использование логики. Но я утверждаю, что сами по себе суждения — это проявления работы *сознания*. Поэтому я полагаю, что, тогда как бессознательные действия мозга происходят в соответствии с алгоритмическими процессами, действие сознания имеет совершенно иную природу, и потому не может быть описано никаким алгоритмом.

Есть некая ирония в том, что многое из того, что я излагаю здесь, является полной противоположностью по отношению к некоторым другим точкам зрения, которые мне приходится довольно часто слышать. Например, утверждают, что именно *сознательный* ум «рационален» и доступен пониманию; тогда как бессознательные действия нередко загадочны и труднообъяснимы. Те, кто работает с искусственным интеллектом, часто считают, что как только мы сможем понять ход осознанной мысли, то можно сразу же будет придумать соответствующий алгоритм для его компьютерной реализации; а вот таинственные бессознательные процессы нашему пониманию (пока!) не доступны. В моем представлении эти процессы вполне могут быть алгоритмическими, но при этом настолько сложными, что их детальный анализ практически невозможен. Четко осознаваемый ход мысли, который может быть разумно объяснен как нечто полностью логичное, в свою очередь может быть (зачастую) переведен на язык алгоритмов — но на *совершенно ином уровне*. Мы сейчас говорим не о внутренних процессах (возбуждении нейронов и т. п.), а о манипулировании законченными мыслями. Иногда оно носит алгоритмический характер (как в случае традиционной логики: древнегреческих силлогизмов, формализованных Аристотелем; или символической логики, разработанной математиком Джорджем Булем; см. Гарднер [1958]); а иногда — неалгоритмический (как в случае с теоремой Геделя или некоторыми примерами, приведенными в главе 4). А как реализовать на компьютере *формирование суждений*, которое я рассматриваю как критерий наличия сознания, — об *этом* разработчики ИИ не имеют даже ни малейшего представления!

Иногда мне возражают, что, поскольку *критерии* для этих суждений не являются в конце концов осознанными, то как я могу приписывать такие суждения сознанию? Однако, тем самым мои оппоненты упускают самую суть тех идей, которые я пытаюсь выразить. Ведь я не требую, чтобы мы осознавали, *как* мы формируем наши сознательные впечатления и суждения. Это привело бы к смешению тех уровней, о которых я только что упоминал. Истинные *основания* наших осознанных впечатлений наверняка будут недоступны сознанию. Они должны были бы рассматриваться на более глубоком (материальном) уровне по сравнению с первопричинами наших явных мыслей, которые мы непосредственно осознаем. (Ниже я осмелюсь предложить на рассмотрение одну интересную гипотезу на этот счет!) Собственно сознательные впечатления и *являются* (неалгоритмическими) суждениями.

Эта тема, затрагивающая вопрос о возможной *неалгоритмической* составляющей механизма нашего осознанного мышления, проходила красной нитью и через все предыдущие главы. В частности, заключительная часть дискуссии в главе 4, особенно касающаяся теоремы Геделя, подводила к мысли о том, что (по крайней мере, в математике) сознательное «вглядывание» иной раз позволяет нам устанавливать справедливость утверждения способом, недоступным для алгоритма. (Подробнее я остановлюсь на этом доводе чуть позже.) Конечно же, сами по себе алгоритмы не способны находить истину! Построить алгоритм, генерирующий только ложные суждения, столь же просто, как и алгоритм, результатом работы которого были бы одни только

истины. Для определения Пригодности или непригодности того или иного алгоритма нам совершенно необходимо своего рода вдохновение, интуитивное прозрение, приходящее извне (далее я еще вернусь к этому вопросу). И я утверждаю, что именно эта способность провидения (или «интуитивного постижения») глубокого различия между истиной и ложью (равно как и между красотой и уродством!) является признаком наличия сознания.

Я, однако, должен сразу же оговориться, что ни в коем случае не имею здесь в виду какое-то мистическое «ясновидение». Сознание абсолютно бесполезно при попытке угадать счастливое число в (честно проводимой) лотерее! Я имею в виду суждения, которые постоянно формируются человеком в сознательном состоянии, когда собираются воедино и сопоставляются все относящиеся к предмету размышлений факты, данные чувственного опыта, воспоминания — а в иную минуту вдохновения даже рождаются мудрые мысли. В принципе, мы располагаем достаточным количеством информации для того, чтобы вынести соответствующее суждение — но процесс его осмысленного формирования путем выделения необходимой информации из трясины фактов может просто не иметь точного выражения на языке алгоритмов (или же подобное выражение существует, но может оказаться при этом бесполезным практически). Возможно, мы находимся в ситуации, что когда суждение уже сделано, некоторый алгоритмический процесс (или просто более простое суждение) проверяет его справедливость, но не его изначальное формирование. В такой ситуации, как мне кажется, сознание «нашло бы себя» в роли создателя подходящих суждений.

Почему я утверждаю, что неалгоритмическое построение суждений является критерием наличия сознания? Отчасти я опираюсь здесь на свой опыт ученого-математика. Я просто не доверяю своим механическим действиям, если они не были сперва придирчиво исследованы сознанием. Часто сам по себе алгоритм, использующийся при определенных вычислениях, не вызывает сомнения — но *тот ли* алгоритм используется для решения данной конкретной задачи? Рассмотрим простой пример: если вас заставят вызубрить алгоритмы перемножения двух чисел и деления одного числа на другое (или даже разрешат использовать запрограммированный карманный калькулятор) — гарантирует ли это, что вы сможете определить в каждом конкретном случае, какое из этих действий приведет к решению поставленной перед вами задачи? Для этого нужно *думать* и строить *осознанное* суждение. (Вскоре мы увидим, почему такие суждения должны быть, по крайней мере иногда, неалгоритмическими!) Разумеется, коль скоро вы решите большое количество однотипных задач, выбор между умножением и делением станет настолько привычным, что будет выполняться совершенно автоматически—не исключено, что при участии одного лишь мозжечка. На этой стадии осознанное восприятие происходящего не является больше необходимым, поэтому можно спокойно позволить своему сознанию занять ум иными проблемами или просто «отпустить его в свободное плавание» — разве что время от времени проверяя ход выполнения алгоритма.

То же самое постоянно происходит на всех уровнях математического мышления. Люди часто стремятся найти адекватные алгоритмы, когда занимаются математикой, но само это стремление отнюдь не кажется алгоритмической процедурой. Как только подходящий алгоритм найден, задача, в некотором смысле, уже решена. Более того, определение с точки зрения математики степени точности или пригодности алгоритма требует значительных усилий со стороны сознания. Нечто подобное имело место при обсуждении формальных систем для математики, которые были описаны в главе 4. Если начать с формулировки нескольких аксиом, то затем из них можно вывести различные математические утверждения. Не исключено, что последняя операция может оказаться алгоритмической, но все же изначально математик должен осознанно решить вопрос об адекватности этих аксиом. Почему это решение с необходимостью будет неалгоритмическим, должно стать ясным из рассуждений, идущих непосредственно после следующего параграфа. Но прежде, чем мы перейдем к этому вопросу, давайте

посмотрим, какая теория возникновения мозга и принципов его деятельности является на сегодняшний день наиболее популярной.

## Естественный отбор алгоритмов?

Если предположить, что умственная деятельность человека — как осознанная, так и нет — это всего лишь выполнение очень сложного алгоритма, то сразу же возникает вопрос: а как, собственно, мог возникнуть такой в высшей степени эффективный алгоритм. Стандартным ответом здесь, разумеется, будет «естественный отбор». Как только стали появляться существа, наделенные мозгом, между ними возникла конкуренция, в которой побеждали (т. е. выживали и производили более многочисленное потомство) те, чей алгоритм оказывался эффективнее. Их потомки тоже имели, как правило, более эффективные алгоритмы поведения, чем их родственники, коль скоро им посчастливилось унаследовать составляющие этих «продвинутых» алгоритмов от своих родителей; так постепенно алгоритмы улучшались — не обязательно стабильно, поскольку могли случаться и значительные колебания в их эволюции — пока не было достигнуто то поразительное совершенство, которое (по всей видимости) можно обнаружить в мозге современного человека (см. Доукинс [1986]).

Даже если судить сообразно моей точке зрения, *некоторая* доля истины в этой картине должна быть, поскольку, как мне представляется, большая часть работы мозга действительно носит алгоритмический характер, и к тому же — как читатель наверняка догадался из предыдущих рассуждений — я являюсь убежденным сторонником (теории) естественного отбора. Но я не понимаю, как естественный отбор сам по себе мог дать рождение алгоритмам, которые позволяли бы делать осознанные выводы касательно *правомерности* применения всех прочих алгоритмов, которыми мы должны, по идее, пользоваться.

Представьте себе обычную компьютерную программу. Как *она* появилась на свет? Ясно, что никак не за счет (непосредственно) естественного отбора! Чтобы это произошло, какой-нибудь программист должен был бы разработать ее и убедиться, что она корректно выполняет те действия, для которых она предназначена. (В действительности большинство сложных компьютерных программ содержат ошибки — как правило, незначительные и малозаметные, которые выявляются только в достаточно редких случаях при необычных стечениях обстоятельств. Наличие таких ошибок не влияет существенно на мои рассуждения.) Иногда компьютерная программа может быть «написана» другой компьютерной программой-«мастером», но тогда та, в свою очередь, с необходимостью должна быть создана человеческим гением; то же самое относится и к тем программам, в состав которых могут входить фрагменты кодов, написанных другими компьютерными программами. Но в любом случае, задача обоснования использования конкретного алгоритма и разработка общей концепции программы «ложится на плечи» (по крайней мере) одного человеческого сознания.

Можно представить себе, конечно, что все могло бы происходить совсем не так, и что по прошествии достаточного количества времени компьютерная программа могла бы самостоятельно эволюционировать в процессе некоего естественного отбора. Если верить, что действия, производимые в сознании программистов, сами являются всего лишь алгоритмами, то надо с необходимостью согласиться с тем, что таким же образом *развились* и алгоритмы. Но меня здесь беспокоит то, что принятие решения о правомерности использования алгоритма отнюдь не является алгоритмическим процессом. Кое-что об этом уже было сказано в главе 2. (Будет или нет машина Тьюринга на самом деле *останавливаться* — вопрос, который не может

быть решен алгоритмическим путем.) Чтобы решить, будет ли алгоритм действительно *работать*, нужно *глубокое понимание*, а не просто еще один алгоритм.

Тем не менее, можно было бы представить себе, что существует определенный процесс естественного отбора, который способен создавать *довольно* эффективные алгоритмы. Лично мне, однако, очень трудно в это поверить. Любой процесс отбора такого рода мог бы оказывать воздействие только *на результаты* выполнения алгоритмов<sup>2</sup>), а не на лежащие в основе этих алгоритмов идеи. И это не только совершенно неэффективно — я думаю, что это не принесло бы вообще никакого результата. Во-первых, нелегко определить, глядя на итог работы алгоритма, что он из себя на самом деле представляет. (Нетрудно сконструировать две простые, но совершенно различные процедуры для машины Тьюринга, выходные ленты которых совпадали бы, скажем, до  $2^{65536}$ -й позиции — и тогда их различие не было бы замечено даже за всю историю вселенной!) Более того, малейшая «мутация» алгоритма (например, небольшое изменение в описании машины Тьюринга или в ее выходной ленте) сделала бы ее полностью бесполезной; поэтому трудно понять, как настоящие *усовершенствования* алгоритмов могут получаться таким вот случайным образом. (Даже *обдуманые* усовершенствования труднореализуемы, когда неизвестен их точный «смысл». Так традиционно получается в тех нередко возникающих ситуациях, когда необходимо внести изменения или исправления в сложную и небрежно задокументированную программу, чей автор находится вне пределов досягаемости или давно умер. Тогда вместо того, чтобы пытаться разобраться в хитросплетениях разнообразных промежуточных значений и неявных подзадач, на которых базируется эта программа, иной раз бывает проще стереть все и начать заново!)

Предположим, что мог бы быть разработан гораздо более «здоровый» метод определения алгоритмов, в отношении которого вышеприведенная критика становилась бы беспочвенной. В некотором роде, это как раз то, о чем я и говорю. «Здоровые» определения — это *идеи*, на которых базируется алгоритм. Но идеям, насколько нам известно, для своего выражения требуется разум, наделенный сознанием. А значит, мы вновь возвращаемся к проблеме определения сознания и тех его свойств, которые отличают обладающих сознанием существ от остального мира — и к вопросу о том, как, черт побери, естественный отбор мог оказаться достаточно «умным», чтобы развивать *такие замечательные* свойства.

Результаты естественного отбора и в самом деле удивительны. Те скромные познания о функционировании человеческого мозга — и, разумеется, мозга любых других существ — которыми я обладаю, просто ошеломляют меня, заставляя испытывать благоговейный трепет. Работа отдельного нейрона поразительна, но все вместе нейроны представляют еще более впечатляющую структуру, с самого момента рождения насчитывающую огромное множество соединений и связей, которые в дальнейшем позволят ей решать практически любые задачи. Замечательно не только само сознание, но и все то, что связано с обеспечением его деятельности!

Если бы однажды нам довелось обнаружить то свойство, которое позволяет физическому объекту обретать сознание, то, изучив его во всех деталях, мы, вероятно, могли бы сконструировать подобные объекты для собственных нужд — хотя они не обязательно расценивались бы как «машины», в современном понимании этого слова. Нетрудно предположить, что такие объекты могли бы многократно нас превосходить, ибо они были разработаны *специально* для этой задачи — т. е. для *обретения сознания*. Им бы не пришлось вырастать из одной клетки. Им бы не пришлось нести на себе «багаж» предков (старые и «бесполезные» части мозга или тела, которые продолжают существовать в нас только благодаря «несчастьям», приключившимся с нашими далекими предками). Можно также представить себе, что благодаря этим преимуществам, такие объекты могли бы *по-настоящему* заменять собой людей там, где (по мнению тех, кто согласен со мной) алгоритмические компьютеры обречены на выполнение обслуживающих функций.

Но вполне возможно, что тема сознания имеет гораздо больше аспектов. Может быть и так, что каким-то образом наше сознание действительно зависит от нашего наследия и от миллионов лет эволюции, лежащих у нас за спиной. Меня не покидает ощущение, что в самой эволюции, в ее явном «нащупывании» пути к какой-то будущей цели есть что-то загадочное и непостижимое. Кажется, что все организовано несколько лучше, чем оно «должно было быть» на основе слепой эволюции и естественного отбора. Вполне возможно, однако, что внешние проявления здесь обманчивы. Возможно, это как-то связано с тем способом, каким действуют физические законы, что позволяет естественному отбору протекать гораздо эффективнее, чем в случае, если бы этот процесс управлялся произвольными законами. Возникающее в результате явно «интеллектуальное нащупывание» — это отдельная интересная тема, к которой я вернусь несколько позже.

### Неалгоритмическая природа математической интуиции

Как я уже указывал ранее, моя уверенность в том, что сознание способно влиять на характер суждений об истинности неалгоритмическим путем, опирается главным образом на результаты теоремы Геделя. Если мы видим, что сознание действует неалгоритмически при формулировании *математических* суждений, где вычисления и строгие доказательства являются неперемнным требованием, то уж наверняка нас нетрудно будет убедить и в том, что эта неалгоритмическая составляющая могла бы являться решающей и для роли сознания при более общих (не связанных с математикой) обстоятельствах.

Вспомним доводы, приведенные в главе 4 в рамках доказательства теоремы Геделя и устанавливающие ее применимость к решению вопроса о вычислимости. Там было показано, что *какой бы* (достаточно сложный) алгоритм ни использовал математик для установления математической истины или, что то же самое  $1^1$ , какую бы *формальную систему* он<sup>3</sup>) ни принял для задания своего критерия истинности — всегда найдутся математические суждения, подобные сформулированному Геделем утверждению  $P_j(t)(f_c)$  для системы (см. с. 98), на которые его алгоритм не сможет дать ответа. Если ум математика работает полностью алгоритмически, то алгоритм (или формальная система), которые он обычно использует для построения своих суждений, оказываются не в состоянии справиться с утверждением  $P^*(k)$ , полученным с помощью его собственного алгоритма. Тем не менее, *мы* можем (в принципе) понять, что  $P^*(f_c)$  на самом деле *истинно*] Этот факт, по всей видимости, должен был бы указать *ему* на противоречие, поскольку *он*, как и *мы*, -не может не заметить его. А это, в свою очередь, может свидетельствовать о не-алгоритмическом характере его рассуждений!

В этом заключается суть довода, предложенного Лукасом [1961] в поддержку точки зрения, согласно которой деятельность мозга не может быть полностью алгоритмической, против которого, однако, время от времени выдвигались различные контрдоводы (см., например, Бенасерраф [1967], Гуд [1969], Льюис [1969, 1989], Хофштадтер [1981], Бови [1982]). В связи с этой дискуссией я должен подчеркнуть, что термины «алгоритм» и «алгоритмический» относятся к чему угодно, что может быть (достоверно) смоделировано на компьютере общего назначения. Сюда включается, конечно, как «параллельная обработка», так и «нейросети» (или «машины с переменной структурой связей»), «эвристика», «обучение» (где всегда заранее задается определенный фиксированный шаблон, по которому машина должна обучаться), а также взаимодействие с внешним миром (которое может моделироваться посредством входной ленты машины Тьюринга). Наиболее серьезным из этих контраргументов является следующий: чтобы *действительно* убедиться в истинности утверждения  $P^*(\&)$ , нам нужно *знать*, какой именно алгоритм использует математик, и при этом быть уверенным в правомерности его использования

в качестве средства достижения математической истины.

Если в голове у математика выполняется очень сложный алгоритм, то у нас *не будет возможности* узнать, что он из себя представляет, и поэтому мы не сможем сконструировать для него утверждение геделевского типа, не говоря уже об уверенности в обоснованности его применения.

Такого типа возражения часто выдвигаются против утверждений подобных тому, которое я привел в начале этого раздела, а именно, что теорема Геделя свидетельствует о неалгоритмическом характере наших математических суждений. Но сам я не нахожу это возражение слишком убедительным. Предположим на мгновение, что способы, которыми математики формируют осознанные суждения о математической истине действительно *являются* алгоритмическими. Попробуем, используя теорему Геделя, доказать абсурдность этого утверждения от противного (*reductio ad absurdum*!).

Прежде всего мы должны рассмотреть возможность того, что разные математики используют *неэквивалентные* алгоритмы для суждения об истинности того или иного утверждения. Однако — и это одно из наиболее поразительных свойств математики (может быть, почти единственной в этом отношении среди всех прочих наук) — истинность математических утверждений может быть установлена посредством абстрактных рассуждений! Математические рассуждения, которые убеждают одного математика, с необходимостью убедят и другого (при условии, что в них нет ошибок и суть нигде не упущена). Это относится и к утверждениям типа геделевского. Если первый математик готов согласиться с тем, что все аксиомы и операции некоторой формальной системы всегда приводят только к *истинным* утверждениям, то он также должен быть готов принять в качестве истинного и соответствующее этой системе геделевское утверждение. Точно то же самое произойдет и со вторым математиком. Таким образом, рассуждения, устанавливающие математическую истину, являются *передаваемыми* И.

Отсюда следует, что мы, говоря об алгоритмах, имеем в виду не какие-то неясные разномастные построения, которые, возможно, рождаются и бродят в голове каждого отдельного математика, а *одну* универсально применяемую формальную систему, которая *эквивалентна* всем возможным алгоритмам, использующимся математиками для суждений о математической истине. Однако мы никак не можем знать, является ли эта гипотетическая «универсальная» система той, которая используется математиками для установления истинности. Ибо в этом случае мы *могли бы* построить для нее геделевское утверждение, и знали бы наверняка, что оно математически истинно. Следовательно, мы приходим к заключению, что алгоритм, который математики используют для определения математической истины, настолько сложен или невразумителен, что даже правомерность его применения навсегда останется для нас под вопросом.

Но это бросает вызов самой сущности математики! Основопологающим принципом всего нашего математического наследия и образования является непоколебимая решимость не склоняться перед авторитетом каких-то неясных правил, понять которые мы не надеемся. Мы должны *видеть* — по крайней мере, в принципе — что каждый этап рассуждений может быть сведен к чему-то простому и очевидному. Математическая истина не есть некая устрашающе сложная догма, обоснованность которой находится вне границ нашего понимания — она строится из подобных простых и очевидных составляющих; и когда они становятся ясны и понятны нам, с их истинностью соглашаются все без исключения.

С моей точки зрения, получить такое явное *reductio ad absurdum* (без применения настоящего математического доказательства) мы даже и мечтать не могли! Основная идея должно быть теперь ясна. Математическая истина — это не то, что мы устанавливаем просто за счет использования алгоритма. Кроме того, я полагаю, что наше *сознание* — это решающая составляющая в нашем понимании математической истины. Мы должны «видеть» истинность математических рассуждений, чтобы убедиться в их

обоснованности. Это «видение» — самая суть сознания. Оно должно присутствовать *везде*, где мы непосредственно постигаем математическую истину. Когда мы убеждаемся в справедливости теоремы Геделя, мы не только «видим» ее, но еще и устанавливаем неалгоритмичность природы самого процесса «видения».

## **Вдохновение, озарение и оригинальность**

Я должен попытаться как-то прокомментировать те внезапные вспышки озарения, которые мы называем вдохновением. Откуда берутся все эти мысли и образы? Может быть, они появляются из нашего бессознательного — или все же сознание существенным образом связано с их рождением? Можно привести множество примеров из воспоминаний великих мыслителей, где они прямо указывали на такие события. Как математик, я особенно интересуюсь теми случаями, когда вдохновение посещало именно математиков, но думаю, что между математикой и другими науками и искусством есть много общего. Эта тема великолепно изложена в небольшой работе «Исследования психологии процессов изобретательства в области математики» — классическом труде выдающегося французского математика Жака Адамара — к которой я и отсылаю читателя. В ней он приводит многочисленные примеры озарения в изложении ведущих математиков и не только. Один из наиболее известных случаев связан с Анри Пуанкаре. В начале Пуанкаре описывает свои напряженные сознательные исследования, связанные с построением так называемых «функций Фукса», которые в конце концов явно зашли в тупик. И вот что он пишет далее:

«...Я покинул Кон, где я жил в то время, чтобы принять участие в геологической экспедиции, организованной Горной школой. Впечатления от поездки заставили меня забыть о моей математической работе. Достигнув местечка Кутонс мы сели в омнибус, чтобы добраться на нем до следующего пункта назначения. В тот момент, когда я ставил ногу на подножку, мне пришла в голову идея, которая, казалось, никоим образом не вытекала из моих прошлых раздумий, что преобразования, используемые мной для определения функций Фукса, были идентичны определенным преобразованиям в неевклидовой геометрии. Я не проверил эту идею. У меня просто не было времени, так как когда я занял свое место в омнибусе, я продолжил прерванную беседу — но я был совершенно уверен в правильности моей догадки. Вернувшись в Кон, я выбрал свободное время и, проверив для собственного спокойствия свое предположение, убедился в его справедливости».

Что поражает в этом примере (как и во многих других, приведенных Адамаром) — это внезапность появления столь сложной и глубокой идеи в сознании Пуанкаре, которое в тот момент было занято совершенно другим; и тот факт, что возникновение этой идеи сопровождалось четким ощущением ее истинности, которую полностью подтвердили последующие расчеты. Тут нужно сразу оговориться, что подобные идеи сама по себе далеко не так просты, чтобы их можно было легко выразить словами. Думаю, что для ясного изложения своих мыслей Пуанкаре потребовалось бы провести примерно часовой семинар для экспертов в этой области. Ясно, что эта идея могла полностью оформиться в сознании Пуанкаре только после долгих часов размышлений, направленных на изучение всех возможных аспектов указанной проблемы. Да, в некотором смысле, идея, осенившая Пуанкаре, когда он сидел в омнибусе, была «единичной» идеей, которую можно было полностью осознать в один момент. Еще более замечательной представляется убежденность Пуанкаре в ее справедливости — убежденность, которая сделала последующую детальную проверку этой идеи почти что излишней.

Пожалуй, мне стоит попытаться соотнести этот случай с моим собственным опытом,

который оказывается в каком-то смысле похожим. На самом деле, я не могу вспомнить ни одной ситуации, когда хорошая идея пришла бы мне в голову «с неба», как это произошло в случае с Пуанкаре (или во многих других известных примерах подлинного вдохновения). Что касается меня, то мне нужно, чтобы выполнялись определенные условия: мне необходимо *думать* о том вопросе, над которым я в данный момент работаю, пусть даже в «фоновом режиме» и не целенаправленно, но обязательно осознанно; вполне возможно, что при этом я буду заниматься чем-то посторонним и успокаивающе-монотонным, например, бриться; вероятно, я заново возьмусь размышлять о проблеме, которая на некоторое время была отложена в сторону. И, конечно же, нужно посвятить не один час упорным сознательным раздумьям, после которых может потребоваться определенное время для того, чтобы вновь переключиться на решаемую проблему. И, тем не менее, ощущение, когда искомое решение возникает при таких условиях подобно «вспышке» — и при этом ты совершенно уверен в его правильности — мне достаточно хорошо знакомо.

Вероятно, стоит привести конкретный пример, который может оказаться небезынтесным и с другой точки зрения. Осенью 1964 года меня занимал вопрос о сингуляр-ностях черных дыр. Оппенгеймер и Снайдер в 1939 году показали, что строго *сферический* коллапс массивной звезды может приводить к образованию центральной сингулярности пространства-времени, выходящей за пределы классической общей теории относительности (см. главу 7, с. 271, 274). Многие считали, что этого неприятного вывода можно было бы избежать, если бы удалось убрать (необоснованное) предположение о *строгой* сферической симметрии. В противном случае получается, что вся коллапсирующая материя стремится к единой центральной точке, где (как вполне закономерно было бы предположить, учитывая симметричность ситуации) возникает сингулярность бесконечной плотности (вещества). Вполне разумным кажется предположение о том, что *без* такой симметрии материя попадала бы в центральную область далеко не так согласованно, вследствие чего сингулярности бесконечной плотности могло бы и не получиться. Возможно даже, что вся материя в этом случае снова «раскрутилась» бы, демонстрируя поведение, совершенно отличное от идеализированной черной дыры Оппенгеймера и Снайдера И.

Стимулом для моих собственных размышлений на эту тему послужил вновь возникший интерес к проблеме черных дыр, связанный со сравнительно недавно открытыми квазарами (в начале 1960-х годов). Физическая природа этих на удивление ярких (с учетом отделяющих их от Земли расстояний) астрономических объектов вызвала у некоторых специалистов предположение о том, что в центре каждого из них может находиться нечто наподобие черной дыры Оппенгеймера—Снайдера. С другой стороны, многие считали, что гипотеза Оппенгеймера—Снайдера о сферической симметрии может привести здесь к совершенно неверным представлениям. Однако мне пришло в голову (в контексте другой работы, которую я выполнял), что можно было бы сформулировать и доказать точную математическую теорему, указывающую на *неизбежность* возникновения сингулярностей в пространстве-времени (в рамках стандартной общей теории относительности) и тем самым подтверждающую наличие черных дыр — при условии достижения коллапсом определенной «точки необратимости». Я не знал, как математически можно было бы определить такую точку (не используя при этом условия сферической симметрии), не говоря уже о формулировке или доказательстве соответствующей теоремы. В то время приехал непостижимым) является способ, которым (иные) художники могут представить себе весь замысел своего творения целиком. Подобный удивительный случай очень живо описан Моцартом (см. Адамар [1945], с. 16):

«Когда я чувствую себя хорошо и нахожусь в добром расположении духа; или когда я предпринимаю поездку или отправляюсь на прогулку после сытной трапезы; или ночью во время бессонницы — мне в голову приходит сколько угодно самых разных идей. Откуда и как они приходят? Я не знаю и ничего не могу с этим поделать. Те, которые мне приятны, я удерживаю в голове и часто напеваю без слов; так, по крайней мере, мне говорили. Когда у меня возникает

тема, сразу же приходит следующая мелодия, соединяясь с первой согласно требованиям композиции в целом: контрапункт, партия каждого инструмента — и, наконец, все музыкальные фрагменты складываются в завершенное произведение. Тогда моя душа горит вдохновением. Произведение растет; я постоянно дополняю его, прорабатываю все более мелкие детали, пока в один прекрасный момент композиция не оказывается полностью сформирована у меня в голове, хотя она может быть и довольно длинной. Тогда мой ум охватывает ее единым взглядом, как красивую картину или прекрасную девушку. Это не последовательный процесс, при котором различные части произведения прорабатываются до мелочей и стыкуются друг с другом (так, как это будет сделано в дальнейшем) — нет, я слышу его целиком, как это позволяет мое воображение».

Мне кажется, что это хорошо согласуется с моей схемой «предложения-отбора». «Предложение» представляется здесь бессознательным («я ничего не могу с этим поделать»), хотя и, несомненно, в высокой степени избирательным, в то время как «отбор» выполняет функцию сознательного судьи вкуса («те, которые мне приятны, я удерживаю...»). Масштабность идеи, рожденной в минуты вдохновения, особенно отчетливо проглядывает в высказывании Моцарта («это не последовательный процесс... нет, я слышу его (произведение) целиком») и Пуанкаре («Я не проверил эту идею. У меня просто не было времени...»). Более того, я готов настаивать, что нашему сознательному мышлению в целом подобная масштабность присуща изначально. К этому вопросу я еще вернусь.

### Невербальность мысли

Одно из главных утверждений, которое Адамар делает в своей работе о творческом мышлении — это убедительное опровержение популярного сегодня тезиса, который гласит, что вербализация необходима для формирования мысли. Вряд ли здесь можно привести возражение более убедительное, чем то, которое содержится в письме Альберта Эйнштейна Адамару:

«Слова или язык, как в устной, так и в письменной форме, по-видимому, не играют никакой роли в механизме моего мышления. Психические сущности, которые, по-видимому, и являются составляющими элементами мысли — это определенные знаки и более или менее отчетливые образы, которые могут „произвольно“ воспроизводиться и комбинироваться по собственному желанию... В моем случае, упомянутые элементы носят визуальный и моторный характер. Общепринятые слова или другие знаки мне приходится подбирать только на второй стадии, когда упомянутые ассоциативные связи приобретают отчетливые очертания и могут быть воспроизведены по моей воле».

Еще здесь стоит процитировать видного генетика Фрэнсиса Гальтона:

«Для меня серьезную трудность представляет письмо, а еще большую — словесное изъяснение, так как размышления в словесной форме даются мне далеко не так легко, как в любой другой. Часто случается, что проделав большую работу и получив результаты, которые мне абсолютно ясны и вполне меня удовлетворяют, при попытке выразить их словами я сталкиваюсь с необходимостью переводить себя в совершенно иную интеллектуальную плоскость. Мне приходится переключать свои мысли на язык, который не слишком-то хорошо им соответствует. Поэтому я вынужден тратить уйму времени в поисках подходящих слов и фраз, и часто осознаю, что, выступая без подготовки, бываю непонятен не из-за неясности содержания высказывания, а только лишь из-за неуклюжести своих вербальных конструкций. Это один из небольших, но досадных моих недостатков».

Нечто сходное пишет и сам Адамар:

«Я утверждаю, что слова полностью отсутствуют в моей голове, когда я действительно *предаюсь* раздумьям, и я нахожу случай Гальтона полностью идентичным моему личному опыту, поскольку и у меня самого даже после прочтения или выслушивания вопроса все слова исчезают в тот самый момент, когда я начинаю их обдумывать; и я полностью согласен с Шопенгауэром, когда он пишет: „Мысли умирают в момент, когда воплощаются в слова“».

Я цитирую эти примеры, потому что они очень хорошо согласуются с моим собственным способом мышления. Почти все мое математическое мышление визуализируется или протекает на уровне не-вербальных понятий, хотя мысли очень часто сопровождаются пустыми и почти бесполезными словесными комментариями, такими как «вот это идет с этим, а это — с этим». (Иногда я могу употреблять слова для выражения простых логических выводов.) Трудности, которые испытывали упомянутые ученые при переводе своих мыслей на язык слов, я часто испытывал и сам. Причиной тому в большинстве случаев служило просто-напросто отсутствие адекватных терминов, способных выразить требуемые понятия. Действительно, я часто веду расчеты, используя специально разработанные диаграммы, которые представляют собой сокращенную запись определенных типов алгебраических выражений (см. Пенроуз и Риндлер [1984]). Необходимость перевода таких диаграмм в слова — это очень трудоемкий процесс, и я это делаю только в случае крайней необходимости, когда нужно подробно объяснить что-то другим. И еще одно наблюдение: я случайно заметил, что если сосредотачиваю все свое внимание на математике и некоторое время занимаюсь только ей, а потом кто-то внезапно обращается ко мне, то в течение нескольких следующих секунд я почти не способен говорить.

Не могу сказать, что я никогда не думаю в словесной форме — просто я нахожу слова почти бесполезными для *математического* мышления. Другие виды рассуждений, возможно, такие, как *философские*, являются, вероятно, гораздо более подходящими для вербального выражения. Может быть, поэтому так много философов считают язык неотъемлемым средством интеллектуальной деятельности и сознательного мышления! Нет сомнения, что каждый человек думает по-своему — это подтверждает и мой собственный опыт, и мнений других математиков. Наиболее полярными стилями математического мышления являются, как кажется, аналитический/геометрический. Интересно, что Ацамар считал себя аналитиком, хотя использовал скорее визуальные, чем вербальные образы в своем математическом мышлении. Что касается меня, то я в значительной степени тяготею к геометрическим методам. Если же говорить обо всех математиках, то разброс здесь окажется весьма широк.

Но коль скоро мы согласились с тем, что значительная часть сознательного мышления, на самом деле, может иметь невербальный характер — а с моей точки зрения к другому выводу приведенные выше соображения привести не могут — тогда, наверное, читателю будет нетрудно поверить также и в то, что подобное мышление может иметь неалгоритмическую составляющую!

Напомню, что в главе 9 (с. 310) я упоминал о часто встречающейся точке зрения, согласно которой только одно полушарие мозга — то, где находится центр речи (левое у большинства людей) — способно также и на сознательное мышление.

После ознакомления с вышеизложенным читателю должно быть ясно, почему я считаю эту точку зрения совершенно неприемлемой. Я не знаю, используют ли, как правило, математики одно полушарие чаще, чем другое; но нет сомнения в том, что для истинного математического мышления необходим высокий уровень сознания. В то время как аналитическое мышление, по всей видимости, сосредоточено в левой половине мозга, геометрическое мышление, напротив, часто приписывают *правой* по-

ловине; так что вполне разумной является предположение о том, что значительная часть *сознательных* математических рассуждений проводится *все-таки* в правом полушарии!

## Сознание у животных?

Прежде чем закончить рассуждения о важности вербализации применительно к сознанию, я должен еще рассмотреть один вопрос, который вкратце уже затрагивался ранее, а именно — могут ли живые существа, отличные от нас, обладать сознанием? Мне кажется, что люди иногда используют неспособность животных говорить как аргумент, отрицающий саму возможность наличия у них достаточно развитого сознания и, как следствие, позволяющий априори отказывать им в самых элементарных «правах». Читатель может легко догадаться, что для меня подобные рассуждения являются неубедительными, поскольку для многих сложных разновидностей сознательного мышления (например, для математического) вербализация и вовсе не требуется. А некоторые по той же причине — т. е. из-за отсутствия речевых способностей — считают, например, что правая сторона мозга обладает сознанием «не более», чем шимпанзе (см. Леду [1985], с. 197-216).

Существуют значительные разногласия относительно способности горилл и шимпанзе выразить свои мысли при помощи *языка знаков* (а не обычной человеческой речи, которую они не могут воспроизводить из-за особенностей строения их голосовых связок; см. статьи Колина Блэйкмора и Сьюзан Гринфилд (Блэйкмор, Гринфилд [1987])). Хотя полемика еще продолжается, один факт уже не вызывает сомнений: эти человекообразные обезьяны могут общаться таким образом, по крайней мере, на некотором элементарном уровне. По моему глубокому убеждению, отказываться признать в этом общении «вербализацию» — это прямое проявление высокомерия со стороны тех, кто придерживается такой точки зрения. Видимо, закрывая обезьянам доступ в «клуб способных к вербализации», надеются автоматически исключить их и из «клуба обладающих сознанием»!

Оставляя пока в стороне вопрос о речи, обратимся к убедительным свидетельствам, которые указывают на способность шимпанзе к подлинному «вдохновению». Конрад Лоренц описывает шимпанзе в комнате, где к потолку был подвешен банан, до которого обезьяна не могла достать, а в одном из углов был поставлен ящик:

«Задача не давала ему покоя, и он возвращался к ней вновь и вновь. Затем внезапно — по-другому и не скажешь — его прежде унылая физиономия „озарилась“. Взгляд шимпанзе перемещался с банана на пустое пространство под ним, оттуда на ящик, потом снова на место под бананом, и оттуда на банан. В следующий момент он издал крик радости и кувыркнулся в сторону ящика, явно пребывая в превосходнейшем настроении. Совершенно уверенный в успехе, он толкнул ящик под банан. Могу поспорить, что никто из видевших его в тот момент не усомнился бы в способности человекообразных обезьян к таким прозрениям, испытав которые, в пору воскликнуть „Эврика!“».

Обратите внимание, что точно так же, как в случае с Пуанкаре, когда тот садился в омнибус, шимпанзе был «совершенно уверен в успехе» еще до того, как он проверил свою идею. И если я прав, утверждая, что подобные суждения требуют участия сознания, то перед нами оказывается неопровержимое свидетельство того, что животные действительно могут обладать сознанием.

Глядя на дельфинов (и китов), мы невольно задаемся одним интригующим вопросом. Как нетрудно заметить, головной мозг дельфинов имеет такие же (или даже

большие) размеры, как и наш собственный; а кроме того, дельфины могут посылать друг другу чрезвычайно сложные звуковые сигналы. Вполне возможно, что такой большой мозг нужен для каких-то иных целей, которые не сводятся к «интеллектуальной» деятельности в человеческом или окологлазочеловеческом понимании. Более того: не имея рук, приспособленных для хватания, они не могут создать «цивилизацию», которую мы были бы способны оценить. И хотя они по той же самой причине не могут писать книг, они вполне способны время от времени превращаться в философов и размышлять о смысле своей жизни! Что, если они иногда передают свое ощущение «самосознания» при помощи этих сложных звуковых сигналов, распространяющихся под водой? Я не встречал ни одного исследования, где бы изучалось, используют ли дельфины какую-то одну определенную сторону мозга для «вербализации» и общения друг с другом. В связи с проведенными на людях операциями по разделению мозга, которые загадочным образом влияли на целостность «я» человека, следует отметить еще одну особенность дельфинов: их полушария никогда не погружаются в сон одновременно — вместо этого каждая сторона мозга спит по очереди. Согласитесь, хорошо было бы выяснить у них, как *они* «ощущают» целостность своего сознания!

### Соприкосновение с миром Платона

Я уже упоминал о том, что разные люди скорее всего мыслят по-разному — и даже у разных математиков мысли при решении математической задачи формируются не одинаково. Я вспоминаю, что, поступая на математический факультет университета, я ожидал, что мои будущие коллеги-математики должны думать примерно так же, как я. В школе мои одноклассники, казалось, думали совсем иначе, чем я, что меня несколько удручало. «Теперь, — думал я с восторгом, — я найду коллег, с которыми общаться мне будет гораздо легче! Некоторые будут мыслить более продуктивно, чем я, а некоторые — менее; но все они смогут настроиться на мою ментальную длину волны!» Как же заблуждался! Думаю, что тогда я познакомился с гораздо большим числом различных способов мышления, чем за все предыдущее время! Да, мой собственный образ мыслей был куда более геометрическим и далеко не столь аналитическим по сравнению с остальными — но у них было и множество других различий в способе мышления. У меня всегда вызывало затруднение понимание словесного описания формулы, в то время как у многих из моих коллег, казалось, с этим не возникало никаких трудностей.

Довольно часто случалось так, что, слушая своего коллегу, пытающегося объяснить мне какую-нибудь математическую выкладку, я практически совсем не улавливал логической связи между следующими друг за другом наборами слов. Однако, в моей голове постепенно формировалась догадка о содержании передаваемых мне идей — причем складывалась она в рамках моей собственной терминологии и, скорее всего, была мало связана с ментальными образами, которыми оперировал мой коллега, обращаясь к данной проблеме, — и тогда я отвечал. К моему удивлению, эти ответы чаще всего воспринимались как адекватные, и беседа продолжала развиваться в таком же ключе, причем к концу становилось ясно, что состоялся поистине позитивный обмен мнениями. Однако сами предложения, которые произносил каждый из нас в ходе беседы, чаще всего оставались не поняты! В последующие годы, будучи уже профессиональным математиком (или физиком-математиком), я пришел к выводу, что ситуация в целом практически не изменилась по сравнению с тем временем, когда я учился на младших курсах. Возможно, с увеличением моего математического багажа я стал несколько лучше разбираться, о чем говорят другие, пытаюсь донести до моего сознания определенную мысль; и, наверное, я научился адаптировать свой стиль изложения, каждый раз подстраиваясь под конкретного слушателя. Однако, в сущности, все осталось по-прежнему.

Для меня часто является загадкой, как вообще возможно *подобное* общение, но теперь я все же осмелюсь дать некоторое объяснение, которое, как мне кажется, могло бы иметь самое непосредственное отношение к уже затронутым ранее вопросам. Суть здесь заключается в том, что

при общении математиков происходит *не* только обмен *фактами*. Чтобы состоялась передача ряда фактов от одного собеседника другому, первому из них необходимо излагать эти факты достаточно понятно, а второму — воспринять каждый из них в отдельности. Но в математике *фактическое* содержание играет второстепенную роль. Математические утверждения являются с необходимостью истинными (или же с необходимостью ложными!), и даже если первый математик своим утверждением только нащупывает искомую истину, то именно эту истину воспримет его собеседник (конечно, если исходное утверждение будет им правильно понято). Ментальные конструкции второго математика могут в деталях отличаться от тех образов, которые возникают у первого, равно как могут отличаться и их словесные описания — но соответствующая математическая идея в результате все-таки будет передана.

Такой тип общения был бы совершенно невозможен, если бы не то обстоятельство, что *интересные* или *глубокие* математические истины растворены (с небольшой плотностью) в массе всех возможных математических истин. Если бы передаваемая истина заключала в себе, скажем, *неинтересное* утверждение наподобие  $4897 \times 512 = 2\,507\,264$ , то второму собеседнику, естественно, придется полностью понять первого, иначе это точное утверждение не сможет быть передано. Но при сообщении математически *интересного* утверждения часто удается понять его интуитивно, даже если для его описания использовались расплывчатые образы и понятия.

Это может показаться парадоксальным, поскольку математика — это предмет, где точность всегда ставится превыше всего. В самом деле, в письменных отчетах большое внимание уделяется точной формулировке и завершенности всех утверждений. Однако, чтобы передать математическую идею (обычно посредством словесного описания), такая точность иной раз является помехой, так что вначале может потребоваться менее четкая описательная форма. А как только будет понята самая суть идеи — тогда можно уже переходить и к деталям.

Как же получается, что математические идеи могут передаваться подобным образом? Лично мне представляется, что всякий раз, когда ум постигает математическую идею, он вступает в контакт с миром математических понятий Платона. (Вспомним, что, по Платону, математические идеи имеют собственное бытие и населяют некий идеальный мир, доступ в который осуществляется только благодаря работе интеллекта (см. с. 88, 134).)

Когда человек «видит» математическую истину, его сознание пробивается в этот мир идей и устанавливает с ним кратковременный прямой контакт (т. е. осуществляет «доступ посредством интеллекта»). Я описал это «видение» в связи с теоремой Геделя, хотя, вообще говоря, здесь заключена сущность математического понимания. Общение математиков становится возможным постольку, поскольку у каждого из них в этот момент есть *прямой путь к истине*, а сознание каждого способно при этом постигать математические истины непосредственно, путем «видения». (В самом деле, часто акт понимания сопровождается словами типа «О, я вижу!»<sup>4</sup>.) Так как каждый математик может установить непосредственный контакт с миром идей Платона, то общение их друг с другом проходит значительно легче, чем это можно было бы ожидать. Ментальные образы, возникающие у каждого из них, когда осуществляется соприкосновение с миром Платона, могут быть существенно различными, но общение тем не менее возможно, поскольку каждый находится в прямом контакте с *одним и тем же* существующим вне нас миром Платона!

В соответствии с этой точкой зрения, наш ум всегда способен на подобный прямой контакт. Но за один раз можно продвинуться лишь на немного. Математическое открытие как раз и состоит в расширении области контакта. Поскольку математические истины

являются с необходимостью истинами, никакой содержательной «информации» в общепринятом смысле этого слова исследователь не получает. Вся информация уже находилась там изначально. Все, что требовалось — это соединить разные части друг с другом и «увидеть» ответ! Это очень хорошо согласуется с представлениями самого Платона о том, что (скажем, математическое) открытие — это всего лишь одна из форм *воспоминания*] В самом деле, меня часто поражало сходство между двумя состояниями, когда ты мучительно стараешься вспомнить чье-то имя — и когда пытаешься найти адекватное математическое понятие. В обоих случаях искомое в некотором смысле *уже присутствует* в голове, хотя во втором случае «вспоминание» математической идеи связано с необычной формой вербализации.

Чтобы такие идеи были полезны для объяснения принципов математического общения, нужно представить себе, что интересные и глубокие математические идеи отличаются способностью к более «основательному» существованию, чем неинтересные или тривиальные. Эти соображения пригодятся нам при рассмотрении ряда умозрительных заключений, приведенных в следующем разделе.

## Взгляд на физическую реальность

Любой подход к вопросу о возникновении сознания в царстве физической реальности неявно подразумевает необходимость определения природы самой физической реальности.

Концепция «сильного» ИИ, например, заключается в том, что «разум» возникает через воплощение достаточно сложного алгоритма, по мере того как этот алгоритм реализуется различными объектами физического мира. При этом, сущность этих объектов, согласно данной теории, значения не имеет. Нервные сигналы, электрический ток, идущий по проводам, винтики, приводные ремни или водопроводные трубы — все это одинаково подходит в качестве «аппаратной части». Алгоритм рассматривается как нечто самодостаточное. Однако, для предоставления алгоритму возможности «существовать» вне зависимости от любой физической реализации совершенно необходимо разделять воззрения Платона на природу математики. Сторонникам «сильного» ИИ было бы нелегко согласиться с альтернативной точкой зрения о том, что «математические понятия существуют только лишь в умах», поскольку при этом образуется замкнутый круг, где для возникновения умов требуется наличие алгоритмов, а для изначального существования алгоритмов — наличие умов! Они могли бы попытаться придерживаться иного мнения, согласно которому алгоритмы могут существовать как знаки на листе бумаги, или как магнитные силовые линии в кусках железа, или как смещения зарядов в памяти компьютера. Но такое упорядочение материала само по себе не есть алгоритм. Чтобы стать алгоритмом, оно должно иметь *интерпретацию*, т. е. должна существовать возможность *декодирования* этого порядка; а это будет зависеть от «языка», на котором написан алгоритм. Снова оказывается, что нужен заранее существующий разум, чтобы «понимать» этот язык, и мы возвращаемся к исходной точке. Предположив затем, что алгоритмы населяют мир Платона, и именно там обитает разум (согласно теории «сильного» ИИ), мы оказываемся перед вопросом отношения между физическим миром и миром Платона. Что, на мой взгляд, является в теории «сильного» ИИ аналогом проблемы «ум—тело»!

Я же придерживаюсь иной точки зрения, поскольку считаю, что (сознательный) ум существенно *неалгоритмичен*. Но меня несколько смущает то, что между взглядами сторонников «сильного» ИИ и моими собственными существует большое число точек соприкосновения. Я уже указывал, что считаю сознание тесно связанным со способностью воспринимать несомненные истины, и тем самым осуществлять прямой контакт с миром математических понятий Платона. Это неалгоритмическая процедура—и отнюдь не алгоритмы должны населять тот мир, который

так важен для нас — и вновь проблема «ум—тело», если следовать этой точке зрения, оказывается тесно связанной с вопросом об отношении мира Платона к «реальному» миру существующих физических объектов.

В главах 5 и 6 мы видели, насколько хорошо реальный физический мир согласуется с некоторыми исключительно точными математическими теориями (ПРЕВОСХОДНЫЕ теории, см. с. 129). И эта поразительная точность неоднократно подчеркивалась многими исследователями (см., в частности, Вигнер [1960]). Мне трудно поверить — хотя многие считают это неоспоримым — в справедливость утверждения о том, что ПРЕВОСХОДНЫЕ теории могли возникнуть просто в результате случайного естественного отбора идей, при котором «выживают» только лучшие. В это трудно поверить потому, что уж *слишком* они хороши для того, чтобы оказаться среди идей, возникших случайным образом. На самом деле должна быть какая-то фундаментальная взаимосвязь между математикой и физикой, т. е. между миром Платона и физическим миром.

Говоря о «мире Платона», мы приписываем ему некоторый вид реальности, которая определенным образом сравнима с реальностью физического мира. С другой стороны, сама реальность физического мира кажется уже менее очевидной, чем она представлялась до появления теорий относительности и квантовой механики (относящихся к ПРЕВОСХОДНЫМ теориям) (см. с. 129, 130 и, в особенности, 232). Сама точность этих теорий обеспечивает почти математический абстрактный уровень существованию нашей физической реальности. Не является ли это своего рода парадоксом? Как может конкретная реальность превратиться в абстрактную, да еще и математическую? Возможно, это обратная сторона вопроса о том, как абстрактные математические понятия могут становиться почти ощутимо реальными в мире Платона. Возможно, в каком-то смысле, эти два мира, на самом деле — *один и тот же* мир? (См. Вигнер [1960], Пенроуз [1979а], Барроу [1988], а также Эт-кинс [1987].)

Хотя я с большой симпатией отношусь к идее такого отождествления двух миров, вопрос этим далеко не исчерпывается. Как я уже упоминал в главе 3, и выше в этой главе, некоторые математические истины в мире Платона кажутся «более реальными» («более глубокими», «более интересными», «более многообещающими?»), чем остальные. Такими должны быть все те истины, которые наиболее тесно связаны с явлениями физической реальности. (Примером здесь может служить система комплексных чисел (см. главу 3), которая является основной составляющей квантовой механики, ибо амплитуды вероятности выражаются через комплексные числа.) Если принимать в расчет возможность такого отождествления, то становится более понятным, как «разум» мог бы играть роль таинственного связующего звена между физическим миром и математическим миром Платона. Вспомним

также (см. главу 4), что есть много областей математического мира — более того, наиболее глубоких и интересных его областей — которым присущ неалгоритмический характер. Поэтому было бы разумным, основываясь на той точке зрения, которую я здесь излагаю, предположить, что неалгоритмические процессы должны играть в физическом мире весьма существенную роль. И я склоняюсь к тому, что эта роль тесно связана с самим понятием «разума».

### Детерминизм и жесткий детерминизм

До сих пор было мало сказано о вопросе «свободы воли», который обычно считается неотъемлемым при рассмотрении *активной* составляющей проблемы «ум—тело». Вместо этого, я уделил основное внимание предположению о наличии существенно *неалгоритмической* составляющей в той роли, которую играет осознанное действие. Обычно тема свободы воли обсуждается в связи с детерминизмом в физике. Вспомним, что в большинстве существующих ПРЕВОСХОДНЫХ теорий типа существует явно выраженный детерминизм: если известно состояние системы в определенный момент времени <sup>15</sup>1, то оно полностью определяется в любой

более поздний (или ранний) момент из уравнений теории. Таким образом, по-видимому, для «свободы воли» не остается места, поскольку будущее поведение системы кажется полностью обусловленным физическими законами. Даже U-часть квантовой механики имеет такой же полностью детерминистский характер. Однако R-часть, связанная с «квантовым скачком», не является детерминистской, внося элемент случайности в эволюцию системы во времени. Был момент, когда исследователи старались найти именно здесь свободу воли, полагая, что действие сознания может непосредственно влиять на «скачок» отдельной квантовой системы. Но если R-часть действительно случайна, то это тоже нам не слишком поможет, если мы хотим конструктивное применение нашей свободе воле.

Моя собственная точка зрения (правда, не очень четко сформулированная в этом случае) заключается в том, что должен быть применен некий новый подход (ПКТГ; см. главу 8), который работал бы на границе между квантовой и классической физикой, интерполируя между U и R (каждая из которых теперь рассматривается как аппроксимация); и этот подход должен содержать *существенно неалгоритмический* элемент. А это подразумевает, что будущее *не будет вычислимым* на основе настоящего, даже если оно им и *определяется*. Я пытался по возможности наиболее ясно определить смысловые различия терминов «вычислимость» и «детерминизм» в главе 5. Мне кажется, что ПКТГ может быть детерминистской, но невычислимой теорией<sup>5</sup>). (Вспомним невычислимую «игрушечную модель», которую я описал в главе 5, с. 144.)

Многие при этом считают, что даже классический (или U-квантовый) детерминизм не является детерминизмом *в полном* смысле этого слова, поскольку исходные условия в принципе не могут быть известны с такой точностью, которая действительно позволила бы просчитать будущее. Иногда совсем небольшие изменения исходных условий могут привести к очень значительным различиям в конечном результате. Именно так возникает «хаос» в (классической) детерминистской системе — явление, приводящее, например, к неопределенностям в прогнозе погоды. Однако очень трудно поверить, что этот вид классической неопределенности может позволить нам сохранять (иллюзорную?) веру в существование свободы воли. Будущее поведение все равно будет *детерминированным* в каждый момент времени, начиная с Большого взрыва, даже если мы окажемся не в состоянии его вычислить (см. с. 147).

То же самое возражение может быть выдвинуто и против моей идеи о том, что *невычислимость* связана скорее с особенностями законов динамики — которые в этом случае считаются исходно неалгоритмическими — чем с нехваткой информации о начальных условиях. *Невычислимое* будущее, согласно этой точке зрения, все равно будет полностью обусловлено прошлым — вплоть до момента Большого взрыва. На самом деле я не настолько привержен догмам, чтобы настаивать на том, что методы ПКТГ должны быть по сути детерминистскими, но невычислимыми. Я полагаю, что искомая теория должна иметь более тонкий характер, вследствие чего подобное грубое описание будет к ней просто неприменимо. Единственное, на чем я настаиваю — так это на необходимости присутствия в ней существенно неалгоритмических элементов.

Завершая этот раздел, я хотел бы упомянуть еще об одном представлении о природе детерминизма, причем из числа весьма радикальных. Я называю его *жестким детерминизмом* (Пенроуз [19876]). Согласно этой теории, не просто будущее предопределяется прошлым — *вся история вселенной* оказывается *раз и навсегда* определенной в соответствии с некоторой точной математической схемой. Такая концепция могла бы привлечь тех, кто склонен каким-нибудь образом отождествлять мир Платона с физическим миром — ибо застывший навеки мир Платона с его однозначной определенностью не оставляет в этом случае вселенной никаких «альтернативных возможностей»! (Я иногда задаю себе вопрос: мог ли Эйнштейн иметь в виду подобную схему, когда он писал: «Что меня

собственно интересует, это следующее: мог ли Бог сотворить мир другим, оставляет ли какую-то свободу требование логической простоты» (письмо Эрнсту Штрауссу; см. Кузнецов [1980], с. 363.)

С одним из вариантов жесткого детерминизма мы сталкиваемся в квантово-механической концепции «множественности миров»\* (см. главу 6, с. 240). В соответствии с ней, вышеупомянутая точная математическая схема определяла бы не *единственную* отдельную историю вселенной, а всю совокупность из мириадом мириадом «возможных» историй вселенной. Несмотря на малопривлекательный характер (по крайней мере для меня) такой схемы и множество проблем и несоответствий, которые она в себе несет, мы все же не имеем права сбрасывать ее со счетов как потенциально возможную.

Мне кажется, что, если принять жесткий детерминизм, но *без* множественности миров, то математическая схема, которая управляет структурой вселенной, вероятно, *должна* быть неалгоритмической<sup>61</sup>. Ибо в противном случае можно было бы, в принципе, просчитать свои будущие действия, а затем вдруг «решить» сделать нечто совершенно другое — получаем очевидное противоречие между «свободой воли» и жестким детерминизмом нашей теории. Вводя не-вычислимость, можно избежать этого противоречия, хотя должен признаться, что я не вполне уверен в адекватности решения такого типа и предвижу в будущем гораздо более тонкое описание «реально действующих» (неалгоритмических) правил, которым подчиняется наш мир!

### Антропный принцип

Насколько важно сознание для вселенной в целом? Могла бы вообще вселенная существовать без населяющих ее сознательных существ? Намеренно ли законы физики задумывались такими, чтобы обеспечить существование сознательной жизни? Является ли наше место во вселенной — как в пространстве, так и во времени — каким-то особенным? Вот вопросы, которые ставит перед нами научная гипотеза, известная как «*антропный принцип*».

Этот принцип может принимать различные формы (см. Барроу, Типлер [1986]). Наиболее приемлемая из них затрагивает только вопрос пространственно-временного расположения сознательной (или «разумной») жизни во вселенной. Это — «*мягкий*» антропный принцип. Он может использоваться для объяснения того, почему условия оказались именно такими, что в современную эпоху стала возможна жизнь на Земле. Ответ прост: ведь если бы не было подходящих условий, то мы должны были бы находиться где-то в другом месте, и в иное (благоприятное) время. Этот принцип был очень эффективно использован Брэндоном Картером и Робертом Диком, чтобы разрешить вопрос, остававшийся для физиков загадкой на протяжении многих лет. Вопрос касается существования определенных числовых соотношений между физическими константами (гравитационная постоянная, масса протона, возраст вселенной и т.д.). Интригующим в этих закономерностях был тот факт, что некоторые из них сложились только в настоящую эпоху истории Земли, тем самым указывая на некую — быть может, случайную — исключительность нашего положения во времени (с точностью до нескольких миллионов лет, разумеется!). Впоследствии Картер и Дик нашли этому следующее объяснение: предположили, что эта эпоха совпадает с временем жизни так называемых звезд главной последовательности, одной из которых является наше Солнце. В любую другую эпоху, согласно их утверждениям, нигде и близко не было бы разумной жизни, чтобы измерить те самые физические константы — так что совпадение *должно* было иметь место просто потому, что разумная жизнь возникла бы только в то время, когда *есть* подобное совпадение!

«Жесткий» антропоцентрический принцип идет еще дальше. В этом случае мы рассматриваем наше уникальное положение в пространстве-времени не только этой вселенной, но и бесконечного множества других *возможных* вселенных. Исходя из этого, мы можем сделать ряд предположений относительно того, почему физические константы — или, в более широком смысле, законы физики — как будто специально были спроектированы так, чтобы разумная жизнь вообще могла существовать. Допустим, что константы (или законы) отличались бы от наблюдаемых — тогда мы просто *не могли бы* появиться в этой вселенной и *должны были бы* оказаться в некоторой другой! По моему мнению, достоинства «жесткого» антропоцентрического принципа несколько сомнительны, и теоретики прибегают к нему всякий раз, когда не находят адекватной теории для объяснения наблюдаемых фактов (в первую очередь, это касается теорий физики частиц, где за отсутствием разумного объяснения массам частиц, предполагается, что если бы их значения отличались от настоящих, то жизнь, вероятнее всего, была бы вообще невозможна, и т.д.). С другой стороны, «мягкий» антропный принцип представляется мне безупречным при условии, что им пользуются крайне осмотрительно.

Взяв на вооружение антропный принцип — либо в «жесткой», либо в «мягкой» формах, — можно попытаться показать, что зарождение сознания было *неизбежно* благодаря тому факту, что сознательные существа, то есть «мы», должны были присутствовать, чтобы наблюдать этот мир — так что *нет* необходимости предполагать, как это делал я, будто способность осознавать дает какое-то преимущество в процессе естественного отбора! По моему мнению, этот довод технически корректен, и доказательство, опирающееся на «мягкий» антропный принцип (по крайней мере), *могло бы* указать на причину, по которой сознание существует в нашем мире независимо от благоволения к нему естественного отбора. С другой стороны, я не могу поверить в то, что антропный принцип и есть та *настоящая* (или единственная) причина, которая обеспечивает эволюцию сознания. Существует достаточно много самых разнообразных свидетельств, способных утвердить меня во мнении, что сознание на самом деле *является* сильным преимуществом в процессе естественного отбора, и что, следовательно, совсем необязательно апеллировать к антропному принципу.

### «Плиточные» структуры и квазикристаллы

Теперь я отойду от масштабных обсуждений последних нескольких разделов и сосредоточусь на обсуждении вопросов, которые, хотя и являются до некоторой степени дискуссионными, все же гораздо более научны и «осязаемы». Возможно, вначале эти рассуждения покажутся отклонением от темы, однако, их важность для нас станет очевидной уже в следующем разделе.

Вспомним примеры «плиточных» замощений, изображенные на рис. 4.12 (с. 118). Эти образцы интересны потому, что они «почти» нарушают общепринятую математическую теорему о кристаллических решетках, которая утверждает, что для кристаллических решеток возможны только симметрии с осью второго, третьего, четвертого и шестого порядков. Под кристаллической решеткой я подразумеваю дискретную систему точек, которая обладает *трансляционной симметрией*. Это означает, что можно определенным образом перемещать решетку без вращения так, чтобы она переходила сама в себя (иными словами, в результате такого сдвига она не изменяется) — а, значит, у такой решетки будет существовать *параллелограмм периодов*. Единственное, что нам здесь важно — это если в нашем распоряжении есть вещество, в котором все атомы расположены в узлах кристаллической решетки с подобной структурой, то оно будет выглядеть, как кристалл, обладая при этом запрещенной симметрией с осью пятого порядка!

В декабре 1984 году израильский физик Дэни Шехтман, работавший вместе с кол-

легами в Национальном бюро стандартов в США, в Вашингтоне, объявил об открытии фазы алюминиево-марганцевого сплава, который был похож на кристаллоподобное вещество — теперь называемое *квазикристаллом* — с осью пятого порядка. На самом деле, у этого квазикристаллического вещества наблюдалась симметрия не только на плоскости, но и в *трех* измерениях — так что в итоге получалась запрещенная икосаэдральная симметрия (Шехтман и др. [1984]). (Икосаэдральный трехмерный аналог моей плоской «плиточной» структуры с осью пятого порядка был открыт Робертом Амманом в 1975 году; см. Гарднер [1989].) Сплавы Шехтмана образовывали только крошечные микроскопические квазикристаллы, достигавшие примерно  $10^{-3}$  мм в поперечном сечении, но позднее были найдены другие квазикристаллические вещества, в частности — алюминиево-литиево-медный сплав, у которого икоса-эдрально симметричные образования могут вырастать до размеров порядка миллиметра, т. е. становятся вполне различимы невооруженным глазом .

Замечательным свойством этих квазикристаллических «плиточных» структур является то, что процесс их составления имеет существенно *нелокальный* характер. Иными словами: при построении подобного покрытия необходимо время от времени проверять состояние кристаллической решетки на расстоянии многих и многих «атомов» от места сборки, чтобы избежать серьезных ошибок при соединении составных частей. (Это чем-то напоминает то почти «сознательное нащупывание», которое я связывал с естественным отбором.) Наличие такого свойства является одной из причин серьезных разногласий, возникающих сегодня в связи с вопросом о квазикристаллических структурах и их выращивании, так что было бы неразумно пытаться делать окончательные выводы до тех пор, пока не будут разрешены некоторые основополагающие проблемы. Тем не менее, никто не запрещает нам выдвигать предположения; поэтому я рискну высказать здесь свою собственную точку зрения. Во-первых, я полагаю, что некоторые из этих квазикристаллических веществ действительно имеют сложное внутреннее строение, и что расположение атомов в их структуре довольно точно повторяет строение тех плиточных структур, которыми я занимался. Во-вторых, отсюда я делаю (всего лишь гипотетическое) заключение о том, что их образование не может совершаться за счет последовательного добавления атомов, как это происходит в рамках *классической* картины роста кристаллов — но с необходимостью должна опираться на *не-локальные* и непременно квантово-механические принципы построения [7].

Механизм такого роста я представляю себе следующим образом: вместо присоединения отдельных атомов к постоянно движущейся линии роста (в случае классического роста кристаллов), происходит квантовая линейная суперпозиция большого числа различных альтернативных сочетаний присоединяющихся атомов (путем квантовой операции U). В самом деле, согласно квантовой механике, все именно так и *должно* (почти всегда) происходить! В каждый момент времени существует не одна возможная структура, но множество альтернативных расположений атомов в сложной линейной суперпозиции. Некоторые из этих структур вырастают в гораздо более крупные образования, так что в определенный момент различия между гравитационными полями альтернативных структур превзойдут «одиогравитонный предел» (или его более подходящий в данном случае аналог; см. главу 8, с. 298). На этой стадии одна из них — или, скорее, это снова будет суперпозиция, но уже в несколько урезанном виде — выделиться в качестве истинной структуры (квантовая операция R). В этот процесс роста, сопровождающийся последовательным отказом от наименее «значимых» на каждом этапе альтернатив, будут вовлекаться все большее и большее количество исходного вещества, пока наконец не сформируется достаточно крупный квазикристалл.

Обычно, когда природа ищет кристаллическую конфигурацию, из всех возможных она

выбирает ту, которая характеризуется *наименьшим уровнем энергии* (считая фоновую температуру нулевой). Нечто аналогичное, по-моему, должно происходить и в процессе роста кристаллов, с той только разницей, что такое состояние с наименьшей энергией гораздо труднее обнаружить, а «наилучшее» расположение атомов не может быть получено просто последовательным добавлением каждый раз одного атома в надежде на то, что индивидуальному атому для этого будет достаточно решить свою *собственную* задачу минимизации. Вместо этого нам предстоит решать эту же задачу для *всей* совокупности атомов, а значит, потребуется их *совместное* усилие. Такое взаимодействие, в моем представлении, должно иметь квантово-механическую природу; и достигаться оно должно при помощи множества различных комбинаций атомных структур, которые одновременно «проверяются» в линейной суперпозиции (примерно так же, как это, вероятно, происходит в квантовом компьютере, упомянутом в конце главы 9).

Условием для выбора подходящего (хотя, возможно, не лучшего) решения задачи минимизации должно быть выполнение «одногравитонного критерия» (или приемлемой в данном конкретном случае альтернативы), что, предположительно, имеет место только При соответствующих физических условиях.

### **Возможная связь с пластичностью мозга**

Позвольте Мне продолжить эти рассуждения и спросить, могут ли они иметь непосредственное отношение к процессам, происходящим в мозге. Насколько я могу судить, наиболее правдоподобно будет выглядеть связь с пластичностью мозга. Давайте вспомним, что мозг, на самом деле, похож скорее не на обычный компьютер, а на компьютер, который постоянно изменяется. За эти изменения, по-видимому, отвечают процессы активации или деактивации синапсов, которые, в свою очередь, происходят вследствие роста или сокращения дендритных шипиков (см. главу 9, с. 319\*, рис. 9.15). Здесь я наберусь смелости и выскажу предположение о том, что этот рост или сокращение теоретически могут подчиняться принципам наподобие тех, которые управляют квазикристаллическим ростом. При этом «тестируется» не одно из возможных альтернативных расположений, а сложная линейная суперпозиция большого числа таких расположений. До тех пор, пока эффект каждой из этих альтернатив не превышает «одногравитонного уровня» (или некоторого его аналога), они будут существовать одновременно (более того: *должны* сосуществовать, коль скоро справедливы законы U-квантовой механики). Пока не превзойден одногравитонный уровень, могут начать одновременно выполняться суперпозиции разных вычислений, что вполне соответствует принципам действия квантового компьютера. Однако, вряд ли такие суперпозиции смогут существовать достаточно долго, поскольку нервные сигналы создают электрические поля, которые должны вносить значительные возмущения в окружающую среду (хотя их миелиновые оболочки являются своего рода изоляторами). Давайте допустим, что такие суперпозиции расчетов все-таки способны существовать в течение определенного минимального времени, которое необходимо для получения какого-нибудь *действительно* важного результата, т. е. что вплоть до этого момента «одногравитонный уровень» (или что-то подобное) в системе не достигается. Успешное завершение такого расчета будет в нашем случае той самой «целью», которая представляет собой аналог более простой «цели» минимизации энергии при квазикристаллическом росте. Таким образом, достижение этой цели будет подобно успешному росту квазикристалла!

В этих рассуждениях, конечно, много неясного и спорного, но я верю, что они описывают принципиально возможную аналогию. Рост кристалла или квазикристалла существенно зависит от концентрации нужных атомов и ионов в окрестности точек роста. Точно

также можно предположить, что процессы роста или сокращения семейств дендритных шипиков, в свою очередь, находятся в прямой зависимости от степени концентрации вокруг них различных нейромедиаторов (например, таких, чья концентрация зависит от испытываемых эмоций). Какие бы расположения атомов в конце концов ни были выделены в качестве реальной структуры получившегося квазикристалла — каждый раз этому должно предшествовать решение задачи минимизации энергии. Тогда я осмелюсь по аналогии предположить, что конкретная мысль, которая возникает на поверхности мозга, тоже возникает в результате решения некоторой задачи, только на сей раз не просто задачи минимизации энергии. Эта задача будет гораздо более сложной, требующей учета желаний и намерений, которые, в свою очередь, напрямую связаны с вычислительными свойствами и функциями мозга. Я полагаю, что сознательное мышление тесно связано с отсеком тех возможных альтернатив, которые прежде входили в линейную суперпозицию. Все это имеет непосредственное отношение к неизвестным (пока!) физическим процессам, которые должны управлять пограничной областью между  $U$  и  $R$ , и которые, я уверен, будут описаны правильной теорией квантовой гравитации — ПКТГ, которую еще предстоит открыть!

Могло бы такое физическое действие быть по своей природе неалгоритмическим? Вспомним, что в общем случае задача о плиточных покрытиях, описанная в главе 4, не имеет алгоритмического решения. Можно предположить, что сходная задача в приложении к атомным структурам имеет такое же свойство «неалгоритмичности». Если эти задачи могут в принципе быть «решены» средствами, о которых я говорю, то тогда есть вероятность, что у рассматриваемого мной типа умственной деятельности действительно существует неалгоритмическая компонента. Однако для того, чтобы это было так, нам необходима определенная неалгоритмичность и в ПКТГ. Конечно, мы сейчас слишком вольно обращаемся с гипотезами — но все же приведенные выше аргументы подсказывают мне, что здесь определенно должно быть *нечто*, имеющее неалгоритмический характер.

Как быстро происходят подобные изменения в мозговых связях? На этот счет у нейрофизиологов нет единого мнения, однако, коль скоро устойчивые отпечатки в памяти могут формироваться за доли секунды, разумно предположить, что указанные изменения происходят примерно за то же время. Чтобы мои собственные идеи получили право на существование, требуется как раз примерно такая быстрота.

## Временные задержки в реакции сознания

Теперь я хочу рассказать о двух экспериментах (описанных в работе Харта [1982]), которые проводились на добровольцах и которые, как мне кажется, имеют прямое отношение к нашим рассуждениям. В ходе этих экспериментов изучалось характерное время, которое требуется сознанию для того, чтобы осуществить определенное действие и отреагировать на внешнее воздействие. Первый из них, таким образом, изучал активную деятельность сознания, а второй — пассивную. Связанные воедино, результаты этих экспериментов представляются еще более поразительными.

Первый эксперимент был проведен Корнхубером и его коллегами в Германии, в 1976 году (Дике и др. [1976]). Ряд испытуемых согласился на запись электрических сигналов, снятых с определенной точки поверхности головы (т. е. на электроэнцефалограмму, или ЭЭГ). Испытуемые должны были внезапно сгибать *по своему усмотрению* указательный палец. Идея заключалась в том, что на электроэнцефалограмме окажутся зафиксированы некие признаки мозговой активности\* которая возникает в голове в момент принятия сознательного решения согнуть палец. Чтобы получить значимый сигнал с дорожек ЭЭГ, нужно было усреднить его по нескольким испытаниям, и результирующий сигнал при этом получался не очень

«показательным». Но вот что оказалось примечательным, так это отмеченное на ленте постепенное нарастание электрического потенциала в течение *целой секунды* — а то и всех полутора — происходящее *до* того момента, как палец действительно, сгибался. Это, по-видимому, означает, что сознательный процесс принятия решения занимает не менее секунды — и только затем следует его исполнение! Это очень сильно отличается по времени от куда более короткого промежутка времени реакции на внешнее раздражение, если способ реакции задан заранее (например, если испытуемый должен сгибать палец не «по собственному хотению», а исключительно в ответ на вспышку светового сигнала). В этом случае нормальной является задержка реакции длительностью около одной пятой секунды, что примерно в пять раз быстрее, чем «волевое действие», отраженное в экспериментальных данных Корнхубера (рис. 10.5).

Во втором эксперименте Бенджамин Либет из Калифорнийского университета, в сотрудничестве с Бертрамом Фейнштейном из Нейрологического института в Сан-Франциско (Либет и др. [1979]) проводили опыты с добровольцами, которым по независимым причинам предстояла хирургическая операция на мозге и которые соглашались на имплантацию электродов в определенные точки соматосенсорной коры своего головного мозга. Эксперимент Либета показал, что между стимуляцией кожи этих пациентов и ее осознанием проходило примерно полсекунды, несмотря на то, что сам мозг должен был получить соответствующий сигнал примерно через одну сотую долю секунды, а заранее запрограммированный «рефлективный» ответ на этот раздражитель (см. выше) мог быть выработан мозгом примерно через одну десятую секунды (рис. 10.6). Более того: несмотря на такое «запоздалое» осознание стимулирующего воздействия, субъективное впечатление, возникавшее у пациентов, позволяло им утверждать, что не было вообще никакой задержки! (Некоторые эксперименты Либета предусматривали стимулирование зрительного бугра, см. с. 307. Результат был тот же, что и при воздействии на сомато-сенсорную кору головного мозга.)

Напомним, что соматосенсорная кора — это область головного мозга, куда поступают сигналы от органов чувств. Так, электрическая стимуляция зоны соматосенсорной коры, соответствующей некоторому определенному участку поверхности кожи, воспринимается испытуемым в точности так, как если бы что-то действительно коснулось его кожи в этом месте. Однако оказалось, что если эта электрическая стимуляция длится менее, чем примерно полсекунды, то испытуемый не получает никакого ощущения вообще. В противоположность этому, прямая стимуляция точки на самой поверхности кожи может ощущаться даже в том случае, когда она мгновенна.

Теперь предположим, что сначала тронули кожу, а затем подвергли электрической стимуляции соответствующую точку в соматосенсорной коре. Что чувствует пациент?

Этот эффект получил название *обратной маскировки*. Каким-то образом стимулирование коры мозга препятствует осознанию обычного ощущения от касания кожи. Сознательное восприятие может быть предотвращено («замаскировано») более поздним событием, если это событие происходит не позднее, чем через полсекунды. Это говорит о том, что сознательное восприятие такого воздействия проявляется примерно через полсекунды после самого воздействия! Однако, по все видимости, испытуемые не осознавали столь долгой задержки своего восприятия. Найти разумное объяснение этому удивительному открытию можно было бы, предположив, например, что «время» всех наших «восприятий» действительно отличается примерно на полсекунды от «реального времени» — как будто наши внутренние часы просто «неверны» и отстают на полсекунды или около того. Время восприятия некоторого события в этом случае должно всегда *отстоять* на полсекунды от того момента, когда указанное событие произошло. Это дало бы возможность представить последовательную, хотя и неудовлетворительно замедленную картину нашего чувственного восприятия.

Возможно, что-то подобное происходило во второй части эксперимента Либета, где он

*вначале* стимулировал кору мозга, продолжая эту стимуляцию гораздо дольше, чем полсекунды — и касался кожи во время этой стимуляции, но ранее, чем через полсекунды после ее начала. И стимуляция коры, и касание кожи воспринимались раздельно, так что пациенту было ясно, какой именно стимул он воспринимал. Однако, когда его спрашивали, какой стимул он ощутил *первым*, он обычно отвечал, что это было касание кожи, тогда как, на самом деле, стимуляция коры всегда предшествовало ему! Таким образом, испытуемый и вправду сдвигал свое восприятие прикосновения к коже *назад во времени* примерно на полсекунды. Однако это представляется не просто глобальной «ошибкой» внутреннего ощущения времени, но более тонкой структурной перестройкой временной шкалы восприятия событий. Поскольку в случае стимуляции коры (учитывая, что она в действительности воспринимается не позднее, чем через полсекунды после ее начала), такая задержка *не* наблюдается.

Опираясь на данные первого из вышеописанных экспериментов, мы, по всей видимости, можем сделать вывод о том, что сознательному действию необходимо примерно около секунды или полутора секунд на то, чтобы быть приведенным в исполнение; а в соответствии со вторым экспериментом — что осознание внешнего события, по-видимому, не происходит раньше, чем через полсекунды после момента события. Представим себе, что происходит, когда человек реагирует на некоторое неожиданное внешнее событие. Предположим также, что ответ требует моментального сознательного действия. Если принять в расчет открытие Либета, то должно пройти полсекунды прежде, чем сознание «включится»; и после этого, как следует из опытов Корнхубера, потребуется еще секунда, а то и более, прежде чем человек «осознанно» отреагирует на это событие. Таким образом, весь процесс — от сенсорного восприятия до моторного отклика — занимает примерно две секунды! Очевидный вывод из этих двух экспериментов, если рассматривать их вместе, напрашивается сам собой: сознание *вообще* не может быть задействовано там, где ответная реакция на внешнее событие должна занимать не более пары секунд!

## Странная роль времени в сознательном восприятии

Можно ли доверять результатам этих экспериментов? Если это так, мы с необходимостью приходим к выводу, что мы действуем как «автоматы», когда чтобы изменить реакцию, требуется менее одной или двух секунд. Становится несомненным, что сознание, по сравнению с другими механизмами нервной системы, работает довольно медленно. Я и сам замечал, как иной раз моя рука продолжала захлопывать дверцу машины еще несколько мгновений спустя после того, как я заметил, что в машине осталось что-то нужное, и как сознательный приказ остановить движение руки патологически не успевает за ним — так медленно все это происходит. Но требует ли это и впрямь целую секунду, а то и две? Такой длительный промежуток времени кажется мне невероятным. Разумеется, мое *сознательное* восприятие забытой вещи в машине вместе с моим воображаемым «свободно-волевым» приказом остановить руку вполне *могли* случиться уже после обоих этих событий. Возможно, сознание — это просто наблюдатель, который воспринимает происходящее не иначе, как повторное исполнения всего спектакля. Аналогично, опираясь на результаты вышеописанных опытов, можно считать, что какая бы роль не отводилась сознанию, например, при отбивании теннисного мяча — не говоря уже об игре в пинг-понг — у него просто не хватило бы времени на ее исполнение! Несомненно, что у опытных игроков все наиболее важные приемы игры наиподробнее образом запрограммированы на уровне мозжечка. Но вот чтобы сознание не играло *никакой* роли в принятии решения относительно того, какой удар должен быть выполнен в конкретный момент — в это я могу поверить с большим трудом. Конечно, многое заложено в интуитивном угадывании следующего движения соперника, и наличии множества заранее просчитанных и подготовленных вариантов ответа на каждое из них — но такой сценарий,

при котором сознание *вообще* не участвует в формировании ответной реакции, кажется мне неэффективным и маловероятным. Эти возражения были бы еще более уместны в случае обычного разговора. Здесь также собеседники могут частично догадываться, что скажет другой, но в ответах оппонента должно достаточно часто присутствовать что-то неожиданное, иначе беседа просто потеряла бы смысл! И вряд ли кто будет спорить, что в обычном разговоре на то, чтобы ответить собеседнику, требуется куда меньше, чем две секунды времени!

Похоже, есть основания сомневаться, что в экспериментах Корнхубера сознанию «действительно» нужно полторы секунды для выполнения задуманного действия. Хотя *усредненная* по всем записям ЭЭГ задержка между возникновением намерения согнуть палец и непосредственным действием дает как раз такую величину, тем не менее может оказаться, что только в *некоторых* случаях намерение проявлялось столь рано — причем часто *не* приводило в действительности к стибанию пальца; тогда как во многих других случаях сознание приводило палец в движение гораздо быстрее. (На самом деле, более поздние эксперименты, см. Ли-бет [1987, 1989], позволяют сделать иные, чем у Корнхубера, выводы. Однако загадки, связанные с временными аспектами сознания, так и остались нерешенными.)

Давайте представим себе на минуту, что результаты обоих экспериментов *справедливы*. Тогда с необходимостью следует признать в том, что мы могли до сих пор идти по глубоко ошибочному пути, используя при изучении работы сознания обычные физические правила для *времени*! В самом деле, есть нечто весьма странное в том, как время входит в наше сознательное восприятие, и я думаю, что для интерпретации этого феномена в рамках наших традиционных представлений может понадобиться совсем иная концепция. Сознание — это, в конце концов, единственное известное нам явление, согласно которому время «течет»! Способ рассмотрения времени в современной физике не отличается по существу от способа рассмотрения *пространства*<sup>6</sup>); так что, на самом деле, «время» в физических процессах не «течет» — вместо этого рассматривается статичное «пространство-время», где фиксируются события, происходящие в нашей вселенной! Однако, мы воспринимаем время *текущим* (см. главу 7). Я полагаю, что и здесь присутствует некая иллюзия, и что, *на самом деле*, время нашего восприятия не течет линейно в одном направлении (что бы это ни значило!). «Кажущуюся» временную упорядоченность воспринимаемых событий мы, по-моему, привносим в наши ощущения сами для того, чтобы как-то согласовать их с единым для окружающего нас физического мира поступательным движением во времени.

Некоторые могли бы усмотреть в подобных замечаниях изрядную долю беспочвенного «философствования» — и их обвинения, конечно же, были бы справедливы. Как можно «ошибаться» относительно того, что ты действительно воспринимаешь? Ясно, что ощущения — это (по *определению*) то, что непосредственно осознается; поэтому они просто не могут быть «неправильными». Тем не менее, я думаю, что, на самом деле, мы все-таки *ошибаемся*, когда воспринимаем время, как движущееся вперед — несмотря на неадекватность доступных мне языковых средств для описание моего убеждения; и что существуют свидетельства, подтверждающие справедливость такой гипотезы (см. Черчланд [1984]).

Исключительно ярким примером (см. с. 342) является способность Моцарта «охватывать единым взглядом» всю музыкальную композицию, даже когда «она бывает довольно длинной». Исходя из описания самого Моцарта, можно предположить, что этот «взгляд» охватывал все существенные стороны произведения — и что, тем не менее, интервал времени (в обычном физическом смысле), необходимый для подобного сознательного восприятия композиции, оказывался заведомо короче того, который потребовался бы для ее исполнения. Кто-то может считать, что все воспринималось совсем по-другому, и Моцарт «видел» свое будущее произведение в форме пространственно-распределенных образов или, допустим, готовой музыкальной партитуры. Но и для внимательного прочтения партитуры таких размеров необходимо довольно много времени — и, к тому же, я сильно сомневаюсь в том, что исходное

восприятие Моцартом своей композиции могло принимать указанную форму (иначе он бы наверняка об этом сказал!). Образное восприятие кажется более вероятным; однако, (как и в большинстве случаев визуализации в математике, с которыми я лично сталкивался) я сильно сомневаюсь, что в сознании Моцарта мог совершаться прямой перевод музыки на язык зрительных образов. Мне кажется, что ин-герпретировать «взгляд» Моцарта правильнее всего с чисто *музыкальной* точки зрения, с четким временным распределением, которое обычно возникает при прослушивании (или исполнении) музыкального произведения. Ведь музыка состоит из звуков, воспроизведение которых требует определенного времени — *времени*, которое, со слов самого Моцарта, «...позволяет мое воображение».

Послушайте четырехчастную фугу из последнего раздела *Искусства фуги* И. С. Баха. Знатоки Баха не могут не переживать стресс, когда музыка останавливается после десяти минут звучания, сразу же после вступления третьей темы. Кажется, что композиция каким-то образом все еще существует «там, вовне» — просто сейчас она внезапно замерла. Бах покинул этот мир, не успев закончить свою работу и не оставив нам ни единого намека на то, как он намеревался продолжить ее. Однако, она начинается с такой уверенностью и бесспорным мастерством, что невозможно представить себе, чтобы у Баха в то время не было ясного представления о всех ключевых моментах своего будущего произведения. Нужно ли ему было мысленно исполнять композицию в обычном темпе, каждый раз «проигрывая» ее заново по мере возникновения новых идей и различных поправок? Я не могу себе представить, что это происходило таким образом. Как и Моцарт, он должен был представлять себе работу целиком, связывая в голове воедино как ее сложнейшую структуру, так и многочисленные замысловатые украшения — все то, без чего не мыслимо создание фуг. При том, что и временные характеристики музыки важны никак не меньше. Ибо как музыка может оставаться музыкой, если она не исполняется «в реальном времени»?

Рождение замысла романа или истории можно было бы рассматривать как аналогичный (хотя, на первый взгляд, и менее непостижимый) процесс. Охватывая внутренним взором всю жизнь персонажа, необходимо продумывать различные события, которые автор, как кажется, просто не сможет вставить в сюжет, не проиграв предварительно в «реальном времени». Однако это далеко не всегда необходимо. Даже сохранившиеся в памяти впечатления от лично пережитых событий оказываются настолько «сжатыми», что их можно мысленно «пережить» вновь за доли секунды!

Видимо, существует определенное (и при том значительное) сходство между сочинением музыки и математическим мышлением. Многие, вероятно, уверены, что математическое доказательство строится в виде цепочки последовательных утверждений, где каждый шаг вытекает из предыдущего. Но, на самом деле, замысел доказательства едва ли когда возникает подобным образом. Общее представление и лишь интуитивно понятное концептуальное содержание — вот что в действительности необходимо для построения математического доказательства; и это едва ли можно соотнести с тем временем, которое потребовалось бы в дальнейшем для его полного последовательного изложения.

Предположим далее, что мы допускаем отсутствие соответствия между внутренней шкалой времени нашего сознания — с одной стороны, и течением времени в окружающем нас физическом мире — с другой. Не рискуем ли мы при этом столкнуться с парадоксом? Предположим к тому же, что в природе наших сознательных действий заложено что-то неуловимо телеологическое, позволяющее будущим впечатлениям от действия в прошлом оказывать влияние на само это действие. Ясно, что *это* могло бы привести нас к противоречию, подобному парадоксальным следствиям из предположения о возможности распространения сигнала со скоростью, превышающей скорость света, которое мы рассматривали — и совершенно обоснованно отвергли — в конце главы 5 (см. с. 175). Я считаю, что никакого парадокса здесь быть *не* должно — как это непосредственно следует из моих утверждений, касающихся самого

понятия сознания и его возможностей. Если вы помните, я выдвигал предположение о том, что сознание, в сущности, есть способность «видеть» непреложную истину; и что оно может представлять собой своеобразный контакт с миром идеальных математических идей Платона; Напомню, что мир Платона сам по себе имеет вневременную природу. Восприятие истины Платона не несет подлинной информации — имея в виду технический аспект понятия «информации», связанный с возможностью ее передачи; так что, на самом деле, не будет никакого противоречия даже в том случае, если бы подобное сознательное восприятие распространялось обратно во времени!

Но даже если мы согласимся с тем, что сознание связано со временем таким причудливым образом — и что благодаря сознанию происходит своего рода контакт между нашим физическим миром и определенной вневременной сущностью — как тогда быть с физически обусловленным и упорядоченным во времени действием материального мозга? И снова мы, по-видимому, вынуждены отводить сознанию роль простого «зрителя» — в противном случае нам придется так или иначе подтасовывать физические законы, чтобы не нарушить естественное развитие событий. Однако я все же *отстаиваю* активную роль сознания, которая дает ему преимущество в ходе естественного отбора. Ответ на эту дилемму, как мне кажется, может заключаться в том странном способе, как должна действовать ПКТГ, разрешая конфликт между двумя квантово-механическими процедурами U и R (см. с. 286, 297).

Вспомним о проблемах со временем, которые возникали в результате наших попыток согласовать R-процедуру со (специальной) теорией относительности (главы 6, 8, с. 232, 301). При описании этой процедуры в обычных пространственно-временных терминах она, кажется, вообще теряет всякий смысл. Рассмотрим квантовое состояние пары частиц. Как правило, такое состояние должно быть *коррелированным* (т. е. описываться *не* простым выражением  $|\Psi\rangle |\chi\rangle$ , где каждый из сомножителей  $|\Psi\rangle$  и  $|\chi\rangle$  описывает только одну частицу, но представляет собой сумму вида  $|\Psi\rangle |\chi\rangle + |a\rangle |\beta\rangle + \dots + |g\rangle |\sigma\rangle$ ). Тогда наблюдение за одной из частиц окажет на другую нелокальное воздействие, которое не может быть описано в обычных пространственно-временных терминах, согласующихся со специальной теорией относительности (ЭПР; эффект Эйнштейна—Подольского—Розена). Подобные нелокальные эффекты должны неявным образом присутствовать в предложенной мною «квазикристаллической» аналогии для роста и сокращения дендритов.

Под «наблюдением» я здесь понимаю усиление действия каждой наблюдаемой частицы до тех пор, пока не достигается некий уровень, соответствующий, например, «одногравитонному критерию» в рамках ПКТГ. В более «стандартной» терминологии «наблюдение» — это крайне нечетко определенное понятие; и, согласитесь, трудно себе представить, как можно начинать теоретические исследования в области квантово-механического описания работы мозга, если приходится считаться с необходимостью рассматривать мозг в качестве объекта, который постоянно «наблюдает сам себя»!

Моя собственная идея заключается в том, что концепция ПКТГ, напротив, должна дать нам *объективную* физическую теорию редукции вектора состояния (R-процедуры), которая никак *не* будет зависеть от нашего взгляда на сознание. У нас пока нет такой теории, но, по крайней мере, можно быть уверенным, что ее создатели не будут спотыкаться о фундаментальные вопросы, связанные с точным определением сознания!

Я полагаю, что именно *после* открытия теории ПКТГ у нас в конце концов появится возможность описания с ее помощью феномена сознания. Вообще говоря, я склонен считать, что априори предполагаемые свойства ПКТГ окажутся на самом деле еще менее удобными для их адекватного описания в обычных пространственно-временных терминах, чем упомянутые выше загадочные явления ЭПР в системе двух частиц. Если я прав, и сознание напрямую связано с будущей теорией ПКТГ, то лишь с большой натяжкой нам удастся применить к нему наши привычные пространственно-временные описания.

## Заключение: точка зрения ребенка

В этой книге я привел множество доводов, призванных показать несостоятельность точки зрения — как выясняется, одной из наиболее распространенных в современной философии — согласно которой наше мышление & основе своей идентично действию очень сложного компьютера. Когда в явном виде делается предположение о том, что простое выполнение алгоритма может привести к возникновению *осознанного восприятия*, то используется терминология концепции «сильного ИИ» Серла. В другие термины (такие, как «функционализм») подчас вкладывают более широкий смысл.

Некоторые читатели, возможно, с самого начала считали «сторонников сильного ИИ» законченной деревенщиной! Разве не «очевидно», что обычные вычисления не могут вызвать удовольствие или причинить боль; что они не способны понимать поэзию, наслаждаться красотой вечернего неба или магией звуков; что они не могут надеяться, любить или отчаиваться; что у них не может возникнуть настоящей независимой цели существования? Однако наука, кажется, заставляет нас поверить в то, что все мы — просто ничтожные частички мира, полностью управляемого (пусть даже только вероятностно) очень точными математическими законами. Равно как и наш мозг, который, казалось бы, контролирует все наши действия, но, в свою очередь, точно также подчиняются тем же самым законам. Возникает ощущение, что вся эта скоординированная физическая активность является, на самом деле, ничем иным, как выполнением некоторого всеобъемлющего (возможно, по своей природе вероятностного) вычислительного процесса — и, следовательно, наш мозг и наш разум нужно рассматривать исключительно в терминах такого вычисления. Может быть, когда степень сложности подобного алгоритма становится чрезвычайно высокой, он приобретает те поэтические или субъективные качества, которые мы привыкли ассоциировать с понятием «разум». Однако трудно избавиться от навязчивого ощущения, что в такой картине всегда будет чего-то не хватать.

В своих рассуждениях я пытался найти обоснование своей уверенности в том, что и вправду должно быть нечто важное и существенное, остающееся за рамками любой «алгоритмической» картины мира. Тем не менее, я по-прежнему связываю свои надежды на разгадку тайны разума с наукой в целом, и математикой в частности. Здесь возникает очевидная дилемма — однако, я старался показать, что из этой ситуации *есть* совершенно естественный выход. Свойство вычислимости — не то же самое, что математическая точность. Сколько тайны и красоты в точном математическом мире Платона — а ведь большая непознанная часть этого мира связана с понятиями, которые находятся за пределами той сравнительно небольшой его части, где располагаются алгоритмы и вычисления.

Сознание представляется мне таким важным явлением, что я просто не могу поверить в возможность его «случайного» возникновения в результате сколь угодно сложных вычислений. Ведь именно благодаря ему мы можем говорить о самом существовании вселенной. Некоторые считают, что вселенная, законы которой не допускают зарождение сознания, вообще не является вселенной. Я бы даже сказал, что все математические описания вселенной, которые до сих пор были сделаны, не должны удовлетворять этому критерию. Только сознание могло вызвать предполагаемую «теоретическую» вселенную к жизни!

Некоторые доводы, приведенные мной в этих главах, могут показаться чересчур сложными для понимания, другие представляются слишком спорными — хотя, по-моему, немало здесь и таких, которые, наоборот, никак нельзя оставить без внимания. Тем не менее, за всеми этими техническими рассуждениями стоит одно — ощущение «очевидности» предположения о том, что разум, наделенный *сознанием*, просто не может работать подобно компьютеру, несмотря на алгоритмическую природу многих составляющих нашей умственной деятельности.

Это тот тип очевидности, который доступен и ребенку — хотя со временем этот ребенок, став уже взрослым, будет вынужден поверить в то, что очевидные проблемы — это «не проблемы»,

что они могут быть сведены на нет при помощи тщательно подобранных рассуждений и удачных определений. Дети иногда ясно видят многие вещи, которые в более зрелые годы теряют для них свою очевидность. Мы часто забываем чувство восхищения, которое мы испытывали в детстве, когда впоследствии на наши плечи ложится груз повседневных забот «мира взрослых». Дети не боятся задавать самые элементарные вопросы из числа тех, которые нам, взрослым, задавать уже «стыдно». Что происходит с каждым из потоков нашего сознания после смерти; где было сознание до нашего рождения; могли бы мы стать, или уже были, кем-то еще; почему мы вообще воспринимаем мир; почему мы здесь; и почему, в конце концов, есть такая вселенная, в которой мы можем существовать? Это загадки, которые /имеют обыкновение возникать в момент пробуждения способности осознать в каждом из нас — и, несомненно, с первыми проблесками подлинного самосознания в любом живом существе.

Я помню свои собственные попытки разрешить для себя многие из этих загадок, когда я был ребенком. Допустим, думал я, мое сознание имеет возможность внезапно поменяться с чьим-то другим — как в таком случае я могу быть уверен, что нечто подобное уже не произошло со мной раньше, предполагая, что каждый человек хранит в памяти только то, что относится к нему лично? Как я мог бы тогда объяснить такой опыт «обмена» кому-то еще? Или все это вообще не имеет никакого смысла? А что если я просто проживаю те же самые десять минут жизни снова и снова — и каждый раз с одними и теми же впечатлениями? Может быть, для меня «существует» только настоящее? Может быть, «я» завтрашнего или вчерашнего дня — это в действительности совершенно иная личность с независимым сознанием? Может быть, я на самом деле живу «задом наперед» во времени, и мой Поток сознания направлен в прошлое, так что моя память говорит не о том, что уже *произошло* со мной, но о том, что еще только *должно произойти* — и неприятности в школе еще впереди и, к сожалению, уже не за горами? Есть ли вообще какое-нибудь значимое различие между *таким* и обычным течением времени, которое позволило бы считать одно из них «правильным», а второе — нет? Для того, чтобы иметь принципиальную возможность получать ответы на подобные вопросы, необходима теория сознания. Но как можно даже *начинать* объяснять сущность таких проблем тому, кто сам не обладает сознанием?..

## Примечания

1. Как мы видели в главе 4 (с. 10S), проверка справедливости доказательства в формальной системе всегда имеет алгоритмический характер. И наоборот, любой алгоритм, который позволяет получать математически истинные утверждения, всегда можно добавить в систему аксиом и правил вывода обычной логики («предикатного исчисления»), тем самым создавая новую формальную систему выведения математических истин.

2. Некоторых читателей может беспокоить тот факт, что в среде математиков действительно но существуют различные точки зрения.

Вспомним рассуждения, приведенные в главе 4. Однако имеющиеся разногласия не так важны для нас. Они относятся только к в высшей степени абстрактным вопросам, касающимся очень больших множеств, в то время как мы вполне можем ограничиться утверждениями арифметического характера (с конечным числом кванторов существования и всеобщности) и применить дальнейшие рассуждения. (Возможно, здесь допущено некоторое преувеличение, поскольку принцип рефлексии, относящийся к бесконечным множествам, может иногда использоваться для вывода утверждений в арифметике.) Что касается крайне догматичного и не желающего соглашаться с Геделем формалиста, для которого такая вещь, как математическая истина, вообще *не существует*, то я его буду просто-напросто игнорировать, поскольку он явно не обладает способностью интуитивного понимания истины, которой посвящены наши рассуждения!

Конечно, математики иногда допускают ошибки. Кажется, сам Тьюринг считал, что именно *это* и есть «лазейка», которая позволяет обойти аргументы геделевского типа в пользу того, что человеческое мышление существенно неалгоритмично. Но лично мне кажется невероятным, что свойство людей ошибаться каким-либо образом связано с нашей способностью к прозрениям! (Между прочим, *генераторы* случайных чисел могут быть успешно реализованы при помощи алгоритмов.)

3. Термин «черная дыра» вошел во всеобщее употребление много позже, около 1968 года (главным образом благодаря пророческим идеям американского физика Джона А. Уилера).

4. Мне кажется, что потребность животных во сне, во время которого они иногда *видят сны* (как это бывает часто заметно у собак), может служить свидетельством того, что они, вполне вероятно, наделены сознанием. Ибо разница между сном без сновидений и сном со сновидениями, по-видимому, во многом определяется как раз наличием сознания.

5. В случае специальной или общей теории относительности под «временами» следует понимать «одновременные пространства» или «пространственно-подобные поверхности» (с. 167, 177).

6. Однако в случае пространственно-бесконечной вселенной есть затруднения, поскольку тогда возникает (как и в случае множественных миров) бесконечное количество копий наблюдателя и его непосредственного окружения! Будущее поведение каждой копии может несколько отличаться, и никто не в состоянии сказать на верняка, какой из приблизительных копий самого себя, смоделированных математическим путем, он мог бы на самом деле «быть»!

7. Даже в ходе реального роста некоторых кристаллов могут возникать подобные проблемы — например, там, где исходная клетка кристаллической решетки содержит не сколько сот атомов (случай так называемых «фаз Фрэнка—Каспера»). С другой стороны, следует упомянуть, что теоретический «почти локальный» (хотя все же не локальный) процесс роста квазикристаллов с осью пятого порядка был предложен Онодой, Стайнхардтом, Ди Винченцо и Соколаром [1988].