

Сандра Блейкли, Джефф Хокинс Об интеллекте



іта

«Об интеллекте»: Издательский дом «Вильямс»; Москва-Санкт-Петербург-Киев; 2007
ISBN ISBN 978-5-8459-1139-1 (рус.), ISBN 0-8050-7456-2 (англ.)

Аннотация

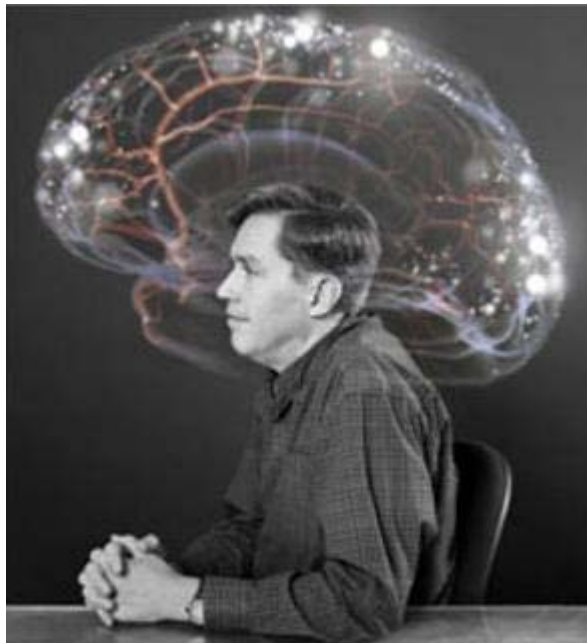
В книге *Об интеллекте* Джефф Хокинс представляет революционную теорию на стыке нейробиологии, психологии и кибернетики, описывающую систему «память-предсказание» как основу человеческого интеллекта. Автор отмечает, что все предшествующие попытки создания разумных машин провалились из-за фундаментальной ошибки разработчиков, стремившихся воссоздать человеческое поведение, но не учитывавших природу биологического разума. Джефф Хокинс предполагает, что идеи, сформулированные им в книге *Об интеллекте*, лягут в основу создания истинного искусственного интеллекта – не копирующего, а превосходящего человеческий разум. Кроме этого, книга содержит рассуждения о последствиях и возможностях создания разумных машин, взгляды автора на природу и отличительные особенности человеческого интеллекта.

Книга рекомендуется всем, кого интересует устройство человеческого мозга и принципы его функционирования, а также тем, кто занимается проблемами разработки искусственного интеллекта.

**Джефф Хокинс
Сандра Блейкли**

Об интеллекте

Об авторах



Джефф Хокинс – один из самых известных предпринимателей и разработчиков компьютеров в Силиконовой долине. Основав в свое время компании *Palm Computing* и *Handspring*, Джефф сейчас работает ведущим инженером в компании *PalmOne*. Кроме того, с целью исследования проблем интеллекта и памяти он создал Редвудский институт нейрологии. Д. Хокинс – действительный член Национальной инженерной академии и ученого совета Cold Spring Harbor Laboratory. Проживает в Северной Калифорнии.

Сандра Блейкли вот уже более тридцати лет занимается подготовкой научных и медицинских публикаций для газеты *New York Times*. Она также является одним из авторов книги *Phantoms in the brain* и вместе с Джудит Уоллерстайн пишет популярные книги о психологии и браке. Живет в Санта-Фе, Нью Мехико.

Благодарность автора

Когда меня спрашивают: «Чем вы зарабатываете на жизнь?», я всегда затрудняюсь ответить. Если честно, то практически ничем. Но мне повезло, потому что я сформировал очень работоспособную команду. Разумеется, мне приходится руководить ими и порой задавать направление их деятельности. Но своим успехом я прежде всего обязан упорному труду и интеллекту моих коллег.

В жизни мне представилась счастливая возможность общаться со многими исследователями, которые меня чему-то научили, все они внесли свою лепту в создание этой книги. Всем им выражаю искреннюю благодарность. Брюно Ольшозен, который работает в Редвудском институте нейрологии (РИН) и Калифорнийском университете, – это настоящая ходячая энциклопедия по нейробиологии. Он постоянно пополняет и углубляет мои познания в данной области. Билл Софски (РИН) был первым, кто обратил мое внимание на уменьшение времени прохождения сигналов в иерархической структуре коры головного мозга и на особенности тонких дендритов. Рик Грейнджер из Калифорнийского университета поднял тему запоминания последовательностей и того, какова роль таламуса в этом процессе. Боб Найт (Калифорнийский университет, г. Беркли) и Кристоф Кох (Калифорнийский технологический институт) сыграли большую роль в становлении Редвудского института нейрологии, а также участвовали в обсуждении многих других вопросов научного характера. Всем сотрудникам РИН я благодарен за конструктивную критику в адрес моих высказываний; многие мои идеи родились в ходе наших заседаний и встреч в РИН. Спасибо всем вам.

Донна Дубински и Эд Коллиган на протяжении более чем десяти лет были моими партнерами по бизнесу. Благодаря их самоотверженной работе и помощи я смог заниматься предпринимательством, параллельно изучая головной мозг, что является довольно-таки сложной задачей. Донна всегда отмечала, что ее роль состоит в поддержке нашего бизнеса, а

моя – в разработке теории мозга. Без Донны и Эда я никогда не достиг бы успеха.

Среди людей, также причастных к созданию книги *Об интеллекте*, я хочу отметить многих моих коллег. Джим Левин, мой представитель, поверил в идею этого труда еще до того, как я начал работу над ним. (Тем, у кого нет такого представителя, как Джим, я не советую писать книги.) Именно Джим познакомил меня с Сандрой Блейкли, моим соавтором. Я хотел, чтобы моя книга была доступна широкому кругу читателей, и Сандра мне чрезвычайно в этом помогла. (В том, что некоторые главы оказались сложными для чтения, полностью моя вина.) Мэтью, сын Сандры, перу которого также принадлежат несколько научных работ, предложил ряд примеров, прекрасно иллюстрирующих положения, представленные в этой книге, а также был инициатором введения термина *модель «Память-предсказание»*.

Я с удовольствием хотел бы упомянуть всех, с кем сталкивался в издательстве *Henry Holt*, а особенно президента Джона Стэрлинга. Я встречался с ним лично лишь однажды, и несколько раз мы общались по телефону. Тем не менее Джон подсказал мне наиболее удачную структуру книги. Он заблаговременно и чрезвычайно точно определил проблемы, с которыми я неизбежно столкнусь при изложении теории интеллекта, помог мне откорректировать стиль написания книги, а также порядок подачи материала.

Выражаю благодарность своим дочерям Энн и Кейт, которые не жаловались, что их папа много выходных дней проводил не с ними, а перед монитором компьютера. И наконец, я признателен своей жене Джанет. Быть моей женой ох как непросто. Но уверяю, Джанет, я люблю тебя больше, чем теорию мозга!

Дж. Хокинс

Пролог

У меня есть два пристрастия, которые послужили источником вдохновения для написания этой книги. Из них же я черпаю смысл всей своей жизни.

Вот уже четверть века я занимаюсь вычислительной техникой. В индустрии высоких технологий я известен как основатель двух компаний – *Palm Computing* и *Handspring*. Кроме того, я участвовал в разработке архитектурных решений для карманных компьютеров PalmPilot и мобильных телефонов Trio.

Другое мое пристрастие еще более значимо, поскольку именно оно предопределило мое увлечение компьютерами. Я просто-таки помешан на изучении человеческого мозга и стремлюсь разобраться в том, как он функционирует. Причем я подхожу к проблеме с позиций не философа, а инженера, изучающего сложную конструкцию, состоящую из множества деталей. Разобравшись в том, что же такое разум и как работает человеческий мозг, я хочу создавать механизмы, действия которых основывались бы на тех же принципах, что и работа мозга. Я верю в то, что создание настоящего искусственного интеллекта не за горами.

Исследователи, как правило, либо стремятся проникнуть вглубь вещей и явлений, либо смотрят очень широко, стараясь постичь события, происходившие миллиарды лет тому назад. Но головной мозг есть у каждого человека, включая вас, мой читатель. Чтобы понять, как рождаются те или иные эмоции и чувства, почему люди воспринимают мир определенным образом, по каким причинам допускают ошибки, что побуждает их к творчеству, почему музыка или искусство бывают источником вдохновения, одним словом, чтобы понять, что означает быть человеком, нужно разобраться в том, как работает мозг. Новые открытия, касающиеся человеческого разума и работы человеческого мозга, не только сулят значительные социальные выгоды (в частности, они позволят добиться прогресса в лечении заболеваний головного мозга), но и станут основой для создания по-настоящему разумных механизмов. Эти механизмы, совсем не похожие на роботов из фантастических романов, будут сконструированы с учетом положений, раскрывающих природу самого эффективного из существующих на нашей планете биологического механизма – человеческого разума. Искусственный интеллект расширит наши знания о мире и о Вселенной и позволит нам добиться большей безопасности существования. И это будет только начало.

К счастью, мы с вами живем в эпоху, когда проблема понимания мозга близка к своему разрешению. У нынешнего поколения есть доступ к накопленному за сотни лет огромному

объему информации о строении и функционировании мозга. И, хотя наука о мозге – нейрология – и сейчас не стоит на месте, целостная теория о разуме и строении мозга пока не разработана. Большинство нейробиологов, увлеченных изучением «подсистем» мозга, не утруждают себя интеграцией приобретаемых знаний. С другой стороны, ни одна из многочисленных попыток создать мыслящий компьютер не увенчалась успехом. И я уверен, что дело не сдвинется с мертвой точки до тех пор, пока исследователи и изобретатели будут игнорировать различия между мозгом и компьютером, их постигнет та же участь, что и их предшественников.

Какая составляющая разума отсутствует в компьютере? Почему двухлетний ребенок способен научиться кататься на трехколесном велосипеде, он не только уверенно ходит, но и бегаёт, подпрыгивает, а современные роботы передвигаются подобно громыхающим зомби? Каким образом к трем годам дитя успеваёт овладеть азами речи, в то время как компьютер, несмотря на столетия стараний лучших программистов, всё ещё на это не способен? Почему человек с первого взгляда – в долю секунды – отличает кошку от собаки, а компьютер, даже самый совершенный, не может этого сделать вообще? На указанные вопросы по сей день не найдено ответов. У нас есть множество догадок, но сейчас мы более всего нуждаемся в том, чтобы выработать критический подход к возможным объяснениям.

У вас, скорее всего, возник вопрос «А почему вдруг разработчик архитектуры компьютеров решил написать книгу о мозге?» Или (как альтернативный вариант): «Если уж он так увлекается устройством мозга, то почему не окунулся с головой в нейробиологию или не нашёл себя в области создания искусственного интеллекта?» На самом деле я пытался. Много раз пытался. Однако мне не хотелось слепо двигаться в русле общепринятого подхода. По моему мнению, нужно рассматривать биологию мозга как ограничитель и своего рода путеводитель для понимания разума, а разум анализировать как вычислительную задачу. Таким образом, решение поставленных задач может быть найдено на стыке биологии и кибернетики. Многие биологи категорически отказываются говорить о мозге в терминах информатики. В свою очередь, специалисты, работающие в сфере информационных технологий, пренебрегают возможностью почерпнуть сведения из нейробиологии. Мир науки более настороженно относится к риску, чем мир бизнеса. В бизнесе высоких технологий человек, предлагающий новую, хорошо обоснованную идею, сможет сделать неплохую карьеру независимо от того, будет ли его разработка успешной или нет. Многие ныне процветающие предприниматели на заре своей деятельности терпели убытки. В академической жизни всё по-другому: если учёный несколько лет занимался исследованием новой теории и его поиски не увенчались успехом, это может стать причиной краха его карьеры. Именно поэтому я решил всерьёз заниматься обоими своими пристрастиями, полагая, что в таком случае процесс окажется взаимодополняемым. Я надеялся, что достижения в сфере информационных технологий позволят мне углубить понимание работы мозга. Реализация данной идеи требовала финансирования, а, кроме того, мне предстояло стать первопроходцем и выяснить, как продавать достигнутые наработки, – всё это я надеялся почерпнуть из работы в Силиконовой Долине.¹

В августе 2002 года я открыл Редвудский институт нейрологии (РИН) – научно-исследовательский центр по изучению мозга. В мире существует много подобных организаций, но ни в одной из них первостепенное значение не уделяется изучению *неокортекса*² – части головного мозга человека, ответственной за интеллект. В РИН же мы именно этим и занимаемся, пытаемся воплотить в жизнь идеи, которые принято считать несбыточными мечтами. Уникальность РИН в том, что это единственное в мире научное заведение, работающее только над изучением новой коры головного мозга – неокортекса. За недолгую историю института у нас побывали почти все ведущие исследователи в этой области.

¹ В Силиконовой долине, которая протянулась на 40 км от г. Сан-Хосе на юго-западе до г. Пало-Алто на северо-западе США, расположены ведущие американские высокотехнологичные фирмы. – Примеч. ред.

² Неокортекс – новые области коры головного мозга, которые у низших млекопитающих только намечены, а у человека составляют основную часть коры. – Примеч. ред.

Я благодарен своей замечательной команде и доволен, что наш труд уже приносит первые плоды.

В основу этой книги была положена следующая программа. Прежде всего, будет изложена общая теория функционирования мозга. Речь пойдет о том, что такое разум и как он формируется. Представляемая мною теория по существу не нова. Многие из изложенных ниже идей были сформулированы до меня, но никогда не приводились к единому знаменателю, который бы объединил их в единое целое. Это не удивительно, ведь недаром говорят, что все новое – это хорошо забытое старое, т. е. «новые идеи» на самом деле оказываются старыми, представленными в новой формулировке и интерпретации. В определенной степени это справедливо и в отношении концепции, представленной в данной книге. Но формулировка и интерпретация – это именно то, что коренным образом отличает хаотическое скопление частных фактов от тщательно обоснованного теоретического подхода. Смею надеяться, что читатели оценят представленную в этой книге теорию по достоинству. Типичной реакцией на нее всегда было: «Звучит вполне правдоподобно. Я никогда не думал о разуме в таком ключе, но сейчас, когда вы все разложили по полочкам, общая картина прояснилась». Благодаря информации, изложенной в этой книге, у вас наверняка повысится осведомленность относительно своего поведения. Вы сможете детально анализировать свои реакции и приходиться к выводу: «Я понимаю, что сейчас произошло у меня в голове». И возможно, кто-то из читателей, вдохновленный моим рассказом, решит посвятить свою жизнь созданию разумных машин, а в основу работы положит принципы, описанные на страницах этой книги.

Излагая свою концепцию, я часто буду подчеркивать, что компьютер никогда не был и не сможет быть адекватной моделью человеческого мозга. Ошибка многих разработчиков искусственного интеллекта состоит в том, что они пытаются запрограммировать компьютеры таким образом, чтобы они функционировали подобно человеческому разуму. Они хотят достичь поставленной цели, обойдя вниманием вопрос о сути разума, о том, что означает слово «понимать». Но ведь этим самым они «выплеснули с водой ребенка» – создавая мыслящие механизмы, забыли о разуме! Но все попытки создания искусственного интеллекта без учета особенностей естественного обречены на провал. (Мы подробно обсудим, почему предшествующие попытки исследования человеческого разума и создания мыслящих компьютеров потерпели фиаско.)

Ключевую идею предложенной теории я назвал прогностической способностью, основанной на запоминании, или моделью «Память-предсказание».

В главе 6 представлены особенности того, как головной мозг дополняет модель прогнозирования памяти, иными словами, как он функционирует. Далее мы коснемся социальной пользы данной концепции. Вполне возможно, что для многих читателей именно эта часть книги станет поводом для наиболее ценных размышлений. В завершение мы обсудим перспективы создания искусственного интеллекта.

Ниже приведен краткий перечень вопросов, ответы на которые вы получите по ходу чтения. Я надеюсь, вы оцените мой труд по достоинству.

Возможно ли создание мыслящего компьютера?

На протяжении десятилетий ученые, занимающиеся разработкой искусственного интеллекта, утверждали, что компьютеры могут стать разумными, если будут достаточно мощными для этого. Я не согласен с таким мнением, и в этой книге я попытаюсь объяснить почему. Дело в том, что между функционированием компьютера и человеческого мозга нельзя поставить знак равенства.

Разве применения нейронных сетей не достаточно для создания разумных механизмов?

Мозг на самом деле состоит из сетей нейронов, но без понимания сути его функционирования применение нейронных сетей принесет не больше пользы в создании разумных механизмов, чем компьютерные программы.

В чем состоит сложность понимания работы мозга?

Большинство ученых утверждают, что мозг очень сложен по своей структуре, и, чтобы постичь ее, требуется очень много времени. Я не могу согласиться с таким утверждением. Сложность – это симптом замешательства, а не его причина. По-моему, у каждого из нас есть

определенные интуитивные (часто неправильные) предположения, вводящие нас в заблуждение. Одна из наиболее вопиющих ошибок подобного рода – вера в то, что разумное поведение предопределяет разум.

Так что же такое разум, если на самом деле он не предопределяется поведением?

Создавая модель мира, мозг задействует большие ресурсы памяти. Эта модель включает весь опыт человека и объем его знаний. Каждый раз, формулируя прогнозы о будущих событиях, мозг обращается к этой основанной на памяти модели. Способность прогнозировать будущее составляет важнейшую функцию мозга. Прогностическая функция мозга детальнейшим образом будет рассмотрена в этой книге, именно она находится в центре внимания моей работы.

Как работает мозг?

Местом «дислокации» разума является неокортекс. Несмотря на богатство функциональных возможностей и невероятную гибкость, структурное строение коры головного мозга на удивление однородно. Все части коры головного мозга – независимо от того, отвечают ли они за слух, зрение, обоняние, тактильную чувствительность или речевую функцию, – работают по одним и тем же принципам, которые можно условно организовать в иерархическую структуру. Мы подробно проанализируем функционирование неокортекса как части коры головного мозга, ответственной за отношения человека с миром. Это будет наиболее «техническая» часть книги, но я думаю, что мои разъяснения будут понятны и широкому кругу читателей, а не только тем, у кого есть соответствующая научная подготовка.

Каково практическое приложение теории, представленной в данной книге?

Основываясь на теории разума, можно найти ответы на множество вопросов: как рождается сознание, в чем состоит процесс обучения, что побуждает человека к творчеству, почему мы склонны к предрассудкам, почему «старые клячи» с большим трудом усваивают новое и т. д. Теория разума позволяет нам лучше понять самих себя, свои реакции, побуждения и поведение.

Возможно ли создание мыслящих механизмов, и как они будут работать?

Да. Мы можем создать и создадим их, возможно, уже в ближайшие десятилетия. Некоторые эксперты высказывают предположение, что механизмы, обладающие разумом, могут причинить людям вред. Я хотел бы полностью опровергнуть данное утверждение. Роботы ни-когда не станут разумнее человека. Гораздо легче разработать компьютер, превосходящий человека в области физики или математики, чем создать передвигающийся и ведущий осмысленные беседы механизм. (Мы часто видим подобное в фантастических фильмах.) В этой книге я поделюсь с вами своими размышлениями о том, в каких невообразимых направлениях предстоит развиваться кибернетике.

Цель данной книги – представить читателям теорию разума и функционирования мозга в доступной и понятной форме. Вначале я изложу основополагающие принципы, которые будут подкрепляться дополнительными аргументами по мере усвоения пройденного материала. Обоснование части из них опирается лишь на здравый смысл и логику, для понимания других нам придется рассмотреть определенные закономерности строения мозга. Я допускаю, что некоторые из моих суждений впоследствии окажутся ошибочными, как это случается практически в любой отрасли науки. Полной и хорошо обоснованной теории требуются годы для того, чтобы достигнуть зрелости, но это не значит, что идея, положенная в основу такой теории, не представляет сама по себе никакой ценности. И, разумеется, я попытаюсь обойтись без профессионального жаргона и наукообразности.

Много лет назад, впервые заинтересовавшись проблемами мозга, я решил взять в библиотеке хорошую книгу, доступно объясняющую, как работает мозг. (Добавлю, что находить ответы на свои вопросы в книгах я научился еще подростком. Тогда меня интересовала теория относительности, черные дыры, магия, математика и многое другое.) Но поиск хорошей книги о человеческом мозге не увенчался успехом. К моему удивлению, не существовало даже самой хлипкой, плохо обоснованной или приблизительной теории мозга. Я был поражен своим открытием! Вот, например, в то время никто не знал, как и почему вымерли динозавры, но, тем не менее, существовало очень много предположений на этот счет, и о каждом из них были написаны ученые труды. Но ничего о мозге. Совершенно ничего. Не в

силах поверить в это, я просто не находил себе места, пока не пришел к выводу, что объяснение, скорее всего, должно быть достаточно простым. Мозг не является чем-то магическим, а значит, ответы на вопросы о нем тоже не будут сверхъестественными. Математик Поль Эрде считал, что простейшие математические доказательства существуют в некоей этерической «книге», и задача математика сводится к тому, чтобы найти и прочесть эту «книгу». Подобно ему, я чувствовал, что объяснение сущности разума где-то здесь, рядом.

Последние двадцать пять лет у меня не выходила из головы та самая маленькая и простая книга, которую мне следовало бы прочесть. Она была для меня как медовый пряник, который получают в награду за хороший поступок. Мысль о ней не оставляла меня все эти годы, и она стала предопределяющей при написании книги, которую вы сейчас держите в руках. Мне никогда не нравилась замысловатость – неважно, в науке или в технологии. Те приборы и устройства, архитектуру которых я разработал, отличаются легкостью и простотой в использовании. Наиболее яркие идеи зачастую просты. Моя книга предлагает простую и доступную теорию разума. Надеюсь, вы оцените ее по достоинству.

Ждем ваших отзывов!

Вы, читатель этой книги, и есть главный ее критик и комментатор. Мы ценим ваше мнение и хотим знать, что было сделано нами правильно, что можно было сделать лучше и что еще вы хотели бы увидеть из-данным нами. Нам интересно услышать и любые другие замечания, которые вам хотелось бы высказать в наш адрес.

Мы ждем ваших комментариев и надеемся на них. Вы можете прислать нам бумажное или электронное письмо либо просто посетить наш Web-сервер и оставить свои замечания там. Одним словом, любым удобным для вас способом дайте нам знать, нравится или нет вам эта книга, а также выскажите свое мнение о том, как сделать наши книги более интересными для вас.

Посылая письмо или сообщение, не забудьте указать название книги и ее авторов, а также ваш обратный адрес. Мы внимательно ознакомимся с вашим мнением и обязательно учтем его при отборе и подготовке к изданию последующих книг. Наши координаты:

E-mail: info@williamspublishing.com

WWW: <http://www.williamspublishing.com>

Информация для писем:

- из России: 115419, Москва, а/я 783
- с Украины: 03150, Киев, а/я 152

1. Искусственный интеллект

В июне 1979 года я закончил факультет электротехники Корнельского университета. На тот момент у меня не было четких планов относительно будущей карьеры, но вскоре меня пригласили на работу в новое подразделение компании *Intel* в Портленде, штат Орегон. Это было начало эры микропроцессорной индустрии, и *Intel* стояла у ее истоков. Моя работа заключалась в том, чтобы анализировать и устранять неполадки, обнаруженные нашими инженерами, которые занимались разработкой основного на тот момент товара, – компьютеров, работающих на одной плате. (Создание единого компьютера на одной электроплате стало возможным совсем недавно благодаря изобретенному в *Intel* микропроцессору.) Еще я занимался подготовкой ленты новостей, а также часто отправлялся в командировки, где встречался с клиентами. Я был молод и отлично проводил время, хотя и скучал по своей подружке из колледжа, которая работала в Цинциннати.

Несколько месяцев спустя я наткнулся на публикацию, которая полностью перевернула мою жизнь. Эта статья, напечатанная в сентябрьском выпуске журнала *Scientific American*, была посвящена функционированию головного мозга. Она с новой силой пробудила во мне возникший еще в подростковом возрасте интерес к данной теме. Статья была протрясающей.

Из нее я узнал о развитии и структуре этого удивительного органа, о нейронных механизмах, обеспечивающих функционирование органов чувств и способность передвижения, а также о биологически обусловленных нарушениях работы мозга. Поистине эта статья оказалась одной из лучших за все существование *Scientific American*. Несколько нейробиологов впоследствии рассказывали мне, что на их выбор карьеры упомянутая статья произвела столь же сильное влияние, что и в моем случае.

Ее написал Фрэнсис Крик – один из ученых, открывших структуру ДНК, – который к тому времени решил применить свой талант в области исследования проблем мозга. Крик утверждал, что, несмотря на несметное количество информации, вопрос о функционировании мозга как единого целого остается для науки открытым. Ученые, как правило, не любят писать о том, чего не знают, но Крика, похоже, это не слишком волновало. Он не боялся выступить в роли мальчишки, напоминающего королю о том, что тот голый. По мнению Крика, все данные о работе мозга были разрозненными и не увязывались в единую теорию. «По-дозреваю, что общая идея отсутствует напрочь», – сетовал Фрэнсис. Я полагаю, что, выразившись изысканным слогом английского джентльмена, исследователь фактически признал, что «наука не имеет ни малейшего понятия о том, как устроен мозг». И с того времени ситуация совсем не изменилась.

Слова Крика стали для меня своего рода призывом. Мое заветное желание познать мозг и создать мыслящий компьютер пробудилось с новой силой. Хотя я совсем недавно закончил колледж и вполне осознавал, что воплощение задуманного мной плана требует времени, это не поколебало моего желания изменить карьеру.

Весной 1980 года я перешел на работу в Бостонский офис компании *Intel*, поскольку хотел быть поближе к своей будущей жене, поступившей в то время в аспирантуру. В Бостоне я занимался обучением клиентов и сотрудников компании созданию систем на базе микропроцессоров. Но я не оставлял своей заветной идеи. Инженер внутри меня понимал, что, лишь познав работу мозга, мы сможем его воссоздать, а естественный путь создания искусственного мозга будет найден в кремнии. Я работал в компании, которая создала микропроцессор и интегральную микросхему памяти. Возможно, она заинтересуется моим предложением? Тогда я смог бы часть своего времени посвятить исследованиям разума и созданию интегральных микросхем памяти, которые по своей структуре были бы подобны человеческому мозгу. Однажды я обратился к главе правления компании Гордону Муру. В сокращенном варианте мое письмо выглядело так:

Уважаемый мистер Мур,

я предлагаю создать научную группу, которая занималась бы исследованиями человеческого мозга. Вначале она может состоять из одного человека, и я готов им стать. Я глубоко убежден, что близок к разгадке, которая позволит осуществить новый качественный переворот в сфере микроэлектроники.

С уважением, Джефф Хокинс

Мур познакомил меня с главным научным сотрудником *Intel* Тэдом Хоффом, который в то время был известен благодаря двум вещам. Об одной я знал – Хофф работал в группе, которая изобрела первый микропроцессор. А вот вторая, о которой я даже не подозревал, состояла в том, что в свое время он занимался теорией нейронных сетей. Хофф не понаслышке знал об искусственных нейронах и о перспективах их применения. К такому я был совершенно не готов. Выслушав мое предложение, он сказал, что, по его мнению, принципы работы мозга не будут раскрыты в ближайшем будущем, таким образом, для *Intel* не представляет интереса финансирование моего проекта. В бизнесе фактор времени – это все. Моему разочарованию не было предела.

От природы мне свойственно искать самые простые пути для достижения поставленных целей. Работать над вопросами мозга в *Intel* было бы путем наименьшего сопротивления. После отказа Хоффа такой вариант отпал, поэтому я взялся за следующий, который пришел мне в голову. Итак, я решил поступить в Массачусетский технологический институт, известный благодаря проводившимся в нем исследованиям в сфере искусственного интеллекта, к тому же

добираться до него тоже было недалеко. Чего еще можно пожелать? У меня была очень хорошая подготовка в области информационных технологий – это плюс. Я мечтал создавать разумные механизмы – еще один плюс. Я хотел изучать мозг для того, чтобы понять, как должны работать эти механизмы... Вот тут и возникала проблема. По представлениям ученых, работающих в лаборатории искусственного интеллекта Массачусетского института, данная цель не могла считаться научно обоснованной.

Мои усилия были сравнимы с ударами головы о каменную стену. Массачусетский институт был колыбелью искусственного интеллекта. Когда я подавал документы, он был пристанищем для десятков одаренных исследователей, которые были одержимы идеей программирования компьютеров, способных мыслить подобно человеку. С точки зрения этих ученых, визуальное восприятие, дар речи, робототехника и математика были попросту задачами программирования. Компьютер мог воспроизвести абсолютно все, что создает мозг, и даже больше, так зачем тогда усложнять себе жизнь, пытаясь постичь компьютер биологический? Изучение мозга только создаст дополнительные ограничения для исследователей, отвлечет их от насущных задач. Таким образом, священная миссия кибернетиков состоит в том, чтобы разрабатывать компьютерные программы, которые вначале будут имитировать, а потом и превзойдут человеческие возможности. Одним словом, маститые ученые придерживались принципа «Цель оправдывает средства» и совершенно не интересовались тем, как работает настоящий мозг. Некоторые из них даже кичились тем, что ни в грош не ставят нейробиологию.

Я интуитивно считал такой подход вопиюще неправильным, по сути – ведущим в никуда. Фундаментальные принципы работы компьютера и функционирования человеческого разума в корне различны. Основой первой является программирование, а второго – процесс самообучения. Компьютер, которым управляет центральный микропроцессор, предназначен для максимально точного исполнения заданных функций, а живой ум, у которого отсутствует единый центр контроля, наделен гибкостью и устойчивостью к возможным неудачам. Список различий не исчерпывается указанными особенностями. Я понял, что основной причиной, препятствующей созданию разумных механизмов, является их транзисторная структура. Именно осознание последнего вселило в меня глубокую уверенность в том, что мозг и компьютер – фундаментально различны. Не имея доказательств, на подсознательном уровне я испытывал непоколебимую уверенность. В конце концов я пришел к выводу, что хотя искусственный интеллект не поможет создать разумный компьютер, зато он вполне пригоден для изобретения других полезных устройств.

С готовностью принимая вызов, брошенный Фрэнсисом Криком, я желал досконально изучить анатомию и психофизиологию мозга и разработать, наконец, единую теорию разума. Мой исследовательский интерес был прикован к неокортексу – большей части головного мозга, развившейся у млекопитающих в процессе эволюции позже всего и отвечающей, как предполагается, за высшие интеллектуальные функции, такие как речь, обучение, память и мышление, одним словом – за умственную деятельность.

К сожалению, преподаватели и студенты Массачусетского института не поддержали меня в этих стремлениях. Мне прямо сообщили, что в сфере создания искусственного интеллекта нет места изучению живого мозга. В 1981 году мои документы были отклонены приемной комиссией.

Согласно распространенному мнению, искусственный интеллект давно существует, и единственная загвоздка в том, что современным механизмам не хватает мощности для его эффективной работы. Кибернетики, работающие в сфере искусственного интеллекта, убеждены в том, что смогут создать мыслящий компьютер, увеличив объем его памяти и ресурсов для обработки данных. Но не тут-то было. Ключевой недостаток искусственного интеллекта – отсутствие зоны, отвечающей за понимание. Чтобы понять, на какой стадии произошло отклонение от правильного пути, давайте обратимся к истории создания искусственного интеллекта.

Искусственный интеллект как подход зародился вместе с появлением цифровых компьютеров. Английский математик Алан Тьюринг, которого считают одним из пионеров идеи искусственного интеллекта, утверждал, что, несмотря на различия, все вычислительные

машины по сути своей одинаковы. Как часть доказательства он предлагал представить компьютер на основе всего лишь трех составляющих: устройства для обработки данных, бумажной ленты и прибора, выполняющего функции считывания и записи на ленту по мере ее продвижения. Бумажная лента предназначалась для записи данных в двужаночной системе исчисления (знаменитые 0 и 1 в современных компьютерных кодах). Замечу, что события происходили задолго до изобретения чипов памяти или дисководов, поэтому для хранения данных Тьюринг предложил использовать бумажную ленту (машина Тьюринга является собой «пределный» случай компьютера, когда сняты ограничения на размеры памяти, ведь лента бесконечна). Устройство для обработки данных (в наши дни его заменил центральный микропроцессор) подчиняется набору определенных правил по считыванию и редактированию данных на бумажной ленте. Тьюринг математически доказал, что если выбрать верный набор правил для центрального микропроцессора и вставить бесконечно длинную бумажную ленту, то процессор сможет выполнить любой заданный набор операций. Это был один из прототипов вычислительных машин, которые впоследствии получили название *универсальные машины Тьюринга*. Независимо от постановки задачи – будь то вычисление квадратного корня, траектории полета пули, участия в играх, редактирования визуального изображения, осуществления банковских операций, – все операции кодировались с помощью нуля и единицы. Последнее в свою очередь означало, что любая машина Тьюринга может быть спрограммирована для выполнения таких операций. Информационная обработка – это информационная обработка данных, которая представляет собой информационную обработку данных, т. е. все цифровые компьютеры логически равны друг другу.

Выводы Тьюринга были правильными и исключительно плодотворными. Они положили начало революции вычислительной техники. Затем Тьюринг обратился к теме создания разумных компьютеров. Прежде всего, следовало предложить формальное определение понятия *разум*, и Тьюринг нашел оригинальный способ проверки наличия интеллекта. Его метод, впоследствии получивший название теста Тьюринга, состоял в следующем: если компьютер сможет обмануть человека, который будет задавать ему различного рода вопросы, так, чтобы последний воспринимал отвечающего не как машину, а как другого человека, тогда компьютер может считаться разумным. Вот таким образом с помощью теста Тьюринга и его оценочных критериев, а также машины Тьюринга как посредника и появились на свет первые ростки истории создания искусственного интеллекта. Основополагающей догмой оставалась следующая: мозг – это просто еще одна разновидность компьютера, а значит, справедливо и обратное: искусственный интеллект должен воспроизводить мышление человека.

Приверженцы теории искусственного интеллекта проводили своего рода параллель между вычислительным процессом и процессом мышления. Они говорили:

«Послушайте, а ведь наиболее выдающиеся изобретения человеческого разума, вне всякого сомнения, возникают в результате использования абстрактной символики. Компьютеры используют символы, как и люди в процессе коммуникации, разговаривая или слушая, оперируют символами, которые мы называем словами (и символы эти подчиняются строгим правилам грамматики). Или возьмем игру в шахматы. Во время партии игроки используют мысленные символы, которые описывают расположение и возможности каждой фигуры на доске. Или же что такое зрение? Зрение – это символическое восприятие объектов, которые мы видим, их названий, свойств, расположения в пространстве. Безусловно, абстрактное мышление человека обусловлено мозгом, а не основано на работе вычислительных машин, но ведь Тьюринг убедительно доказал, что способ применения и особенности комбинирования символов не играют никакой роли. Мы можем создать машину из гаек и проводов, используя систему электронных переключателей или используя сети нейронов. Какая в сущности разница, главное – чтобы созданная вещь по своей функциональности стала эквивалентом универсальной машины Тьюринга».

В 1943 году в одном из влиятельных научных изданий была опубликована статья нейропсихолога Уоррена Мак-Калоба и математик Уолтера Питтса. Ученые описали механизм осуществления нейронам: цифровых функций, т. е. того, как нервные клетки могут

воссоздавать формальную логику, положенную в основу работы компьютеров. Нейроны вполне могут выполнять функцию логических входов/выходов как это принято называть у разработчиков электронно-вычислительной техники. Логические входы служат для выполнения операций «И» «ИЛИ», «НЕ». Микропроцессоры вычислительных машин состоят из миллионов логических входов, которые связаны между собой в очень четкую и сложную схему. Центральный микропроцессор – это не что иное как совокупность логических входов.

Мак-Калох и Питтс подчеркивали: вполне возможно, что нейроны тоже связываются между собой для выполнения тех или иных логических функций. Вероятно, они получают данные друг от друга, затем обрабатывают их и на основе обработки подают выходящие данные. Это означает, что нейроны являются живыми логическими входами. Таким образом, – продолжали развивать мысль двое ученых, – предположительно мозг по своему строению состоит из "И"-входов, «ИЛИ»-входов, а также других логических элементов, образованных нейронами, что само по себе аналогично структуре схем в цифровой электронной плате. Из публикации ясно, что исследователи лишь предполагали, что мозг так устроен, но не были в этом уверены. Рассуждая логически, описанный выше принцип работы нейронов не выглядит абсурдным. Теоретически нейроны вполне могут выполнять цифровые функции. С другой стороны, ни у Мак-Калоха, ни у Питтса не возник вопрос: а действительно ли нейроны взаимодействуют между собой именно таким способом? Несмотря на недостаток биологических доказательств, они выдвинули постулат, что мозг является разновидностью компьютера.

Следует отметить, что философия искусственного интеллекта пребывала под значительным влиянием бихевиоризма, наиболее влиятельного направления психологической науки первой половины XX века. Бихевиористы утверждали, что мозг является непознаваемым, и проводили аналогию с черным ящиком. С другой стороны, в исследованиях поведения животных была установлена связь между внешними факторами, влияющими на их поведение, и ответной реакцией. Исходя из этого, бихевиористы утверждали, что мозг включает рефлекторные механизмы, и, для того чтобы побудить животное к определенному типу поведения, нужно использовать методы поощрения и наказания. Если последовательно придерживаться упомянутого подхода, то нет необходимости изучать мозг, в особенности такие беспорядочные субъективные ощущения, как страх или голод, а также функции сознания. По вполне понятным причинам данное направление потеряло популярность во второй половине XX века, в то время как теория искусственного интеллекта продолжала развиваться.

После окончания Второй мировой войны электронные цифровые вычислительные машины получили более широкое применение. Первопроходцы создания искусственного мозга были полны энтузиазма. Научить компьютер переводить с одного языка на другой? Никаких проблем! Это напоминает взлом кода. Единственное, что для этого нужно, – соотнести каждый символ системы А с его аналогом в системе Б. Визуализация? Ни малейших трудностей. Человечеству давно известны теоремы о ротации, изменении масштабов, перемещении в пространстве, а значит, на их основе можно построить компьютерные алгоритмы, т. е. мы на верном пути. Ученые мужи, бьющиеся над созданием искусственного интеллекта, делали громкие заявления о том, как искусственный мозг сначала сравняется по своему потенциалу с человеческим разумом, а со временем и превзойдет его.

По иронии судьбы программисты смогли создать программу, которая была очень близка к тому, чтобы пройти тест Тьюринга. Программа, разработанная Джозефом Вейзенбаумом из Массачусетского технологического института, называлась «Элиза» и копировала поведение врача-психотерапевта при первичном опросе пациента³. Например, если девушка вводит предложение: «Я и мой парень больше не разговариваем», то Элиза спросит: «А расскажи-ка мне побольше о своем парне», или: «Как по-твоему, почему ты и твой парень больше не

³ Многие психотерапевты стараются говорить как можно меньше, ограничиваясь лишь краткими вопросами и общими наблюдениями, что позволяет пациенту самому рассказывать о своих переживаниях, поэтому в «Элизе» были заложены минимальные познания – информацию, необходимую для построения замечаний, имитирующих врача, программа извлекала из ответов пациента. – Примеч. ред.

общаетесь?»

Хоть автор программы заявил, что изобрел пародию на поведение психотерапевта, предназначенную лишь для исследования иллюзии понимания, которая часто возникает в разговоре между людьми. Некоторых «Элиза» приводила в изумление.

Более серьезным достижением были программы, подобные «Фигурному миру» – виртуальной комнате, смоделированной из геометрических фигур разных форм и цветов. Вы можете, например, спросить компьютер: «Стоит ли зеленая пирамида на большом красном кубе?», или же приказать ему: «Положи синий кубик на маленький красный кубик» и т. п. Программа ответит на ваш вопрос или выполнит ваш приказ. Все это было смоделировано и тем не менее работало! Единственным недостатком было то, что работа программы ограничивалась своим собственным мирком геометрических фигур. Программисты не смогли расширить ее функциональность, добившись решения более полезных задач.

Тем временем широкая публика была в полном восторге от кажущихся поразительными успехов в области разработки искусственного интеллекта. Одна из программ, вызвавших особенно бурные отклики, могла доказывать математические теоремы. Со времен Платона пошаговый метод математической дедукции считался венцом человеческого разума, а значит, рукотворный разум наконец-то сорвал джекпот! Однако ликование оказалось преждевременным. Как и в случае «Фигурного мира», прикладные возможности программы были несущественными. Она могла доказывать лишь очень простые теоремы, которые и так были давно известны и доказаны. Затем всеобщий интерес вызвали так называемые экспертные системы. Это были базы данных, которые могли давать ответы на вопросы человека. Например, медицинская экспертная система могла поставить диагноз больному, если в нее предварительно были введены списки возможных симптомов. Но, как и в предыдущем случае, применение экспертных систем было достаточно ограниченным, поскольку они не могли даже приблизиться к функциональной универсальности живого мозга. Компьютеры могли отлично играть в шахматы – всем известен случай, когда компьютер Deep Blue, детище компании *IBM*, выиграл партию у чемпиона мира по шахматам Гарри Каспарова. Все приведенные примеры не могут свидетельствовать о серьезных успехах в разработке искусственного интеллекта. Компьютер выиграл у чемпиона мира вовсе не потому, что он был умнее своего соперника, а потому, что скорость его реакций в миллионы раз превышает таковую у человека. У компьютера нет интуиции. Профессионал в шахматах смотрит на позиции фигур на доске, и в его уме возникает целостная картина. Он видит потенциально выигрышные и проигрышные зоны, а также предугадывает действия соперника. Компьютер же не обладает такой внутренней интуицией, поэтому вынужден анализировать намного большее количество ходов. Deep Blue не проводил ретроспективный анализ партии, ему ничего не было известно о своем противнике. Компьютер играл в шахматы, ничего в них не смысля, подобно тому, как калькулятор выполняет арифметические действия, не зная математики.

Все попытки создания искусственного интеллекта заканчивались изобретением очередной программы, обеспечивающей выполнение только одной четко определенной функции. Компьютеры так и не научились обобщать или проявлять гибкость. Даже создатели программ отмечали, что их детища не могут мыслить, как человек. Некоторые проблемы искусственного интеллекта поначалу казались простыми, но со временем так и не были найдены способы их решения. Даже сегодня ни один компьютер не может воспринимать речь так же легко, как трехлетний ребенок, или видеть так же хорошо, как, скажем, мышь.

После многолетних усилий, несбыточных надежд и отсутствия какого-то осязаемого прогресса идея искусственного интеллекта потеряла былую привлекательность. Ученые, отдавшие годы жизни исследованиям и разработкам в данной области, сменили сферы научных интересов. Предпринимательские инициативы, основанные на применении мыслящих компьютеров, тоже терпели неудачи. Финансирование подобных проектов значительно сократилось. Складывалось впечатление, что воссоздание таких базовых человеческих функций, как восприятие, речь или осмысленная деятельность, вообще не представляется возможным. Это продолжается и по сей день. Как я уже отмечал, у идеи искусственного интеллекта остались некоторые сторонники, но большинство ученых считает его создание невозможным.

Не стоит возлагать вину за провал на первопроходцев этой теории. Алан Тьюринг был гением, его универсальная машина, без преувеличения, совершила революцию в мире технологий, хотя и не помогла создать искусственный интеллект.

Мой скептицизм в отношении создания искусственного интеллекта соотносится с периодом, когда я подавал документы для поступления в Массачусетский технологический институт. Джон Сирл, один из наиболее именитых преподавателей философии Калифорнийского университета (Беркли), в свою очередь утверждал, что у компьютера разума нет и быть не может. В доказательство своих взглядов в 1980 году он предложил провести эксперимент, который назвал «Китайской комнатой». Эксперимент состоял в следующем.

Представьте комнату со щелью в стене. За столом в комнате сидит человек, который владеет исключительно английским языком. На столе перед ним лежат карандаши, большой учебник и стопка листов бумаги. Полистав учебник, испытуемый выясняет, что это написанное на английском языке руководство, подробно разъясняющее, как писать, различать и сравнивать китайские иероглифы. Заметьте, что значения иероглифов не приводятся, речь идет лишь о правилах воспроизведения символов.

Другой человек находится вне описанной комнаты. Время от времени он просовывает сквозь щель в стене записки. На них по-китайски написан рассказ и вопросы к нему. Испытуемый, не имеющий никакого представления о китайском языке, по условиям эксперимента должен взять руководство к действию и поработать над полученными записками. Он старается и пыхтит, четко выполняя инструкции из книги. Иногда в инструкциях стоит указание написать определенные символы на листике бумаги, иногда он должен переставить или убрать символы. Шаг за шагом, неуклонно следуя инструкциям, человек пишет одни иероглифы, убирает или переставляет другие, пока не выполнит задание до конца. Так создается целая страница новых символов, которые являются ответами на вопросы, заданные на китайском языке. Разумеется, сам автор страницы этого не осознает. В книге содержится указание отдать исписанный листок бумаги сквозь щель в стене. Участник эксперимента так и делает, недоумевая, в чем же суть столь странного задания.

Полученный от испытуемого лист с иероглифами дают прочитать китайцу, который отмечает, что все ответы – верны, а некоторые из них впечатляют своей глубокомысленностью. Если мы спросим у китайца, считает ли он, что ответы написал разумный человек, который понял прочитанный рассказ, он ответит утвердительно. Как вы думаете, он прав? Кто понял рассказ? Однозначно, не человек, сидящий в комнате, ведь он не знает китайского, а значит, не имеет ни малейшего представления, о чем шла в исходной записке. Так на какой стадии было достигнуто понимание? Ответ Сирла: ни на какой. Все, что происходило в комнате, – это бездумное листание страниц и рисование. А сейчас проведем параллель: «Китайская комната» – это аналог цифрового компьютера, руководство по написанию иероглифов есть не что иное, как «компьютерная программа», соответственно, ее авторы – «программисты». Таким образом, все старания смоделировать искусственный интеллект напрасны. Участник эксперимента, подобно компьютеру, манипулирует символами, но не может придать им какого бы то ни было смысла. Следовательно, заключил Сирл, машина, или физическая система способная выполнять определенные функции, не может стать разумной и действовать осознанно. Сирл отмечал, что он затрудняется дать определение понятию разума, однако полон уверенности, что, независимо от определения, компьютеры разумом не обладают и обладать не могут.

Выводы Сирла спровоцировали войну между философами и сторонниками проблем искусственного интеллекта. Последние выдвигали десятки контраргументов. Вот, например один из них: хотя ни одна из составляющих «Китайской комнаты» не понимала китайского, сама комната как целое понимала язык, т. е. испытуемый понимал китайский язык, сам того не осознавая. Что касается меня самого, то я вполне разделяю точку зрения Сирла и считаю, что в эксперименте под названием «Китайская комната» налицо лишь манипуляция символами, но не истинное понимание. Давайте определим, что же такое «понимание».

Чтобы понять рассказ, человеку достаточно его прочесть. С другой стороны, если я не подам виду, что понял рассказ, то вы не сможете заключить из моего поведения, владею ли я языком, на котором написан рассказ или нет, а также усвоил ли я предложенную информацию. Чтобы выяснить это, вы можете задать вопросы. Однако заметьте, что мне придется отвечать на

них сейчас, а рассказ-то я понял намного раньше, еще во время чтения. Ключевой тезис моей книги: понимание не поддается оцениванию на основе наблюдения внешних реакций. Из следующих глав вы узнаете, что понимание является внутренней матрицей, которая формируется в мозге при запоминании информации и в дальнейшем используется для построения прогнозов. Компьютер Deep Blue, испытанный в «Китайской комнате», и многочисленные компьютерные программы не способны осознавать выполняемые ими функции. А единственным способом оценки разумности машины является ее «поведение».

И, наконец, последний аргумент в пользу искусственного интеллекта. Теоретически компьютеры могли бы имитировать работу всего живого мозга. Возможно, что когда-то наука дойдет до создания компьютерной модели, имитирующей работу нейронов и взаимодействие между ними. Если бы такое действительно произошло, то человеческий разум и искусственный интеллект можно было бы назвать равноценными. Хотя в реальной жизни вряд ли возможна столь совершенная имитация живого ума, но с теоретической точки зрения она выглядит вполне корректно. Беда в том, что исследователи, занимающиеся созданием искусственного интеллекта, не пытаются имитировать его живой прототип, а программы, которые они создают, по сути своей не могут проявлять разум. Не понимая того, как работает живой мозг, ни один ученый не сможет создать его искусственный аналог.

Получив отказы в компании *Intel* и Массачусетском технологическом институте, я пребывал в растерянности. Решив придерживаться одной из моих жизненных стратегий – не зная, как лучше действовать, лучше ничего не менять до тех пор, пока не прояснятся возможные варианты выбора, – я продолжал работать в индустрии информационных технологий. Мне нравился Бостон, но жена настояла на переезде в Калифорнию, что и свершилось в 1982 году (данное событие также вполне соответствовало моему жизненному кредо – всегда избирать путь наименьшего сопротивления). Я нашел работу в Силиконовой долине, в новой компании *Grid Systems*. Именно в этой фирме был создан первый портативный компьютер – отличное устройство, ставшее первым экспонатом в коллекции Музея современного искусства в Нью-Йорке. Сначала я работал в отделе маркетинга, потом – в инженерном отделе. Я разработал новый язык программирования под названием *GridTask*, принесший компании большой успех, и моя карьера быстро пошла в гору.

Но, несмотря ни на что, меня все так же преследовала загадка человеческого мозга и идея создания «разумных» компьютеров. Снедаемый желанием изучать живой мозг, я поступил на дистанционный курс «Физиология человека» (к счастью, на дистанционную форму обучения охотно принимают всех желающих). Получив базовые знания по биологии, я решил поступить в аспирантуру и заняться изучением разума с позиций биологической науки. Раз уж мир информационных технологий отверг начинания теоретика, изучающего проблемы мозга, то, возможно, мир биологии будет более гостеприимным в отношении ученого-кибернетика. В то время не существовало такого направления, как теоретическая биология, теоретической нейробиологии не было и в по-мине, и наиболее близкой кругу моих интересов я посчитал биофизику. Я тщательно подготовился, составил резюме, собрал рекомендательные письма, сдал вступительные экзамены, и – о счастье! – поступил в аспирантуру по специальности «Биофизика» в Калифорнийский университет (Беркли).

Я сгорал от нетерпения. «Наконец можно будет вплотную заняться изучением теории разума», – думал я. Работа в *Grid Systems* осталась позади, и в мои планы не входило возвращение в индустрию информационных технологий, что неминуемо означало ухудшение финансово-го благосостояния. Моя жена как раз созрела для того, чтобы подумать о продолжении рода, и тут я со счастливой миной бросаю свою прежнюю работу и перестаю быть кормильцем семьи. Трудно назвать такое решение путем наименьшего сопротивления, но я избрал свой путь сознательно и считал принятое решение наилучшим. К тому же меня поддержала и жена.

Джон Элленби, основатель *Grid Systems*, прощаясь со мной, сказал: «Я понимаю, что ты не намерен возвращаться в *Grid* или в кибернетику вообще, но человек предполагает, а Бог располагает. Вместо того чтобы сжигать все мосты, возьми лучше длительный отпуск за свой счет. И если все же через год-два тебе захочется вернуться в компанию, ты придешь на ту же должность и заработную плату плюс та же доля акций нашей компании». Это был очень

благородный жест с его стороны. Я согласился, хотя и предчувствовал, что навсегда покидаю компьютерный бизнес.

2. Нейронные сети

В январе 1986 года я приступил к учебе в аспирантуре. Прежде всего я решил составить обзор теорий разума и функций мозга. Я перечитал сотни работ анатомов, физиологов, лингвистов, философов, психологов, а также ученых-кибернетиков. В каждой из перечисленных сфер деятельности предлагалась своя терминология и особый взгляд на функционирование мозга. Все описания показались мне достаточно поверхностными и неопределенными. Лингвисты писали о разуме в терминах синтаксиса⁴ и семантики⁵, утверждая, что мозг и, соответственно, интеллект, можно полностью исследовать посредством языка. Изучавшие зрительное восприятие писали о двухмерном и трехмерном измерениях. Они полагали, что основой разума и функционирования мозга служит визуальное распознавание объектов окружающей среды. Компьютерщики писали о схемах и структурах – понятиях, специально введенных для описания информации, но не утруждали себя изучением живого разума. С другой стороны, анатомы и нейрофизиологи очень подробно описывали строение мозга и функционирование нейронов, но при этом даже и не пытались создать единую всеохватывающую теорию. Упорядочить это невероятное количество подходов и подкрепляющих их экспериментальных данных представлялось невыполнимым.

И вот в сфере разработки искусственного интеллекта появилось новое многообещающее направление, а именно – нейронные сети, о которых заговорили еще в начале шестидесятых годов XX века. В умах руководителей организаций, осуществляющих финансирование исследований, нейронные сети и искусственный интеллект занимали конкурирующие позиции. Причем 800-фунтовой гориллой, которая легко вытесняла соперника с ринга, был искусственный интеллект. Исследования нейронных сетей финансировались плохо и на несколько лет были даже занесены в черный список. Но это не остановило немногочисленных энтузиастов, не перестававших заниматься нейронными сетями. Их счастливая звезда возшла в середине 1980-х годов. Трудно сказать наверняка, чем именно был вызван внезапный интерес к нейронным сетям, но, вне всякого сомнения, свою роль сыграла череда неудачных попыток создания искусственного интеллекта. Начался поиск альтернатив, и выбор пал на нейронные сети.

Создатели нейронных сетей (НС) оказались далеко впереди своих коллег, бившихся над разработкой искусственного интеллекта. НС были построены, хоть и весьма приблизительно, по принципу биологической нервной системы. Вместо того чтобы заниматься программированием, исследователи нейронных сетей, так называемые коннекционисты⁶, сосредоточились на изучении того, какие типы поведения генерируют различные нейронные комбинации. Мозг состоит из нейронов, стало быть, мозг – это нейронная сеть. Задачи коннекционистов состояли в том, чтобы изучить неуловимые свойства разума путем изучения взаимодействия нейронов. Они рассчитывали, правильно воссоздав связи между группами нейронов, тем самым приблизиться к решению задач, которых не смог одолеть искусственный интеллект. Нейронные сети отличаются от компьютеров тем, что у них нет центрального микропроцессора и они не сохраняют информацию в центральном блоке памяти. Информация, занесенная в память нейронной сети, сосредоточена в связях – точно так же, как и в головном мозге человека.

⁴ Синтаксис – здесь: способы соединения слов и символов в логическое высказывание. – Примеч. ред.

⁵ Семантика – лингвистическая дисциплина, изучающая значения языковых единиц, их функционирование в языке и речи. – Примеч. ред.

⁶ Другое название нейронных сетей – коннекционистские сети. – Примеч. ред.

На первый взгляд, разработка нейронных сетей полностью соответствовала сфере моих научных интересов. Однако на тот момент я четко видел три фактора, критических для понимания работы мозга.

Во-первых, в исследования мозга следует включать временной критерий, ведь скорость обработки потока информации чрезвычайно высока. Данные, поступающие в мозг и исходящие из него, никогда не пребывают в статическом состоянии. Во-вторых, мозг насквозь пронизан обратными связями. Например, обмен между неокортексом и таламусом, главным подкорковым центром, направляющим импульсы всех видов чувствительности (температурной, болевой и др.) к стволу мозга, подкорковым узлам и коре больших полушарий, построен таким образом, что количество обратных связей превышает количество исходящих почти в десять раз! Это значит, что на каждое волокно, подающее информацию в неокортекс, приходится десять волокон, отправляющих обратную информацию к органам чувств. Обратная связь также является преобладающей формой связи между нейронами внутри неокортекса. Роль обратной связи пока что до конца не изучена, но с уверенностью можно сказать, что эта связь вездесуща. Для нас это очень важно.

Наконец, в-третьих, любая модель (или теория мозга) должна соответствовать биологическому строению живого мозга. У неокортекса очень сложное строение, которое, как мы увидим позже, представляет собой повторяющуюся иерархию. Любая нейронная сеть, не имеющая таковой, не сможет воссоздать работу мозга.

Первые нейронные сети представляли собой крайне упрощенные v модели, которые не удовлетворяли ни одному из трех описанных выше требований. Большинство из них были трехслойными. Входной слой нейронов служил для ввода значений входных переменных. Нейроны этого слоя были связаны с нейронами промежуточного слоя, так называемыми скрытыми элементами. Скрытые элементы были связаны с последним слоем нейронов – элементами выхода. Связи между нейронами имели переменную силу. Это означало, что активность внутри одного нейрона могла усилить активность внутри второго и ослабить активность внутри третьего – в зависимости от силы связей. Изменяя силу связей, можно было «обучить» нейронную сеть соотношению входящих данных с исходящими.

Простейшие нейронные сети были предназначены для обработки статических данных, в них не была задействована обратная связь, и они не имели ничего общего с живым мозгом. Типичный пример нейронной сети – сеть с прямой передачей сигнала, в которой элементы выхода транслируют сигнал об ошибке, и этот сигнал подается на элементы входа. Может показаться, что трансляция ошибок является обратной связью, но на самом деле это не так. Обратная трансляция ошибок происходила только на стадии обучения, а в процессе собственно функционирования нейронной сети потоки информации всегда передавались одним и тем же способом. Никакой обратной связи между элементами входа и выхода не наблюдалось. Кроме того, модель не учитывала временной фактор. Статические входящие сигналы превращались в статические исходящие сигналы – и ничего более. Потом подавался следующий сигнал, и все повторялось сначала. Нейронная сеть не сохраняла никаких данных о произошедших событиях, даже о тех, что случились совсем недавно. Структура нейронных сетей не шла ни в какое сравнение со сложным иерархическим строением человеческого мозга.

Я рассчитывал, что разработчики нейронных сетей в будущем перейдут от простых моделей к более совершенным и реалистичным, однако этого не произошло. Складывалось впечатление, что из-за подтвержденной функциональности простейших нейронных сетей все исследователи на долгие годы готовы остановиться на достигнутом уровне. Был найден новый и очень занятный инструмент; тысячи ученых, инженеров и студентов получали гранты, защищали диссертации, издавали книги о нейронных сетях. При помощи нейронных сетей составлялись прогнозы для фондового рынка, проводилась обработка документации для кредитования, осуществлялись графологические экспертизы и сотни других операций. Кто знает, возможно цели изобретателей нейронных сетей были гораздо более широкими, но в конечном счете эту сферу оккупировали те, кто меньше всего интересовался пониманием работы мозга и устройством человеческого разума.

Различия между нейронными сетями и работой мозга вопиющим образом нивелировались средствами массовой информации. В газетах, журналах, научных телепередачах нейронные

сети называли «подобными мозгу человека» или «функционирующими совершенно по такому же принципу, что и человеческий мозг». В отличие от программируемого искусственного интеллекта нейронные сети обучались на примерах, и на основании этого их считали носителями разума. В качестве наиболее яркого примера могу привести NetTalk – нейронную сеть, способную переводить последовательности буквенных символов в звуки человеческой речи. Обученная работать с печатным текстом, нейронная сеть обрела компьютерный голос, читающий слова, и в федеральных новостях ее тут же окрестили «читающей машиной». В действительности механизм ее действия очень прост: она соотносит буквенные комбинации с предопределенными речевыми сигналами.

Попробую объяснить, используя аналогию, насколько далеки нейронные сети от настоящего мозга. Представьте себе, что, вместо того чтобы разбираться в работе мозга, мы решили исследовать цифровой компьютер. Потратив на это годы, мы приходим к выводу, что компьютер состоит из сотен миллионов транзисторов, которые связаны между собой в очень точную и сложную схему. В то же время мы не поймем, ни как работает компьютер, ни почему транзисторы собраны именно в такую схему. И, если однажды мы соберем несколько транзисторов в одну схему, чтобы понять, как они работают, окажется, что три транзистора, собранные в определенную схему, преобразуются в усилитель. Слабый на одном конце сигнал многократно усиливается на другом. Такие усилители сейчас используются в радиоприемниках и телевизорах. Это важное открытие в свое время произвело фурор. Начался промышленный выпуск транзисторных радиоприемников, телевизоров и другой электроники с использованием усилителей транзисторного типа. Все это конечно, замечательно, но не вносит ясности в понимание устройства компьютера. Усилитель и компьютер состоят из транзисторов, но, тем не менее, у них нет практически ничего общего. Точно так же человеческий мозг и простейшая нейронная сеть состоят из нейронов – и это все, что их объединяет.

Летом 1987 года произошел случай, еще больше охладивший мой пыл в отношении нейронных сетей. Я участвовал в работе конференции, посвященной нейронным сетям, и побывал на презентации компании *Nestor*. Целью презентации была продажа приложения, созданного на основе нейронных сетей и предназначенного для распознавания рукописного текста. Стоимость продукта, безусловно, завоевавшего мое внимание, составляла 3 млн. долл. Хотя *Nestor* представляла приложение к нейронным сетям как значительно усложненное и модернизированное и говорила о нем как о блестящем научном открытии, я чувствовал, что проблема распознавания рукописного текста может быть решена не столь изощренным, а более традиционным способом. В тот день я вернулся домой, непрестанно раздумывая над полученными впечатлениями. Два дня спустя я создал устройство для распознавания рукописного текста – быстродействующее, небольшое по размеру и функциональное. В моем изобретении не использовались нейронные сети, да и принцип его работы не был имитацией функционирования головного мозга человека. И, хотя та конференция пробудила во мне интерес к созданию компьютеров с интерфейсом, включающим пишущий элемент (что в конечном счете спустя десятилетие привело к созданию *PalmPilot*), я убедился в том, что традиционные методы ни в чем не уступают нейронным сетям.

Система распознавания рукописного текста, созданная мною, в конце концов стала основой для системы ввода текста, получившей название *Graffiti* (последняя применялась в первых версиях «палмов»). А компания *Nestor*, насколько мне известно, просуществовала недолго.

Море внимания столь простому явлению, как нейронные сети! Большинство их возможностей легко воспроизводятся при помощи иных методов. И вот, наконец, вся эта шумиха вокруг НС потихоньку начинает спадать. Надо отдать должное изобретателям и исследователям нейронных сетей – они хотя бы не утверждали, что их детищу присущ интеллект. Первые нейронные сети были весьма простыми устройствами и значительно уступали программам, базирующимся на искусственном интеллекте. (Я вовсе не хочу, чтобы у вас сложилось впечатление, будто нейронные сети представляют собой упрощенную трехслойную структуру.) Ряд исследователей продолжили изучать нейронные сети различных модификаций. В настоящее время термин *нейронные сети* применяется для описания широкого диапазона моделей, в разной степени соответствующих биологическому «шаблону».

Но не было предпринято даже малейших попыток воспроизвести общий принцип функционирования и архитектуру неокортекса.

По моему мнению, фундаментальной проблемой большинства нейронных сетей является следующая их особенность, присущая также программам, основывающимся на принципе искусственного интеллекта. И программы, базирующиеся на принципах искусственного интеллекта, и нейронные сети отягощены акцентом на поведении. Называют ли его «ответами», «моделями» или «выходными сигналами», предполагается, что именно в поведении, моделируемом НС или ИИ, заключается их «разумность». Об успешности компьютерной программы или нейронной сети судят по тому, выдает она правильный или желаемый выходной сигнал. Со времен Алана Тьюринга между интеллектом и поведением ставят знак равенства.

Но интеллект – это не просто разумный способ действий или разумное поведение. Действительно, поведение является проявлением интеллекта, но ни в коем случае не его ключевой характеристикой и не основным аспектом определения данного понятия. Это легко доказать в одно мгновение: вы остаетесь разумным, лежа в полной темноте, размышляя и осозная себя самого и окружающий мир. Отрицая важность того, что происходит у вас в голове, и сосредоточиваясь, наоборот, на видимом поведении, ученые создали себе непреодолимое препятствие на пути познания человеческого интеллекта и создания по-настоящему разумных машин.

Прежде чем мы сформулируем новое определение интеллекта, я хочу рассказать вам о другом, коннекционистском, подходе, последователи которого подошли к пониманию работы мозга гораздо ближе. Основная проблема состоит в том, что важность указанного направления познания осознают лишь немногие ученые.

В то время как нейронные сети попали в центр внимания исследователей, небольшая отколовшаяся группа теоретиков разработала НС, положив в ее основу не поведение. Изобретение, названное *автоассоциативной памятью*, также состояло из простых «нейронов», сообщающихся друг с другом и генерирующих возбуждение при достижении определенного порога. В отличие от нейронных сетей предыдущих поколений (в которых информация проходила только в одном направлении) нейроны автоассоциативной памяти были соединены большим числом обратных связей. Каждый выходящий сигнал в таких сетях преобразовывался в исходящий – как в телефонном разговоре. Благодаря этой цепи обратных связей ученые получили возможность наблюдать ряд интересных моментов.

Усвоив определенную модель активности, искусственные нейроны запомнили ее. Автоассоциативная сеть ассоциировала модели сами с собой, отсюда ее название – автоассоциативная память.

Результат подобной связи, на первый взгляд, может показаться смехотворным. Для воспроизведения модели, сохраненной в такой памяти, следовало предоставить указанную модель. Проиллюстрирую данный вывод наглядным примером. Допустим, вы пришли в магазин с намерением купить связку бананов. Когда продавец просит вас оплатить покупку, вы говорите ему, что расплатитесь... бананами. «Что же ценного тогда в этом открытии?» – спросите вы. А то, что автоассоциативная память обладает несколькими важными особенностями, присущими головному мозгу человека. Самая важная из них такова: нет необходимости предъявлять системе целую модель, которую вы желаете получить от нее, – достаточно ее фрагмента или видоизмененной по сравнению с желаемой модели. Автоассоциативная память способна воспроизвести сохраненную модель в неизменном виде, даже если все, что у вас есть в наличии, – ее искаженный вариант.

Вернемся к нашему примеру. Представьте, что вы приходите в магазин с несколькими надкушенными бананами и получаете взамен целую связку. Или обращаетесь в банк с просьбой обменять надорванный и испорченный чек, и ваше желание удовлетворяют: вы получаете новенький, хрустящий документ.

В отличие от большинства нейронных сетей автоассоциативная память может быть разработана таким образом, что будет сохранять целые последовательности моделей, или временные паттерны. Такая ее особенность достигается путем добавления временной задержки к обратной связи. Благодаря этой задержке вы можете предоставлять автоассоциативной

памяти последовательности моделей (что подобно звучанию мелодии), и она будет сохранять их.

Таким образом, я могу предоставить автоассоциативной памяти первые несколько нот песни *Twinkle Twinkle Little Star*, а система воспроизведет ее целиком. Имея в наличии лишь часть усвоенной последовательности, автоассоциативная память воссоздаст ее всю.

Как мы увидим ниже, именно таким способом обучаются люди – усваивая последовательности паттернов. И я предполагаю, что мозг использует эти цепи подобно тому, как это происходит в автоассоциативной памяти.

Принцип действия автоассоциативной памяти навел нас на мысль о потенциальном значении обратной связи и изменяющихся во времени входящих сигналов. Но в подавляющем большинстве машин, обладающих искусственным интеллектом, и в нейронных сетях фактору времени и обратной связи не придается должного значения. Подобную ошибку допускают и специалисты, практикующие в области когнитивной психологии.

Немногим отличаются от них ученые, занимающиеся проблемами головного мозга (нейробиологи, нейрологи). Они знают о том, что существует обратная связь (ведь они сами ее и открыли), но не предлагают теории, которая бы шла дальше неопределенных высказываний о «фазах» и «модуляциях» и которая бы демонстрировала существенную роль данного феномена в функционировании головного мозга в целом. Они склонны описывать мозг в терминах локализации тех или иных процессов, но почти не касаются того, в каких случаях и каким образом нейроны взаимодействуют друг с другом.

Это упущение частично объясняется ограниченностью существующих в данное время экспериментальных техник. Одной из наиболее предпочитаемых техник 1990-х годов (как вы помните, «Десятилетия мозга») была техника функционального изображения. Соответствующие механизмы строили изображения мозговой активности у людей. Однако они не были высокочувствительными – не выделяли быстрых изменений. Тогда ученые предложили участникам эксперимента сосредоточиться на единственной задаче: на протяжении довольно длительного времени их просили замереть, как перед объективом фотокамеры, но фиксировалось не изображение их внешности, а мыслей. В результате мы собрали много информации о том, *где именно* в мозге локализуются зоны, ответственные за решение разных задач, но очень мало знаем о том, *какой* путь в головном мозге проходят входящие сигналы, отражающие реальность и изменяющиеся во времени.

Техника функционального изображения позволяет нам понять, *где именно* в головном мозге происходят текущие процессы, но не предоставляет возможность осознать, каким образом мозговая активность меняется на протяжении времени. Ученые хотели бы получить такие данные, но у них нет подходящих методов.

Итак, многие ведущие нейробиологи впадают в так называемое заблуждение входа-выхода. Вы предоставляете системе входящий сигнал и смотрите, *что* получите на выходе. Диаграммы прохождения информационных потоков в головном мозге показывают, как потоки сигналов поступают вначале в первичные сенсорные зоны мозга (воспринимающие визуальные, звуковые, тактильные, обонятельные и вкусовые сигналы), а затем, посредством обратной связи, отправляют команды к мышцам. Вы чувствуете – и действуете.

Я не утверждаю, что никто из исследователей не уделял внимания факторам времени и обратной связи. В столь обширной сфере практически каждая идея имеет своих приверженцев. В последнее время интерес к проблемам обратной связи, временного фактора и прогностической функции мозга растет. Но акцент на разработке моделей искусственного интеллекта и нейронных сетей привел к тому, что другие подходы обесценились – им уделялось гораздо меньше внимания, чем они того заслуживают.

* * * * *

Не трудно понять, почему люди – как специалисты, так и непрофессионалы – считают, что поведение определяет интеллект.

Как минимум, на протяжении двух столетий люди проводили аналогии между работой головного мозга и часовым механизмом, затем насосом, трубопроводом, паровым двигателем и,

наконец, компьютером.

Многие десятилетия научная фантастика купалась в идеях искусственного интеллекта – от законов роботехники Айзека Азимова до «Звездных войн».

Идея разумных машин, которые *осуществляют действия*, прочно укоренилась в нашем сознании. Все механизмы, созданные людьми или существующие всего лишь в воображении, разработаны для выполнения каких-либо функций. У нас нет механизмов, которые думают, но есть такие, которые действуют. Даже наблюдая за другими людьми, мы обращаем внимание прежде всего на их поведение, а не на скрытые мысли.

Казалось бы, мысль о том, что именно поведение должно служить мерилom интеллекта, вполне разумна. Однако, оглядываясь на историю развития науки, мы приходим к выводу, что мнение, лежащее на поверхности, – именно то, что чаще всего мешает увидеть правду. Научные рамки понимания зачастую сложны, поскольку интуитивные, но неверные предположения затрудняют поиск правильного ответа.

Астрономы, предшественники Н. Коперника (1473–1543), ошибочно полагали, что Земля неподвижна и расположена в центре Вселенной. Ведь они чувствовали, что земля под ногами никуда не движется, поэтому им казалось, будто они находятся в центре Вселенной. Им казалось очевидным, что звезды являются частью гигантской вращающейся сферы с Землей в центре. Того, кто осмелился бы предположить, что Земля вращается вокруг своей оси и движется вокруг Солнца со скоростью приблизительно 30 километров в секунду (это почти в 100 раз быстрее скорости звука), не говоря уже о том, что звезды расположены за триллионы миль от нас, назвали бы сумасшедшим.

Но именно эти еретические рамки познания были правильными. То, что интуитивно казалось очевидным, на самом деле было ошибочным.

До Ч. Дарвина (1809–1882) казалось очевидным, что дифференциация живых видов жесткая и неизменная. Представление о том, что виды эволюционируют, противоречила не только религиозным учениям, но и здравому смыслу, ведь, принимая концепцию эволюции, нам следует согласиться с тем, что мы, люди, имеем общих родственников со всеми живыми существами на планете, включая червей и комнатные растения, цветущие на подоконнике.

Я привел все эти примеры, поскольку убежден в том, что созданию разумных машин мешает ряд интуитивных предположений, в действительности затрудняющих поиск. Когда вы задаете себе вопрос: «Что должна делать система, обладающая искусственным интеллектом?», то интуитивно ищете ответ на него в поведенческих терминах. Мы проявляем человеческий интеллект посредством устной речи, владения письменной речью и действий, не так ли? Да, но только в определенной степени. Интеллект – нечто, происходящее в нашей голове. Разумное поведение – это не обязательная составляющая. Данный факт нельзя назвать очевидным, но, с другой стороны, его не так трудно признать.

Весной 1986 года я дни и ночи напролет просиживал в своем кабинете, читая научные статьи и занимаясь разработкой моей теории интеллекта. Также я изучал существовавшие на то время теории искусственного интеллекта и нейронных сетей, пока наконец не погряз в деталях. Материалов по этой теме, которые следовало бы прочесть и изучить, было великое множество. Но у меня никак не складывалась цельная картина того, как функционирует головной мозг.

Причина заключалась в том, что сама наука о головном мозге погрязла в деталях. Мало что изменилось и сегодня. Ежегодно публикуются тысячи отчетов о проведенных исследованиях, но они лишь добавляют путаницы вместо того, чтобы внести ясность. До сих пор не существует единой теории, позволяющей досконально разобраться, в чем именно заключается работа головного мозга и как он осуществляет ее.

Я попробовал представить себе, каким может быть решение этой задачи. Должно ли оно быть крайне сложным, поскольку мозг – сложное устройство? Придется ли описывать работу мозга с помощью сотен страниц математических формул? Придется ли разработать сотни тысяч схем, прежде чем прийти к полезному выводу? Я так не думаю. Из истории нам известно, что лучшие научные открытия и решения просты и изящны. Хотя отдельные элементы знания могут выглядеть непостижимыми, а путь структурирования теории невероятно сложен, окончательная концепция обычно очень проста.

Без ключевого объяснения, которым можно было бы руководствоваться в поиске,

ученые-нейробиологи не смогут связать отдельные факты и имеющиеся данные в цельную картину. Мозг – невероятно сложное, обширное образование. На первый взгляд, он напоминает кастрюлю, наполненную вареными спагетти. Можно его назвать и ночным кошмаром электрика, но при ближайшем рассмотрении мы обнаруживаем, что устройство мозга не столь беспорядочное, как клубок спутанных между собой проводов. Мозг имеет свою структуру, но она слишком сложна, чтобы рассчитывать понять ее на интуитивном уровне.

Мы терпим неудачу за неудачей не только потому, что нам не хватает данных или умения их сопоставить; больше всего на данном этапе мы нуждаемся в изменении перспективы. Сформулировав правильные рамки познания, мы убедимся, что непонятные ранее подробности легко складываются в цельную картину, подобно тому как разно-цветные кусочки стекляшек в калейдоскопе образуют поразительной красоты узор.

Ниже я привожу еще одну аналогию, благодаря которой вы осознаете смысл моих слов.

Представьте, что спустя тысячелетия человеческая раса вымрет. И однажды с другой планеты, на которой обитает высокоразвитая цивилизация, прибудут исследователи, которые захотят изучить, как жили люди. Наибольшее изумление у них вызвала бы наша сеть дорог. Для чего может быть нужна такая забавная система? Задавшись целью узнать ответ, инопланетяне начали бы скрупулезное изучение земного рельефа, подключив также изображения, получаемые со спутников. Будучи искусными археологами, инопланетные исследователи изучили бы каждый фрагмент асфальтового покрытия, каждый дорожный указатель, упавший и частично разрушившийся, каждую мелочь, оставшуюся от человеческой цивилизации. Они обратили бы внимание, что сети дорог отличаются друг от друга. Некоторые из них – извилистые, узкие и, похоже, расположены хаотично (вероятно, возникли сами по себе). Другие образуют строгую сетку, а на некоторых участках, становясь широкими и ровными, буквально разрезают пустыни.

Так инопланетные исследователи собрали бы множество сведений о дорогах, но все эти данные ровным счетом ничего для них не значили бы. Не теряя надежды, они продолжали бы собирать все новые и новые факты в надежде, что вот-вот найдут разгадку. Но все напрасно. Так продолжалось бы до тех пор, пока один из исследователей не воскликнул бы: «Эврика! Кажется, я понял. Эти существа не могли передвигаться с помощью телепортации, как мы с вами. Им приходилось передвигаться в буквальном смысле с места на место, возможно, на забавных подвижных платформах». Теперь, на фоне сделанного открытия, многое стало бы на свои места. Исследователи поняли бы, что узкие петляющие улочки сохранились с древнейших времен, когда средств передвижения было немного и они были медленными. Широкие дороги были предназначены для перемещения на большие расстояния с большими скоростями (наконец нашлось и объяснение, почему числа на знаках, расположенных на этих дорогах, были другими). Ученые начали бы различать жилые и промышленные зоны, очертили бы взаимодействие коммерческой и транспортной инфраструктур и т. д. С другой стороны, часть собранных исследователями данных оказались бы несущественными. Таким образом, при том же наборе сведений инопланетным ученым удалось бы увидеть общую картину.

Нет никаких сомнений в том, что подобного рода открытие позволит и нам разобраться в работе мозга и составить в цельную картину все имеющиеся о нем данные.

К сожалению, не все верят в то, что мы сможем познать работу мозга. Поразительное число людей, включая ряд нейробиологов, полагают, что мозг и интеллект непознаваемы по сути. А некоторые уверены, что даже если мы поймем глубинные процессы, то все равно не сумеем создать машины, которые работали бы по принципу мозга, поскольку интеллект невозможен без человеческого организма, нейронов и непостижимых законов физики.

Выслушивая приведенные аргументы, я представляю «интеллектуалов» прошлого, которые возражали против исследований (скажем, пытались запретить анатомам вскрывать трупы с целью изучения того, как устроен человеческий организм). «Даже не пытайтесь познать это, все равно не получится ничего хорошего, и если даже вы разберетесь в том, как оно работает, это не принесет вам никакой пользы» – подобные увещания привели к развитию *функционализма* – философского направления, последней нашей остановки в этой короткой истории размышлений о мышлении.

Согласно положениям функционализма, наличие интеллекта или обладание разумом –

исключительно свойство организации, не имеющее ничего общего с составными элементами. Разум присущ любой системе, составные части которой, будь то нейроны, кремниевые чипы или что-то еще, взаимодействуют друг с другом.

По правде говоря, подобный взгляд заслуживает того, чтобы его придерживался какой бы то ни было разработчик разумных машин. Только подумайте! Стала бы шахматная игра менее реальной оттого, что утерянную фигуру коня заменили бы солонкой? Очевидно, нет. Функционально солонка полностью соответствует настоящему «коню», поскольку ее можно переставлять по клеточкам доски согласно тем же правилам, что и «коня». Таким образом, шахматная партия остается шахматной партией, а не чем-то иным. Подумайте также, изменился бы смысл этого предложения, если бы я прошелся по нему курсором, удаляя, а затем восстанавливая каждый символ.

Приведу еще один пример. Известно ли вам, что каждые несколько лет ваше тело полностью обновляется, вплоть до последнего атома. Несмотря на это, вы остаетесь собой, а роль атомов не меняется, хотя меняются они сами. То же самое справедливо и в отношении головного мозга: если бы какой-то сумасшедший ученый вздумал заменить каждый из ваших нейронов функционально эквивалентной микросхемой, вы остались бы сами собой и даже не ощутили бы, что нечто в вас изменилось.

Согласно изложенному принципу, искусственная система, имеющая ту же архитектуру, что и биологическая (например, построенная по образу и подобию живого мозга), будет по-настоящему разумной. Защитники идеи искусственного интеллекта, коннекционисты и я сам являемся функционалистами, поскольку все мы уверены в том, что интеллект обеспечивает человеку нечто отнюдь не мистическое в головном мозге.

Все мы уверены в том, что научимся когда-то создавать разумные машины, возможно, не очень скоро.

Но существуют разные определения функционализма. Наряду с тем, что я только что представил вам как роковую ошибку, приведшую к провалу идеи искусственного интеллекта и парадигмы коннекционистов, – ошибка, основанная на последовательности «входящий-исходящий сигнал», – следует назвать еще несколько причин, по которым мы на сегодняшний день не способны создавать разумные машины. Пока сторонники идеи искусственного интеллекта придерживались подхода, определенно обреченного на провал, коннекционисты, на мой взгляд, преимущественно оставались в тени. Ученые, занятые в сфере искусственного интеллекта, спрашивают: «Почему мы, инженеры, должны ограничивать себя решениями, которые случайно возникли в ходе эволюции?!» В некотором роде они правы. Биологические системы, такие как головной мозг и гены, неизящны. Общеизвестная метафора – механизм Руба Гольдберга, названный в честь выдающегося американского карикатуриста (инженера по образованию), лауреата Пулитцеровской премии Рэувена (Руба) Люциуса Гольдберга. В машинах Гольдберга (разумеется, вымышленных) соединенные причудливым образом десятки блоков, рычагов, белок в колесе, электрических вентиляторов, противовесов, воздушных шариков и клеток с канарейками взаимодействуют сложным образом, чтобы, например, заточить карандаш или надеть шляпу на изобретательного хозяина.

Существуют, однако, энтузиасты, которые строят, так сказать, настоящие машины Гольдберга в металле, и последние потом выставляются на всевозможных конкурсах и выставках.

Среди инженеров-разработчиков программного обеспечения распространен термин *клюдж*, которым обозначают программы, созданные случайным образом, содержащие массу неточностей, бесполезных, усложняющих аспектов, часто непонятных самим программистам, создавшим их. Ученые, работающие в сфере искусственного интеллекта, опасаются того, что мозг представляет собой такой клудж, которому уже несколько миллионов лет и который полон рудиментов. Если так, рассуждают они, то почему бы не отбросить всю ерунду и не начать сызнова?!

Многие философы и психологи, придерживающиеся когнитивной традиции, соглашаются с таким видением. Они принимают метафору, описывающую разум как программное обеспечение, «установленное» в головной мозг – органический эквивалент компьютерного «железа». В компьютерах «железо» и программное обеспечение существуют по-рознь. Одна и

та же программа может быть установлена на любой универсальной машине Тьюринга. В ваших силах установить WordPerfect на ПК, или на Macintosh, или на Cray supercomputer (несмотря на то, что перечислены три принципиально разные конфигурации). И конфигурация не имеет для вас значения, если вы осваиваете WordPerfect. Согласно нашей аналогии, мозг не может научить нас тому, что такое интеллект.

Защитники идеи искусственного интеллекта любят приводить в качестве примера исторические факты, когда научные открытия радикально отличались от биологических аналогов.

Например, как нам удалось создать летательные аппараты? Не путем ли имитации летательных движений крыльев птиц? Отнюдь. Крылья самолетов неподвижны, а летит он потому, что у него есть двигатель. Рукотворная конструкция, совсем не похожая на природную, работает – причем гораздо лучше, чем последняя.

И созданные нами наземные средства передвижения, способные обогнать гепарда, имеют не четыре конечности, а колеса. Несмотря на то что в процессе эволюции не было создано ничего, хотя бы отдаленно напоминающего колесо, последнее, тем не менее, является революционным изобретением, позволяющим великолепно передвигаться по ровным поверхностям.

Философы порой обращаются к метафоре «когнитивного колеса», отражающей возможное решение проблемы создания искусственного интеллекта, вероятно, принципиально отличное от того, как функционирует мозг. Другими словами, программа, благодаря которой машина воспроизводит (или даже превосходит) человеческое исполнение какой-либо задачи в узком, но полезном аспекте, подобна мозгу.

Я убежден, что данная интерпретация, которую можно было бы назвать «Цель оправдывает средства», вводит теоретиков искусственного интеллекта в глубокое заблуждение.

Как показал эксперимент Сирла «Китайская комната», поведенческого соответствия недостаточно. Поскольку интеллект изначально присущ головному мозгу, нам следует заглянуть внутрь, чтобы понять, *что* представляет собой разум.

В наших исследованиях мозга, и особенно неокортекса, нам следует соблюдать осторожность в определении того, какие особенности являются всего лишь рудиментами прошлого, потому что даже в машинах Руба Гольдберга можно найти кое-что полезное. Но, как мы убедимся далее, здесь, возможно, присутствует изящество, превосходящее по мощности лучшие компьютеры, и оно ожидает того, что его обнаружат в нервной системе.

Коннекционисты интуитивно чувствовали, что мозг – не компьютер, и разгадки его секретов не надо искать в сообщении нейронов между собой. Это хорошее начало, но обычно дело дальше не шло. Хотя тысячи разработчиков бились над созданием трехслойных нейронных сетей и продолжают заниматься этим по сей день, нам далеко до разработки такой их модификации, которая в полной мере отражала бы работу мозга.

На протяжении пятидесяти лет мы старались поместить интеллект в компьютер. На сегодняшний день мы имеем текстовые процессоры Word, базы данных, видеоигры, мобильные телефоны. Но до создания по-настоящему разумных машин нам еще очень далеко. Чтобы достичь успеха, нам надо отталкиваться от устройства биологического «двигателя» интеллекта – неокортекса. Нам надо отделить разум от мозга. Это единственно возможный путь к успеху.

3. Мозг человека

В чем состоит принципиальное различие между человеческим мозгом и программируемым искусственным интеллектом, а также нейронными сетями? Как мы увидим в последующих главах, архитектура мозга может многое рассказать о его функционировании и о принципиальных отличиях от компьютера.

Предлагаю начать наш анализ с рассмотрения самого органа. Представьте, что перед нами на столе лежит мозг, и мы его разрезаем пополам. Первое, что вы отметите, – внешняя оболочка мозга выглядит очень однородной. Она напоминает головку цветной капусты, испещренную многочисленными извилинами и бороздами. Это кора головного мозга, в которой преобладает

неокортекс (около 90%) – новая кора, которая впервые появилась у млекопитающих. Тончайшая ткань нейронов обволакивает более старые филогенетические отделы головного мозга. Наше внимание будет приковано в первую очередь к неокортексу. Все функции, связанные с работой разума (восприятие, речь, воображение, способности к математике, искусству, музыке, прогнозированию и так далее) связаны с этим отделом головного мозга. Чтение данной этой книги – тоже отражение одного из процессов, происходящих в неокортексе.

Должен сознаться, что меня можно обвинить в шовинизме в отношении неокортекса. Отдавая себе отчет в том, что далеко не все разделяют мои взгляды, хочу обосновать свою позицию, пока мы еще не углубились в дебри. У каждой части человеческого мозга есть поклонники, посвящающие свое время ее изучению, поэтому гипотеза о том, что вся сущность разума может быть постигнута благодаря изучению только лишь неокортекса, будет встречена критикой со стороны инакомыслящих. Они возразят: «Вы не разберетесь в работе неокортекса, не включив в свой анализ некую зону X человеческого мозга, потому что они тесно взаимосвязаны, а зона X выполняет такие-то и такие-то функции». Не могу не согласиться. Безусловно, все отделы головного мозга важны для того, чтобы быть человеком. (Любопытно, что исключением из этого правила является часть мозга, которая включает наибольшее количество клеток, – мозжечок. Человек, у которого поврежден или от рождения отсутствует данный отдел мозга, все равно сможет вести практически полноценную жизнь. Большинство других зон мозга жизненно важны для человека.) Мой контраргумент – я же не собираюсь создавать людей! Моим намерением является понять разум и создать мыслящий компьютер. Существует огромная разница между тем, чтобы быть человеком и обладать интеллектом. У разумного компьютера нет пульса, мускулов, человекоподобного тела, сексуальных потребностей, он не чувствует голода и не испытывает эмоций. Человек – это намного больше, чем разумный компьютер. Мы являемся биологическими созданиями со всем необходимым, иногда нежелательным, багажом, собранным за многие годы эволюционного развития. Если вы захотите создать разумные механизмы, деятельность которых была бы неотличима от человеческой, способные без малейших трудностей пройти тест Тьюринга, тогда вам придется воссоздать и другие признаки, совокупность которых делает человека человеком. Если же ваша цель – создание мыслящей машины, не обязательно во всем подобной человеку, тогда имеет смысл сосредоточиться на изучении той части мозга, которая отвечает за разум.

Всем оскорбленным моей однобокой заинтересованностью неокортексом хочу сказать следующее: я с готовностью признаю важность других составляющих человеческого мозга (таких, как ствол, подкорковые узлы или миндалевидное тело), но настаиваю на том, что зоной действия интеллекта является именно неокортекс, а также две других значимых для моего исследования области мозга – таламус и гиппокамп (мы обсудим их немного позже). В долгосрочной перспективе нужно понять функциональную роль всех зон мозга. Однако начать следует с неокортекса, к чему мы и приступим.

Возьмите шесть визитных карточек (или игральных карт) и сложите их в стопку. (Лучше всего, если вы действительно выполните это задание, а не просто прочтете описание.) В руках у вас простейшая модель коры головного мозга. Толщина вашей стопки будет составлять около двух миллиметров и поможет вам составить представление о толщине слоев коры головного мозга. Как и стопка визиток или карт, кора головного мозга имеет толщину около двух миллиметров и состоит из шести слоев, каждый из которых имеет примерно такую же толщину, как одна визитная карточка или игральная карта.

Площадь коры головного мозга примерно составляет около 2200 см², что вдвое превышает площадь поверхности стандартной клавиатуры или соответствует размеру большой салфетки. У других млекопитающих кора головного мозга имеет гораздо меньшие размеры: у крысы она размером с почтовую марку, у обезьяны – как конверт. Независимо от размеров, общей чертой, характерной для коры головного мозга большинства млекопитающих, является шестислойное строение. У человека слои коры головного мозга ничуть не толще и не содержат каких-то особенных «разумных» клеток. Все дело в соотношении площадей неокортекса и тела. Это соотношение у человека гораздо выше, чем у других млекопитающих.

Кроме того, для размещения такого большого мозга в теле человека природа была

вынуждена модифицировать его общее анатомическое строение. Чтобы зрелый плод мог появиться на свет через родовые пути, у женщины в процессе эволюции изменилось строение таза. (Палеоантропологи считают, что это произошло также по причине перехода предков человека к прямохождению – передвижению на двух конечностях.) Однако этого было недостаточно, поэтому эволюция свернула неокортекс в черепной коробке, подобно тому, как лист бумаги можно скомкать и поместить в бокал.

Кора головного мозга образована нервными клетками, или нейронами. На площади, равной крошечному квадрату со стороной один миллиметр (вдвое меньше, чем буква *o* в этом тексте), содержится примерно сто тысяч нейронов. Подсчитать точное количество нервных клеток в коре головного мозга практически невозможно, тем не менее некоторые анатомы утверждают, что их число составляет порядка 30 млрд., хотя никто не удивился бы, если фактическая цифра оказалась намного меньшей или большей.

Стало быть, в вашей голове живут 30 млрд. нейронов, хранящих ваши знания, навыки, накопленный жизненный опыт. После 25 лет размышлений о мозге данный факт кажется мне не менее поразительным, чем раньше. Тончайшая пленка, состоящая из нервных клеток, видит, чувствует, творит наше мировоззрение. Это просто невероятно! Наслаждение теплотой летнего дня и смелые мечты о будущем – все создается этими клетками. Много лет спустя после публикации статьи в журнале *Scientific American* Фрэнсис Крик написал книгу под названием *Поразительная гипотеза*. Поразительная гипотеза состояла в том, что разум создается нейронами. Ничего другого не существует: никакой магии, никакого специального соуса, только нейроны, исполняющие информационный танец. Надеюсь, вы прониклись тем, насколько важным было это открытие. Существует большой философский канал, соединяющий нервные клетки и опыт нашего сознания, однако разум и мозг – это одного поля ягоды. Называя свой тезис гипотезой, Крик просто проявил политическую корректность. То, что клетки мозга – эти тридцать миллиардов нейронов – творят разум, совсем не предположение, а научный факт. Чтобы разобраться в том, как кора головного мозга формирует сознание, мы рассмотрим ее структуру более подробно.

Давайте вернемся к нашему секционному столу и еще раз взглянем на мозг в разрезе. Осматривая кору головного мозга невооруженным глазом, мы отметим, что она довольно однородна. Лишь большая продольная борозда разделяет два полушария головного мозга, а глубокая центральная борозда – его лобную долю и теменную. Других различных разграничительных линий не существует. Одинаков и цвет всех видимых невооруженным глазом отделов мозга.

Тем не менее людям давно известно, что такие границы все же существуют. Еще до того как нейрологи обнаружили наличие обратных связей в коре головного мозга, им было известно, что за разные психические функции отвечают определенные отделы головного мозга. Человек, у которого повреждено правое полушарие, может лишиться чувствительности левой половины своего тела или способности воспринимать внешнюю среду слева от себя. При поражении левой лобной доли, где расположен центр Брока, возникает афазия Брока, которая характеризуется невозможностью объединения отдельных речевых движений в единый речевой акт (при этом словарный запас и способность понимать слова остаются неизменными). Поражение веретенообразной борозды, пролегающей по нижней поверхности височной доли, может лишить человека способности распознавать лица – он не узнает свою мать, жену, детей и даже свое собственное лицо на фотографии. Наблюдая за такими трудно вообразимыми мозговыми расстройствами, нейрологи пришли к осознанию того, что кора головного мозга состоит из многих функциональных зон. Каждая зона полунезависима и специализируется на определенных аспектах восприятия или мышления. Возникает ассоциация с лоскутным одеялом, причем почти одинаковым у большинства людей. Функциональная организация головного мозга имеет форму отраслевой иерархии.

Понятие «иерархия» очень важное, поэтому я хочу обсудить его подробнее и дать ему четкое определение, на которое буду ссылаться в последующих главах книги. Любая иерархическая система характеризуется тем, что одни элементы расположены выше, а другие – ниже. В иерархии делового мира, например, менеджер среднего звена расположен выше клерка и ниже вице-президента компании. Иерархическое расположение и физическая позиция в

пространстве не тождественны: даже если кабинет менеджера находится этажом ниже той комнаты, в которой работает клерк, последний все равно стоит ниже в иерархической структуре компании. Я подчеркиваю это различие, чтобы в дальнейшем у вас не возникало сомнений, что я имею в виду, когда говорю, что какая-либо функциональная зона является высшей или низшей. Физическое местонахождение в структуре мозга в данном случае роли не играет. Все функциональные зоны головного мозга обитают в одной и той же ткани коры. Одна зона будет «выше» или «ниже» другой в зависимости от того, как они связаны и взаимодействуют друг с другом. Так, низшие зоны поставляют информацию в высшие через определенные нейронные каналы связи. В свою очередь, высшие зоны используют совершенно другие нейронные каналы связи для передачи обратных сигналов⁷. Кроме этого, еще существуют вторичные связи между отдельными областями иерархии – по аналогии с коммуникацией коллег одного уровня, но работающих в разных филиалах одной и той же компании. Двое ученых – Дэниэль Феллеман и Дэвид ван Эссен разработали схему коры головного мозга обезьяны, на которой изображены десятки областей, связанных между собой в сложную иерархию. Можно предположить, что похожая иерархия существует и в коре головного мозга человека.

Первичные сенсорные зоны, в которые непосредственно поступает информация об окружающем мире, являются низшими функциональными зонами. Эти области занимаются обработкой первичной информации на самом простом, базовом уровне. Например, зрительная информация поступает в кору головного мозга через первичную зрительную зону, назовем ее *V1*. Зона *V1* обеспечивает зрительные свойства низшего уровня, такие как восприятие мелких контурных сегментов, простых составляющих движения, бинокулярное рассеивание (для стереоизображения), основные цвета, информация о контрастности. Зона *V1* посылает информацию в зоны *V2*, *V4*, *IT* (о них мы расскажем позже), а также в другие зоны коры головного мозга. Каждая из этих зон отвечает за более узко специализированные или абстрактные аспекты восприятия визуальной информации. Например, нервные клетки, образующие зону *V4*, позволяют нам воспринимать объекты средней сложности, такие как формы звезд различных цветов. Зона *MT* специализируется на восприятии движущихся объектов. На более высоких иерархических уровнях расположены зоны, отвечающие за запоминание всевозможных визуальных объектов (людей, животных, предметов и т. д.) и ассоциативные связи между ними.

Похожая иерархическая структура существует и в других отделах мозга, ответственных за получение информации об окружающем мире по другим сенсорным каналам. Так, есть первичная слуховая зона *A1* и иерархия слуховых областей, расположенных над ней, а также соматосенсорная (ответственная за восприятие физических ощущений) зона *S1* и иерархия соматосенсорных областей, расположенных над ней.

Наконец, сенсорная информация поступает в *ассоциативные зоны* (данное название используют для описания тех областей коры головного мозга, которые получают и оценивают информацию, исходящую от разных рецепторов). Например, есть зоны, получающие информацию от органов зрения и осязания. Именно благодаря этим ассоциативным зонам вы понимаете связь между видом мухи, ползущей по вашей руке, и щекочущим чувством на коже руки. Большинство ассоциативных областей получают значительно переработанную информацию от нескольких органов чувств, но их функции до сих пор остаются невыясненными. Позже мы еще вернемся к иерархической структуре коры головного мозга.⁸

⁷ В отечественной науке выделяют такие зоны коры: первичные (проекционные), вторичные, третичные (интегративные, ассоциативные). – Примеч. ред.

⁸ Считается, что первичные зоны коры головного мозга обеспечивают элементарные ощущения; вторичные – целостные образы (люди, бабочки, звуки и т. д.); третичные (называемые еще ассоциативной корой) обеспечивают совместную работу различных анализаторов и формирование знаков и символов; при раздражении третичных зон у человека во время нейрохирургических операций на мозге могут возникнуть сценородные галлюцинации. (Например, человек может «видеть» и «слышать», как смеются его друзья, находящиеся на большом расстоянии от места операции. При этом, больной сам удивляется происходящему, так как осознает, где находится.)

Такая иерархичность организации зон является продуктом долгой эволюции животного мира. Например, у ежа и крысы первичная и вторичная зоны почти не дифференцированы, а третичных зон нет вообще. У человека,

В лобных долях коры головного мозга расположены моторные зоны, также имеющие иерархическое строение. Низшая моторная зона *M1* посылает сигналы в спинной мозг и непосредственно управляет мускулами. Высшие зоны осуществляют обратную связь, посылая сложные моторные команды в зону *M1*. Иерархия моторных зон и иерархия сенсорных зон удивительно похожи, словно построены по одной и той же модели. В моторной области информация поступает из зон высшего порядка в зону *M1* низшего порядка приводит мускулы в движение; в сенсорных областях органы чувств посылают информационные сигналы вверх по иерархической лестнице. В действительности же информационные сигналы передаются в обоих направлениях. То, что является обратной связью для сенсорных областей, одновременно является выходной информацией для моторных областей, и наоборот.

Большинство схематических изображений мозга представляют собой упрощенные отображения потоков информации и иерархических отношений зон коры головного мозга. Согласно таким описаниям, сенсорные сигналы из органов чувств (зрение, слух, осязание) поступают в первичные сенсорные зоны и по мере передвижения вверх по иерархии подвергаются обработке, потом проходят через ассоциативные зоны, поступают в лобные доли коры головного мозга и, видоизмененные, передаются назад в моторные зоны. Я не оспариваю такие представления. Действительно, когда вы читаете вслух, зрительная информация на самом деле поступает в зону *V1*, передается вверх к ассоциативным зонам, потом поступает в моторные зоны, расположенные в лобных долях коры головного мозга, и преобразуется в команду, заставляющую мышцы речевого аппарата сокращаться и воспроизводить звуки. Но не все так просто. Я хотел бы предостеречь вас от такого излишне упрощенного подхода, в котором сложные процессы передачи информации считаются односторонними, как будто сигналы всегда передаются в одном и том же направлении, подобно деталям автомобиля при сборке на конвейере. Я утверждаю, что информационные сигналы в коре головного мозга одновременно передаются и в нисходящем направлении, причем обратные информационные потоки, поступающие от зон высшего порядка к низшим, имеют большую информационную насыщенность. Возвращаясь к примеру с чтением вслух, высшие зоны коры головного мозга *посылают* к первичным зонам зрительного восприятия намного больше информации, чем *получают* ее от взгляда, бегущего по строкам книги! Ниже мы еще коснемся вопроса обратной передачи информации. А теперь – внимание: хотя иерархия строения коры головного мозга действительно существует, не стоит считать, что информационные потоки всегда движутся одними и теми же путями.

Вернемся к нашему воображаемому секционному столу. Допустим, у нас есть очень мощный микроскоп. Мы сделали небольшой срез коры головного мозга, нанесли краску на несколько нервных клеток и рассматриваем их под микроскопом. Если бы мы окрасили все нейроны, то увидели бы однородную черную массу, поскольку клетки очень плотно прилегают друг к другу. Но, окрасив лишь небольшую их часть, мы сможем увидеть шестислойную структуру, о которой упоминалось выше. Слои различаются по типам и плотности составляющих их нейронов, а также по характеру связей между ними.

Рассмотрим строение нейрона. Любая нервная клетка состоит из тела клетки, или сомы, и двух типов внешних древоподобных ветвей: *аксона* («передатчика») и *дендритов* («приемников»). Тело клетки включает ядро, которое содержит информацию о наследственных свойствах, и плазму, обладающую молекулярными средствами для производства необходимых нейрону материалов. Нейрон получает сигналы (импульсы) от других нейронов через дендриты и передает сигналы, сгенерированные телом клетки, вдоль аксона, который в конце разветвляется на волокна. На окончаниях этих волокон находятся *синапсы*. Синапсы (от греч. *synapsis* – соединение, связь) – это специализированные функциональные контакты между возбудимыми клетками, служащие для передачи и преобразования сигналов.

Нервный импульс – это процесс распространения возбуждения по аксону от тела клетки

напротив, первичные зоны занимают небольшое место, они как бы вытеснены вторичными, а третичные зоны мозговой теменно-височно-затылочной коры занимают подавляющую часть места и наиболее развиты. – Примеч. ред.

до окончания аксона. Некоторые аксоны имеют обратное действие, таким образом подавляя возбуждение клетки. Итак, по функциональному значению синапсы могут быть возбуждающими и тормозящими – в зависимости от того, активируют они или подавляют деятельность соответствующей клетки. В зависимости от поведения двух клеток сила синапса может изменяться. Наиболее простая форма синаптического обмена имеет место, когда два нейрона создают возбуждение почти одновременно, а сила взаимодействия между ними возрастает. Исследователь нейронных сетей Дональд Хебб предположил, что синаптическая связь, соединяющая два нейрона, будет усиливаться, если в процессе обучения нейронной сети оба нейрона согласованно испытывают возбуждение либо торможение. Простой алгоритм, реализующий такой механизм обучения, получил название *правила Хебба*, к которому мы вернемся позже. Помимо переменной силы синапса, существуют также доказательства того, что в результате взаимодействия двух нейронов могут возникнуть совершенно новые синапсы. Возможно, данный процесс происходит постоянно. Научные доказательства такого факта носят противоречивый характер. Независимо от того как изменяется сила синапсов, с уверенностью можно утверждать, что формирование и усиление синапсов – это то, от чего зависит процесс запоминания.

В коре головного мозга существует много типов нейронов, но 80% из них являются пирамидальными. Они называются так потому, что тела их клеток напоминают форму пирамид. За исключением верхнего слоя шестислойной коры головного мозга, которая состоит из километров аксонов, но очень небольшого количества клеток, каждый последующий слой состоит из пирамидальных клеток. Каждый пирамидальный нейрон связан со многими соседними нейронами, и каждый из них посылает длинный аксон к более отдаленным зонам коры головного мозга или к низшим зонам мозга, таким, например, как таламус.

Типичная пирамидальная клетка имеет несколько, тысяч синапсов. Из-за исключительно высокой плотности и малых размеров очень трудно установить их точное количество. Причем это количество различно для разных клеток, слоев и зон мозга. Если бы мы заняли консервативную позицию, утверждающую, что средняя пирамидальная клетка состоит из тысячи синапсов (на самом деле это число ближе к 5 или 10 тысячам), тогда кора головного мозга в общей сложности состояла бы из 30 миллиардов синапсов. Это астрономически большое число, которое невозможно охватить человеческим воображением. Такого количества синапсов было бы вполне достаточно, чтобы запомнить все, чему мы учимся на протяжении своей жизни.

Поговаривают, будто Альберт Эйнштейн утверждал, что открыть теорию относительности было очень просто. Ее можно было сформулировать даже из одного наблюдения: скорость света является постоянной для всех наблюдателей, даже если они передвигаются с различной скоростью. На первый взгляд, это идет вразрез с очевидным. Это настолько же нелепо, как и утверждать, что скорость брошенного мяча всегда одинакова независимо от силы броска или того, насколько быстро передвигаются игроки или наблюдатели. Разве по отношению к каждому из них мяч передвигается с одной и той же скоростью, независимо от обстоятельств? На первый взгляд, такое кажется невозможным. Тем не менее для света это был достоверный факт. И тут Эйнштейн спросил себя о том, каковы последствия столь удивительного факта. Он методически размышлял о последствиях неизменной скорости света и пришел к еще более неожиданным расчетам особенной относительности. Время начинает идти медленнее, если вы двигаетесь быстрее, в то время как масса и энергия остаются неизменными. Книжки о принципе относительности объясняли, как он работает в повседневной жизни. Теория сама по себе не была сложной, но она шла вразрез с интуицией.

Подобное открытие имело место и в нейрологии, касалось оно коры головного мозга. Многие нейрологи отказывались признавать его только потому, что не понимали, какую пользу из него можно извлечь. И все же это никоим образом не умаляет научную ценность данного поразительного факта. А если внимательно и методично изучить последствия данного открытия, то мы значительно продвинемся в познании свойств и особенностей работы коры головного мозга. Итак, о чем же речь?

Выдающийся ученый Вернон Маунткэстл – нейрофизик, работавший в Университете

Джона Хопкинса, Балтимор, – в 1978 году опубликовал работу под названием *Организирующий принцип мозговой функции*. В своей публикации автор подчеркнул, что кора головного мозга очень однородна по внешнему виду и строению. Зоны коры, отвечающие за зрительное восприятие, внешне не отличаются от зон, отвечающих за осязание, зон управления двигательной активностью, речевых зон и т. д. Все они выглядят почти одинаково. А если эти зоны выглядят одинаково, значит, заключил Маунткастл, вполне возможно, что они выполняют одну и ту же операцию! Он предположил, что при выполнении разных функций кора головного мозга использует один и тот же инструмент.

Все анатомы в то время, точно так же как и десятки лет до Маунткастла, знали, что кора головного мозга однородна. Но, вместо того чтобы задуматься над смыслом этой однородности, они тратили время и силы на поиск различий между зонами коры головного мозга. Они полагали, что если одна зона отвечает за зрительное восприятие, а другая – за слух, то между ними непременно должны существовать различия. Действительно, определенные различия можно заметить. Зоны коры головного мозга отличаются по толщине, плотности нервных клеток, относительной пропорции разных типов клеток, длине горизонтальных связей, плотности синапсов и многим другим параметрам, которые довольно сложно исследовать. Один из слоев первичной зрительной зоны V1 – из числа наиболее изученных зон мозга – как оказалось (к радости ученых), действительно имеет деления. Ситуация очень напоминала работу биологов XIX века, которые посвящали свое время изучению малейших различий между видами. Например, большим успехом они считали открытие того, что две внешне совершенно одинаковые мыши на самом деле принадлежали к двум различным видам. Много лет такой тропой следовал и Дарвин, изучая моллюсков. Но Дарвин проявил неординарную проницательность, поставив вопрос о причине поразительной общности всех видов. Общность видов была для него куда более интересной и невероятной, чем их различия.

Так произошло и с Маунткастлом. В то время как именитые анатомы бились над поиском ничтожных различий между зонами коры головного мозга, он обратил внимание именно на ее однородность. Одни и те же слои, одни и те же типы клеток, одни и те же связи существуют везде. Различия, в свою очередь, настолько неуловимы, что профессиональные анатомы не могут прийти к единому мнению в отношении того, в чем они заключаются.

Маунткастл заключил, что причина внешних отличий различных зон коры головного мозга заключается в том, с чем эта зона взаимодействует, а не в различии базовых функций. Таким образом, существует общая функция, общий алгоритм, выполняемые всеми зонами коры головного мозга. Механизм зрительного восприятия не отличается от слухового восприятия, а слуховое восприятие – от двигательной функции. По предположению Маунткастла, гены человека предопределяют характер связи между зонами коры головного мозга, что является очень специфическим параметром для разных видов, однако во всех случаях ткань коры головного мозга сама по себе выполняет одну и ту же функцию.

Давайте остановимся на этом подробнее. Зрение, слух и осязание кажутся мне совсем непохожими, а порой – принципиально различными. Зрение позволяет судить о цвете, составе, форме, глубине исследуемых объектов. Слух позволяет оценивать ритм, тембр, громкость звуков. В чем может заключаться их сходство? Но Маунткастл не настаивает на сходстве, он лишь утверждает, что кора головного мозга использует одни и те же принципы обработки слуховых, зрительных и двигательных сигналов.

Большинство ученых и инженеров оставили без внимания теорию Маунткастла. Пытаясь создать компьютер, который смог бы «видеть», кибернетики использовали терминологию, характеризующую визуальное восприятие, и старались воссоздать его механизм. Они размышляли в терминах форм и трехмерных изображений. Те же из них, кто намеревался разработать говорящий компьютер, создавали алгоритмы, использующие правила синтаксиса, грамматики и семантики. Если бы они прислушались к Маунткастлу, то поняли бы, что такие подходы лишь уведут прочь от создания искусственного интеллекта, а потому обречены на провал. Следуя логике этого непризнанного гения, алгоритм коры головного мозга един, независимо от того, из какого органа чувств обрабатывается информация или о каком типе восприятия идет речь. Чтобы человек мог слышать и видеть, мозг следует одному и тому же алгоритму, выполняет универсальную функцию.

Впервые наткнувшись на публикацию Манткастла, я был поражен. Да это же ключ к разгадке нейробиологии – одна работа, единая теория давала ключ ко всем самым невероятным загадкам человеческого мозга! В один миг я понял всю тщетность предыдущих попыток понять и воспроизвести поведение человека как набор различных операций. Надеюсь, вы сможете по достоинству оценить, насколько революционным и в то же время изысканным был подход, предложенный Маунткастлом. Впрочем, это не удивительно, ведь все лучшие научные открытия были просты, изысканны и неожиданны. Подход Маунткастла как нельзя лучше доказал последнее, и я считаю его революционным в области нейробиологии. После сказанного мною отказ большинства ученых-кибернетиков поверить в модель Маунткастла выглядит не просто досадным упущением, а преступлением против науки.

* * * * *

Игнорирование разработчиками искусственного интеллекта гипотезы Маунткастла отчасти объяснимо нехваткой соответствующего инструментария, который позволил бы изучить передвижение потоков информации в шестислойной коре головного мозга. Доступные ученым методы позволяют локализовать зоны, в которых сосредоточены те или иные способности человека, но не дают возможности досконально исследовать их механизмы. Например, в наши дни научно-популярные публикации в области нейробиологии косвенно пропагандируют мысль о том, что мозг является совокупностью высокоспециализированных модулей. Технологии функциональной визуализации, наподобие магнитно-резонансной томографии (МРТ), и позитронное эмиссионное томографическое сканирование (ПЭТ-сканирование) позволяют сосредоточиться исключительно на топографических картах мозга и функциональных областях, которые мы рассматривали раньше. Как правило, подобные экспериментальные исследования проводятся в положении пациента лежа, так, что его голова находится под сканирующим устройством. Пациент выполняет мыслительное задание (например, спряжение глаголов, повторение предложений, вызывание образов, запоминание списков, решение логических задач) или осуществляет заданное движение. Диагностическое устройство фиксирует активность различных зон мозга при выполнении того или иного задания и наносит цветные мазки на изображение («карту») человеческого мозга. На основании результатов тысяч экспериментов по функциональной визуализации ученые-физиологи установили, какие зоны мозга выполняют те или иные функции. Им известно местонахождение областей распознавания человеческих лиц, осуществления математических операций или восприятия музыки и многих других. Казалось бы, вполне естественно предположить, что различные функции мозг выполняет разными способами.

Но существуют неоспоримые доказательства правоты Маунткастла, подтверждающие невероятную гибкость коры головного мозга. Мозг любого человека, если ему обеспечить правильное питание и соответствующие условия, способен освоить любой из тысяч разговорных языков, язык знаков, язык музыки, язык математики, языки программирования и язык тела. Мозг может приспособиться к условиям арктической мерзлоты и испепеляющей жары пустыни. Благодаря мозгу человек становится экспертом в шахматах, рыбной ловле, сельском хозяйстве, физике и других областях деятельности. Особая зрительная зона коры головного мозга специализируется на считывании визуальных символов – прежде всего, букв и цифр. Означает ли последнее, что у человека уже при рождении присутствует языковая зона для обработки букв и цифр? Вряд ли. Письменный язык – это очень молодое приобретение человека как вида, слишком молодое для того, чтобы соответствующий механизм его обработки мог передаваться генетически. Таким образом, кора головного мозга сама себя подразделяет на специфические функциональные зоны, руководствуясь исключительно опытом человека. Человеческий мозг обладает редкостным даром – способностью к обучению и адаптации к изменениям. Это признак очень гибкой системы, а не такой, у которой есть тысяча решений на тысячу задач.

Нейрологи также отмечают, что соединения в коре головного мозга очень пластичны, т. е. они могут видоизменяться в зависимости от входных информационных сигналов. Например, мозг новорожденного хорька можно путем хирургического вмешательства изменить так, что

глаза зверька будут посылать сигналы в зону коры головного мозга, отвечающую за слуховое восприятие. Результаты просто удивительны: у хорьков развиваются каналы визуальной передачи в зоне слухового восприятия. Подобные эксперименты проводились также с другими органами восприятия и зонами мозга. Например, частички зрительной коры головного мозга у крысы после ее рождения трансплантируют в зоны коры головного мозга, отвечающие за осязание. По мере взросления крысы трансплантированные ткани занимаются переработкой осязательных сигналов в большей степени, чем визуальных. Нервные клетки универсальны, они не специализируются на зрительном восприятии, осязании или же слуховом восприятии.

Кора головного мозга человека в высшей степени податлива. Мозг взрослых людей с врожденной глухотой перерабатывает зрительную информацию в зонах, которые в обычных условиях отвечают за слуховое восприятие. Слепые от рождения люди используют большую часть зоны, которая у зрячих отвечает за зрительное восприятие, в процессе освоения и использования шрифта Брайля. Но ведь шрифт Брайля в первую очередь связан с осязанием, и закономерно было бы ожидать, что его освоение прежде всего должно активизировать зоны, отвечающие за осязание! Однако, судя по всему, такого, чтобы какая-либо из зон коры головного мозга ни за что не отвечала, не бывает. Когда, вопреки начальным ожиданиям, часть коры головного мозга, отвечающая за зрительное восприятие, не получает зрительных сигналов, она начинает искать другие информационные сигналы, хотя бы и предназначенные для других зон.

Все приведенные примеры показывают, что зоны мозга развивают специализированные функции в зависимости от типа поступающей к ним информации. Кора головного мозга не является жесткой структурой, части которой предназначены для выполнения разных функций согласно разным алгоритмам, подобно тому, как разделение поверхности Земли на страны не было «запрограммировано» в ходе эволюции. Организация вашей коры головного мозга, как и политическая карта мира, могла бы стать совсем иной при других обстоятельствах.

Гены предопределяют общее строение коры головного мозга, включая характер взаимосвязи различных зон, но ее внутреннее строение отличается необычайной гибкостью.

Маунткастл был прав – каждая зона коры головного мозга следует единому мощному алгоритму. Объединенные в соответствующую иерархическую структуру и под воздействием потока информационных сигналов на входе, эти зоны исследуют окружающую среду. Механизм действия коры головного мозга может быть использован новаторским образом в рукотворном мозге, причем так, что искусственный интеллект станет реальностью.

С гипотезой Маунткастла связана тема, которая столь же неожиданна, как и сама гипотеза. Входные сигналы, поступающие в вашу кору головного мозга, по сути своей одинаковы. Возможно, вы полагаете, что ваши функции восприятия являются совершенно разными. В конце концов, ведь звуковое восприятие обусловлено чувствительностью к колебанию воздушных волн, зрительное восприятие – чувствительностью к свету, а осязание зрительное – чувствительностью к давлению на кожу. Звук (колебания воздушных волн) характеризуется продолжительностью, зрительные образы – своими визуальными свойствами и положением в пространстве, осязание – в первую очередь соматосенсорными ощущениями. Что может быть более несхожим, чем блеяние козы, вид яблока и ощущение мяча у вас в руках?

Давайте поразмыслим. Зрительная информация от внешнего мира передается в головной мозг посредством миллионов волокон в вашем зрительном нерве. Ненадолго задержавшись в таламусе, она поступает в первичную зону коры головного мозга, отвечающую за зрительное восприятие.

Представим эти потоки сигналов в виде скопления электропроводов или оптических волокон. Возможно, вам доводилось видеть лампы с оптическими волокнами, где кончики каждого волокна светятся разными цветами. В нашем случае волокнам соответствуют аксоны, передающие к телу нейрона нервные импульсы. Как только сигналы, поступающие в головной мозг от различных органов чувств, превращаются в нервные импульсы, последние сразу становятся равноценными.

Например, когда вы смотрите на собаку, набор сигналов поступает по волокнам зрительного нерва в зрительную зону коры головного мозга. Когда вы слышите лай собаки,

другой набор сигналов посредством слухового нерва поступает в слуховые зоны коры головного мозга. Когда вы гладите собаку, набор осязательных сигналов проходит от вашей руки по волокнам спинного мозга и поступает в зоны коры головного мозга, отвечающие за осязание. Каждый набор сигналов – увидеть собаку, услышать собаку, почувствовать собаку – воспринимается по-разному, потому что передается в соответствующие зоны коры головного мозга по разным каналам. Однако на абстрактном уровне все сенсорные сигналы одинаковы, потому что перерабатываются шестислойной корой головного мозга по одному и тому же алгоритму. Вы слышите звук, видите изображение, чувствуете давление, но для вашего мозга не существует принципиальных различий между типами подачи информации. Нервный импульс – это нервный импульс. Он одинаков независимо от того, *что* послужило его причиной. Все, чем оперирует ваш мозг, – это сигналы.

На этих сигналах и основано ваше восприятие и знание мира. В вашей голове нет света. Там темно. Звук не попадает в ваш мозг. Там тихо. В действительности мозг – это единственная часть вашего тела, не обладающая чувствительностью. В нем нет болевых рецепторов. Хирург мог бы засунуть палец вам в мозг, и вы бы даже не почувствовали этого. Вся информация, поступающая в ваш мозг, попадает туда в виде пространственных или временных сигналов и передается по аксонам.

Что я подразумеваю под пространственными и временными сигналами? Рассмотрим поочередно каждое из наших основных ощущений. Зрительное восприятие включает как пространственную, так и временную информацию. Пространственные сигналы – это сигналы, совпадающие по времени. Они возникают, когда несколько рецепторов одного органа чувств стимулируются одновременно. Для зрительного восприятия органом чувств является сетчатка глаза, в которой формируется пространственный сигнал. Через зрительный нерв и проводящие пути центральной нервной системы этот сигнал передается в зрительные центры головного мозга, где возникает зрительное ощущение. Люди склонны считать, что в зрительные зоны попадает немного искаженная картинка мира, но на самом деле это совсем не так. Нет никакой картинки и нет никакого изображения. Есть только электрические разряды, вспыхивающие в сигналах. Визуальные качества быстро исчезают, как только кора головного мозга начинает перерабатывать полученную информацию, передавая сигналы вверх-вниз от зоны к зоне, фильтруя и просеивая их.

Не следует забывать, что визуальные сигналы постоянно меняются. Если пространственный аспект зрительного восприятия интуитивно понятен, то разобраться в его временном аспекте немного труднее. Порядка трех раз в секунду ваши глаза делают быстрое скачкообразное движение – саккаду⁹. При каждой саккаде изображение на сетчатке ваших глаз меняется. В свою очередь, это означает, что сигналы, поступающие в ваш мозг, полностью меняются при каждом скачкообразном движении глаз. Даже когда вы полагаете, что недвижимо сидите и наблюдаете одну и ту же сцену, ваше тело пребывает в непрерывном движении, а внешняя среда вокруг вас тоже постоянно меняется. В вашем сознании существует впечатление, что мир неизменен, а уследить за людьми и объектами, находящимися вокруг вас, не так уж и сложно. Однако такое впечатление возникает только благодаря тому, что ваш мозг обладает поразительной способностью обрабатывать стремительный поток поступающих из органа зрения изображений, сигналы о которых практически никогда не повторяются. Естественное зрительное восприятие, передаваемое как поток сигналов в мозг, напоминает реку. Можно также провести аналогию с песней, но не картиной.

Скачкообразные движения, совершаемые глазами, мало интересуют исследователей зрительного восприятия. Зато весьма популярны эксперименты над животными, которых подвергают наркозу, а затем изучают особенности зрительного восприятия при фиксации на одной точке. Так устраняется переменная времени. Разумеется, исключение переменных положено в основу большинства научных исследований. Однако в описанном случае не принимается в расчет главный компонент зрительного восприятия, собственно то, из чего оно и состоит. Временной фактор должен занимать центральное место в нейробиологических

⁹ Саккады – очень быстрые движения глаз длительностью от 10 до 80 мсек. – Примеч. ред.

исследованиях зрительного восприятия.

Что касается способности слышать, то мы привыкли рассматривать ее во временных рамках. Очевидно, что любые звуки, будь то разговорная речь или музыка, меняются с течением времени. Так, песня «существует» определенный промежуток времени, но воспринимать ее как набор пространственных сигналов мы не привыкли. Этот пример является обратной ситуации со зрительным восприятием – акцент ставится на временной аспект. Но в способности слушать трудно найти пространственный компонент.

У звукового восприятия тоже есть пространственный компонент. Звуки превращаются в нервные импульсы, проходя через улитку, или переднюю часть ушного лабиринта. Крошечная светонепроницаемая спиралевидная улитка, помещенная в наиболее твердую кость человеческого тела, была открыта более полусотни лет тому назад венгерским физиком Джорджем ван Бекси. Создавая модели внутренней части уха, исследователь выявил, что разные типы воспринимаемых человеком звуков продуцируют колебания разных частей улитки. Высокочастотные звуки вызывают вибрацию твердой основы улитки, а низкочастотные – вибрацию в более гибкой и широкой наружной части. Звуки средней частоты вызывают вибрацию промежуточных сегментов улитки. Каждая часть улитки буквально усеяна нейронами, которые активизируются при колебаниях. В повседневной жизни, когда вы непрерывно подвергаетесь одновременному действию самых различных частот, улитка пребывает в постоянном движении. Таким образом, по всей длине улитки постоянно возникают сигналы стимуляции иного типа, а именно – пространственные. Каждую секунду в зрительный нерв поступает новый пространственный сигнал. Итак, мы снова убедились, что сенсорная информация передается пространственно-временными сигналами.

Мы не привыкли считать осязание временным феноменом, тем не менее и это сенсорное ощущение включает как пространственный, так и временной компонент. Чтобы убедиться в этом, проведите простой опыт. Попросите своего друга протянуть руку вперед (ладонью кверху) и закрыть глаза. Положите ему на ладонь какой-нибудь небольшой предмет (например, кольцо или ластик) и попросите опознать его, не шевеля при этом ни одной частью руки. Располагая лишь приблизительной информацией о массе и размере предмета, ваш друг, скорее всего, не сможет дать правильный ответ. Но предложите ему, не открывая глаз, провести пальцем по предмету, и он тотчас его опознает! Разрешив шевелить пальцами, вы добавили составляющую времени к сенсорному восприятию осязания. Чувствительность клетчатки сравнима с чувствительностью подушечек пальцев. Значит, осязание тоже подобно песне. Постоянно меняющиеся во времени сигналы обретают особое значение в случаях, когда не обойтись без тонкого осязания (например, при застегивании рубашки или открывании двери в потемках).

Детей учат тому, что у человека есть пять органов чувств: зрение, слух, осязание, обоняние и вкус. Но их подвидов намного больше. Зрение, например, состоит из трех компонентов – восприятия движения объектов, их цвета и контрастности. Осязание включает регистрацию корой головного мозга давления, температуры, вибрации и боли. У нас также имеется целая сенсорная система, контролирующая положение тела, – проприоцептивная система¹⁰. Чувствительные окончания органов проприоцептивной системы рассеяны повсеместно в составе мышечных пучков, в сухожилиях и суставных сумках. Эти рецепторы сигнализируют мозгу о положениях звеньев тела, о суставных углах, о напряжениях в тех или иных мышцах и т. д. У нас также есть вестибулярная система. Расположенная во внутренней части уха, она обеспечивает нам ощущение равновесия.

Есть ощущения более или менее яркие, неуловимые или очевидные, но все они передвигаются по аксонам нейронов и попадают в мозг в виде потоков временно-пространственных сигналов. Кора головного мозга не познает и не чувствует окружающий мир непосредственно. Единственное, чем она оперирует, – это сигналы, передвигающиеся по входным и выходным аксонам. Ваше видение мира и себя возникает на основе сигналов.

¹⁰ Проприоцепция буквально означает «восприятие себя». – Примеч. ред.

Нейробиологи, изучавшие закономерности формирования образа тела, выявили, что наше восприятие самих себя является куда более гибким, чем это может показаться на первый взгляд. Например, я вам дам маленькую кочергу, чтобы вы использовали ее для хватательных функций вместо руки. Вскоре вы почувствуете, что кочерга словно стала частью вашего тела. Мозг приспособился к новым входным сигналам тактильного ощущения. Кочерга фактически стала частью образа вашего тела.

Гипотеза о том, что сигналы, поступающие в головной мозг от разных органов чувств, по сути одинаковы, конечно, очень неожиданна. Чтобы обосновать ее, приведу еще несколько примеров. Первый эксперимент вы сможете провести в домашних условиях с помощью друга. Вам понадобятся картонный экран и искусственная рука. Для первого раза будет неплохо, если вы сможете достать резиновую искусственную руку вроде тех, которые продают на Хэллоуин. Если такой нет, тогда обведите контур своей руки на чистом листе бумаги. Свою настоящую руку вытяните на некоторое расстояние, причем так, чтобы она повторяла положение искусственной руки (кончики пальцев должны смотреть в одном и том же направлении, ладонь в обоих случаях повернута вверх либо вниз). Затем поставьте между ними картонный экран так, чтобы вы могли видеть только искусственную руку. Итак, вы смотрите на искусственную руку, а задача вашего друга состоит в том, чтобы одновременно поглаживать/прикасаться к обеим рукам в одинаковых местах. Некоторое время спустя зоны мозга, в которых встречаются зрительные и соматосенсорные сигналы, – ассоциативные зоны, о которых мы упоминали ранее, – придут в замешательство. Вы ощутите искусственную руку как свою собственную.

Другой, не менее захватывающий пример такой «сигнальной эквивалентности», называется сенсорным замещением. Это может полностью перевернуть жизнь людей, которые потеряли зрение в детском возрасте, и даже незрячих от рождения. Возможно, он поможет создать технологии нового машинного интерфейса для всех остальных, зрячих.

Понимая, что работа мозга по сути своей не что иное как передача сигналов, Пол Бачирита, профессор, специалист в области биомедицинских инженерных технологий университета штата Висконсин, задумался над тем, нельзя ли транслировать на новые участки коры головного мозга сигнал, допустим, с видеокамеры, чтобы человек, потерявший зрение, вновь начал видеть?

Теперь Бачирита знает ответ. Он воплощён в виде комплекта электроники, именуемой «Машинно-мозговой интерфейс», а конечным узлом комплекса является «языковой дисплей» (*Tongue Display Unit*, сокращенно *TDU*). Это устройство с десятками электродов, создающих слабое покалывание на языке сообразно яркости пикселей на картинке, снимаемой видеокамерой. При помощи «Языкового дисплея» незрячий человек учится «видеть» через ощущения языка.

Дисплей функционирует следующим образом.

На лоб испытуемому помещают небольшую видеокамеру. Зрительные образы попиксельно транслируются как точки давления на язык. Зрительная сцена, которая отображается сотней пикселей на телеэкране, превращается в сигналы сотен крошечных точек давления на языке человека. Мозг быстро учится правильно интерпретировать сигналы.

Эрик Вайенмайер, атлет мирового уровня, одним из первых надел «языковое устройство». Он ослеп в возрасте тринадцати лет и часто выступал с лекциями о том, что слепота не должна лишать человека полноценной жизни. В 2002 году он совершил восхождение на гору Эверест, став единственным незрячим альпинистом, взявшимся за подобное дело.

В 2003 году Вайенмайер испробовал устройство Бачириты и впервые за долгие годы увидел изображение. Он смог совершить перехват мяча, катающегося по полу, взять со стола бутылку лимонада, сыграть в игру «Камень-ножницы-бумага». Затем он спустился вниз по лестнице, прошел сквозь дверной проем, рассмотрел дверь и ее раму и заметил, что на ней висит какая-то табличка. Изображения, по сути бывшие ощущениями на языке, в скором времени стали восприниматься как пространственные.

Эти примеры еще раз показывают, что мозг очень пластичен, а входящая информация, которая поступает в него, является не чем иным, как сигналами.

Все приведенные примеры подтверждают, что на самом фундаментальном уровне работа мозга описывается языком сигналов. Независимо от того, насколько отличаются друг от друга

функции зон коры головного мозга, базовый алгоритм остается одним и тем же. Коре головного мозга безразлично, какой из органов чувств посылает сигналы. Для нее не имеет никакого значения, поступают сигналы от одного органа чувств или от четырех. Ничего не изменилось бы, воспринимай вы мир через звуковые, радарные или магнитные поля, если бы у вас были щупальца вместо рук, или даже если бы вы существовали не в трехмерном, а в четырехмерном пространстве.

Все перечисленное означает, что базовой основой интеллекта не являются сенсорные каналы взаимодействия или их комбинации. Элен Келлер не могла видеть и слышать, тем не менее она выучила язык и стала знаменитой писательницей, что не под силу большинству зрячих и хорошо слышащих людей¹¹. Невероятная гибкость мозга позволила Элен, лишенной двух из основных функций, обеспечивающих восприятие окружающего мира, жить такой же полноценной жизнью, как и обычные люди.

Именно эта поразительная гибкость человеческого мозга питает во мне очень большие надежды на создание искусственного интеллекта. Задумываясь о создании разумных компьютеров, я задаю себе вопрос: «А зачем, собственно, привязываться только к известным нам пяти чувствам?» Как только мы расшифруем алгоритм коры головного мозга и создадим науку сигналов, то сможем применять ее к любой системе, которую захотим наделить разумом. Причем одно из наиболее замечательных свойств заложенной в мозге обратной связи состоит в том, что нам не нужно быть исключительно умными, чтобы ее программировать. Подобно тому как слуховая зона хорька «переходит» в зрительную, как зрительная часть коры находит себе альтернативное применения у незрячих людей, система, работающая на основе алгоритма коры головного мозга, будет мыслящей независимо от того, сигналы какого типа буду в нее поступать. Конечно, нам нужно быть достаточно сведущими, чтобы задать параметры системы, а также чтобы обучить ее. Но во всем остальном миллиарды нейронов, дающие мозгу способность генерировать сложные творческие мысли, позаботятся сами о себе, точно так же как они делают это в организме человека. Когда я сижу в комнате вместе со своими друзьями, откуда я знаю, что они здесь или что они реальны? Мой мозг получает набор сигналов, подобных тем, которые получал в прошлом. Сигналы соотносятся со знакомыми мне людьми, их лицами, голосами, их привычным поведением и всевозможными фактами из их жизни. Я лишь считываю определенные комбинации сигналов.

Все наши знания – это модель, построенная на сигналах. Уверены ли мы, что мир существует? Наверняка, этот вопрос показался вам странным. Но им задавались некоторые авторы фантастических фильмов и книг. Речь не о том, что люди или объекты окружающего мира не существуют. Они существуют, но наша уверенность в существовании мира полностью построена на сигналах и том, как мы их интерпретируем. Иллюзорно лишь непосредственное восприятие. Как вы помните, мозг – это темный звуконепроницаемый ящик, которому неведомо ничего, кроме переменных входящих сигналов. Ваше восприятие мира создается исключительно на основе этих сигналов.

Тогда возникает вопрос о связи галлюцинаций и реальности. Если у вас возникают галлюцинационные ощущения от искусственной руки и вы можете «видеть» через точечную стимуляцию языка, то, может, вас так же дурачат, когда вы чувствуете собственной рукой или видите своими глазами? Можем ли мы верить в реальность того мира, который ощущаем? Да. Мир действительно существует в абсолютной форме, очень приближенной к той, которую мы воспринимаем. Однако наш мозг не в состоянии познать этот мир непосредственно.

Ощущения, призванные сканировать какие-то срезы, отдельные аспекты абсолютного мира, сворачиваются в сигналы, поступающие в кору головного мозга, а в результате переработки стандартным алгоритмом они создают модель мира. И хотя письменный и разговорный язык воспринимаются органами чувств совершенно по-разному, на уровне неокортекса они работают почти одинаково. Точно так же модель мира Элен Келлер была очень близка к нашей с вами, хотя она располагала значительно меньшим набором ощущений.

¹¹ Элен Келлер с детства была слепой и глухонемой, но, несмотря на это, закончила университет и оставила после себя богатое литературное наследство. – Примеч. ред.

Благодаря сигналам кора головного мозга создает модель, очень близкую к реальному миру, а потом эту модель фиксирует в памяти. Запоминание – вот что происходит с сигналами после их попадания в кору головного мозга. Это и станет темой нашей следующей главы.

4. Память

Когда вы читаете книгу, идете по улице, заполненной прохожими, слушаете симфонию, утешаете плачущего ребенка, все органы чувств посылают в ваш головной мозг пространственно-временные сигналы. Окружающий мир – это океан постоянно меняющихся сигналов. Что же происходит с ними при попадании в кору головного мозга?

Еще на заре индустриальной революции ученые рассматривали мозг как своего рода биологический механизм, или программируемый компьютер. Разумеется, никто не сомневался в отсутствии гаек и шурупов в голове человека, но данная метафора представлялась большинству исследователей самой точной.

Как мы отмечали в главе 1, этой позиции придерживались и разработчики искусственного интеллекта. Считалось, что неспособность создать искусственный интеллект напрямую связана с тем, что компьютеры слишком малы и медленны по сравнению с человеческим мозгом. Современный компьютер может считаться эквивалентом мозга букашки, говорили изобретатели, а вот когда мы создадим более быстрые и объемные машины, они будут такими же умными, как человек.

Аналогия между живым мозгом и вычислительной машиной некорректна. По сравнению с транзисторами нейроны головного мозга действуют гораздо медленнее. Нейрон собирает из синапсов входные сигналы, объединяет их и принимает решение. Он либо генерирует импульс – потенциал действия, который будет восприниматься синапсами соседних нейронов, – либо воздерживается от этого. Обычный нейрон может осуществить подобную операцию и вернуться в исходное состояние за пять миллисекунд, т. е. 1/200 секунды. Современный компьютер на кремниевой основе может осуществлять миллиард операций в секунду. Значит, основная компьютерная операция осуществляется в 5 миллионов раз быстрее, чем базовая операция в человеческом мозге! Каким же образом, вопреки столь существенной разнице, мозг способен действовать быстрее, чем наиболее быстрые цифровые компьютеры? «Да это же очевидно, – скажут сторонники идеи создания вычислительных и управляющих систем по образу и подобию мозга, – мозг – это параллельный компьютер. У него миллиарды клеток, занимающихся одновременным вычислением. Этот параллельный процесс во много раз увеличивает мощность биологического мозга».

Я считаю данный аргумент заблуждением и готов опровергнуть его при помощи простого эксперимента, так называемого «правила ста шагов». Человек способен выполнять существенные задания за намного меньший промежуток времени, чем одна секунда. Например, я мог бы показать вам фотографию и спросить, есть ли на снимке кошка. Если вы увидите кошку, а не медведя или луковицу, то должны будете нажать кнопку. Для современного компьютера такая задача является очень сложной или вообще невыполнимой, а человек даст правильный ответ за полсекунды или даже быстрее. Но, поскольку нейроны малоподвижны, это значит, что за полсекунды информация, поступившая в ваш мозг, может пройти цепочку не более чем из сотни нейронов. То есть мозг «вычисляет» ответ за сто шагов или даже меньше, независимо от того, какое количество нейронов вовлечено в процесс передачи информации. С момента, когда ваши глаза фиксируют изображение, и до момента, когда вы нажимаете кнопку, сигналы прошли цепочку из ста нейронов. Цифровой компьютер для решения той же задачи использует около миллиарда шагов. Сотни шагов для компьютера будет недостаточно даже для того, чтобы передвинуть один знак на дисплее, не говоря о более сложных задачах.

Но разве миллионы одновременно работающих нейронов не похожи на параллельный компьютер? Не совсем. Принцип параллельности используется и мозгом, и компьютером, но это все, что их роднит. Параллельные компьютеры объединяют значительное количество быстродействующих компьютеров для выполнения сложных задач, таких как составление прогноза погоды. Чтобы предвидеть погоду, нужно просчитать физические условия во многих

частях планеты одновременно. Каждый компьютер может работать над одной задачей. Но даже если сотни или тысячи машин объединить в параллельную систему, отдельному компьютеру для обработки информации и выполнения своей задачи понадобятся миллиарды или миллионы шагов. Какой большой параллельный компьютер мы бы ни собрали, каким бы быстродействующим он ни был, он не сможет сделать ничего полезного за сто шагов.

Приведу пример: представьте, что я попросил вас перенести сотню каменных глыб через пустыню. За один раз вы можете перенести только один камень, а чтобы пересечь пустыню, вам нужно сделать миллион шагов. Понимая, что справиться с заданием в одиночку будет непросто, вы нанимаете сто рабочих для параллельного выполнения задачи. Дело станет продвигаться в сто раз быстрее, однако тот факт, что для пересечения пустыни нужно пройти миллион шагов, остался неизменным. Наем новых работников, даже тысячи, ничего изменить не может. Независимо от количества нанятых работников задание может быть выполнено только за миллион шагов. То же самое справедливо и для параллельных компьютеров. Добавление дополнительных процессоров не меняет дела. Независимо от количества и мощности процессоров и быстродействия компьютер не может решить сложную задачу за сто шагов.

Но как же мозгу удается решать сложные задачи за сто шагов, в то время как даже самый большой параллельный компьютер ничего подобного сделать не в состоянии за миллион или миллиард шагов? Дело в том, что мозг на самом деле не «вычисляет» решения задач, он извлекает их из памяти, представляющей, по сути, хранилище готовых решений. Соответственно, для того чтобы их извлечь, нужно всего несколько шагов. Медленных нейронов для этого более чем достаточно, ведь они сами и составляют память. Можно утверждать, что мозг, точнее, неокортекс, который является «интеллектуальной» частью мозга, – это единое запоминающее устройство, а вовсе не компьютер.

* * * * *

С вашего позволения, я продемонстрирую разницу между *вычислением* решения и *использованием памяти* для решения задачи. Пусть нашей задачей будет поймать мяч. Кто-то бросает мяч, и он летит прямо на вас. У вас есть меньше секунды, чтобы схватить его еще в воздухе. Задача не ахти какая сложная, но лишь до тех пор, пока вы не захотите запрограммировать на ее выполнение робота. Когда инженеры и разработчики программного обеспечения берутся за такое задание, они сначала пытаются вычислить траекторию полета мяча и определить его пространственное расположение в момент контакта с рукой. Вычисления требуют использования ряда уравнений, которые можно почерпнуть из курса физики высшей школы. После этого все сгибы руки робота нужно отрегулировать так, чтобы конечность могла занимать определенную позицию. Здание требует составления и решения математических уравнений другого типа, на сей раз более сложных, чем предыдущие. И наконец, всю операцию нужно повторить несколько раз, поскольку по мере приближения мяча робот получает более точную информацию о его положении в пространстве и траектории полета. Если робот начнет движение только после определения точной цели полета мяча, он не успеет его поймать. Робот должен начать движение с целью поймать мяч гораздо раньше, не обладая полной информацией, а потом постоянно корректировать свои действия по мере приближения мяча. И хотя для выполнения этой задачи можно запрограммировать компьютер, для успешного решения ему понадобится миллион шагов. Правило «ста шагов» подскажет нам, что мозг решает эту задачу иначе. Он обращается к памяти.

Как вы поймаете мяч, используя при этом память? В вашем мозге сохраняется память о моторных командах (точно так же, как и многих других типов поведения), необходимых для поимки мяча. Когда бросают мяч, происходят три события. Во-первых, вид мяча автоматически активизирует соответствующие воспоминания. Во-вторых, в памяти оживает временная последовательность моторных команд. В-третьих, вызванное воспоминание соотносится с обстоятельствами конкретной ситуации, т. е. приспособливается к фактической траектории мяча и положению вашего тела. Воспоминание о том, как нужно ловить мяч, не было запрограммировано в вашем мозге. Вы обучились этому в процессе жизни, соответственно информация не вычисляется, а сохраняется в нейронах коры вашего головного мозга.

«Минуточку! – возразите вы, – Каждый бросок отличается от другого. Как было только что сказано, вызванное воспоминание приспособливается к положению мяча при каждом броске. Разве это не требует решения уравнений – того, чего мы пытались избежать?» На первый взгляд, это действительно так, но природа решила проблему переменных обстоятельств иным, гораздо более совершенным способом.

В этой главе мы увидим, что кора головного мозга создает *инвариантные представления*, которые компенсируют изменчивость окружающего мира. Представьте, например, что вы садитесь на водную кровать. При этом произвольно перемещаются все расположенные на ней объекты – скажем, подушки и другие люди. Кровать не высчитывает, насколько высоко или низко нужно поднять тот или иной край; функция приспособливания осуществляется автоматически за счет физических свойств воды и поверхности матраца. В следующей главе мы увидим, что структура шестислойной коры головного мозга обладает приблизительно таким же свойством по отношению к поступающему в нее из внешнего мира потоку информации.

Итак, принципы действия неокортекса и компьютера различны. В место вычисления решений и программирования поведения кора головного мозга использует память.

У компьютеров тоже есть память в форме жестких накопителей и чипов памяти. Но четыре особенности памяти неокортекса коренным образом отличают ее от памяти компьютера:

- неокортекс запоминает последовательности элементов, а не отдельные элементы окружающего мира;
- неокортекс вспоминает последовательности автоассоциативно;
- неокортекс запоминает последовательности в инвариантной форме;
- неокортекс сохраняет последовательности иерархически.

Первые три принципа мы рассмотрим в данной главе. Концепция иерархичности коры головного мозга была представлена в главе 3, а в главе 6 мы более подробно остановимся на ее прикладном значении.

В следующий раз, рассказывая собеседнику о каком-нибудь происшествии, попробуйте абстрагироваться от происходящего и проанализировать последовательность вашего повествования. Вы не можете рассказать все, что случилось, одновременно, даже если ваша речь будет очень быстрой, а собеседник – внимательным. Вы выдаете порции информации последовательно, одну за другой. Это происходит не только потому, что разговорная речь сама по себе является серийной. Любой рассказ – будь то письменный, устный или рисованный в картинках – всегда носит серийный характер. Это объясняется тем, что память о событиях сохраняется в вашем мозге в серийной форме и может быть извлечена оттуда тоже в серийной форме. Вы не можете вспомнить все, случившееся с вами, одновременно. Собственно, невозможно думать о чем-то сложном вне рамок серии событий или мыслей.

Некоторые люди в разговорах подолгу не могут перейти к сути. Они топчутся на месте, нагромождая в своем рассказе скучные подробности и многочисленные отступления от темы. Если вам доводилось общаться с такими собеседниками, то вы наверняка вспомните, насколько раздражает подобная манера повествования. Вам так и хочется крикнуть: «Да скажите же наконец о главном!» Но человек продолжает в деталях описывать хронологию произошедших с ним событий. По-другому у него просто не выходит.

Приведу другой пример. Закройте глаза и представьте свой дом. Вообразите, что вы подходите ко входной двери. Постарайтесь рассмотреть ее во всех подробностях. Откройте ее. Войдите внутрь. Посмотрите налево. Что вы видите? Посмотрите направо. Затем пройдите в ванную комнату. Что вы видите на стене справа от вас? Слева? Что лежит в верхнем правом ящике туалетного столика? Какие еще вещи вы храните в ванной? Вы знаете ответы на эти и тысячи подобных вопросов, вы можете вспомнить их со всеми подробностями. Вся эта информация хранится в коре вашего головного мозга. Можно сказать, что ответы на перечисленные вопросы – часть воспоминания о вашем доме. Но вы не способны думать обо всем сразу. В вашей памяти эта информация действительно связана, но не существует способа, который бы позволил вам увидеть в своем воображении все одновременно. Чтобы вызвать из памяти полное и устойчивое воспоминание о своем доме, вам придется пройти через серию последовательных фрагментов в таком же порядке, как вы их переживали в действительности.

И таковы все воспоминания. Воспоминание об одном эпизоде (приближение к двери) активизирует следующий (войти в дверь), который, в свою очередь, повлечет за собой следующий (подняться или спустится по лестнице) и так далее. Все это – последовательности, которые вы проходили раньше. Конечно, приложив определенные усилия, можно сменить порядок описания своего дома – скажем, перепрыгнуть от описания входной двери к описанию второго этажа. Но, как только мне вздумается рассказать собеседнику, как выглядит моя комната или любой элемент интерьера, я возвращусь к последовательному способу описания. По-настоящему беспорядочных мыслей вообще не существует. Воспоминания всплывают в памяти путем ассоциирования.

Вы знаете алфавит. Но попробуйте-ка произнести его в обратном порядке и убедитесь, насколько это трудновыполнимая задача. Кроме того, вы на собственном опыте поймете, каково учить алфавит маленькому ребенку. В вашей памяти алфавит представлен как последовательность сигналов. Как и другие воспоминания, данное не является чем-то «законсервированным» в вашем мозге и не поддается моментальному извлечению или извлечению в произвольном порядке. Точно так же обстоит дело и с днями недели, месяцами в году, вашим номером телефона и множеством другой информации. Запоминание песни – еще один отличный пример временной последовательности. Припомните какой-нибудь известный вам мотив. Скажем, моя любимая песня – *Somewhere over the Rainbow*, а вы выберите другую мелодию по своему вкусу. Вы не сможете вспомнить песню всю сразу, только в последовательности. Вы начнете с начала или с припева, но будете придерживаться определенной последовательности. Вам не удастся припомнить слова в обратном порядке или охватить своим вниманием песню целиком. В реальности прослушивание песни занимало у вас определенный промежуток времени, и вы сможете вспомнить ее только в том порядке, в каком заучили.

То же самое справедливо и для низшего уровня сенсорных запоминаний. Возьмем, к примеру, тактильную память о разных типах поверхностей. В вашем мозге хранятся воспоминания об ощущениях, вызванных касанием к гравию, поглаживанием вельветовой ткани, нажиманием клавиш пианино. Эти запоминания основаны на последовательностях точно так же, как алфавит или песня. Различие в том, что более короткие последовательные ряды тактильных ощущений будут разбиваться на секундные промежутки вместо минутных. Если бы во время вашего сна я поместил бы вашу руку в ведро с гравием, то, проснувшись, вы бы легко определили, к чему прикасается ваша рука, пошевелив пальцами. Ваше воспоминание о тактильном ощущении при касании к гравию основано на последовательности сигналов о вибрации и давлении, воспринимаемых нейронами вашей кожи. Эти последовательности сигналов очень отличаются от тех, которые возникли бы у вас, будь ваша рука зарыта в песок или сухие листья. Скрип и перекатывание гальки, сопровождающее даже самые незначительные движения вашей руки, продуцируют контрольные последовательности сигналов, которые приводят в действие соответствующие воспоминание в соматосенсорной части неокортекса.

В следующий раз, выходя из душа, обратите внимание на то, как вы вытираетесь полотенцем. Однажды я сделал для себя открытие, что всякий раз этот процесс состоит из одних и тех же движений и положений тела. Наблюдая за своей женой (очень приятный эксперимент!), я заметил точно то же. Думаю, в вашем случае результат окажется сходным: вы отметите полужесткую последовательность движений и поворотов тела. Но попробуйте-ка ее изменить! Приложив некоторые усилия, вы сможете это проделать, но лишь при условии предельной сосредоточенности на своем занятии. Как только отвлечетесь, сразу почувствуете, что вернулись к старой модели поведения.

Все наши воспоминания хранятся в синаптических связях между нейронами. Если учесть, что неокортекс содержит огромное количество информации, но в каждый конкретный момент времени мы можем вспомнить лишь малую ее толику, то можно предположить, что любое воспоминание обеспечивается лишь ограниченным количеством нейронов и синаптических связей между ними. Когда вы начинаете вспоминать об интерьере своего дома, то сначала активизируется один набор нейронов, потом он приводит в действие следующий набор и так далее. Объем памяти неокортекса невероятно вместителен. Тем не менее в каждый

конкретный момент времени мы можем погрузиться лишь в несколько воспоминаний, вызываемых лишь как последовательность ассоциаций.

Выполните еще одно задание. Погрузитесь в прошлое и попытайтесь вспомнить места, где вы жили, в которых вам довелось побывать, людей, с которыми вы были знакомы. Оказывается, в процессе такого упражнения в памяти всплывают факты, казалось бы, давно позабытые. Нейронные синапсы хранят тысячи подробнейших воспоминаний, используемых очень редко или не используемых вовсе. В каждое отдельное мгновение мы извлекаем из памяти лишь крошечную часть того, *что* в ней хранится. Большинство же воспоминаний, образно говоря, сидят и бьют баклуши в ожидании сигнала.

Память компьютера не хранит последовательности сигналов. С помощью разных программных доработок можно этого достичь (как, например, в случае, сохранения и воспроизведения песни), но память компьютера не способна на автоматическое выполнение такого задания. В то же время запоминание последовательностей символов – это неотъемлемое свойство системы памяти коры головного мозга, данное ей от природы.

А сейчас предлагаю рассмотреть второе ключевое свойство человеческой памяти – ее автоассоциативную природу. В главе 2 мы выяснили, что значение данного термина подразумевает ассоциирование сигналов самих с собой. Автоассоциативная система способна воспроизвести полную последовательность на основе неполной или искаженной входной информации. Последнее касается как пространственных, так и временных сигналов. Так, заметив, что из-за шторы выглядывают ботинки вашего сына, вы автоматически представите его в полный рост. Вы создали полный образ из неполной информации о нем. Или представьте, что вы увидели человека, ожидающего автобус, но при этом большая часть его тела скрыта от вас кустом. Ваш мозг не растеряется. Хотя глаза увидят только часть целого, перед вами возникнет полный образ. Причем он будет настолько отчетливым и ярким, что вы вряд ли будете отдавать себе отчет в том, что это всего лишь предположение.

Мозг способен дополнять и временные образы. Припомните незначительную подробность из давно минувших событий, и вам на ум придет вся соответствующая последовательность. Марсель Пруст в известном собрании романов, объединенных названием «В поисках утраченного времени» открывает свой первый рассказ воспоминанием о том, как пахнет печенье «Мадлен», а затем разворачивает повествование на тысячу с лишним страниц. Общаясь в каком-либо шумном месте, мы зачастую не слышим все слова собеседника. Ничего страшного, ведь наш мозг сам дополнит то, *что* он ожидает услышать. Общеизвестно, что мы на самом деле не слышим полностью всех слов, которые воспринимаем. Привычку некоторых людей «читать мысли» собеседника и заканчивать чужие высказывания считают дурным тоном. Но наш мозг поступает подобным образом постоянно, причем, дополняя не только окончания наблюдаемых явлений, но и их начало и середину. Подобная склонность заниматься постоянным додумыванием фактов редко осознается человеком, однако она является фундаментальной характеристикой памяти. В любой момент времени часть может активизировать целое – в этом состоит суть автоассоциативных воспоминаний.

Неокортекс – это сложная биологическая автоассоциативная система памяти. В каждый момент времени каждая функциональная зона зорко бдит, не появились ли на входе знакомые элементы или их фрагменты. Появление в поле вашего зрения знакомого человека мгновенно прервет поток самых глубоких размышлений – ваши мысли переключатся на приятеля. Визуальный сигнал – появление знакомого – заставляет мозг включиться в процесс вспоминания других сигналов, ассоциируемых с ним. Это неизбежно. А после такого внезапного отвлечения внимания мы часто ломаем голову: «Так о чем же я думал?» Обеденный разговор с друзьями тоже протекает в ассоциативных рамках. Скажем, беседа начинается с обсуждения блюд. Вы отмечаете, что такой же салат ваша мать приготовила вам на свадьбу, и это наводит кого-то на воспоминания о его свадьбе, затем о медовом месяце и путешествии в экзотическую страну, о политических проблемах в той части мира и так далее. Мысли и воспоминания ассоциативно взаимосвязаны – случайностей здесь нет и быть не может. Входные сигналы, поступающие в мозг, автоассоциативны по отношению к самим себе, они дополняют новые модели и автоассоциативно связаны с тем, что произойдет дальше. Такая цепочка воспоминаний называется *мыслью*, и, хотя ее путь не предначертан, полного контроля над ней мы тоже не

имеем.

Теперь обратимся к третьей особенности памяти неокортекса – формированию инвариантных представлений. Здесь мы обсудим их общие черты, а в главе 6 поговорим о том, как мозг создает эти модели.

Память компьютера устроена так, чтобы сохранять информацию в максимально неизменном виде. Когда вы копируете программу с компакт-диска, то на жесткий диск вашего компьютера переносится каждый байт информации со 100-процентной точностью. Незначительное расхождение между оригиналом и копией может вызвать фатальный сбой в работе компьютерной программы. Память неокортекса действует иначе. Мозг не запоминает с абсолютной точностью все увиденное, услышанное или почувствованное, причем совершенно не из-за того, что деятельность коры головного мозга или нейронов беспорядочна, что чревато возникновением ошибок. Так происходит потому, что мозг запоминает важные взаимосвязи внешнего мира, а не привязывается к отдельным его элементам. Давайте рассмотрим несколько примеров.

Как упоминалось в главе 2, простейшие автоассоциативные модели создаются учеными уже несколько десятков лет, а мозг, как я только что объяснил, вызывает воспоминания автоассоциативным способом. Тем не менее существует большая разница между автоассоциативными воспоминаниями, реализуемыми в нейронных сетях, и воспоминаниями неокортекса. В искусственных автоассоциативных воспоминаниях не задействованы инвариантные представления, и поэтому в них возможны существенные упущения.

Я могу сохранять в искусственной автоассоциативной памяти множество снимков лиц, состоящих из черно-белых точек. Если вы предъявите лишь часть одного из изображений (например, глаза), то система дополнит недостающие части (такие эксперименты проводились неоднократно). Однако, если каждую точку картинки передвинуть на пять пикселей вправо, то машина не сумеет распознать лицо. Для искусственной автоассоциативной памяти это уже совершенно новая модель, потому что расположение пикселей в ней не совпадает с исходной моделью. А вы и я без труда опознали бы сдвинутое изображение как одно и то же лицо.

Искусственная автоассоциативная память не в состоянии распознать изображение, если оно было подвергнуто перемещению, ротации, изменению масштаба, какому угодно другому из тысяч возможных превращений. В отличие от нейронных сетей человеческий мозг с легкостью справляется с такой задачей. Так почему же мы воспринимаем факты, явления или обстоятельства как неизменные, если входные сигналы, их описывающие, поменялись или изменились? Давайте рассмотрим еще один пример.

Сейчас, скорее всего, вы держите в руках книгу. Если вы повернете книгу, измените освещение или положение своего тела на стуле, зафиксируете взгляд на разных частях книги, то световые сигналы, попадающие на сетчатку вашего глаза, полностью изменятся. Зрительные сигналы, которые вы получаете, различны в каждое отдельное мгновение. Входные сигналы, поступающие в ваш мозг, не повторились бы ни разу даже в том случае, если бы вы держали книгу в руках на протяжении ста лет. Однако вы ни на миг не усомнитесь, что держите в руках все ту же книгу. Внутренняя модель «этой книги», которой располагает ваш мозг, не изменяется даже в условиях информационной переменчивости. Такие внутренние репрезентации мозга описывает термин *инвариантная репрезентация*.

Другой пример: представьте лицо своего друга. Вы узнаете его каждый раз при новой встрече. Это происходит менее, чем за одну секунду, независимо от того, находитесь ли этот человек в двух шагах от вас, в пяти метрах или в противоположном конце комнаты. Когда он рядом, то его образ занимает большую часть сетчатки, а когда он рядом – лишь малую ее часть. Он может стоять к вам лицом, быть повернутым в профиль или под углом. Он может улыбнуться, прищуриться или зевнуть. Вы могли бы увидеть его при ярком дневном свете, в тени или в разноцветных бликах дискотеки. В каждом из перечисленных случаев доля света, попадающего на вашу сетчатку, будет не такой, как в остальных случаях, однако в каждом отдельном случае вы сразу же узнаете друга.

Давайте приподнимем занавес и посмотрим, что происходит в вашем мозге, что позволяет ему проявлять столь необыкновенные свойства. Экспериментально установлено, что модель деятельности нейронов в зрительной зоне V1 неокортекса, принимающей визуальные

стимулы, будет отличаться для каждого положения лица вашего друга. Всякий раз, когда лицо меняет свое положение или происходит новая фиксация ваших глаз, модель активности нейронов первичной зрительной зоны меняется в соответствии со сменой изображения на вашей сетчатке. Однако если мы проанализируем деятельность ваших клеток в зоне распознавания лиц – функциональной зоне, расположенной на порядок выше V1 в иерархии коры головного мозга, – то обнаружим полную стабильность. Значит, определенная часть нервных клеток зоны распознавания лиц остается активной все то время, пока лицо друга пребывает в поле вашего зрения (или даже до тех пор, пока вы рисуете его в своем воображении), независимо от размеров, положения, ориентации, масштаба и выражения лица. Эта стабильность нервных клеток высших функциональных зон коры головного мозга и является инвариантной репрезентацией.

Описанный процесс кажется настолько примитивным, что его даже трудно назвать задачей. Он происходит так же автоматически, как и дыхание. Мы можем назвать его тривиальным, поскольку не отдаем себе отчет в том, что происходит. В определенном смысле это действительно тривиальный процесс, ведь наш мозг выполняет его очень быстро (вспомните правило «ста шагов»). Однако вопрос о том, как кора головного мозга формирует инвариантные представления, остается одной из загадок, по сей день не разгаданных наукой. «Насколько сложной загадкой?» – спросите вы. «Насколько, – отвечу я, – что ни один ученый из числа обладающих самыми современными исследовательскими инструментами, не смог ее разгадать». И совсем не по причине отсутствия стараний.

У данной проблемы очень древние корни – о ней говорили еще во времена Платона, т. е. две тысячи триста лет тому назад. Задаваясь вопросом о том, как люди могут мыслить и познавать мир, Платон обратил внимание, что в реальной жизни нет двух одинаковых предметов или людей, и, кроме того, все они всегда несовершенны по своей форме. Например, у вас есть представление о том, что такое правильная окружность, тем не менее в реальной жизни она почти не встречается. Все изображения круга несовершенны. Даже если я начерчу окружность с помощью циркуля, его границы будут обозначены темной линией, в то время как идеальный круг не должен включать никаких утолщений. Как же в таком случае у вас сформировалось представление об идеальной окружности? Возьмем другой пример из повседневной жизни – например, ваше представление о собаке. Любой когда-либо увиденный вами пес отличается от остальных своих собратьев, а кроме того, любой новый взгляд на одну и ту же собаку также неповторим. Соответственно, каждый раз вы видите новую форму. Однако все впечатления о собаках, полученные вами в процессе жизненного опыта, объединились в вашем мозге под общим понятием «собака», которое остается неизменным. Платон задавался вопросом: как в мире бесконечно меняющихся форм и быстротечных ощущений люди способны усваивать и использовать неизменные представления?

Решением Платона стала его теория форм. Он считал, что наш высший разум связан с трансцендентной суперреальностью, в которой неизменные стабильные идеи («Формы», с прописной буквы) существуют в безвременном совершенстве. Наши души приходят из этого мистического места, где они обитают еще до нашего рождения. Именно там, в этом мистическом месте, человеческие души и постигают, что такое Формы. Согласно представлениям Платона, при рождении у нас уже имеются латентные знания о Формах. Обучение и понимание возможны только благодаря тому, что несовершенные формы реального мира напоминают нам соответствующие идеальные Формы. Мы узнаем окружность и собаку, потому что их образы вызывают воспоминания об Окружности и Собаке.

С современной точки зрения теория Платона чересчур эксцентрична, но если отвести пелену метафизики, то становится понятно, что Платон, по сути, имел в виду инвариантность. Даже называя его систему толкований чудаковатой, мы вынуждены признать, что интуитивное осознание древним мыслителем того, что поставленный вопрос является одним из наиболее важных в понимании природы человека, было попаданием в десятку.

Чтобы у вас не сложилось представление, что инвариантность присуща только зрительному восприятию, давайте обратимся к другим типам ощущений – например, к тактильному. Когда вы засовываете руку в бардачок вашего автомобиля в поисках солнцезащитных очков, вашим пальцам достаточно прикосновения, чтобы определить, что вы

их нашли. Не имеет значения, какая часть руки контактирует с предметом (большой палец, указательный, любая точка другого пальца или вся внутренняя поверхность ладони) и с какой именно его частью (со стеклом, оправой, дужкой, петлей). Для распознавания очков вашему мозгу достаточно одной секунды соприкосновения любой части вашей ладони с любой частью предмета. В каждом случае поток пространственно-временных сигналов, поступающий от осязательных рецепторов, различен, поскольку задействованы различные зоны вашей кожи, разные части предмета. Тем не менее вы, не раздумывая, возьмете свои очки.

Или же давайте рассмотрим такой пример сенсорномоторной задачи: вы вставляете ключ в замок зажигания автомобиля. Каждый раз положение вашего сидения, тела, руки немного отличаются, однако вам кажется, что это одно и то же движение, повторяющееся изо дня в день. И все это только потому, что ваш мозг обладает инвариантной репрезентативностью. Создание робота, который проделывал бы ту же операцию, забираясь в машину и вставляя ключ, практически невозможно. Даже заставив робота абсолютно точно воспроизводить одну и ту же позицию и класть ключи точно в то же место и точно так же, как и в предыдущий раз, вы непременно столкнетесь с неразрешимой проблемой: программа будет работать только для одного-единственного автомобиля. Роботы и компьютерные программы, как и искусственная автоассоциативная память, обладают одним и тем же недостатком – вариативные задачи для них неразрешимы.

Еще один интересный пример – ваша подпись. В моторной зоне коры головного мозга, в ее лобной доле, хранится репрезентация вашего автографа. Каждый раз, выводя свою подпись, вы выполняете одну и ту же последовательность действий: и в случаях, когда вы ставите аккуратную подпись, когда даете вычурный автограф в стиле Джона Хэнкока, когда у вас локоть на весу или когда вы держите ручку пальцами ног. Естественно, каждый раз ваша подпись выглядит немного по-другому (особенно в последнем случае). Но, независимо от орудия написания, размеров подписи и комбинации ее частей, для ее написания вы всегда выполняете одну и ту же абстрактную «моторную программу».

Из данного примера видно, что инвариантная репрезентация в моторной зоне коры головного мозга – это в каком-то отношении зеркальное отражение инвариантной репрезентации в сенсорной зоне коры головного мозга. В сенсорной зоне различные входные сигналы приводят в действие устойчивый набор нейронных клеток, соответствующих определенным абстрактным моделям (в приведенных выше примерах – лицо друга и солнцезащитные очки). В моторной зоне устойчивый набор нейронных клеток, соответствующий абстрактной моторной команде (хватание мяча, начертание подписи), может выразить себя посредством включения множества комбинаций мускульных групп с учетом всех вероятных ограничений. Если в разных зонах коры головного мозга всегда действует один и тот же алгоритм, как предполагал Маунткастл, то именно такой симметрии между восприятием и действием следовало ожидать.

Давайте вернемся к сенсорной зоне коры головного мозга и рассмотрим последний пример, касающийся восприятия музыки (я часто использую пример с музыкой, потому что он позволяет рассмотреть все аспекты функционирования неокортекса). Инвариантные представления в музыке обеспечивают способность узнавать мелодию, исполненную в разных тональностях. Выбирая тональность, вы автоматически предопределяете остальные ноты мелодии. Любую мелодию можно сыграть в разных тональностях. Это означает, что каждая аранжировка одной и той же мелодии – по сути, совершенно другая последовательность нот! Каждая новая аранжировка стимулирует совершенно разные наборы чувствительных слуховых рецепторов, заставляя разные наборы пространственно-временных сигналов устремляться к слуховой зоне неокортекса. Тем не менее каждый раз вы воспринимаете услышанное как одну и ту же мелодию. Не обладая абсолютным слухом, вы даже не различите два варианта одной и той же песни, исполненной в разных музыкальных тональностях.

Вспомните песню *Somewhere over the Rainbow*. Скорее всего, вы впервые услышали ее в исполнении Джуди Гарланд в фильме *Волшебник страны Оз*, но, не имея совершенного слуха, вы вряд ли вспомните тональность, в которой она была спета. Если я сяду за пианино и исполню эту песню в такой тональности, в которой вы ее никогда не слышали, для вас она прозвучит та же самая песня. Вы даже не заметите, что все ноты отличаются от тех, которые вы

слышали когда-то. Это, в свою очередь, означает, что в вашей памяти песня была запечатлена в форме, игнорирующей высоту звука. Память сохраняет важные связующие песни, а не отдельные ноты. В данном случае важными являются соотношение высоты нот и интервалы между ними. Песня *Somewhere over the Rainbow* начинается с верхней октавы, затем снижается на полтона, затем еще на один тон ниже и так далее. Интервальная структура мелодии сохраняется неизменной вне зависимости от того, в какой тональности она исполняется. Ваша способность без затруднений распознать песню в любой тональности указывает на то, что ваш мозг сохранил ее в форме инвариантного по отношению к высоте звуков представления.

Точно так же лицо вашего друга сохраняется в памяти в инвариантной форме, независимо от угла видения. Вы распознаете его лицо на основе относительных измерений, относительных цветов и относительных пропорций, а не на основе того, каким вы его увидели во вторник за обедом. Между чертами его лица существуют «пространственные интервалы», подобно тому, как в песне существуют неизменные интервалы между нотами. Его лицо является широким по отношению к его глазам, по отношению к расстоянию между глазами, у него короткий нос. Цвет его глаз и цвет волос остаются приблизительно в одном и том же соотношении при разном освещении, даже когда оно существенно меняется. Запоминая его лицо, вы запоминаете именно эти важные признаки.

По моему мнению, подобная абстрактность форм свойственна всем зонам коры головного мозга, т. е. это общее свойство неокортекса. Воспоминания сохраняются в форме, охватывающей существенные связи между элементами целого, а не преходящие детали. Когда вы видите, чувствуете или слышите что-то, кора вашего головного мозга получает специфические информационные сигналы, которые она сохраняет в инвариантной форме. Именно инвариантная форма хранится в вашем мозге, и именно с ней впоследствии сравниваются новые входные сигналы. Запоминание, припоминание и распознавание – все это происходит на уровне инвариантных форм. Компьютеры на такое не способны.

Все сказанное выше подводит нас к весьма интересному вопросу. В следующей главе я попытаюсь вас убедить, что важнейшей функцией неокортекса является прогнозирование на основе воспоминаний. Но если мозг действительно сохраняет инвариантные формы, то как же он тогда может составлять прогнозы для конкретных случаев? Я хотел бы привести некоторые примеры и предложить возможные варианты их решения.

Представьте себе 1890 год, вы находитесь в небольшом городке на западе и ожидаете свою подружку, которая должна прибыть к вам с востока. Разумеется, вы хотите встретить ее на вокзале, поэтому за несколько недель до даты ее приезда начинаете следить за расписанием прибытия-отправления поездов. Строгого графика не существует, и, по вашим наблюдениям, поезд никогда не прибывает и не отправляется в одно и то же время на протяжении дня. Вам начинает казаться, что вы никак не сможете предвидеть, когда именно приедет ваша подружка. И вдруг вы замечаете определенную закономерность. Поезд с востока прибывает через четыре часа после отправления поезда в сторону востока. Оказывается, четырехчасовой интервал остается постоянным всегда, хотя время отправления и прибытия меняется день ото дня. Наконец наступает день встречи. Вы засекаете время отправления поезда на восток. Ровно через четыре часа приходите на вокзал и встречаете свою девушку. Это пример задач, с которыми сталкивается неокортекс, и пример того, как он находит их решение.

Внешний мир, который воспринимают ваши органы чувств, не бывает статичным. Он подобен поездам, прибывающим и отправляющимся в разное время суток. Единственный способ, которым человек может познать этот изменчивый мир, – найти инвариантную структуру для переменного потока информации. Однако инвариантная структура сама по себе не является достаточной базой для частных прогнозов. Знания того, что поезд прибывает через четыре часа после отправления состава в обратном направлении, явно недостаточно, чтобы явиться на вокзал вовремя и встретить свою подружку. В каждом конкретном случае мозг сопоставляет инвариантную структуру с текущими данными. Чтобы составить прогноз времени прибытия конкретного поезда, недостаточно вывести правило четырехчасового интервала из расписания прибытия-отправления поездов. Нужно также применить его к точному времени отправления конкретного поезда в восточном направлении.

Когда вы слушаете, как кто-то исполняет знакомую вам мелодию на пианино, ваш мозг

прогнозирует следующую ноту еще до того, как она будет сыграна. Воспоминание песни сохраняется в инвариантной форме. Ваша память подсказывает вам следующий интервал, но она не знает, какая конкретно нота будет следующей. Чтобы спрогнозировать следующую ноту, нужно сочетать следующий интервал с последней нотой.

Когда вы видите лицо вашего друга, неокортекс мгновенно восполняет пробелы и прогнозирует детали внешнего вида вашего приятеля в данный момент времени. Он определяет, что это именно те глаза, тот нос, те губы и те волосы. Прогноз неокортекса отличается завидной точностью. Он может предугадать малейшие особенности лица вашего друга, хотя никогда не видел его именно под таким углом или в такой окружающей обстановке. Если вы знаете относительное расстояние между глазами и носом, структуру лица, то вы правильно спрогнозируете, где должны находиться губы. Если вам известно, что в лучах заката кожа вашего друга кажется оранжевой, то вы можете спрогнозировать и цвет его волос. Ваш мозг сочетает инвариантную структуру лица с особенностями непосредственно воспринимаемой ситуации.

Из трех вышеописанных только пример с поездом является аналогией того, что происходит в вашей коре головного мозга. В случаях с мелодией и лицом речь идет о сочетании инвариантных структур и непосредственных сигналов. Это вездесущий процесс, происходящий во всех без исключения зонах неокортекса. Благодаря ему вы можете составлять прогноз о комнате, в которой сейчас находитесь. Благодаря ему вы можете спрогнозировать не только слова, которые сейчас будут произнесены, но и тон, ударения, а также в какой части комнаты будут произнесены эти слова. Благодаря ему вы в точности знаете, в каком месте ваша нога коснется пола и какие ощущения вызовет подъем на лестничную площадку этажом выше. Благодаря ему вы можете поставить подпись, держа ручку пальцами ноги, или поймать летящий мяч.

Три особенности памяти неокортекса, рассмотренные нами в данной главе (сохранение последовательностей символов, автоассоциативное запоминание, инвариантные представления), являются необходимыми для прогнозирования будущего на основе воспоминаний о прошлом. В следующей главе сосредоточимся на доказательстве того, что сущностью разума является составление прогнозов.

5. Новые рамки понимания интеллекта

В 1986 году я подолгу размышлял над вопросом, что означает «понимать что-либо». Я стремился найти ответ на терзавший меня вопрос: если мозг не генерирует поведение, чем он тогда занимается?

Задумайтесь и вы: чем занимается ваш мозг, когда вы пассивно слушаете чью-то речь? А непосредственно сейчас, во время чтения данной книги? Информация поступает в ваш мозг, но не выходит оттуда. Что с ней происходит? Ваше внешнее поведение в данный момент остается на базовом уровне – вы дышите и совершаете движения глазами яблоками. В то же время, как вы понимаете, ваш мозг проделывает колоссальную работу, благодаря которой возможны чтение и понимание прочитанного. Понимание является результатом деятельности нейронов. Но какой деятельности? Чем именно занимаются нейроны в процессе осознания?

Зайдя сегодня в свой офис, я увидел знакомые стулья, окна, вазоны, ручки и так далее. Я увидел сотни самых различных предметов и объектов вокруг. Когда я осматривался, то одно лишь зрительное восприятие не побудило меня ни к какому действию. Тем не менее я «воспринял» комнату и находящиеся в ней объекты. Я совершил то, чего не могла сделать «Китайская комната», и мне не нужно было ничего подавать через щель в стене. Я понял, но не совершил никакого действия, доказывающего этот факт. Что же означает «понял»?

Я долго бился над этой дилеммой, и тут на меня снизошло озарение. Я задался вопросом: что бы случилось, если бы в комнате появился какой-либо новый предмет, которого я здесь раньше никогда не встречал. Например, синяя кофейная чашка.

Ответ кажется очевидным – я бы заметил новый объект как не принадлежащий к привычной обстановке. Мне не пришлось бы сознательно себе задавать вопрос о том, является

данный объект новым или нет. Осознание факта происходит само по себе. В основе кажущегося тривиальным ответа лежит очень важная концепция. Чтобы заметить изменения в обстановке, некоторые нейроны моего мозга, до этого пребывавшие в пассивном состоянии, должны были активизироваться. Откуда этим нейронам стало известно, что синяя чашка – новый объект, а сотни других предметов – старые? Я не устаю восхищаться ответом на данный вопрос. Наш мозг использует сохраненные воспоминания для того, чтобы постоянно осуществлять прогноз относительно всего, что мы видим, слышим, чувствуем. Когда я осматриваюсь в комнате, мой мозг формирует прогноз того, что он ожидает воспринять, причем делает это еще до того, как я оценю реальную обстановку комнаты. Подавляющее большинство прогнозов продуцируются неосознанно. Каждая часть мозга как бы ведет беззвучный диалог: «Стоит ли компьютер на столе? Да. Он черного цвета? Да. Лежит ли словарь на том же месте, где его оставили? Да. Стоит ли лампа в правом углу стола? Да. Имеет ли окно в комнате прямоугольную форму, вертикально ли расположены стены? Да. Попадает ли солнечный свет из направления, обычного в такое время суток? Да». А когда на сцене появляется визуальный объект, память о котором в данном контексте отсутствует, происходит нарушение ожидаемой модели. Мое внимание сразу сосредоточивается на ошибке.

Конечно, мозг не разговаривает сам с собой, составляя прогнозы, и он не составляет прогнозы, в серийном виде. Кроме того, он не составляет прогнозы относительно каких-то отдельных объектов, как, например, кофейная чашка. Мозг составляет прогнозы о самой структуре мира, в котором мы живем, и делает это, используя параллельные формы. Он с одинаковой готовностью распознает непривычную поверхность, деформированный нос или неожиданное движение. Сразу становится понятным, насколько распространенными являются такого рода неосознательные прогнозы и почему на них так долго не обращали никакого внимания. Они происходят без малейших заминок, автоматически, и нам непросто уловить, *что* происходит внутри нашего черепа. Надеюсь, это открытие поразит вас не меньше, чем меня. Прогностическая функция настолько органична для мозга, что наше восприятие мира не основывается исключительно на сигналах, которые мы непрерывно получаем от органов чувств. На самом деле восприятие действительности является комбинацией наших ощущений и прогнозов, составляемых мозгом на основе воспоминаний.

Пораженный своим открытием, минутой позже я придумал мысленный эксперимент, его подтверждающий, и назвал его «Опытом измененной двери». Вот описание эксперимента.

Возвращаясь домой, вы каждый день проходите через входную дверь. Вы протягиваете руку, беретесь ею за ручку, заходите внутрь и закрываете дверь за собой. Это глубоко укоренившаяся привычка, вы делаете это постоянно, совершенно не задумываясь. Представьте, что, пока вас не было дома, я проник в ваш дом и изменил что-то в вашей двери. Изменение может быть каким угодно. Я мог заменить круглую ручку на прямоугольную (или медную на хромовую). Я мог бы изменить вес двери – скажем, вместо тяжелой дубовой поставить железную полу, или наоборот. Я мог бы установить скрипящие плохо смазанные шарниры вместо хорошо смазанных и работающих безупречно. Я мог бы покрасить дверь в другой цвет, повесить колокольчик вместо электрического звонка, я мог бы изменить ширину дверной коробки или добавить глазок. Другими словами, я мог бы предпринять тысячи всевозможных изменений, о которых вы пока не знаете. Когда вы придете домой и попытаетесь открыть дверь, то быстро поймете: что-то не так. Размышления о том, *что* именно изменилось, могут занять у вас всего лишь несколько секунд, но в любом случае вы очень быстро найдете ответ. Например, протянув руку к ручке, вы поймете, что она сдвинута в другое место. Или заметите глазок на том месте, где его никогда не было. Если изменилась масса двери, вы толкнете ее слишком сильно или слишком слабо и очень удивитесь. Я веду к тому, что вы в несколько мгновений распознаете любое из тысячи возможных изменений.

Как вам это удастся? Как вы замечаете изменения? В качестве возможного решения разработчики искусственного интеллекта предложили бы составить список всех свойств двери и сформировать базу данных. Каждое поле соответствовало бы определенной характеристике двери, а вводимые данные могли бы служить ее описанием. В момент вашего приближения к двери компьютер запрашивал бы базу данных относительно формы ручки, ширины дверного проема, цвета двери, положения ручки, веса, звуков и так далее. На первый взгляд, это очень

похоже на то, как биологический мозг сопоставляет все мириады своих прогнозов относительно реальной обстановки, но в действительности здесь присутствует большая разница, влекущая значительные расхождения. Стратегия, используемая искусственным интеллектом, ненадежна. Во-первых, невозможно внести в базу данных все качества, которыми может обладать дверь (список атрибутов имеет реальные шансы стать бесконечным). Во-вторых, в случае реализации такой программы нам нужны были бы подобные списки свойств каждого объекта, который мы встречаем на своем пути ежесекундно, – очевидно, что эта задача не осуществима. В-третьих, ничего из того, *что* нам известно о мозге и нейронах, не наводит на мысль, что именно так они и работают. И наконец, последний аргумент: нейроны попросту не столь быстродействующие, чтобы эффективно использовать базы данных, аналогичные компьютерным (в приведенном примере с дверью вам бы пришлось потратить добрых двадцать минут, а не двадцать секунд, чтобы заметить изменения).

Вашу реакцию на измененную дверь можно объяснить только следующим образом: каждый миг ваш мозг формирует низкоуровневые сенсорные прогнозы относительно того, *что* он ожидает увидеть, услышать, почувствовать, причем эти прогностические процессы протекают параллельно. Все зоны неокортекса одновременно прогнозируют свои ожидаемые ближайшие ощущения. Зрительные зоны составляют прогнозы о формах, расположении и перемещении объектов окружающей среды. Слуховые зоны предугадывают источники, высоту, громкость и другие характеристики звуков. Соматосенсорные зоны прогнозируют тактильные ощущения, а следовательно – поверхность, контуры и температуру предметов. Вернемся к примеру с дверью. В процессе прогнозирования нейроны, задействованные в восприятии двери, активизируются еще до того, как получают входные сенсорные сигналы. При получении сенсорных сигналов мозг сравнивает их с ожидаемыми. По мере приближения к двери кора головного мозга формирует ряд прогнозов на основе вашего предыдущего опыта. Когда вы протягиваете руку, она прогнозирует, *что* вы почувствуете пальцами, когда прикоснетесь к двери, каким будет угол касания. Когда вы начинаете толкать дверь, ваш мозг прогнозирует, какую силу имеет сопротивление двери и какой звук она издаст. Если все эти ожидания оправдаются, то вы войдете в дверь, даже не осознавая, что прогнозы были проверены. Однако, если предположения окажутся ошибочными, вы обратите на дверь более пристальное внимание. Правильные прогнозы обеспечивают понимание: с дверью все в порядке. Некорректные прогнозы приводят мозг в замешательство, тем самым вынуждая вас задуматься: в чем же дело? Возможно, задвижка на двери не там, где должна быть, дверь слишком светлая, дверь смещена от центра, форма ручки изменилась и так далее. Мы пребываем в процессе постоянного параллельного формирования прогнозов для всех органов чувств.

Но это еще не все. Я рискую высказать еще более радикальное предположение. Прогнозирование, по моему мнению, – это не просто одна из функций коры головного мозга. Это *первичная функция* неокортекса и основа интеллекта. Кора головного мозга является органом предвидения. Если мы хотим понять, что такое разум, что такое творчество, как работает наш мозг и как научиться создавать разумные машины, нам нужно постичь природу прогнозов и понять, каким образом кора головного мозга их формирует. Даже поведение можно лучше всего представить как промежуточный продукт процесса прогнозирования.

Не знаю, кто первым выдвинул предположение, что прогнозирование является основой понимания интеллекта. В науке и производственной сфере не бывает совершенно новых открытий. Чаще всего люди просто изменяют угол зрения. Составляющие нового подхода, как правило, витают в воздухе научного дискурса еще до их непосредственного открытия. Новой является та обобщающая форма, в которую облачают уже известные мысли. Точно так же и мысль о том, что первичной функцией коры головного мозга является прогнозирование, не нова. Тем не менее до сих пор никто не отводил ей надлежащее ключевое место в теории интеллекта.

По иронии судьбы некоторые из пионеров искусственного интеллекта представляли себе, что компьютеры должны создавать модель мира и использовать ее для формирования прогнозов. Например, в 1956 году Д. М. Макай высказал предположение, что разумные машины должны располагать «внутренним механизмом соответствия», предназначенным для того, чтобы «сверять полученную информацию». Он не употребил терминов «память» или

«прогнозы», но мыслил как раз в таком ключе.

Начиная с середины девяностых годов XX века в научной номенклатуре появились такие термины, как «инференция», «генеративные модели» и «прогнозирование». Все они взаимосвязаны. Например, Родольфо Линас (Медицинская школа Нью-Йоркского университета) в своей книге *i of the vortex*, изданной в 2001 году, пишет: «Способность предвидеть результаты предстоящих событий, что является критическим для успешного передвижения, скорее всего, является первичной и наиболее универсальной из всех фундаментальных функций мозга». Такие ученые, как Дэвид Мамфорд (Браунский университет), Раджеш Рао (Вашингтонский университет), Стивен Гроссберг (Бостонский университет), а также многие другие высказывали мысль о значении обратной связи и прогнозирования. Существует целый раздел математики, посвященный сетям Байеса. Названы они были в честь Томаса Байеса, английского министра, жившего в XVIII веке, которого считают пионером статистики. Составление прогнозов в сетях Байеса осуществляется на основе теории вероятности.

Чего не хватало всем нашим предшественникам, так это собрать все крупницы знаний в одни логически обоснованные теоретические рамки. Доселе никто этого не сделал, и именно эта задача является целью написанной мною книги.

Прежде чем перейти к подробному рассмотрению прогностической функции коры головного мозга, давайте обсудим еще несколько примеров. Чем больше вы будете размышлять над указанной идеей, тем больше преисполнитесь убеждения, что прогнозирование вездесуще, оно является основой нашего понимания мира.

Сегодня я жарил блины. В какой-то момент я протянул руку, чтобы открыть дверцу кухонного шкафчика. Интуитивно, не поднимая глаз к дверце, я знал, что сейчас почувствую и когда именно это произойдет. Так и случилось: я действительно ощутил, когда прикоснулся к ручке шкафчика. Я покрутил крышку контейнера для молока, зная, что откручу ее. Я поставил сковородку на плиту, приложив ровно столько силы, сколько необходимо, чтобы ее поднять. Я повернул выключатель на плите, ожидая через пару секунд увидеть вспышку газового пламени. Находясь на кухне, я ежесекундно совершал сотни движений, каждое из которых состояло из множества прогнозов. Я знаю это, ведь если бы какое-либо из привычных движений вызвало не соответствующий моим ожиданиям результат, то я бы это заметил.

Каждый раз во время ходьбы пешком ваш мозг прогнозирует, в какое мгновение ваша нога прекратит движение и какое давление вы почувствуете, ступив на определенную поверхность. Если, спускаясь по лестнице, вы вдруг не попадете ногой на ступеньку, то мгновенно почувствуете: что-то не так. Когда вы опускаете ногу и там, где ожидали ощутить ступеньку, ее не оказывается, вы сразу осознаете, что у вас проблемы. Ваш мозг составил прогноз, который не оправдался. Компьютеризированный робот однозначно с грохотом полетел бы вниз, не осознавая того, что чего-то не хватает. Вы же, напротив, осознаете это через долю секунды после того, как ваша нога минует точку, где, согласно прогнозам мозга, она должна была остановиться.

Наслаждаясь звучанием знакомой мелодии, вы слышите следующую ноту еще до того, как она прозвучала. Прослушивая любимый сборник песен, вы слышите начало каждой следующей песни за несколько секунд до того, как оно действительно воспроизводится техникой. Что происходит? Нейроны вашего мозга передают импульс о звуках мелодии заблаговременно (в ответ на воспоминание о музыкальном произведении), еще до того, как вы их действительно слышите. Именно таким образом вы «слушаете» песню в своей голове. Такие воспоминания могут быть на удивление продолжительными. Мозг автоматически воспроизводит следующую песню после окончания предыдущей даже в случае, когда вы не слушали музыкальный сборник много лет. Поэтому прослушивание любимого компакт-диска в режиме произвольного проигрывания создает чувство приятного ожидания. Ведь вы никак не можете предугадать, какая мелодия станет следующей.

В беседе вы часто можете предвосхитить (или, по крайней мере, думаете, что можете) то, что ваш собеседник скажет дальше. А порой мы слушаем, но не слышим: не вникаем в смысл слов другого человека, просто слышим то, что нам хочется услышать (со мной в детстве это приключалось настолько часто, что моя мать даже дважды показывала меня

врачу-отоларингологу, чтобы убедиться, что с моим слухом все в порядке). Отчасти подобное происходит потому, что люди часто используют в устной речи общие фразы и выражения. Если я скажу: «Ну вот, опять двадцать...», ваш мозг активизирует нейроны, отвечающие за слово «пять» еще до того, как я его произнесу. Разумеется, мы не всегда знаем, *что* человек произнесет дальше. Прогнозы не всегда сбываются. Точнее, наш мозг составляет вероятностные прогнозы того, *что* должно случиться. Иногда мы знаем наверняка, *что* произойдет, иногда наши ожидания рассеяны между несколькими возможными вариантами событий. Если мы с вами обедаем за одним столом и я вас попрошу: «Простите, вы не могли бы мне подать...», ваш мозг будет готов услышать не только «соль», но и «горчицу», и «перец». В определенном смысле он спрогнозировал все три варианта сразу. Но скажи я вам: «Простите, вы не могли бы мне подать тротуар для пешеходов», вы мгновенно озадачились бы.

Вернемся к музыке. Здесь тоже действуют вероятностные прогнозы. Даже при прослушивании незнакомой песни, ваш мозг формирует достаточно сильные ожидания. От западной музыки я ожидаю правильного такта, повторяющегося быстрого ритма, куплетов из одной строфы и окончания песни на высоком тоне. Вы можете не знать, *что* означают все эти понятия, но, если вы слышали подобную музыку, ваш мозг автоматически составляет прогнозы о такте, повторяющихся ритмах, окончании музыкальных фраз, окончании всей песни. Если в новой песне указанные принципы были нарушены, вы сразу поймете: что-то не так. Задумайтесь об этом на мгновение. Вы слушаете песню, которую раньше никогда не слышали, ваш мозг получает сигналы, которые никогда не получал раньше, тем не менее вы составляете прогнозы, и можете определить «расхождения с правилами». Основой этих преимущественно бессознательных прогнозов является набор воспоминаний, хранящихся в коре вашего головного мозга. Ваш мозг не может в точности предвидеть звучание незнакомой песни, однако он прогнозирует, какие звуковые сигналы являются потенциально вероятными, а какие – нет.

Каждому из нас когда-либо приходилось замечать за собой, что мы не слышим постоянного фонового шума¹² – например, стука дятла, сопровождающего вашу прогулку в лесу, или какого-нибудь заигранного мотивчика в маленькой кофейне. С другой стороны, как только он прекращается, мы сразу же это замечаем. Наши слуховые зоны спрогнозировали, что звук будет непрекращающимся, поэтому, когда он не менялся, вы не обращали на него ровным счетом никакого внимания. И вот звук исчез, ваши ожидания не оправдались, и это сразу привлекло ваше внимание. Приведу исторический пример: сразу после того, как в Нью-Йорке были отменены наземные поезда, многие жители города посреди ночи звонили в полицию и жаловались, что их разбудило грохотание. Как правило, звонки раздавались в те часы, когда наземный поезд должен был бы проходить мимо их домов.

Часто говорят, что верить и видеть сродни друг другу. Однако мы видим то, что ожидали увидеть, только тогда, когда действительно это видим. Наиболее ярким примером может служить явление, названное исследователями *заполнением*. Возможно, вы обращали внимание, что в каждом вашем глазу есть темная точка – здесь ваш зрительный нерв выходит из сетчатки через отверстие, называемое *зрительным диском*. В этой зоне нет фоторецепторов, поэтому она является постоянной слепой зоной в поле вашего зрения. Существуют две причины, по которым вы, как правило, не замечаете этих зон. Первая причина состоит в том, что эти два пятна никогда не пересекаются, поэтому один глаз компенсирует то, чего не увидел другой.

Но что любопытно – вы не замечаете темного пятна, когда у вас открыт только один глаз. Ваша зрительная система «заполняет» пробелы. Если вы закроете один глаз и будете разглядывать богато расшитый турецкий ковер, то целые узлы ковра будут ускользать из поля зрения вашей сетчатки, попадая в ваше слепое пятно. Однако ваш опыт цельного рисунка на ковре поможет вам создать непрерывный поток прогнозов, и вы легко восстановите цельное изображение.

Заполнение происходит во всех частях визуального изображения, а не только в слепых пятнах сетчатки глаза. Например, я покажу вам картинку побережья со скалами, на которых лежит бревно. Граница между скалами и бревном очень четкая и хорошо различимая. Увеличив

¹² Иначе он называется белым шумом. – Примеч. ред.

изображение, мы увидим, что бревно и скалы практически неразличимы по цвету в точках их пересечения. При достаточно большом увеличении масштаба провести между ними границу вовсе не представляется возможным. Видимая четкость контуров бревна при взгляде на исходное изображение объясняется тем, что мы выделили их по отношению к остальной части рисунка. Когда мы смотрим на мир, то видим четкие контуры и границы объектов, хотя фактические сигналы, поступающие в зрительные зоны, преимущественно нечеткие и неоднозначные. Кора нашего головного мозга заполняет пробелы и выделяет четкие контуры, основываясь на своих ожиданиях. Мы же в конечном счете воспринимаем все как реальную картину.

Прогнозирование в зрительном восприятии является функцией движений ваших глаз. В главе 3 я упоминал о саккаде. Три раза в секунду ваши глаза фиксируются на одной точке, а потом перемещаются в другую. Вы не осознаете этих движений и не контролируете их сознательно. Всякий раз, когда ваши глаза фиксируются на новой точке, сигналы, поступающие в зрительную зону неокортекса, отличны от тех, которые сопровождали фиксацию на предыдущей точке. Получается, трижды в секунду ваш мозг видит новое изображение. Саккады глаз нельзя назвать совершенно произвольными. Когда вы рассматриваете лицо другого человека, то, как правило, ваш взор сначала останавливается на одном его глазу, потом на другом, затем перепрыгивает туда и обратно, время от времени фиксируясь на носу, ушах, рте и других чертах. Вы воспринимаете «лицо», но ваши глаза смотрят на один глаз, второй глаз, нос, рот, снова глаз и так далее. Скорее всего, вы этого совсем не осознаете. Вам кажется, что вы видите непрерывную картину мира, но первичные данные, поступающие в ваш мозг, настолько же беспорядочны, как и съемка на плохо настроенной цифровой видеокамере.

А теперь представьте себе, что вы встретили кого-то со вторым носом на месте одного глаза. Ваши глаза сначала фиксируются на одном глазу, потом перепрыгивают к другому, но там вместо глаза вы видите нос. Вы сразу понимаете: что-то не в порядке. Ваш мозг сформировал определенное ожидание, или прогноз, того, что вы должны были увидеть, но оно не оправдалось. Вы предвидели еще один глаз, а вместо этого увидели нос, и это привлекает ваше внимание. Именно поэтому так трудно не пялиться на людей, обладающих физическими особенностями. Легко ли вам было бы отвести свой взгляд от человека с двумя носами? Естественно, если бы вы прожили вместе с ним некоторое время, то привыкли бы к двум носам и не воспринимали бы эту его особенность как нечто из ряда вон выходящее.

А сейчас проведите самоанализ. Какие ожидания формируются у вас в данный момент, в процессе чтения этой книги? Когда вы переворачиваете страницы, у вас есть ожидания касательно того, как они будут переворачиваться. Ожидания относительно страниц книги отличаются от ожиданий от переворачивания обложки книги. Если вы сидите, значит, у вас есть ожидание, что ощущение давления на тело не исчезнет. Если сидение под вами вдруг станет мокрым, или начнет отодвигаться назад, или же с ним начнут происходить другие непредвиденные изменения, вы отвлечетесь от чтения до тех пор, пока не выясните, что происходит. Понаблюдав за собой таким образом, вы наверняка поймете, что ваше восприятие и понимание мира самым непосредственным образом связано с прогнозированием. В вашем мозге хранится модель мира, которая подвергается постоянному сопоставлению с реальностью.

Прогнозирование действует не только в отношении сигналов сенсорной информации низшего порядка. До сих пор я ограничивал наш анализ такими примерами, потому что это самый простой путь найти рамки понимания интеллекта. Если придерживаться принципа Маунт-кастла, принципы восприятия информации зонами низшего порядка справедливы и для высших зон коры головного мозга. Мозг человека обладает более высоким интеллектом, чем мозг животных, поскольку способен составлять прогнозы для более абстрактных типов сигналов и более длинных их последовательностей. Чтобы спрогнозировать, что жена скажет мне при встрече, я должен знать, что она говорила мне в прошлом, а также учитывать, что сегодня пятница и что мусор нужно выносить в контейнер в пятницу вечером, чего я не сделал на прошлой неделе, а ее лицо в таких случаях принимает характерное выражение. Еще до того как она откроет рот, у меня есть обоснованные предположения о том, что она скажет. Разумеется, я не уверен стопроцентно, но с большой долей вероятности могу утверждать, что она будет напоминать мне, чтобы я не забыл выбросить мусор в контейнер. Высший интеллект

по признаку процесса формирования не отличается от интеллекта перцептивного. Они оба базируются на одной и той же системе памяти коры головного мозга и на общем прогностическом алгоритме.

Обратите внимание, что наши тесты на определение коэффициента интеллекта (IQ) по своей сути являются прогностическими задачами. Предназначенные для вычисления IQ людей всех возрастных категорий – от дошкольников до аспирантов, – эти тесты оценивают прогностическую способность. Вам предлагают определить, какое число будет следующим в представленной вам последовательности. Вам предлагают три разные проекции одного предмета и спрашивают, какой из предложенных рисунков является еще одной проекцией того же предмета. Или предлагают найти слово, которое относится к слову В так же, как и слово А относится к слову Б.

По сути, даже наука основана на прогнозировании. Мы расширяем свои знания о мире путем формулирования гипотез и их проверки. Эта книга является прогнозом того, что такое интеллект и как функционирует мозг. Даже разработка дизайна товара тоже является прогностическим процессом. При создании нового фасона одежды или типа мобильных телефонов дизайнеры и инженеры пытаются предвидеть запросы потребителей, возможные действия конкурентов, стоимость нового проекта и прочее.

Интеллект определяется способностью запоминать и предвидеть различные аспекты внешнего мира, включая язык, математику, физические свойства объектов, социальные ситуации. Ваш мозг получает сигналы из внешнего мира, сохраняет их в форме воспоминаний, и на основе того, что случилось раньше, и того, что происходит сейчас, составляет прогнозы о будущем.

Сейчас вы, наверное, подумаете: «Ладно, допустим, я согласен, что мой мозг непрерывно занят прогнозированием и может проявлять интеллект, даже когда я лежу в потемках и ни о чем не думаю. Как вы отметили, мне не надо действовать, чтобы осознавать или проявлять интеллект. Но разве подобная ситуация не является исключением? Вы в самом деле полагаете, что интеллект и разумное поведение совершенно независимы? В конце концов, разве не поведение, а прогнозирование является главным проявлением нашего разума? Ведь именно поведением определяется возможность выживания».

Такие вопросы закономерны. В конечном счете для выживания животного поведение действительно является определяющим. Прогнозирование и поведение не являются независимыми, но связь между ними очень тонкая и трудноуловимая. Во-первых, кора головного мозга появилась у млекопитающих лишь тогда, когда их поведение развилось до уровня достаточно сложных форм. Таким образом, влияние неокортекса нужно рассматривать с точки зрения усовершенствования заведомо существовавших моделей поведения животных. Разумное поведение возникло прежде интеллекта. Во-вторых, подавляющее большинство наших ощущений в значительной мере зависят от того, *что* мы делаем и *как* передвигаемся в этом мире. Прогнозирование и поведение тесно связаны друг с другом. Давайте остановимся на этом более подробно. У млекопитающих развилась большая кора головного мозга, потому что это создавало дополнительное преимущество для выживания, и это преимущество в конечном счете закрепилось в поведении. Однако вначале кора головного мозга служила для извлечения наибольшей пользы из существующих моделей поведения, а не для создания принципиально новых моделей. Чтобы лучше понять данный тезис, нам стоит рассмотреть, как развивался человеческий мозг.

Простейшие нервные системы развились у многоклеточных еще сотни миллионов лет назад. Но вот история появления настоящего разума начинается со времен наших предшественников – рептилий. Рептилии очень успешно завоевывали все новые и новые территории. Они распространились по всем континентам и стали родоначальниками множества видов. У них были хорошо развиты органы чувств и мозг, обеспечивающий их сложными моделями поведения. Их прямые потомки – современные рептилии – обладают точно такими же характеристиками. У аллигатора, например, органы чувств настолько же хорошо развиты, как у вас или у меня. Он использует сложные модели поведения, включая умение плавать, прятаться, охотиться, загорать, обустривать гнездо и спариваться.

Насколько велика разница между мозгом человека и мозгом рептилии? С одной

стороны, невелика, с другой – огромна. Невелика, потому что в грубом приближении все, что имеется в мозге рептилии, присутствует и у человека. Огромна, потому что кора головного мозга человека гораздо больше, чем у пресмыкающегося.

Доводилось ли вам слышать упоминания о «старом», или «простейшем», мозге? Эта часть мозга находится в основании черепа и отвечает за функции жизненно важные функции организма: репродуктивные, инстинкт самосохранения, циркуляцию крови, дыхание, сон, сокращение мышц в ответ на внешнюю стимуляцию¹³. Когда вы стоите, передвигаетесь или держите равновесие, то в значительной мере это происходит благодаря «старому» мозгу. Так чем же тогда занята кора головного мозга, если она не является незаменимой для того, чтобы видеть, слышать и передвигаться?

Млекопитающие умнее рептилий благодаря коре головного мозга – неокортексу, развившемуся десятки миллионов лет назад, позже всего в процессе эволюции (именно поэтому его называют «новым мозгом»). Неокортекс присутствует только у млекопитающих, причем стремительное увеличение коры головного мозга у человека произошло всего лишь пару миллионов лет тому назад. Кора головного мозга у разных видов в целом однородна. Слой вашего неокортекса имеет ту же толщину и почти ту же структуру, что и ткань коры головного мозга ваших млекопитающих сородичей. Любые значительные эволюционные изменения, происходящие в короткий промежуток времени, обеспечиваются заимствованием существующих структур. Существует распространенное, но ошибочное мнение, что человеческий мозг является кульминационным результатом миллиардов лет эволюционного развития. Подобное мнение верно для нервной системы в целом, но кора головного мозга является относительно новой структурой и существует не так давно, чтобы стало возможным продолжительное эволюционное усовершенствование. Мы умнее за счет увеличения количества повторяющихся элементов универсального алгоритма коры головного мозга.

В этом и состоит сущность моего подхода к пониманию неокортекса – к признанию системы памяти и прогностической функции главными ключами к тайне разума. В мозге рептилии отсутствует неокортекс. Согласно эволюционной теории, появление в головном мозге коры знаменует переход к умению прогнозировать. Представьте себе, что «старый» мозг рептилий занимается тем же, что и всегда, но теперь сенсорные сигналы одновременно поступают и в кору головного мозга. Кора головного мозга сохраняет полученную сенсорную информацию в своей памяти. В будущем, если животное сталкивается с такой же или подобной ситуацией, память распознает сигналы как похожие на полученные в прошлом и вспоминает, что случилось. Вызванное воспоминание сравнивается с потоком поступающей сенсорной информации. Оно дополняет настоящий поток сигналов и прогнозирует, *что* случится дальше. Сравнивая поступающие сенсорные сигналы с вызванным воспоминанием, животное не только понимает происходящее, но и может предвидеть, *что* случится в будущем. А теперь представьте себе, что кора головного мозга не только помнит то, что увидело животное, но помнит также и то, какую модель поведения использовал «старый» мозг в подобных ситуациях. Не следует считать, что кора головного мозга проводит различие между ощущениями и поведением. Для неокортекса как первое, так и второе – всего лишь сигналы. Когда животное попадает в такую же или подобную ситуацию, оно не только прогнозирует будущее, но и вспоминает, какая модель поведения привела к подобному прогнозу будущего. Таким образом, память и прогнозирование позволяют животному использовать существующие модели поведения (хранящиеся в «старом» мозге) более разумно.

Рассмотрим еще один пример. Представьте крысу, впервые попавшую в лабиринт и пытающуюся найти из него выход. Возбужденное голодом и неопределенностью животное использует врожденные качества своего «старого» мозга для исследования непривычной среды, т. е. слушает, смотрит, обнюхивает, осторожно передвигается вдоль стен. Получаемая сенсорная информация не только используется старым мозгом, но попадает в кору головного мозга, где и сохраняется. Когда некоторое время спустя крыса помещается в тот же лабиринт,

¹³ Биологи называют «старый» мозг «мозгом рептилии», т. к. аналогичную анатомическую структуру имеют все позвоночные – от рептилий до млекопитающих. – Примеч. ред.

ее кора распознает текущие сигналы как уже знакомые и вспоминает сохраненные сигналы, которые описывают то, что случилось в прошлом. По сути, кора помогает крысе увидеть наиболее краткий путь в будущее. Если бы животное могло говорить, то, скорее всего, сказало бы: «О, узнаю этот лабиринт, и вот этот угол я тоже хорошо помню». Крыса помнит, где лежит сыр и как она добралась до него в прошлый раз. «Если я поверну здесь, то знаю, какой поворот будет следующим. В конце того коридора лежит кусочек сыра. Я вижу его в своем воображении». Пробираясь через лабиринт вслепую, переставляя лапки или шевеля усиками, крыса полагается на старые примитивные структуры мозга. Благодаря своей относительно большой коре головного мозга крыса способна запомнить места, в которых побывала, и узнать их в будущем, а также спрогнозировать дальнейшие события. У ящерицы отсутствует кора головного мозга, поэтому ее способности запоминать прошлое намного более ограничены. Сколько бы раз не оказалась ящерица в лабиринте, ей приходится искать выход по-новому. Крыса же благодаря системе памяти коры головного мозга имеет образ мира и прогнозирует будущее. В мозге животного всплывают яркие образы опасностей и вознаграждений, стоящих за каждым принятым решением, поэтому крыса, обладая способностью предвидения, существует в мире более эффективно.

Обратите внимание: крыса не следует никакой сверхсложной или принципиально новой модели поведения. Она не сконструировала себе дельтаплан, чтобы долететь до вожделенного сыра в конце коридора. Кора головного мозга крысы составляет прогнозы о сенсорных сигналах, которые позволяют ей предвидеть будущее, но палитра моделей поведения животного остается неизменной. Ее способности подпрыгивать, красться, исследовать помещение почти такие же, как и у ящерицы.

По мере развития и увеличения в процессе эволюции кора головного мозга обрела способность запоминать все большие и большие объемы информации о мире, соответственно, совершенствовалась ее прогностическая функция. Сложность воспоминаний и прогнозов тоже возросла. Но произошло еще одно чрезвычайно важное событие, которое привело к появлению у человека истинного разумного поведения.

Человеческое поведение значительно превосходит старый базовый репертуар ориентирования, свойственный крысе. Эволюция коры головного мозга человека вышла на качественно новый уровень. Только человеческие существа в состоянии были создать разговорную речь и письменность. Только люди готовят себе пищу, шьют одежду, летают на самолетах, сооружают небоскребы. Наши моторные способности и умение планировать намного превосходят способности наших ближайших животных сородичей. Как удастся коре головного мозга, изначально предназначавшейся для составления сенсорных прогнозов, генерировать сложные модели поведения, присущие лишь человеку? Откуда могли появиться эти модели столь неожиданно? Есть два возможных ответа на поставленный вопрос. Первый: алгоритм функционирования коры головного мозга чрезвычайно мощный и гибкий. Путем незначительных изменений способа соединения, что присуще только человеку, она может создавать новые сложные модели поведения. Второй: поведение и прогноз являются двумя сторонами одной и той же медали. Кора головного мозга действительно может предвидеть будущее, но ее сенсорные прогнозы будут отличаться точностью лишь при учете текущих моделей поведения.

Вспомните пример с крысой, мечущейся по лабиринту в поисках сыра. На любом перекрестке у животного есть выбор: повернуть направо или налево. Только благодаря одновременному воспоминанию о месте расположения сыра и правильной модели поведения оно составляет правильный прогноз, который и приводит к сыру. Конечно, это примитивный пример, но он прекрасно демонстрирует тонкую связь между поведением и сенсорными прогнозами. Наше поведение оказывает огромное влияние на то, *что* мы слышим, видим и чувствуем. Большинство наших сиюминутных ощущений зависят от наших действий. Проведите рукой перед лицом. Чтобы спрогнозировать вид руки, мозг учитывает, что он приказал руке двигаться. Если ваш мозг увидит движущуюся руку, не давая такой моторной команды, вы будете очень удивлены. Сначала кора головного мозга приказала руке двигаться, а потом спрогнозировала то, что должна увидеть. Я полагаю, такое объяснение неверно. По-моему, кора головного мозга прогнозирует увидеть руку, именно этот прогноз заставляет

моторные команды осуществить ожидания. Сначала вы думаете, это заставляет вас действовать, и таким образом ваши мысли осуществляются.

А сейчас давайте разберемся, какие изменения привели к значительному расширению репертуара моделей поведения человека. Существуют ли физические различия между корой головного мозга обезьяны и человека, которые могли бы объяснить, почему только у людей существует язык, а также другие сложные модели поведения. Объем человеческого мозга превышает объем мозга обезьяны приблизительно в три раза. Однако дело здесь вовсе не в том, что больше – всегда лучше. Ключом к пониманию качественного прыжка человеческого поведения является форма связи между зонами неокортекса и «старым» мозгом. У человека и животных эти формы связи различны.

Давайте рассмотрим данную особенность более подробно. Вам наверняка известно, что головной мозг делится на правое и левое полушария. Но существует и другая форма разделения, используя которую мы лучше сможем понять, чем отличается человек от остальных животных.

Головной мозг, особенно если он достаточно большой, разделяет кору головного мозга на переднюю и заднюю часть, *антериорную* и *постериорную*. Их разделяет извилина, которую называют *центральной бороздой*. Заднюю часть коры головного мозга образуют зоны сенсорного восприятия, получающие входные сигналы от органов зрения, слуха и осязания. Передняя часть коры головного мозга включает зоны, осуществляющие мышление и планирование высокого уровня сложности. Сюда также входит моторная зона коры головного мозга, отвечающая за контроль над мышцами, т. е. за непосредственное поведение.

В процессе эволюции кора головного мозга (особенно ее передняя часть) у людей и приматов увеличилась. По сравнению с другими приматами и ранними гоминидами¹⁴ у нас непропорционально большой лоб, основное назначение которого – служить вместилищем для очень большой передней части коры головного мозга. Но одного этого факта недостаточно, чтобы объяснить улучшение наших моторных способностей по сравнению с другими живыми существами. Способность человека осуществлять сложнейшие движения связана с тем, что моторная зона коры головного мозга у *homo sapiens* имеет намного больше связей с мышцами тела, чем у других млекопитающих. Поведение большинства животных генерирует «старый» мозг. А у человека передняя часть неокортекса узурпировала большую часть моторного контроля. Если вы повредите моторную зону коры головного мозга крысы, она не почувствует существенных неудобств. Если вы повредите моторную зону коры головного мозга человека, он будет парализован.

Меня часто спрашивают о дельфинах: правда ли, что у них очень большой мозг. Да, это правда. С более простым строением, чем у человека (три слоя вместо шести), но в целом мозг дельфина довольно велик. Вполне возможно, что дельфин может запоминать и понимать большое количество информации, распознавать лица, сохранять в памяти события своей жизни. Не исключено, что он запоминает все уголки океана, в которых ему когда-либо случалось побывать. Но, несмотря на способность проявлять сложные модели поведения, возможности дельфина далеки от человеческих. Таким образом, приходим к выводу, что даже у этих удивительных созданий кора головного мозга не оказывает столь существенного влияния на поведение, как у человека. Кора головного мозга в первую очередь развивалась с целью запоминания мира. Животные с достаточно большой корой головного мозга воспринимают мир ничуть не хуже нас с вами. Однако то, что делает человека уникальным, – это доминирующая роль коры головного мозга в формировании поведения. Именно по этой причине у людей существует система речи, именно поэтому создаются замысловатые инструменты. Именно поэтому мы можем писать романы, находить информацию в Интернете, посылать исследовательские зонды на Марс и строить корабли, в то время как животные совершенно не способны на подобные виды деятельности.

Вот мы и подошли к тому, чтобы нарисовать полную картину. Природа создала рептилий

¹⁴ Гоминиды – род из отряда приматов, включающий ископаемых предков людей (человека умелого и человека прямоходящего) и современных людей. – Примеч. ред.

со сложными органами чувств и сложными, но относительно устойчивыми моделями поведения. Потом она сделала открытие: если дополнить их мозг системой памяти, к которой подключить поток сенсорной информации, животное сможет запоминать свой прошлый опыт. Когда животное попадает в такую же или подобную ситуацию, происходит вызов воспоминания из памяти, а это ведет к прогнозированию того, что, вероятнее всего, должно случиться. Таким образом, разум и понимание начались с системы памяти, которая посылала прогнозы в поток сенсорного восприятия. Эти прогнозы являются сущностью понимания. Знать что-либо означает, что вы можете составлять об этом предположение.

Кора головного мозга развивалась в двух направлениях. Во-первых, она увеличилась, и, соответственно, научилась хранить более сложные воспоминания. Она могла запоминать больше информации и составлять прогнозы на основе более сложных связей. Во-вторых, кора начала взаимодействовать с моторной системой «старого» мозга. Чтобы спрогнозировать, *что* будет дальше, ей нужно учитывать текущие действия. В результате управление большей частью моторного поведения человека перешло к неокортексу. Уже не ограничиваясь составлением прогнозов на основе поведения, диктуемого «старым» мозгом, неокортекс человека управляет его поведением для удовлетворения своих ожиданий.

Кора головного мозга человека особенно велика, поэтому обладает большой запоминающей способностью. Она постоянно составляет прогнозы того, *что* вы почувствуете, услышите, увидите, причем вы этого не осознаете. Прогнозы – это наши мысли, а в сочетании с сенсорными входными потоками информации – наше восприятие. Я назвал такое видение мозга *запоминающе-прогностическими рамками* интеллекта.

Если бы «Китайская комната» включала подобную систему памяти, которая могла бы прогнозировать, какой иероглиф появится следующим, мы с полной уверенностью могли бы сказать, что она поняла китайский и поняла рассказ. Теперь нам понятна ошибка Алана Тьюринга. Прогнозирование, а не поведение является свидетельством наличия разума.

Теперь мы готовы к тому, чтобы углубиться в подробности идеи о запоминающе-прогностических рамках интеллекта. Для прогнозирования будущих событий коре головного мозга нужно сохранять последовательности сигналов. Для вызова соответствующего воспоминания ей нужно соотнести поступающие сигналы с подобными сигналами в прошлом (автоассоциативное воспоминание). Кроме того, воспоминания должны сохраняться в инвариантной форме, чтобы знание прошлых событий было применимо к новым подобным, но не обязательно идентичным ситуациям. В следующей главе мы поговорим о том, как кора головного мозга выполняет эти задачи, а также более подробно рассмотрим ее иерархическое строение.

6. Как работает кора головного мозга

Пытаться разгадать, как работает кора головного мозга, все равно, что собирать пазлы – головоломку, в которой картинку складывают из множества отдельных кусочков. Пазлы можно собирать двумя способами. Первый – от целого к частному – предполагает, что у вас есть образец картинки, которая должна получиться. Начав собирать с какой-либо части изображения, вы подыскиваете подходящие кусочки и шаг за шагом восстанавливаете его целиком. Второй подход – от частного к целому – подразумевает сосредоточенность на отдельных кусочках. Вы ищете в отдельных фрагментах картины общие детали и пытаетесь соотнести их друг с другом. При отсутствии подсказки в виде готового изображения единственно возможным способом собирания пазлов является второй.

Головоломка под условным названием «Как понять мозг» особенно трудна. Поскольку единого подхода к пониманию мозга долгое время вообще не существовало, исследователи были вынуждены прибегать к методу от частного к целому. С таким уровнем сложности, как мозг человека, данная задача была по силам лишь Геркулесу. Только представьте себе: каждый из многих тысяч фрагментов – двухсторонний, причем не существует однозначных его интерпретаций, и только одна из сторон каждого кусочка имеет отношение к головоломке. Края всех кусочков неровные (словно они были неаккуратно вырезаны), поэтому вы не можете

судить с полной уверенностью, правильно ли соединяете каждую пару. Многие из имеющихся у вас фрагментов вообще не нужны для составления этого гигантского изображения, их наличие в общей куче лишь усложняет вам задачу. Кроме того, по почте вам постоянно присылают все новые и новые кусочки, причем некоторые из них предназначены для того, чтобы заменить старые. Автор головоломки как бы говорит вам: «Я знаю, что вы несколько лет бились над тем, как соединить все эти фрагменты, но оказалось, что они не имеют отношения к моей картинке. Вы уж извините. Вот вам новые, до следующего раза». И самое главное: вы даже не представляете себе, каким должен быть конечный результат. Или того хуже – у вас есть определенные соображения, но они ошибочны.

Аналогия с головоломкой-пазлами прекрасно иллюстрирует трудности, с которыми мы сталкиваемся на пути создания единой теории неокортекса и интеллекта. Кусочки головоломки – это биологические и поведенческие данные, накопленные учеными за сто с лишним лет. Результаты новых исследований добавляют в головоломку новые кусочки. Иногда сведения, предложенные разными учеными, крайне противоречивы. А поскольку любые данные можно истолковывать по-разному, то единого мнения не существует практически ни по какому вопросу. Ввиду невозможности подхода от целого к частному цель научных изысканий зачастую остается невыясненной, и, что не менее важно, ученые порой не имеют представления о том, как интерпретировать массивы накопленной информации. Изучая мозг с помощью подхода от частного к целому, мы попросту зашли в тупик. Очевидно, необходимо очертить рамки для подхода от целого к частному.

Таковыми рамками может послужить модель «Память-запоминание». Она подскажет нам, как сложить вместе все кусочки головоломки в целую картинку. Для осуществления прогностической функции неокортексу нужна определенная технология запоминания и сохранения информации о последовательностях событий. Для прогнозирования вероятных событий кора головного мозга должна использовать (а значит, формировать) инвариантные представления. Ваш мозг должен создавать и сохранять модель реального мира так, чтобы при восприятии последнего в разных обстоятельствах изменения компенсировались. Разобравшись в том, чем должна заниматься кора головного мозга, мы сможем понять ее строение, особенно шестислойную иерархическую структуру.

В данной книге такое видение рамок исследования интеллекта будет представлено впервые. Иногда я буду забираться в научные дебри, что, возможно, некоторым читателям покажется слишком сложным. Многие из концепций будут незнакомы даже специалистам-нейрологам. И все же, приложив определенные усилия, любой читатель, как минимум, сможет понять основы нового подхода. (Главы 7 и 8 гораздо меньше насыщены техническими подробностями – в них речь пойдет о более широком применении предложенной теории.)

Прежде всего мы рассмотрим биологические предпосылки, на которых основана запоминающе-прогностическая гипотеза. Это поможет нам отбросить многие ненужные кусочки воображаемых пазлов, тем более что нам известно: лишь немногие из элементов действительно необходимы для воссоздания полной картины. Как только мы поставим перед собой четкую цель и разберемся, *что* же нужно искать, задача станет разрешимой.

Хочу отметить, что предлагаемые мной рамки исследования далеко не окончательны. У меня до сих пор нет ответов на многие вопросы. Но, основываясь на дедуктивном методе поиска решений, результатах многочисленных экспериментов, проведенных мною во многих лабораториях, а также на знаниях анатомии, мне удалось обнаружить интересные факты, которыми я намерен поделиться с вами.

На протяжении последнего десятилетия ученые-нейробиологи занимались исследованием идей, подобных моим. Обычно они использовали другую терминологию и, насколько мне известно, не пытались создать на основе полученных разрозненных сведений общую теорию. В отчетах об исследованиях они также упоминают обработку данных по принципу от целого к частному и от частного к целому, говорят о том, как нервные сигналы передаются между сенсорными областями мозга, и отмечают важную роль инвариантных репрезентаций. Так, Габриэль Крейман и Кристофер Кох, нейробиологи, работающие в Калифорнийском технологическом институте, и Айзек Фрейд из Калифорнийского университета в

Лос-Анджелесе открыли нервные клетки, посылающие разряд в том случае, когда человек видит лицо Билла Клинтона. Одна из моих целей – объяснить вам, откуда появились эти *нейроны Билла Клинтона*. Естественно, создание каждой теории сопряжено с формулированием гипотез, которые следует проверять в лабораторных условиях. (В приложении к данной книге я перечислю несколько таких гипотез.) Теперь, когда у нас есть цель и мы знаем, *что* ищем, невероятно сложная, на первый взгляд, система уже не будет казаться нам такой уж замысловатой.

Ниже в этой главе мы детально углубимся в модель «Память-предсказание», характеризующую работу неокортекса. Вначале рассмотрим структуру и наиболее общие функции коры головного мозга, а затем двинемся дальше – обратимся к частным вопросам и проанализируем, как они вписываются в целостную картину.

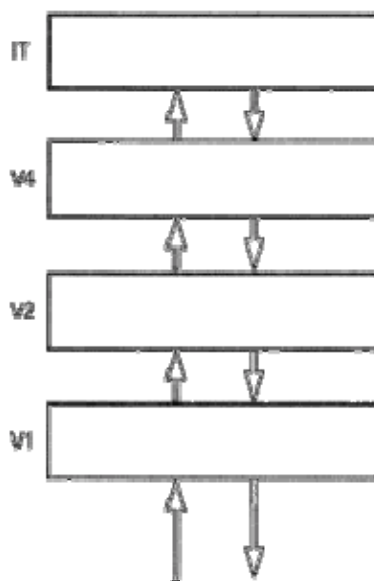


Рис. 6.1. Первые четыре зоны в распознавании объектов

Инвариантные представления

Помните ли вы, как я назвал кору головного мозга тканью из нервных клеток размером со столовую салфетку и толщиной в шесть визитных карточек, в которой связи между разными зонами создают определенную иерархическую структуру? Сейчас я хочу проиллюстрировать иерархическую структуру неокортекса иначе. Представьте, что мы разрезали эту салфетку на несколько частей, соответствующих зонам, специализирующимся на выполнении определенных задач, и наложили полученные кусочки друг на друга. У нас вышел своеобразный слоеный, или «блинный» пирог. Разрезав «пирог», получим срез (см. рис. 6.3). Разумеется, в действительности неокортекс выглядит по-другому, но благодаря этой картинке вы составите четкое представление о передвижении потоков информации. На рис. 6.1 показано, как информационные сигналы поступают в низшую зону коры головного мозга и передвигаются вверх от зоны к зоне. Обратите внимание, что информационные потоки циркулируют в обоих направлениях.

Рис. 6.1 представляет первые четыре зрительные зоны, вовлеченные в распознавание объектов окружающего мира. С их помощью вы узнаете кошку, церковь, свою мать, Большую китайскую стену и так далее. Биологи обозначают эти зоны V1, V2, V4 и IT. Входной визуальный сигнал показан стрелкой под зоной V1. Зрительная информация с сетчатки ваших глаз передается к V1. Этот входной поток информации можно представить как непрерывно изменяющиеся последовательности сигналов, которые передаются приблизительно по миллиону аксонов, составляющих ваш зрительный нерв.

Ранее мы уже упоминали о пространственно-временных сигналах, однако сейчас стоит освежить эти знания, поскольку мы будем часто к ним обращаться. Кора головного мозга

состоит из функциональных зон, специализирующихся на отдельных задачах. Эти зоны связаны друг с другом большими пучками аксонов, или нервных волокон, передающих информацию из одной зоны в другую. Каждую секунду какой-то пучок волокон посылает электрический разряд, называемый потенциалом действия, или импульсом, в то время как другие пучки молчат. Общая активность пучка волокон называется *паттерном* (*последовательностью сигналов*). Паттерн, полученный зоной V1, будет пространственным, если ваши глаза на мгновение остановились на объекте, и временным, если они его разглядывают.

Примерно трижды в секунду ваши глаза осуществляют скачкообразное движение – *саккаду*, после чего следует пауза – *фиксация*. Если бы вы участвовали в эксперименте и исследователь при помощи специального устройства отслеживал движения ваших глаз, вы бы удивились, узнав, насколько резки саккады, при том что воспринимаемое изображение кажется вам постоянным и устойчивым. На рис. 6.2 а показано, как двигались глаза одного из участников эксперимента, когда он рассматривал изображение лица. Обратите внимание, что фиксации вовсе не беспорядочны. А теперь представьте, что вы могли бы видеть паттерн, поступающий к зоне V1 этого человека. Он полностью меняется с каждой саккадой. Значит, несколько раз в секунду зрительная зона коры головного мозга получает совершенно новую последовательность сигналов.



Рис. 6.2. а) Траектория саккад при рассматривании лица человека

Возможно, вы подумаете: «Ну ладно, ведь все равно это лицо остается одним и тем же, несмотря на сдвиги». Доля истины в этом есть, но намного меньшая, чем вы себе представляете. Световые рецепторы на сетчатке глаза распределены неравномерно. Их плотность, очень высокая в центральной ямке, постепенно уменьшается по направлению к периферии глаза. А нервные клетки коры головного мозга, напротив, распределены равномерно. В результате изображение, переданное с сетчатки в первичную зрительную зону V1, очень искажено. Если вы, рассматривая одно и то же лицо, направите взгляд на нос, а потом сконцентрируете его на глазе, зрительный входной сигнал будет очень отличаться, как будто вы смотрите на это лицо через немилосердно дергающийся «рыбий глаз»¹⁵. Тем не менее при взгляде на лицо оно не кажется вам искаженным; изображение не «прыгает» у вас перед глазами. Скорее всего, вы вообще не осознаете, что получаемые от сетчатки визуальные сигналы изменяются, более того – так сильно изменяются. Вы просто видите «лицо» (на рис. 6.2 б показан эффект взгляда на прибрежный пейзаж через линзу «рыбий глаз»). По сути, это переформулирование загадки об инвариантном представлении, о которой мы беседовали в главе 4.

¹⁵ Фотообъектив типа «рыбий глаз» имеет поле зрения, близкое к 180 градусам, внося искажения, которые можно контролировать только пользуясь зеркальным видеоискателем. – Примеч. ред.

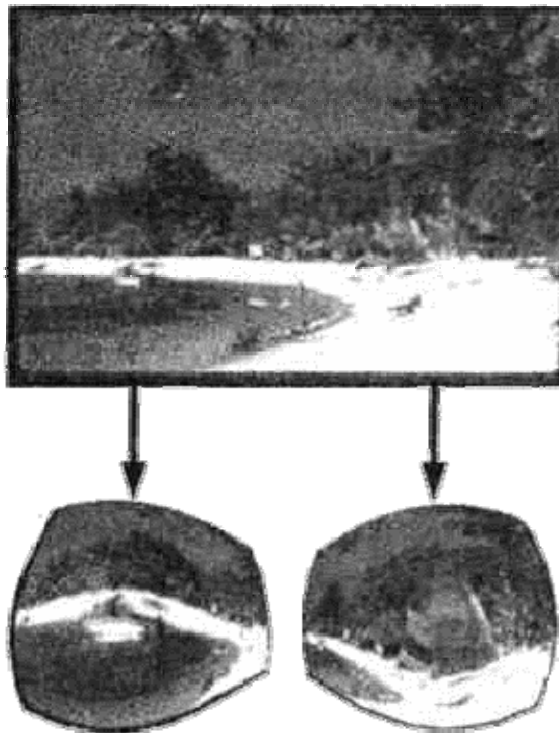


Рис. 6.2. б) искажение, вызванное неравномерным распределением рецепторов в сетчатке глаза

Возможно, вы подумаете: «Ну ладно, ведь все равно это лицо остается одним и тем же, несмотря на сдвиги». Доля истины в этом есть, но намного меньшая, чем вы себе представляете. Световые рецепторы на сетчатке глаза распределены неравномерно. Их плотность, очень высокая в центральной ямке, постепенно уменьшается по направлению к периферии глаза. А нервные клетки коры головного мозга, напротив, распределены равномерно. В результате изображение, переданное с сетчатки в первичную зрительную зону V1, очень искажено. Если вы, рассматривая одно и то же лицо, направите взгляд на нос, а потом сконцентрируете его на глазе, зрительный входной сигнал будет очень отличаться, как будто вы смотрите на это лицо через немилосердно дергающийся «рыбий глаз»¹⁶. Тем не менее при взгляде на лицо оно не кажется вам искаженным; изображение не «прыгает» у вас перед глазами. Скорее всего, вы вообще не осознаете, что получаемые от сетчатки визуальные сигналы изменяются, более того – так сильно изменяются. Вы просто видите «лицо» (на рис. 6.2 б показан эффект взгляда на прибрежный пейзаж через линзу «рыбий глаз»). По сути, это переформулирование загадки об инвариантном представлении, о которой мы беседовали в главе 4.

Вы воспринимаете не то, что видит зона V1. Откуда ваш мозг знает, что вы смотрите на одно и то же лицо, и почему вашему восприятию совершенно не мешает, что входные сигналы постоянно меняются, более того, они очень искажены?

Поместив в зону V1 электрод и наблюдая за «поведением» отдельных клеток, мы обнаружим, что каждый нейрон посылает разряд только в ответ на зрительный сигнал из крошечной части сетчатки. Такой опыт по изучению зрительной функции проводили много раз. Каждый нейрон зоны V1 привязан к своему рецептивному полю, которое является очень ограниченным по сравнению с общим полем зрения. Нервные клетки зоны V1 ничего не знают о лицах, автомобилях, книгах или других объектах, постоянно встречающихся на вашем пути. Все, что они знают, – это крошечная, как укол булавки, часть всего мира перед вашими глазами.

Каждая клетка зоны V1 настроена на восприятие определенных видов входных сигналов.

¹⁶ Фотообъектив типа «рыбий глаз» имеет поле зрения, близкое к 180 градусам, внося искажения, которые можно контролировать только пользуясь зеркальным видеоискателем. – Примеч. ред.

Например, некий нейрон бурно реагирует, когда в его рецептивное поле попадает линия или угол в тридцать градусов. Сам по себе такой угол не имеет особенного значения. Он может быть частью какого угодно предмета – доской паркета, обломком дерева, частью буквы М. При каждой новой фиксации рецептивное поле клетки исследует новую часть видимого пространства. При разных фиксациях поток импульсов, посылаемый клеткой, может быть сильнее или слабее, а может и вовсе отсутствовать, т. е. при каждой саккаде активность клеток в зоне V1 изменяется.

А вот если мы введем электрод в верхнюю зону IT (см. рис. 6.1), то обнаружим нечто совершенно невероятное. Мы увидим, что некоторые клетки указанной зоны возбуждаются и остаются активными, когда в поле зрения человека появляются целые объекты. Например, мы можем найти клетку, энергично реагирующую каждый раз, когда в поле зрения появляется лицо. Эта клетка будет активной до тех пор, пока лицо присутствует в любой точке поля зрения, под любым ракурсом и при любом освещении. Она не включается-выключается при каждой последующей саккаде, как это делают клетки зоны V1. Рецептивное поле такой клетки покрывает большую часть зрительного пространства, и она возбуждается всякий раз, когда человек видит *лица*.

Остановимся на этом подробнее. Какие изменения характеризуют четыре уровня зон коры головного мозга? По мере продвижения от сетчатки до зоны IT изменчивые, пространственно специфические, ориентированные на распознавание мелких деталей нейроны сменяются нейронами высших зон – пространственно неспецифическими, постоянно активными и способными распознавать целые объекты. Нейрон из зоны IT сообщает нам, что где-то в поле нашего зрения присутствует лицо. Эту нервную клетку называют клеткой лица, и она активизируется независимо от того, стоит ли хозяин лица прямо, склонил ли он голову или повернулся к вам в профиль. Это часть инвариантной репрезентации для объекта «лицо».

Да, описать данную схему несложно. Проходим быстренько четыре стадии, и вот вам лицо. Но ни одна компьютерная программа, ни одна математическая формула не в состоянии обеспечить решение подобной задачи с той же надежностью и универсальностью, как это делает человеческий мозг. Нам известно, что мозг решает эту задачу за несколько шагов, значит, разгадка не должна быть слишком сложной. Одной из задач данной главы является поиск объяснения возникновения этой «клетки лица» (а также «клетки Билла Клинтона» и любой другой). Мы достигнем своей цели, но прежде нам следует обработать массу информации. Как говорится, всему свое время.

Как видно из рис. 6.1, по сети обратных связей информация поступает из высших зон неокортекса в низшие. Эти обратные связи представлены пучками аксонов, которые проходят от высшей зоны IT к нижшим зонам V4, V2, V1. Примечательно, что количество обратных связей в зрительной коре головного мозга превышает количество восходящих связей.

Долгие годы ученые-нейробиологи игнорировали эти обратные связи, что вполне объяснимо: если вы ограничиваете свои задачи изучением того, как мозг принимает входную информацию, подвергает ее обработке, а потом на основе этого моделирует поведение, то в рассмотрении обратных связей, казалось бы, совсем нет необходимости. В таком случае фокус внимания исследователей прикован к прямым связям – проведению нервных импульсов от сенсорных зон головного мозга в моторные зоны. Однако если принять за основу идею, что ключевой функцией неокортекса является прогнозирование, то без изучения обратных связей не обойтись. Мозгу необходимо отсылать информацию из высших нервных центров обратно к тем зонам, от которых он ее получил: для того чтобы делать прогнозы, надо сравнивать реальные события и ожидания. То, что на самом деле произошло, идет снизу вверх, а то, чего вы ожидаете, спускается сверху вниз.

Подобный круговой процесс происходит во всех сенсорных зонах коры головного мозга. На рис. 6.3 изображены аналогичные «слоеные пироги» для тактильного, слухового и зрительного восприятия. Также показаны высшие или ассоциативные зоны коры головного мозга, которые получают и анализируют информационные стимулы от нескольких органов чувств, в данном случае – слуха, зрения и осязания.



Рис. 6.3. Формирование инвариантных репрезентаций при осязании, слуховом и зрительном восприятии

Если рис. 6.1 представляет общеизвестную взаимосвязь четырех хорошо изученных зон коры головного мозга, то рис. 6.3 – это условная диаграмма, целью которой *не* является отображение фактических зон неокортекса. В действительности десятки зон коры головного мозга человека взаимосвязаны друг с другом самыми различными способами. Большая часть неокортекса образована ассоциативными зонами. (Я использую схематические рисунки, чтобы помочь вам лучше разобраться в излагаемом материале. И надеюсь, что они не введут вас в заблуждение.)

Переход от быстроменяющихся к медленноменяющимся и от пространственно-специфических к пространственно-инвариантным нейронам был доказан многочисленными результатами исследований зрительного восприятия. В отношении других сенсорных зон данных значительно меньше, но нейробиологи утверждают, что так же устроены и все остальные сенсорные зоны коры головного мозга.

Рассмотрим, например, слуховое восприятие. Когда вы слышите устную речь, изменения звуковых колебаний происходят очень быстро. Так же быстро изменяются паттерны, поступающие в первичную слуховую зону А1. Но если бы мы могли поместить электрод в высшую зону неокортекса, ответственную за слуховое восприятие, то обнаружили бы инвариантные клетки, реагирующие на определенные слова и даже фразы. Вполне вероятно, что в вашей слуховой зоне коры головного мозга существует группа клеток, активизирующихся всякий раз, когда вы слышите слово «спасибо», или группа клеток, реагирующих на приветствие «доброе утро». Такие клетки остаются активными на протяжении всего времени, пока звучит фраза, при условии, конечно, что вы ее воспринимаете. Последовательности сигналов, попадающие в низшую слуховую зону, могут очень сильно отличаться друг от друга. Одно и то же слово можно произнести с разными акцентами, разным тоном и в разном темпе. Чем выше мы поднимаемся по иерархии слуховых зон коры головного мозга, тем менее важными становятся эти низкоуровневые различия. Слово идентифицируется согласно его значению, независимо от акустических эффектов. То же самое справедливо и в отношении музыки. Вы можете услышать одну и ту же песню, сыгранную на фортепиано, на кларнете или в исполнении ребенка. В каждом из перечисленных случаев ваша зона А1 получает совершенно различные звуковые сигналы. Тем не менее электрод, введенный в высшую зону слухового восприятия, подтвердил бы наличие клеток, посылающих нервный разряд всякий раз, когда ваши органы слуха воспринимают данную песню, независимо от инструмента или особенностей исполнения. Отмечу, что такой эксперимент никогда не проводился, это было бы слишком жестоко по отношению к человеку. Но, признавая существование универсального алгоритма функционирования неокортекса, следует признать и наличие подобных клеток. В слуховом восприятии существует точно такая же обратная связь – прогнозирование, построенное на инвариантных представлениях, – как и в зрительном.

Осязание работает по тому же принципу. На людях (по вполне понятным причинам)

соответствующие эксперименты никогда не проводились. В настоящее время ученые занимаются подготовкой к проведению таких опытов с использованием исследовательских инструментов, сканирующих мозг обезьяны.

Сейчас, сидя за столом, я держу в руке ручку. Я могу дотронуться до ее колпачка или провести пальцами вдоль металлической скрепки, с помощью которой ручку можно закрепить в нагрудном кармане. По мере того как я проделываю эти движения разными пальцами или даже губами, последовательности сигналов, поступающие в соматосенсорную зону коры головного мозга от рецепторов осязания, постоянно меняются. Будучи столь непохожими, они попадают в разные части первичной соматосенсорной зоны. Тем не менее в зонах, высших по иерархии, наш воображаемый датчик обнаружил бы клетки, которые реагируют на ручку инвариантно. Эти клетки пребывают в активном состоянии до тех пор, пока я прикасаюсь к ручке, независимо от того, какими пальцами (или даже частями тела) я это делаю.

Вдумайтесь: ваш мозг не способен распознать объект на основании единственного мгновенного сигнала. Паттерн, поступающий от ваших органов чувств в какой-то один момент времени, содержит лишь крошечную толику информации, которой недостаточно для того, чтобы определить, *что* же вы услышали или *к чему* прикоснулись. Чтобы воспринять серию слуховых паттернов (например, мелодию, слово, хлопок закрывающейся двери) или понять, что за объект у вас в руках – та же ручка, – вам необходимо время. Вы не можете узнать мелодию по одной ноте и распознать предьявленный вам объект с первого прикосновения. Следовательно, нейронная деятельность, обеспечивающая сознательное восприятие объектов окружающей действительности, должна быть более продолжительной по времени, чем каждая отдельно взятая последовательность сигналов. Таким образом, мы опять пришли к выводу: чем выше расположена зона коры головного мозга, тем меньше в ней наблюдается временных изменений.

Зрительное восприятие – зависимый от фактора времени информационный поток, действующий по тому же принципу, что и звуковое восприятие или осязание. Но у зрения есть особенность, которая вносит некоторую путаницу: мы все-таки способны распознавать отдельные объекты с одной фиксации. Это умение распознавать пространственные паттерны «с первого взгляда» долгие годы вводило в заблуждение ученых, занимающихся исследованиями зрительного восприятия животных. Они, как правило, не принимали в расчет фактор времени. Действительно, в лабораторных условиях как люди, так и животные способны распознавать объекты таким образом, однако это не является нормой. Нормальное зрительное восприятие, как, например, чтение этой книги, предполагает постоянные движения глазных яблок.

Интеграция разных видов сенсорного восприятия

А как же ассоциативные зоны? Мы уже подробно рассмотрели закономерности прохождения информационных потоков в пределах одной сенсорной зоны коры головного мозга. Сигналы обратной связи дополняют текущие сигналы и дают возможность прогнозировать будущие события. Подобные процессы происходят перекрестно и между разными сенсорными каналами восприятия, т. е. между зрением, слухом и осязанием. Например, на основе услышанного я составляю прогноз о том, *что* сейчас увижу или почувствую. Сейчас, печатая эти строки, я нахожусь в своей спальне. На ошейнике моей кошки Кео есть бубенчик, который позванивает, когда она ходит. И вот я слышу приближающийся звук бубенчика. На основании этого звукового сигнала я догадываюсь, что сейчас увижу свою кошку. И действительно, повернув голову в направлении двери, я вижу Кео. Я ожидал ее увидеть на основе услышанных звуков. Если бы Кео не вошла или вошло бы какое-то другое животное, я бы очень удивился. В приведенном примере звуковой сигнал вызвал звуковое распознавание Кео. Информационный поток прошел все уровни иерархии слуховых зон неокортекса и поступил в ассоциативную зону, связывающую слуховое и зрительное восприятие. На основании инвариантного представления сформировался прогноз, который по каналу обратной связи поступил назад – в зрительную и слуховую зоны. Этот процесс схематически изображен на рис. 6.4.

Мозг постоянно продуцирует полисенсорные прогнозы такого типа. Я отгибаю скрепку на

ручке и ожидаю, что, стоит мне отпустить пальцы, прозвучит характерный щелчок, вызванный ударом скрепки по корпусу ручки. Не услышав этого щелчка, я бы очень удивился. Мой мозг точно прогнозирует, когда я услышу звук и каким именно он будет. Для того чтобы такой прогноз осуществился, информация пересекает по иерархии все соматосенсорные зоны коры головного мозга и по обратной связи поступает назад по иерархии и слуховой зон. В результате описанного процесса возникает ожидание услышать и почувствовать щелчок.

Приведу другой пример: несколько раз в неделю я добираюсь на работу велосипедом. В такие дни я иду утром в гараж, беру свой велосипед, разворачиваю его и вывожу через калитку. Пока я проделываю все указанные манипуляции, в мой мозг поступает множество зрительных и слуховых сигналов. Ежесекундно формируются полисенсорные прогнозы – предположения неокортекса относительно того, *что* я услышу и почувствую в следующее мгновение. Я вижу, как велосипед переезжает порог, и это вызывает у меня ожидание характерного звука. Задев ногой педаль, я наклоняю голову и ожидаю увидеть ее именно около моей ноги. Прогнозы настолько точны, что в случае возникновения малейших отклонений от ожидаемого сценария развития событий я бы тотчас заметил эти отклонения. Информационный поток движется в сенсорной иерархии одновременно в двух направлениях, таким образом формируется унифицированное восприятие, неотъемлемой частью которого является прогнозирование.

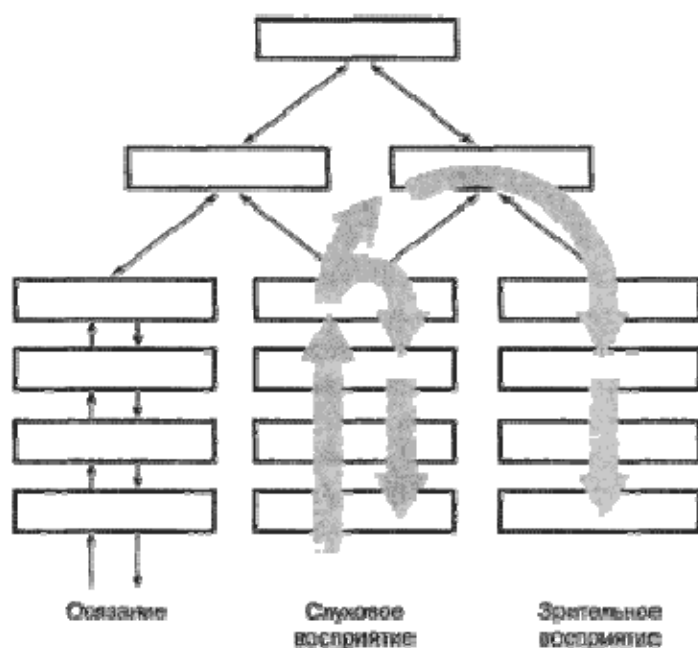


Рис. 6.4. Потоки информационных сигналов передвигаются вверх по иерархически организованным сенсорным зонам, а затем возвращаются к нижним зонам. Таким образом формируется объединенный сенсорный опыт и осуществляется прогностическая функция

Предлагаю вам провести следующий опыт: оторвитесь от чтения и что-нибудь сделайте. Что угодно, лишь бы вы двигались и манипулировали каким-то предметом. Можете, например, подойти к раковине и повернуть ручку крана. Делая это, попробуйте обратить внимание на каждый звук, каждое осязательное ощущение, каждое изменение визуальной «картинки». Будьте внимательны, помните, что каждое движение тесно связано с созерцанием, слухом и осязательными ощущениями. Покрутите ручку крана вперед-назад. Ваш мозг ожидает почувствовать давление на кожу и сопротивление ваших мускулов. Вы ожидаете увидеть и почувствовать прокручивание крана, вы ожидаете увидеть и почувствовать воду. Когда вода долетит до раковины, вы ожидаете услышать другой звук – звук падающей воды, увидеть и почувствовать на своей коже обжигающе горячие (или ледяные – в зависимости от того, какой кран включаете) брызги.

При ходьбе вы всегда – осознанно или нет – ожидаете услышать звук каждого шага. Даже то, что вы просто держите эту книгу в руках, тоже вызывает множество сенсорных прогнозов.

Представьте себе, что вы почувствовали, будто книга закрывается, хотя зрение подсказывает вам совершенно другое: книга остается раскрытой. Наверняка вы испытали бы полное замешательство, вплоть до шока. Помните наш воображаемый эксперимент с изменением двери, описанный в главе 5? Тогда мы убедились, что мозг постоянно занят прогнозированием, причем это касается абсолютно всех сенсорных модальностей. Каждый раз, сосредоточиваясь на всех своих текущих ощущениях, вплоть до самых незначительных, я просто поражаюсь, насколько интегрированными являются все наши перцепционные прогнозы. Хотя они могут казаться простыми, даже тривиальными, надо помнить, насколько они всеобъемлющи. Они могут формироваться только на основе мощных скоординированных потоков информации, непрерывно циркулирующих в двух противоположных направлениях по иерархии зон коры головного мозга.

Осознав столь тесную взаимосвязь чувств, мы делаем вывод, что вся кора головного мозга, все ее сенсорные и ассоциативные зоны действуют как единое целое. Да, у нас есть зрительная область коры, но она – лишь часть единой всеобъемлющей сенсорной системы, объединяющей восприятие изображения, звуков, прикосновений и многое другое. И в этой системе информационные потоки циркулируют вверх и вниз по сложной разветвленной иерархии.

Следует также отметить, что в основе всех наших предположений лежит непосредственный опыт. Я ожидаю, что отведенная скрепка ручки при отпускании ударится о корпус ручки и издаст характерный звук только потому, что так случилось в прошлом. Звук велосипеда предсказуем по той причине, что я вывожу его из гаража далеко не в первый раз. Ни один бит подобной информации не дается нам от рождения. Все эти сведения сохраняются в мозге только благодаря его невероятно огромной способности запоминать входящую информацию. Если среди информации, поступающей в мозг, встречаются постоянные последовательности сигналов, неокортекс сохраняет их, а в будущем использует для прогнозирования событий.

Хотя на рис. 6.3 и 6.4 не изображена моторная зона коры головного мозга, у нее такая же иерархическая структура, как и у описанных выше сенсорных зон. Ее можно представить как подобную сенсорную систему, но соединенную с органами чувств посредством ассоциативных зон (только с более тонкими связями с соматосенсорной зоной, что обеспечивает способность двигаться). Таким образом, моторная зона коры головного мозга функционирует подобно сенсорным зонам; кроме того, они активно взаимодействуют. Сигнал, поступающий от любого сенсорного анализатора, перемещается вверх к ассоциативной зоне, что вызывает перемещение другого импульса вниз, к моторной зоне коры головного мозга, в результате чего возникает определенное поведение. Подобно тому, как зрительный сигнал может вызвать поступление импульса в осязательную и слуховую зоны, он также может активизировать моторную зону коры головного мозга. В первом случае мы интерпретируем обратные сигналы как прогнозы, во втором – как двигательные команды. Как заметил Маунткастл, моторные зоны коры устроены так же, как и сенсорные. Кора генерирует сенсорные прогнозы так же, как моторные команды.

Вскоре мы убедимся, что в неокортексе нет «чисто сенсорных» и «чисто моторных» зон. Сенсорные потоки одновременно поступают отовсюду и, возвращаясь потом по иерархии зон, формируют прогнозы или генерируют моторные команды. Хотя у моторной зоны коры головного мозга есть свои особенности, вполне корректно считать ее частью общей иерархической запоминающе-прогностической системы. Ее можно считать еще одним органом чувств. Зрение, осязание, слух и действие тесно переплетены друг с другом.

Новый взгляд на зону V1

Продолжая разбираться в запутанной архитектуре коры головного мозга, попробуем по-новому взглянуть на зоны неокортекса. Нам уже известно, что высшие зоны коры головного мозга формируют инвариантные представления, но почему столь важная функция должна осуществляться только в высших зонах? Не забывая о понятии симметрии, сформулированном Маунткастлом, я занялся изучением различных способов связи между зонами коры головного мозга.

На рис. 6.1 показаны четыре зоны зрительного восприятия: V1, V2, V4 и IT. Как правило, каждая из них считается отдельной и непрерывной: все клетки зоны V1, обрабатывая каждая свою область зрительного поля, делают одно и то же, все клетки зоны V2 выполняют одинаковые задачи, все клетки из V4 одинаково специализированы.

Как же выглядит распознавание лица с точки зрения традиционного подхода? Когда изображение лица фиксируется клетками, образующими зону V1, вначале создается грубый «эскиз» лица, состоящий из линий и других простейших элементов. Затем «набросок» поступает в зону V2, где происходит более сложный анализ черт лица. Модифицированная информация передается в зону V4 и так далее. Распознавание объекта осуществляется только в зоне IT, которой свойственна инвариантность представлений.

К сожалению, при таком подходе возникает масса вопросов. Почему инвариантные репрезентации действуют только в зоне IT? Если все зоны коры головного мозга выполняют сходные функции, чем именно примечательна эта? И еще: лицо может быть воспринято правой или левой частью зоны V1, и в обоих случаях мозг его распознает. В то же время экспериментальным путем было установлено, что между этими частями зоны V1 нет непосредственного взаимодействия. Левая часть зоны V1 не знает, что видит правая. Давайте подумаем, какой вывод можно сделать. Судя по всему, разные части зоны V1 выполняют сходные функции, поскольку участвуют в распознавании лица. И в то же время физически они независимы друг от друга. Подзоны, или кластеры зоны V1, физически не связаны, однако выполняют одну и ту же функцию.

И наконец, результаты исследований показывают, что все высшие зоны коры головного мозга получают объединенные сигналы от двух и более сенсорных зон, расположенных ниже в иерархии (рис. 6.3). Данный рисунок является схематическим, в действительности каждая ассоциативная зона может получать сигналы из десятка и более низших зон. Согласно традиционным представлениям, взаимосвязь между низшими зонами (например, V1, V2, V4) иная. Каждая из них якобы имеет единственный источник входящей информации, единственный поток, поступающий из низшей зоны, который не имеет отношения к потокам информации, поступающим из других зон. (Например, принято считать, что зона V2 получает информационный поток от V1, и ее работа состоит только в этом.) Но почему тогда одни зоны коры головного мозга получают объединенную информацию, а другие – нет? Это еще одно несоответствие с гипотезой универсального алгоритма неокортекса, сформулированной Маунткастлом.

По этой и другим причинам я пришел к выводу, что не следует рассматривать области V1, V2, V4 как цельные зоны. Каждая из них образована более мелкими подзонами. Давайте вернемся к воображаемой обеденной салфетке (аналогии с распластанной корой головного мозга). Допустим, мы решили отметить ручкой различные функциональные зоны. Самой большой будет зона V1 – первичная зрительная зона. Следующей по размеру – зона V2. По сравнению со всеми остальными зонами эти две просто огромны. Я веду к тому, что V1 нужно рассматривать как эквивалент многих малых зон. Вместо того чтобы обвести большую зону на нашей салфетке, мы нарисуем множество маленьких зон, которые и займут территорию, принадлежащую V1. Другими словами, V1 состоит из многих малых зон коры головного мозга, которые связаны друг с другом лишь опосредованно, через более высокие зоны иерархии. В V1 входит наибольшее количество подзон по сравнению с другими зонами зрительного восприятия. V2 состоит из меньшего количества зон, но они имеют большие размеры, чем подзоны V1. То же самое утверждение действительно и для V4. Вот когда вы доберетесь до IT, она действительно окажется однородной и унифицированной. Этим объясняется тот факт, что именно клетки зоны IT формируют общую картину всего визуального восприятия окружающего мира.

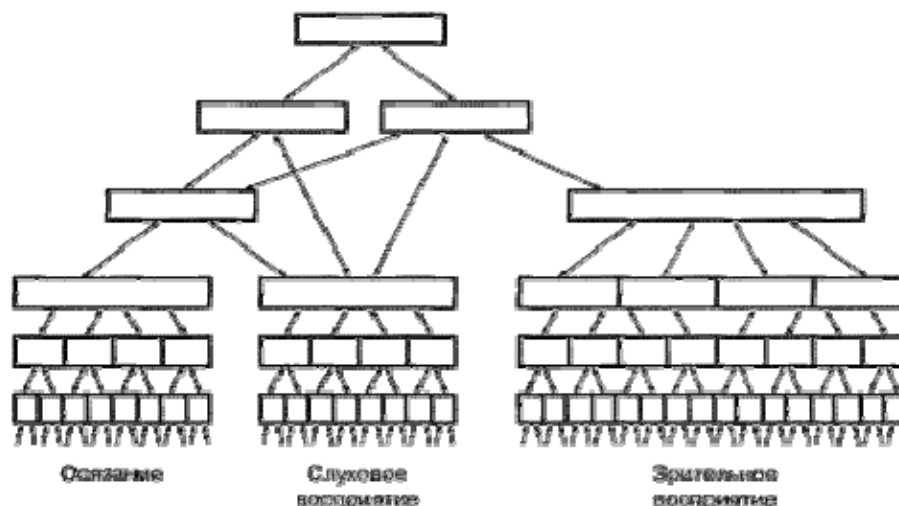


Рис. 6.5. Альтернативный подход к рассмотрению иерархии коры головного мозга

Взгляните на рис. 6.5. Не правда ли, занятная симметрия? Здесь изображена та же иерархия, что и на рис. 6.3, но с учетом только что приведенной информации. Обратите внимание, что кора головного мозга везде устроена одинаково. Любая из зон получает информацию от многих других зон низшего уровня. Эта же получающая зона направляет вниз по иерархии сигналы о предполагаемых событиях. Высшие ассоциативные центры интегрируют информацию, полученную от нескольких органов чувств, например, зрительную и осязательную. Зоны низшего уровня, например, подзоны, входящие в зону V2, интегрируют информацию, воспринимаемую несколькими отдельными подзонами, входящими в V1. Зона в целом не интерпретирует – она и не может этого делать – каждый из отдельно взятых сигналов. Подзоне, входящей в V2, совершенно необязательно знать, что она получает входные сигналы из нескольких частей зоны V1. Ассоциативной зоне совершенно необязательно различать входящие в нее сигналы от зрительного и слухового анализаторов. Задача каждой зоны коры головного мозга состоит, скорее, в том, чтобы выяснить, как связаны между собой входящие сигналы, запомнить последовательность взаимосвязей между ними и в будущем использовать эту информацию для прогнозирования. Кора головного мозга всегда остается корой головного мозга. Это универсальный алгоритм функционирования неокортекса: во всех зонах происходят одни и те же процессы.

Новая иерархическая схема поможет нам понять, как формируются инвариантные представления. На низшем уровне обработки визуальных сигналов левая часть зрительного поля отличается от правой части зрительного пространства так же сильно, как зрительное восприятие отличается от слухового. Левая и правая части зоны V1 формируют похожие репрезентации только потому, что получили сходный предшествующий «жизненный опыт». Их можно рассматривать как два независимых сенсорных потока, такие же, как слуховое и зрительное восприятие, которые объединяются в зонах более высокого порядка.

Подобно этому, более мелкие зоны в составе V2 и V4 являются ассоциативными зонами зрительного восприятия. (Подзоны могут пересекаться, но это не меняет функционирование зон радикально.) Рассмотрение коры головного мозга под таким углом никоим образом не противоречит нашим знаниям о ее анатомическом строении. Информационные потоки передвигаются по всем восходящим и нисходящим ветвям иерархического дерева системы памяти. Сигнал, поступивший в левую часть зрительного поля, может привести к формированию прогнозов для правой части зрительного поля, как звук колокольчика на шее кошки стал причиной визуального прогноза о том, что она войдет в спальню.

Главная польза нашего нового схематического изображения заключается в том, что теперь мы можем быть уверены: в формировании инвариантных репрезентаций участвует каждая зона коры головного мозга. Согласно традиционным взглядам, о полных инвариантных представлениях речь не шла до тех пор, пока сигнал не попадал в высшие зоны неокортекса (такую, например, как зрительная зона IT, в которой формируется представление об общей

наблюдаемой визуальной картине). Теперь мы установили, что на самом деле инвариантные репрезентации вездесущи – они формируются каждой зоной коры головного мозга. Инвариантность – не чудо, возникающее в высших зонах. Каждая зона формирует инвариантные репрезентации на основе данных, получаемых из зон, расположенных ниже по иерархической лестнице (каждая из которых «видит» лишь небольшую часть окружающего мира). Можно сказать, что уровень восприятия в низших зонах значительно беднее такового в высшей зоне IT, однако все они выполняют сходные функции: V4, V2, V1 выполняют ту же работу, что и IT Ассоциативные зоны, расположенные выше IT, формируют инвариантные представления, основанные на информации, полученной от нескольких сенсорных анализаторов. Следовательно, каждая из зон коры головного мозга создает инвариантные представления исходя из того, что находится ниже в иерархии. Это красиво.

Наша головоломка изменилась. Нам больше не нужно выяснять, каким образом за четыре шага формируются инвариантные представления. Но теперь мы должны разобраться, как формируются инвариантные репрезентации во всех зонах коры головного мозга. Если мы признаем существование универсального алгоритма коры головного мозга, то такая постановка вопроса обоснованна. Если в одной зоне хранятся последовательности сигналов, то и в остальных зонах тоже. Если одна зона создает инвариантные репрезентации, то и остальные зоны тоже. Новая схема иерархии коры головного мозга, представленная на рис. 6.5, поможет нам глубже понять суть данного подхода.

Модель окружающего мира

Почему кора головного мозга устроена иерархически?

Вы обладаете способностью размышлять о мире, передвигаться в окружающем пространстве и прогнозировать будущее, потому что в вашем неокортексе сформирована модель внешнего мира. Одна из наиболее важных концепций данной книги состоит в том, что иерархическая структура коры головного мозга хранит модель иерархического строения внешнего мира. Вложенная структура реального мира отображается вложенной структурой коры вашего головного мозга.

Что я подразумеваю под вложенной, или иерархической, структурой? Обратимся к музыке. В ней ноты объединяются в группы, которые, в свою очередь, образуют мелодические фразы. Так рождается мелодия или песня. Из песен состоят альбомы. В качестве второго, не менее удачного, примера можно привести письменность. Буквы складываются в слоги, которые сочетаются в слова. Из слов состоят предложения и фразы. Хотите еще один пример? Вспомните, *что* у вас расположено по соседству – дороги, школы, дома? В каждом доме есть комнаты. В каждой комнате есть стены, потолок, пол, дверь, одно или больше окон. Каждая из этих составляющих, в свою очередь, состоит из более мелких частей. Окно состоит из стекла, рамы, ручки, щеколды. Щеколда состоит из еще более мелких частей, таких как винты.

Осмотрите вокруг. Сигналы от вашей сетчатки, поступающие в первичную зону зрительного восприятия, сочетаются в линейные сегменты. Линейные сегменты объединяются в более сложные формы. Эти формы объединяются в еще более сложные образы, такие, например, как нос. Носы вкупе с глазами и ртами головной мозг объединяет в человеческие лица. Из лица, дополненного остальными частями тела, у вас складывается образ человека, сидящего напротив.

Все объекты окружающего мира состоят из подобъектов, составляющих одно целое. Эти составляющие отдельных образов и являются их определениями. Называя что-либо определенным словом, мы подразумеваем, что существует набор характеристик, связанных с этим определением. Лицо называется лицом, потому что на нем всегда присутствуют глаза, рот, нос. Глаз является глазом, потому что в нем всегда присутствуют зрачок, радужная оболочка, ресницы, веки и так далее. То же самое можно сказать и о стульях, машинах, деревьях, парках, странах. Песня называется песней благодаря определенной последовательности звуков.

Весь мир, окружающий нас, – это своего рода песня. Каждый объект в нем состоит из множества более мелких, а большинство малых объектов являются частью больших. Именно это я и подразумеваю под вложенной структурой, характерной практически для всех объектов

окружающей среды. Аналогичным образом, информация об объектах и способы их представления сохраняются в иерархической структуре коры головного мозга. Информация о том, что такое здание, хранится не в единственной зоне неокортекса, а в иерархии зон. Соответственно, эта информация отражает иерархическую структуру понятия «здание». Фундаментальные взаимосвязи сохраняются в верхней части иерархии зон, а более специфичные – ближе к «подножию» иерархической лестницы. Иерархические связи внешнего мира подтверждаются в ходе процесса обучения. Никому из нас не даются от природы знание языка, образов окружающего мира, музыки. Но в кору нашего головного мозга встроен разумный самообучающийся алгоритм, который естественным образом выявляет иерархические структуры и запоминает их. Если никакой структуры нет, мы испытываем смятение и считаем сложившуюся ситуацию хаотичной.

В каждый отдельно взятый момент времени вы можете воспринимать только определенную часть внешнего мира. Вы можете находиться только в одной комнате вашего дома и смотреть только в одном направлении. Благодаря иерархическому строению коры головного мозга вы в состоянии осознавать, что находитесь дома, в своей гостиной, и что вы смотрите в окно, даже если в данный момент ваш взгляд остановился на щеколде. Высшие зоны коры головного мозга поддерживают инвариантную репрезентацию вашего дома, более низкие – репрезентацию комнаты, а первичные – анализируют непосредственные визуальные символы, в данном случае – происходящее за окном. Точно так же иерархическое строение неокортекса позволяет вам осознавать, что вы прослушиваете отдельную песню из определенного музыкального альбома, несмотря на то что в каждый отдельно взятый момент вы слышите только одну ноту, которая сама по себе ничего для вас не значит. Благодаря иерархическому строению неокортекса вы понимаете, что беседуете со своим лучшим другом, хотя в данный конкретный момент вы смотрите только на его ладонь. Высшие зоны коры головного мозга выполняют стратегическую функцию – следят за ходом развития событий в крупных масштабах, в то время как более низкие зоны сосредоточены на восприятии скоротечных аспектов каждой конкретной ситуации.

Поскольку в некий отдельно взятый момент времени вы можете видеть, слышать и ощущать только очень небольшую часть внешнего мира, информация поступает в мозг в виде последовательностей сигналов. Кора головного мозга особенно чувствительна к тем последовательностям, которые постоянно повторяются, и стремится их запомнить. В некоторых случаях, таких как, например, звучание мелодии, последовательность сигналов, поступающая в ваш головной мозг, определенным образом организована. Надеюсь, что концепция последовательностей сигналов понятна любому из вас. Но я собираюсь использовать термин *последовательность* в более широком смысле, приближающемся к математическому термину *ряд*. Последовательность – это ряд сигналов, которые следуют друг за другом, но не всегда в строгом порядке.

Для большей наглядности предлагаю вам рассмотреть примеры. Когда я смотрю на ваше лицо, последовательность входных сигналов, поступающих в кору головного мозга, не остается неизменной, она определяется саккадами моих глаз. Один раз мой взгляд движется по пути «глаз-глаз-нос-рот», а через мгновение порядок изменяется: «рот-глаз-нос-глаз». Части лица – это последовательности. Они статистически связаны и встречаются близко во времени, хотя их порядок может меняться. Если вы воспринимаете «лицо», фиксируя взгляд на «носу», наиболее вероятным следующим паттерном будет «глаз» или «рот», а не «ручка» или «машина».

В каждую область коры поступает поток таких паттернов. Если зона коры может предсказать, какой из паттернов последовательности будет следующим, она формирует постоянное представление этой последовательности – запоминает ее. Запоминание последовательностей это самый главный компонент формирования инвариантных репрезентаций объектов реального мира. Эти объекты могут быть конкретными, как ящерица, лицо или дверь, или абстрактными, как слово или теория. Мозг рассматривает абстрактные и конкретные объекты одинаково. И те, и другие – всего лишь предсказуемые последовательности паттернов, которые время от времени повторяются. Именно благодаря повторяемости зона коры головного мозга определяет, что она имеет дело с объектом реального мира.

Предсказуемость – это самое сердце реальности. Зона коры может обнаружить, что во время физических движений (таких как саккады глаз или движения пальцев) паттерны гарантированно и предсказуемо сменяют друг друга, или же она может их предсказывать по мере того, как они разворачиваются во времени (как, например, звуки, из которых состоит слово или мелодия). В этом случае мозг решает, что здесь есть причинная связь. Вероятность того, что многочисленные паттерны будут неоднократно встречаться вместе чисто случайно, ничтожно мала. Предсказуемая последовательность паттернов должна быть частью реального объекта, который существует на самом деле. Поэтому возможность надежно предсказывать – «железный» способ узнать, что какие-то события в мире на самом деле связаны. Нос, глаза, уши и рот есть у любого лица. Если мозг воспринимает глаз, после саккады – еще один, после следующей саккады – рот, значит, он точно имеет дело с лицом.

Если бы зоны коры могли говорить, мы бы услышали что-то вроде: «Я воспринимаю много разных паттернов. Иногда я не могу предсказать, какой из них будет следующим. Но между ними точно есть связь. Я всегда наблюдаю их вместе и уверенно могу перескакивать от одного к другому. Поэтому я дала этим событиям общее имя и употребляю его всякий раз, когда мне попадается любое из них. Именно это групповое имя, а не разрозненные паттерны, я передаю высшим зонам коры».

Следовательно, можно сказать, что мозг хранит последовательности последовательностей. Каждая зона коры выделяет последовательности, дает им то, что я называю «именами», и передает эти имена зонам, которые расположены выше в иерархии.

Последовательности последовательностей

Мы видим все меньше и меньше временных изменений информации по мере того, как она поднимается по иерархии зон коры головного мозга. В первичных визуальных зонах вроде V1 совокупность активных клеток быстро изменяется по мере того, как несколько раз в секунду на сетчатку поступают новые паттерны. А вот в визуальной зоне IT возбужденные клетки более стабильны. Что же происходит? Каждая область коры владеет определенным репертуаром последовательностей, подобным репертуару певца. Последовательности-песни могут быть о чем угодно: как шумит прибой, как выглядит лицо вашей матери, как пройти от вашего дома к магазину на углу, как произносится слово «попкорн», как тасовать колоду карт.

Мы даем имена песням, и каждая зона коры тоже дает имена всем последовательностям, которые ей известны. Это «имя» – группа клеток, чья коллективная активность обозначает объекты последовательности (пока не думайте о том, как выбирают клетки для того, чтобы обозначать последовательность, – с этим мы разберемся чуть позже). Означенные клетки будут активны, пока длится последовательность, и именно это «имя» передается вверх по иерархии. До пор пока входные сигналы являются частью прогнозируемой последовательности, зона коры головного мозга посылает к высшим зонам одно и то же «имя».

Зона, принимающая входные сигналы, словно отправляет вверх сообщение: «Вот имя последовательности, которую я вижу, слышу или осязаю. Отдельные ноты, контуры или качество поверхности не имеют значения. Я обязательно вам сообщу, если произойдет что-то новое или непредвиденное». Так, например, высшая зона зрительной иерархии сообщает ассоциативной зоне: «Я вижу лицо. Каждая саккада дает фиксацию на различных частях лица. Я вижу различные части лица в определенной последовательности, но это точно одно и то же лицо. Я вам сообщу, если встретится что-то новое». Таким образом, предсказуемой последовательности событий соответствует «имя» – постоянный паттерн импульсов от возбужденных клеток. Продвигаясь вверх по иерархической пирамиде, мы будем видеть это снова и снова. Одна зона может заниматься распознаванием звуков, из которых состоят фонемы: она посылает сигнал, представляющий фонемы, – в более высокую зону иерархии. Следующая зона распознает последовательность фонем и формирует из них слова. Следующая, более высокая, зона распознает слова и формирует из них фразы и так далее. Не забывайте, что в более низких зонах «последовательность» может быть относительно простой, например, кромка чего-либо, передвигающаяся в пространстве.

Давая имена прогнозируемой последовательности сигналов в каждой зоне иерархии, мы

получаем все более и более высокую стабильность наверху иерархии. Так создаются инвариантные репрезентации.

Обратный процесс происходит, когда сигнал перемещается по обратной связи вниз по иерархии: устойчивые репрезентации «разворачиваются» в последовательности сигналов. Предположим, вы решили вспомнить содержание Геттисбергской речи Линкольна, которую учили в седьмом классе. В одной из верхних зон коры вашего головного мозга хранится модель, которая является репрезентацией известной речи Линкольна. Сначала модель разветвляется на последовательности фраз. Зоной ниже каждая фраза разворачивается в последовательности слов. На этом этапе сигнал расщепляется и движется в зрительную и моторную зоны коры головного мозга. Спускаясь по моторному пути, каждое слово расщепляется на усвоенную последовательность фонем. В конце концов, на самом низком уровне, каждая фонема расщепляется на последовательность артикуляционных команд, обеспечивающих произнесение звуков. Чем ниже по иерархии мы спускаемся, тем быстрее меняются сигналы. Одна цельная репрезентация, хранящаяся в высшей зоне иерархии, преобразуется в продолжительную и сложную последовательность звуков речи.

По мере того как информационный поток спускается вниз по иерархии, инвариантность приобретает особое значение. Если вы хотите напечатать Геттисбергскую речь, не произнося ее вслух, то начнете с той же самой репрезентации из высшей зоны. В следующей зоне сигнал разделится на фразы, потом – на слова. До этого момента никаких различий не наблюдается. Но далее моторная зона коры головного мозга избирает иной путь. Слова расщепляются на буквы, а затем формируются мышечные команды, поступающие к вашим пальцам с тем, чтобы вы начали печатать. Итак, усвоенная вами речь Линкольна представлена в головном мозге в виде инвариантной репрезентации. Нет никакой разницы в том, будете ли вы ее произносить, печатать на компьютере или записывать от руки. Обратите внимание: вам не нужно запоминать речь дважды (раз для письменного воспроизведения и второй – для устного произношения). Единоразово усвоенная информация о речи Линкольна может привести к различным поведенческим проявлениям. В любой зоне коры головного мозга инвариантный сигнал может разветвляться и передвигаться вниз по иерархии различными путями.

Как еще одно проявление эффективной работы головного мозга, репрезентации простых объектов из нижней части иерархии корковых зон могут использоваться повторно для формирования разных последовательностей сигналов в высших зонах. Например, нам не нужно запоминать один набор слов и фонем для воспроизведения вслух Геттисбергского обращения и совершенно другой набор – для речи Мартина Лютера Кинга «У меня есть мечта», хотя эти два выступления имеют некоторые общие слова. Иерархия вложенных последовательностей обеспечивает повторное использование объектов более низкого уровня (например, слов, фонем, букв). Это исключительно эффективный способ хранения информации о внешнем мире и его структуре, и он очень сильно отличается от принципов работы компьютера.

В сенсорных областях происходит такое же разворачивание последовательностей, как и в моторных зонах. Этот процесс позволяет вам ощущать и воспринимать все многообразие объектов окружающей среды. Когда вы подходите к холодильнику за порцией мороженого, зрительная зона коры головного мозга активизируется на многих уровнях. На высшем уровне вы постоянно «видите» холодильник. На более низких уровнях зрительное ожидание подразделяется на серию более специфических, локализованных на отдельных объектах зрительных сигналов. Зрительное восприятие холодильника складывается из фиксаций взгляда на ручке дверцы, самой дверце, прикрепленных магнитах, детских рисунках и так далее. За несколько миллисекунд, которые проходят в период саккады от одного свойства холодильника к другому, прогнозы результатов каждой саккады устремляются вниз по зрительной иерархии. До тех пор пока прогнозы от саккады до саккады оправдываются, высшие зоны зрительного восприятия твердо уверены в том, что вы действительно видите перед собой свой холодильник. Обратите внимание, что в отличие от строгого порядка слов в Геттисбергском обращении, та последовательность паттернов, с которой вы имеете дело в случае с холодильником, может оказаться изменчивой. Поток входных сигналов и прогнозы о них зависят от ваших действий. Значит, в случаях, подобных этому, разветвляющиеся сигналы не являются стабильной последовательностью, но по сути своей процесс все равно остается одним и тем же:

стабильные, мало меняющиеся репрезентации высшего уровня, которые разветвляются в быстро меняющиеся паттерны низших корковых зон.

Тот способ, которым мы запоминаем последовательности и представляем их под каким-либо именем во время передвижения информационных потоков вниз и вверх по иерархии коры головного мозга, чем-то напоминает иерархию армейских команд. Генерал отдает приказ: «На зиму перебросьте военные части во Флориду». Простая команда высшего уровня по мере ее передвижения вниз по армейской иерархии расщепляется на все более детализированные последовательности команд. Подчиненные генерала знают, что команда требует определенной последовательности шагов, таких как приготовления к передислокации, транспортировка во Флориду, подготовка к обустройству на новом месте. Каждый из этих шагов разбивается, в свою очередь, на еще более специфические задания, которые адресуются подчиненным. На нижней ступени иерархии будут стоять тысячи рядовых солдат, предпринимающих десятки тысяч физических шагов для того, чтобы военные части были переведены во Флориду. Что касается обратной связи, то на каждом уровне формируются отчеты о проделанной работе. В результате генерал получает доклад: «Передислокация во Флориду осуществлена успешно», но при этом не вникает во все подробности инициированного им процесса.

У данного правила есть исключение. Если возникает проблема, которая не может быть разрешена подчиненными на соответствующем низком уровне, запрос передается к более высоким уровням иерархии до тех пор, пока решение не будет найдено. Но офицер, который знает, как решить возникшую проблему, не рассматривает ее как исключение. То, что является неожиданной проблемой для его подчиненных, для него самого является следующим по списку заданием – он лишь отдает подчиненным новые команды. Кора головного мозга функционирует по аналогичному принципу. Когда происходит неожиданное событие (поступают непрогнозируемые сигналы), информация передается в более высокие по иерархии зоны коры головного мозга до тех пор, пока возникшая трудность будет разрешена. Если более низкие зоны не в состоянии спрогнозировать поступающие сигналы, они классифицируют их как ошибку и передают в высшие «командные центры». Процедура повторяется до тех пор, пока следующая зона не сможет распознать поступающий сигнал.

Каждая зона коры головного мозга устроена так, что старается сохранять и активизировать последовательности сигналов. Однако такое представление о мозге является слишком упрощенным. Давайте немного усложним нашу модель.

Сигналы поступают в зоны коры головного мозга по тысячам или даже миллионам аксонов. Эти аксоны идут от самых разных областей и по ним передаются самые разные последовательности сигналов. Потенциальное количество сигналов на одной тысяче аксонов превышает количество молекул во Вселенной. Каждая зона коры головного мозга на протяжении всего времени своего существования воспринимает лишь небольшую часть этого невероятного количества сигналов.

Возникает вопрос: из чего состоят последовательности сигналов, сохраняющиеся в различных зонах неокортекса? Я предполагаю, что сначала зона классифицирует входные сигналы на основе ограниченного числа возможных вариантов, а потом формирует последовательности.

Представьте себе, что вы – зона коры головного мозга. Ваша задача состоит в том, чтобы рассортировать цветные листы бумаги. У вас есть десять ведер, каждое из которых помечено определенным цветом. Одно ведро предназначено для листов зеленого цвета, другое – для красных, следующее – для желтых и так далее. Каждый из предлагаемых вам листов немного отличается по цвету. Поскольку количество существующих в мире цветов не ограничено, то не найдется даже двух листов бумаги одинакового цвета. Иногда очень легко определить, в какое из ведерок нужно отправить лист, а порой наоборот – очень трудно. Лист, окрашенный в промежуточный между красным и оранжевым цвет, можно бросить и в то, и в другое ведро, но вам нужно выбрать только одно, даже если вы сделаете это наугад. Задача данного упражнения состоит в том, чтобы показать, как мозг должен классифицировать сигналы. Зоны коры головного мозга именно этим и занимаются, с той лишь разницей, что у них не имеется контейнеров, в которые они бы помещали сигналы.

А сейчас вам выпадает шанс найти последовательность. Вы замечаете, что листы очень часто следуют в таком порядке: красный, красный, пурпурный, оранжевый, зеленый. Вы присваиваете этой последовательности имя «ккпоз». Обратите внимание, что распознавание какой бы то ни было последовательности было бы невозможным, если бы вы сначала не рассортировали все листы по десяти ведам. Не рассортировав листы по десяти из возможных категорий, вы бы не смогли определить, что две последовательности одинаковы.

Сейчас вы собираетесь проанализировать все входные сигналы – цветные листы бумаги, соответствующие в нашей аналогии сигналам, получаемым из низших зон коры головного мозга. Вы сортируете их и пытаетесь обнаружить последовательности. Оба шага – сортировка и формирование последовательностей – являются необходимыми для создания инвариантных репрезентаций, и любая зона коры головного мозга именно этим и занимается.

Процесс формирования последовательностей окупается с лихвой, когда входные сигналы неоднозначны, как, например, лист бумаги, который можно назвать и красным, и оранжевым. Допустим, вам предстоит определиться, какое ведро выбрать для этого листа, даже если вы не уверены наверняка в своем выборе. Зная наиболее вероятную последовательность для серии входных сигналов, вы сможете использовать эти знания для классификации неоднозначных входных сигналов. Если вы считаете, что попали на «ккпоз»-последовательность, поскольку вам попались два красных, зеленый и пурпурный цвета, то у вас формируется ожидание, что следующий цвет будет оранжевым. Но оказывается, что следующий лист совсем не оранжевый, он, скорее, нечто между красным и оранжевым. Вполне может быть, что он даже более близок к красному цвету, чем к оранжевому. Вы же ожидали «ккпоз»-последовательность, вы знакомы с ней, поэтому положите лист в ведро, предназначенное для оранжевых листов. В сомнительных случаях вы наверняка воспользуетесь контекстом известных последовательностей.

Этот феномен постоянно подтверждается в повседневной жизни. Когда люди разговаривают, то часто «глотают» отдельные слова, понять которые вне контекста было бы невозможно. Тем не менее такая особенность устной речи не мешает собеседникам прекрасно понимать друг друга. Бывает, невозможно прочесть отдельно взятое написанное от руки слово, смысл которого проясняется при прочтении целой фразы. В подавляющем числе таких случаев вы не осознаете, что восполняете недостающую информацию, исходя из усвоенных ранее последовательностей сигналов. Вы слышите то, что ожидаете услышать, и видите то, что ожидаете увидеть, тогда, когда услышанное и увиденное вами соответствует прошлому опыту.

Обратите внимание, что запоминание последовательностей не только облегчает интерпретацию неоднозначных входных сигналов, но и позволяет спрогнозировать, какой сигнал окажется следующим. Высшие зоны коры головного мозга сообщают зонам более низкого порядка, чего им следует ожидать дальше. Возвращаясь к примеру с сортировкой бумаги, вы могли бы сказать человеку, передающему вам листы: «Послушай, если ты не можешь никак решить, какого цвета лист подавать мне следующим, то, насколько я помню, это должен быть оранжевый». Распознавая последовательности сигналов, вышестоящая зона коры предсказывает, каким будет следующий сигнал, и сообщает зонам нижестоящим, чего им следует ожидать.

Зона коры головного мозга не только усваивает определенные последовательности, но и модифицирует уже имеющиеся классификации. Допустим, вы начали заполнять ведра зеленого, желтого, красного, пурпурного и оранжевого цветов. Вы уже приготовились к поступлению листов в последовательности «ккпоз» (или похожей комбинации этих цветов). Но что, если ваши ожидания не оправдаются? Что, если всякий раз, когда вы столкнетесь с подобием данной последовательности, вместо пурпурного вам предоставят лист цвета индиго? Скорее всего, вы измените надпись на соответствующем ведре с «Пурпурного» на «Индиго». Так вы приспосабливаетесь к меняющимся обстоятельствам и уходите от двусмысленности. И все это благодаря поразительной гибкости неокортекса.

Подобные изменения уже сложившихся классификаций происходят в зонах коры головного мозга на протяжении всей вашей жизни. Данный процесс осуществляется на основе взаимодействия последовательностей сигналов; он же положен и в основу обучения. Все зоны коры головного мозга очень пластичны, т. е. информация, сохраняющаяся в них, модифицируется под влиянием прошлого опыта и восприятия текущей ситуации. Создавая

новые классификации и новые последовательности, вы запоминаете мир.

И наконец, давайте обсудим, как эти классификации и прогнозы взаимодействуют с высшими зонами коры головного мозга. Здесь следует представить вам еще одну функцию коры головного мозга – трансляцию воспринятого в зону на уровень выше (в нашем случае это будет передача вверх по иерархии листа с отметкой «ккпоз»). Для следующего уровня буквы сами по себе мало что означают. Название – это просто сигнал, подлежащий последующей обработке и сопоставлению с остальными сигналами, классификации и последующему размещению в последовательности более высокого порядка. Точно так же, как и вы, эта высшая зона следит за последовательностями, которые воспринимает. В какой-то момент она может обратиться к вам: «Эй, послушайте, если вы не знаете, *что* мне подавать следующим, то, насколько мне известно, это должна быть последовательность «жжкзж». Фактически, это прямое указание на то, *что* вы должны искать в данный момент в своем потоке входящих сигналов. Вы будете изо всех сил стараться распознать составляющие компоненты известной вам последовательности.

Поскольку многие из нас сталкивались с понятием классификации сигналов, принятым в теории искусственного интеллекта и используемым при описании зрительного восприятия, давайте рассмотрим, чем отличаются эти сигналы от функционирования коры головного мозга. Пытаясь обучить компьютер распознавать объекты, исследователи, как правило, создают шаблон, например, изображение чашки или какую-то прототипную форму чашки, а потом программируют машину на поиск соответствия входящей информации шаблону. В случае значительного сходства выдается ответ, что обнаружена чашка. Отличие реального интеллекта состоит в том, что в нашем мозге нет таких шаблонов, а сигналы, поступающие в верхние зоны коры головного мозга, совсем не похожи на картинки. Вы не запоминаете моментальных изображений, воспринимаемых вашей сетчаткой, или ушами, или тактильными рецепторами, расположенными по всей поверхности вашей кожи. Иерархическое строение коры головного мозга обеспечивает распределение информации об объекте по всей иерархии – воспоминание о том или ином явлении или событии, предмете или человеке не сохраняется в каком-то одном месте. Кроме того, каждая зона из образующих иерархию формирует инвариантные запоминания, а в определенном поле неокортекса возникает последовательность инвариантных репрезентаций, которые, в свою очередь, тоже являются последовательностями инвариантных запоминаний. В головном мозге не сохраняются изображения чашки или каких-либо других объектов внешней среды.

В отличие от камеры ваш мозг запоминает окружающую среду такой, какая она есть, а не такой, какой она выглядит. Когда вы думаете о мире, то вспоминаете последовательности сигналов, соответствующих свойствам объектов, их функционированию и характеристикам ощущений в определенный момент времени. Последовательности, в рамках которых вы воспринимаете объекты окружающей действительности, отражают инвариантную структуру самого мира. Порядок, в котором вы воспринимаете составляющие внешнего мира, предопределяется его структурой. Например, вы можете попасть в самолет по трапу, но не напрямую от кассы. Последовательностями, в которых вы воспринимаете мир, представлена его реалистичная структура, именно она и формируется в коре головного мозга.

Не забывайте, однако, что инвариантная репрезентация в любой зоне коры головного мозга может превратиться в подробный прогноз того, какие ощущения возникнут в органах чувств, с одновременной ретрансляцией сигнала вниз по иерархии. Точно так же инвариантная репрезентация в моторной коре головного мозга может перейти в подробные ситуационно-специфические моторные команды посредством передачи сигнала вниз по иерархии.

Строение зон коры головного мозга

А сейчас давайте рассмотрим отдельную зону коры головного мозга, одну из тех, что изображены блоками на рис. 6.5.

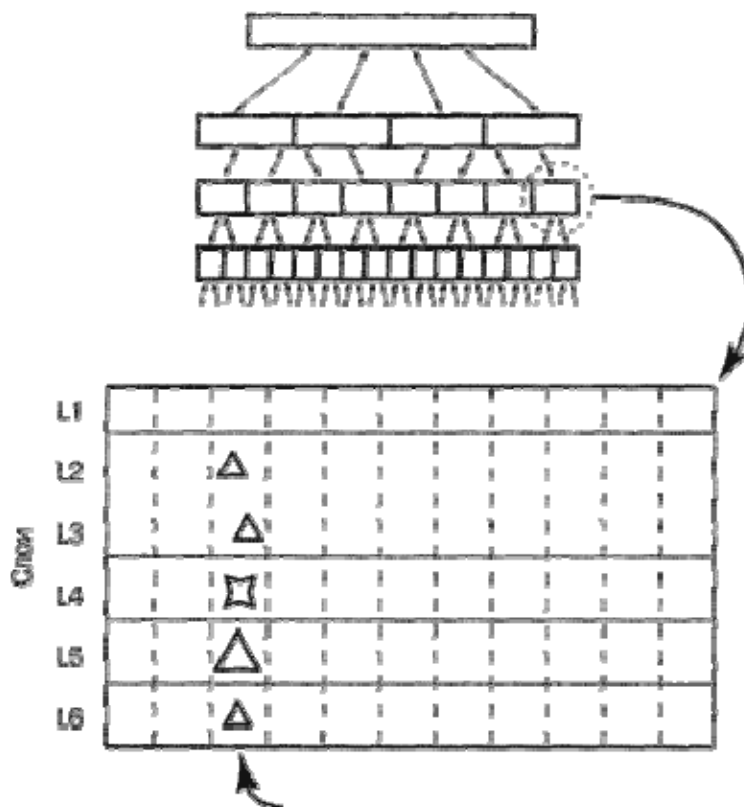


Рис. 6.6. Слои и колонки в зоне коры головного мозга

На рис. 6.6 приведено схематическое строение зоны коры головного мозга. В этом разделе я покажу, как клетки зоны коры головного мозга распознают и вызывают из памяти последовательности сигналов, обеспечивая тем самым формирование инвариантных репрезентаций и составление прогнозов. Начнем с описания структуры, общей для всех зон коры головного мозга. Прежде всего, между ними имеются значительные расхождения в размерах. Наиболее обширными являются первичные сенсорные зоны. К примеру, зона V1 занимает в задней части мозга область, приблизительно соответствующую площади паспорта. Как я уже упоминал выше, она состоит из множества подзон, которые могут быть не больше букв на этой странице. Но давайте пока предположим, что типичная зона коры головного мозга размером с мелкую монету.

Вспомните шесть визиток, с помощью которых в главе 3 мы представили слои ткани неокортекса. Что это за слои? Если мы рассмотрим корковую зону размером с монету под микроскопом, то увидим, что при передвижении снизу вверх плотность и форма нервных клеток различаются. Верхний слой 1 сильнее всего отличается от остальных слоев. Он состоит из густого переплетения аксонов, идущих параллельно поверхности коры, нейронов в этом слое почти нет. Строение слоев 2 и 3 сходно – они густо «населены» большим количеством пирамидных клеток. Слой 4 состоит из нервных клеток звездчатой формы. В слое 5 мы увидим как пирамидные клетки обычных размеров, так чрезвычайно большие. Несколько типов нейронов есть и в нижнем слое 6.

Визуально мы воспринимаем горизонтальные слои, но большее значение имеют колонки клеток, располагающиеся перпендикулярно. Можете считать колонки вертикальными объединениями клеток, которые работают вместе. (Термин «колонка» вызывает много дебатов в научных кругах. Нейробиологи спорят об их размере, функциях и роли в целом. Для удобства будем мыслить в рамках их архитектуры, поскольку то, что у коры столбчатая архитектура, признается всеми.) Слои в пределах каждой колонки связываются аксонами, которые проходят вверх-вниз и создают синапсы. Отмечу, что наши колонки не имеют четких очертаний – в коре нет ничего простого. То, что колонки существуют, доказывается следующими положениями.

Положение 1. Вертикально расположенные клетки одной колонки активизируются под

воздействием одних и тех же стимулов. Так, при подробном рассмотрении колонок зоны V1 мы обнаружим, что некоторые из них соответствуют линейным сегментам с правым уклоном (/), а другие – линейным сегментам с левым уклоном (\). Клетки из одной колонки тесно связаны, именно поэтому вся колонка реагирует на один и тот же стимул. Например, активная клетка слоя 4 активизирует расположенные над ней клетки слоев 3 и 2, что, в свою очередь, активизирует клетки, входящие в слои 5 и 6. Активность распространяется вниз и вверх в пределах колонки.

Положение 2. Наличие колонок нейронов в зонах неокортекса объясняется особенностями формирования коры головного мозга в эмбриогенезе. Эмбриональные клетки-предшественники мигрируют из внутренней мозговой впадины к месту образования коры головного мозга. Каждая из указанных клеток делится, образуя приблизительно сотню нейронов, называемых *микрколонками*. Эти нейроны связаны друг с другом вертикально так, как я описал. Термин «колонка» часто используется достаточно свободно для описания разных явлений. Под колонкой может подразумеваться общая вертикальная связность в коре или же вполне определенная группа клеток, происходящих от одной клетки-предшественницы. В рамках последнего определения мы можем утверждать, что в коре головного мозга есть несколько сотен миллионов микрколонок.

Чтобы вам легче было представить эту колончатую структуру, вообразите микрколонку в человеческий волос толщиной. Возьмите тысячи волосков и разрежьте их на небольшие сегменты, имеющие высоту строчной буквы T. Если поставить их бок о бок очень плотно друг к другу и склеить, получится густая щетина. Теперь создайте покров из длинных и очень тонких волосков. Они будут представлять аксоны слоя 1. Приклейте их горизонтально к поверхности нашей щетины. Этот щетиноподобный коврик является упрощенной моделью зоны коры головного мозга размером с монету. Потоки информации преимущественно передвигаются преимущественно вдоль волосков: горизонтально в слое 1 и вертикально в слоях 2-6.

Рассмотрим еще одну важную особенность колонок, а затем перейдем к анализу их назначения. При пристальном рассмотрении мы увидим, что не менее 90% синапсов клеток из одной колонки – это связи с клетками, расположенными вне пределов самой колонки: некоторые соединения – с соседними колонками, некоторые – вообще с другими частями мозга. На основании чего же мы говорим о важности колонок, если большинство связей распределено по всему мозгу, а не локализовано?

Ответ на этот вопрос нам даст модель «Память-предсказание». В 1979 году Вернон Маунткастл не только заявил о существовании единого алгоритма коры головного мозга, но также предположил, что колонка нейронов в коре головного мозга является базовой единицей обработки информации. Правда, он не знал, какова именно функция колонки. Я считаю, что последняя является базовой единицей прогнозирования. Чтобы колонка нейронов могла прогнозировать, когда ей следует активизироваться, нервным клеткам необходимо знать, *что* происходит в других частях мозга. Именно в этом и заключается смысл обширных синаптических связей.

Ниже мы обсудим этот интересный момент подробнее, а сейчас я хочу дать вам общее представление о назначении подобных соединений нейронов в коре головного мозга. Чтобы предсказать следующий звук мелодии, вы учитываете ее название, ту часть, которая звучит в данный момент времени, время, прошедшее с момента последнего услышанного вами звука, и то, каким он был. Большое количество синапсов, связывающих нервные клетки одной колонки с остальными частями мозга, обеспечивает контекст, в котором и осуществляется прогностическая функция коры головного мозга.

Теперь обратимся к рассмотрению потоков информации, циркулирующих вверх-вниз по иерархии корковых зон. Восходящий поток сигналов проходит относительно прямой путь, который схематически отображен на рис. 6.7. Представьте, что перед нами зона коры головного мозга, состоящая из тысяч колонок. Увеличим одну из них. Входные сигналы из нижних зон попадают в слой 4, который является основным входящим слоем. По дороге они «отмечаются» в слое 6 (скоро вы поймете, почему это так важно). Клетки слоя 4 передают сигналы вверх – к клеткам слоев 2 и 3, находящимся в пределах той же колонки. Далее клетки слоев 2 и 3 передают информацию по аксонам еще выше – к входному слою следующей в иерархии зоны.

Таким образом информация передвигается от зоны к зоне вверх по иерархии в пределах одной колонки нейронов.

Путь нисходящего потока информации (рис. 6.8) более сложен. Клетки слоя 6 посылают информацию вниз, в слой 1 области, расположенной ниже в иерархии. Аксоны нейронов слоя 1 распространяются на очень большие расстояния. Следовательно, информация, передвигающаяся вниз, потенциально может активизировать многие колонки нейронов. Хотя «своих» клеток в слое 1 очень мало, сигналы, проходящие через этот слой, могут активизировать клетки слоев 2, 3 и 5, поскольку у последних есть дендриты в слое 1. Аксоны, выходящие из клеток слоев 2 и 3, покидая кортекс, формируют синапсы в слое 5 и, как предполагается, возбуждают клетки слоев 5 и 6. Итак, мы приходим к выводу, что информация, передвигаясь вниз по иерархии, проходит более запутанный путь, чем в случае движения вверх. Благодаря слою 1 она может распространяться в самых разных направлениях. Обратная информация генерируется в клетках слоя 6 высшей зоны. Затем она распространяется в слое 1 низшей зоны. Клетки слоев 2, 3 и 5 низшей зоны возбуждаются, некоторые из них, в свою очередь, возбуждают клетки слоя 6, который проецируется в слой 1 более низких в иерархии областей и так далее. (При помощи рис. 6.8 понять описанные процессы намного проще.)

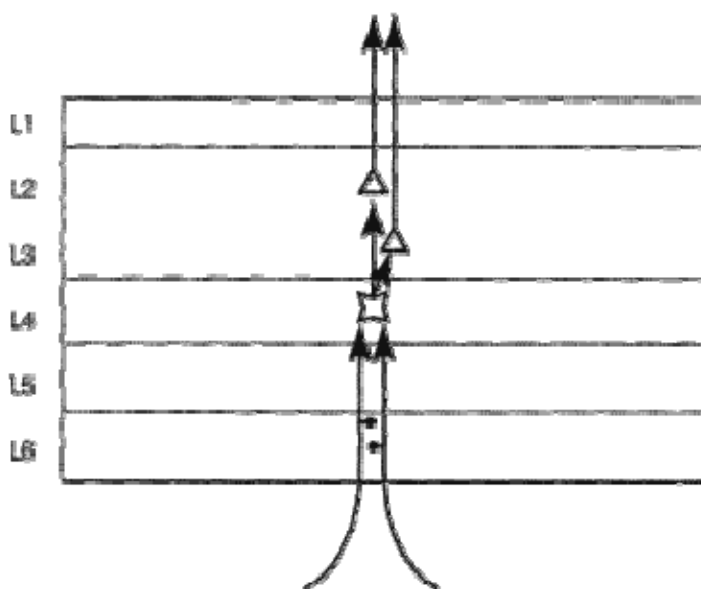


Рис. 6.7. Восходящий информационный поток

Приведу предварительное объяснение того, зачем информация распределяется в слое 1. Чтобы трансформировать инвариантную репрезентацию в какой-то прогноз, каждую секунду надо решать, куда должен идти сигнал, направленный вниз по иерархии. Слой 1 обеспечивает возможность конвертирования инвариантного представления в более специфическую и конкретную репрезентацию. Например, вы можете воспроизвести Геттисбергскую речь как в устной, так и в письменной форме. Общая репрезентация трансформируется двумя разными способами: первый – для устной, а второй – для письменной формы. Горизонтальный поток активности в слое 1 обеспечивает механизм для осуществления такого процесса. Чтобы инвариантные представления высокого уровня могли транслироваться вниз и принимать форму конкретных прогнозов, мы должны располагать механизмом, позволяющим потоку сигналов разветвляться в каждом слое. На эту роль как нельзя лучше подходит слой 1. Даже не зная о его существовании, мы могли бы предвидеть, что такой слой есть.

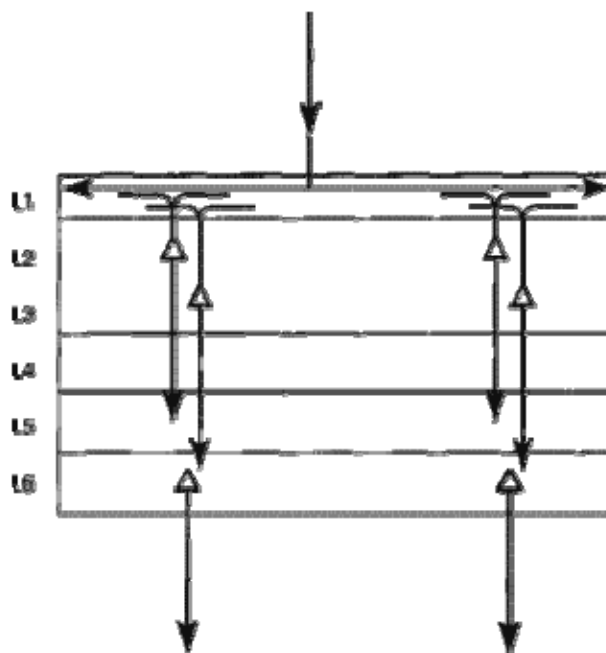


Рис. 6.8. Нисходящий информационный поток

И еще немного сведений из анатомии: аксоны, покидающие слой 6, окутаны оболочкой из белого жирового вещества, которое называется *миелином*. Это белое вещество выполняет ту же роль, что и изоляция на электрических проводах, – препятствует смешиванию сигналов и повышает скорость их передвижения до 320 километров в час. Аксоны избавляются от белого защитного вещества, попадая в новую колонку нейронов слоя 6 в коре головного мозга.

И наконец, есть еще один не прямой способ связи зон коры головного мозга.

Но прежде я хотел бы вам снова напомнить об автоассоциативном свойстве системы памяти (о котором шла речь в главе 2). Как вы, вероятно, помните, автоассоциативные запоминания служат для сохранения последовательностей сигналов. Когда выходной сигнал группы искусственных нейронов передается назад с целью формирования входного потока для всех нейронов (и при этом обратная связь дополняется задержкой), то таким образом сигналы учатся последовательно двигаться друг за другом. По моему мнению, кора головного мозга использует подобный механизм для сохранения последовательностей, но в нем присутствуют еще некоторые особенности. Вместо того чтобы формировать автоассоциативные последовательности из искусственных нейронов, она формирует их из колонок коры головного мозга. Исходящие данные из всех колонок передаются назад, в слой 1. Выходит, слой 1 располагает свежей информацией о том, какие колонки зоны коры головного мозга только что были активными.

Рассмотрим элементы, изображенные на рис. 6.9. Ученым давно известно, что нейроны особо крупных размеров слоя 5 моторной коры (зона M1) непосредственным образом связаны с мышцами и моторными зонами спинного мозга – они обеспечивают управление вашими мышцами и передвижение. Каждый раз, когда вы говорите, печатаете, выполняете любые самые сложные действия, эти клетки посылают высококоординированные импульсы, заставляя сокращаться ваши мышцы.

Совсем недавно ученые открыли, что крупные нейроны слоя 5, возможно, выполняют какие-то функции и в других частях коры головного мозга, а не только в моторных зонах. Например, крупные нейроны слоя 5 зрительной зоны имеют проекции в зоне, отвечающей за движения глаз. Сенсорные зрительные зоны V2 и V4 не только занимаются обработкой зрительных входных сигналов, но также помогают предопределять движение глаз, а следовательно, управляют тем, что вы видите. Крупные нейроны слоя 5 присутствуют в каждой зоне коры головного мозга, что наводит на мысль об их немаловажной роли во всех типах движений.

Аксоны этих нейронов не только участвуют в формировании поведения, они

разветвляются надвое. Одно ответвление идет к таламусу, изображенному на рис. 6.9 в виде овальной фигуры. Таламус человека находится в центре головного мозга, под корой, сверху «старого мозга», и окружен белым веществом. Размером и формой он похож на два маленьких птичьих яйца. Все зоны коры головного мозга направляют многочисленные аксоны в таламус, от него тоже исходят аксоны в обратном направлении. Науке известно достаточно об этих связях, но сам таламус очень сложен по своей структуре, и его роль до сих пор мало изучена. Таламус необходим для того, чтобы вести нормальный образ жизни: его повреждение приводит к устойчивому «растительному» состоянию.

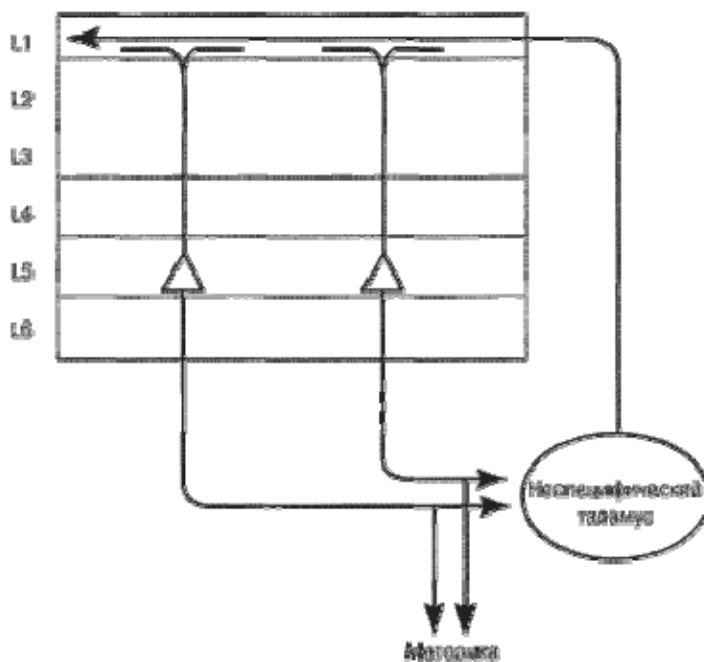


Рис. 6.9. Как посредством таламуса сообщаются текущее состояние и текущее моторное поведение

Существуют несколько путей от таламуса к коре головного мозга, но сейчас лишь один из них представляет для нас интерес. Он начинается с крупных нейронов слоя 5, проецирующихся на группу клеток таламуса, считающихся функционально неспецифическими. Аксоны этих неспецифических клеток идут назад, в слой 1 множества различных зон коры головного мозга. Например, нервные клетки слоя 5 зон V2 и V4 посылают аксоны к таламусу, а таламус, в свою очередь, отправляет информацию назад к слою 1 зон V2 и V4. Аналогичный процесс происходит и в других зонах коры головного мозга. Клетки слоя 5 множества корковых зон посылают сигналы в таламус, который, в свою очередь, отправляет назад информацию в слой 1 этих же самых и связанных с ними зон. Я считаю, что такой круговорот полностью дублирует обратную связь с задержкой, которая позволяет автоассоциативной памяти усваивать последовательности сигналов.

Только что я упомянул о двух источниках входной информации, поступающей в слой 1. Высшие зоны коры головного мозга возбуждают активность в слое 1 более низких зон коры. Активные колонки в пределах одной зоны возбуждают активность в пределах слоя 1 той же зоны через таламус. Эти входящие в слой 1 сигналы можно представить как название песни – считать словами песни (входящие сверху сигналы) и знание, в каком месте песни мы сейчас находимся (задержка в деятельности активных колонок в пределах одной зоны). Таким образом, слой 1 содержит большое количество информации, необходимой нам для прогнозирования времени активизации колонки: название последовательности и наше местонахождение в ней. Используя эти два сигнала слоя 1, зона коры головного мозга в состоянии обучаться и запоминать огромное количество последовательностей сигналов.

Как работает корковая зона: подробности

Не забывая о наших трех схемах (входные сигналы, поступающие вверх по иерархии; разветвляющиеся сигналы, спускающиеся вниз по иерархии; обратная связь с задержкой в таламусе), приступим к рассмотрению того, как зона коры головного мозга выполняет свои функции. Нам необходимо выяснить следующее.

Как зона мозга классифицирует входящие сигналы (вспомните сортировку цветных листов бумаги по ведрам).

Как она усваивает последовательности сигналов (вспомните последовательности интервалов в песне и восприятие лиц посредством комбинаций типа «глаз-нос-глаз»).

Как корковая зона формирует постоянные паттерны, или «имена» последовательностей.

Как она составляет специфические прогнозы (встреча поезда в определенное время или прогнозирование следующего звука мелодии).

Допустим, что колонки одной зоны – это те же ведра, которые мы использовали для воображаемой сортировки получаемых цветных листов бумаги. Каждая колонка помечена точно так же, как и ведро. К клеткам слоя 4 в каждой колонке протянуты волокна ввода из нескольких зон, расположенных ниже по иерархии. Всякий раз, получая правильную комбинацию входных сигналов, они выдают импульсный разряд. Возбуждаясь, клетка слоя 4 «голосует» за то, что полученные сигналы соответствуют своему ярлыку. Точно так же, как при сортировке листов бумаги, входные сигналы могут быть неоднозначными (например несколько колонок могут подходить для единственного входного сигнала). Мы хотим, чтобы зона коры нашего головного мозга сделала выбор: лист бумаги может считаться либо красным, либо оранжевым, но ни в коем случае не тем и другим сразу. Колонка, которой входной сигнал соответствует более всего, должна предотвратить активизацию остальных колонок.

Этим занимаются так называемые *тормозные нейроны*. Они предотвращают активизацию соседних нервных клеток, эффективно заботясь о том, чтобы был только один «победитель». Эти тормозные нейроны влияют исключительно на зону вокруг «своей» колонки. Другими словами, несмотря на наличие механизма подавления, существует возможность одновременной активизации нескольких колонок. (В мозге ничто не бывает представлено одним единственным нейроном или единственной колонкой.) Для простоты понимания можем представить, что зона выбирает одну и только одну колонку-«победителя», однако на будущее запомним, что на самом деле одновременно активизируются множество колонок. Процесс, используемый корой головного мозга для классификации входных сигналов, сложен и плохо изучен. Я не собираюсь заводить вас в дебри. Вместо этого я предлагаю принять предположение, что зона коры головного мозга отсортировала входной сигнал как активизацию в наборе колонок. Тогда мы сможем сосредоточиться на формировании последовательностей и присвоении им названий.

Каким образом зона коры головного мозга сохраняет последовательность отсортированных сигналов? Я уже излагал свои соображения по этому поводу и сейчас хотел бы углубиться в подробности. Представьте себе, что вы – колонка нервных клеток, а сигнал, поступивший из зоны низшего порядка, заставил активизироваться одну из ваших клеток слоя 4. Вслед за этим возбуждаются клетки в слоях 2 и 3, потом – в слое 5, а затем – в 6-м. Постепенно активизируется вся колонка нейронов. У каждой из клеток слоев 2, 3 и 5 есть тысячи синапсов в слое 1. Если какие-то из этих синапсов активны во время разряда нервных клеток слоев 2, 3 и 5, они усиливаются. Если такое происходит часто, то синапсы слоя 1 становятся достаточно сильными для того, чтобы вызвать вспышки клеток в слоях 2, 3 и 5 даже тогда, когда клетки слоя 4 молчат. Таким образом, некоторые части колонки могут активизироваться, даже если они не получили входного сигнала от более низкой по иерархии зоны коры головного мозга. Именно так клетки слоев 2, 3 и 5 учатся предвосхищать собственную активизацию на основе сигнала, поступающего из слоя 1. «До обучения» колонка может активизироваться исключительно через клетки слоя 4. «После обучения» колонка может частично активизироваться через память. Когда колонка активизируется через синапсы слоя 1, то входящие в нее нейроны предвидят, что произойдет активизация снизу, – это и есть прогноз. Если бы колонка могла разговаривать, она бы сказала: «Когда я активизировалась в прошлом, то как раз этот набор моих синапсов слоя 1 был активен. Как только я увижу такой набор

синапсов, сразу же начну генерировать импульс в предвкушении сигнала».

Вспомните, что половина входных сигналов слоя 1 происходит из слоя 5 клеток соседних колонок и других зон коры головного мозга. Это информация о том, *что* случилось мгновение тому назад. Она говорит нам, какие колонки были активными до того, как активизировалась наша колонка. Она говорит нам, каким был последний интервал мелодии, *что* я видел последним, *что* я чувствовал последним, какую фонему устной речи я услышал последней и так далее. Если порядок возникновения указанных сигналов является последовательным во времени, то колонки усваивают последовательность. Они будут возбуждаться одна за другой в определенной последовательности.

Вторая половина сигналов поступает в слой 1 из клеток слоя 6 зон, расположенных выше в иерархии. Эта информация является более стабильной и представляет «имя» последовательности, с которой вы непосредственно сталкиваетесь в данный момент времени. К примеру, если колонки представить как музыкальные фразы, то название мелодии будет «именем» последовательности. Если фонемы будут колонками, то «именем» будет слово, которое вы слышите. Если колонками будут произносимые вами слова, «именем» будет речь, которую вы произносите. Следовательно, в слое 1 содержится и информация об «имени» последовательности, и о последнем сигнале в ней. Благодаря этому одна колонка может использоваться многими последовательностями без возникновения недоразумений. Колонки «учатся» активизироваться в нужном контексте и в правильном порядке.

Я должен подчеркнуть, что синапсы слоя 1 не являются единственными синапсами, участвующими в процессе обучения тому, когда должна активизироваться колонка. Как я уже упоминал ранее, клетки получают и отсылают сигналы ко многим соседним колонкам. Ведь более 90% всех синапсов – соединения с клетками за пределами колонки, и большинство из них расположены не в слое 1. Например, у клеток слоев 2, 3 и 5 есть тысячи синапсов в слое 1, но, кроме того, они имеют и тысячи синапсов в собственных слоях. Общий принцип состоит в том, что нейроны стремятся получать любую информацию, которая поможет им спрогнозировать возникновение импульса снизу. Как правило, активность соседних колонок отличается высоким уровнем корреляции, поэтому существует много связей между соседними колонками. Например, если в поле вашего зрения находится прямая линия, это активизирует колонки одну за другой. Но достаточно часто для прогнозирования активности колонки нужна более обширная информация. Именно тогда вступают в игру синапсы слоя 1. Нервным клеткам или колонкам нейронов не известно, что означает тот или иной синапс, но посредством синапсов они прогнозируют, когда нужно активизироваться.

А сейчас давайте рассмотрим, как зона коры головного мозга формирует имя выученной последовательности. Опять представьте себя зоной коры головного мозга. Активность ваших колонок нейронов меняется при каждом новом входном сигнале. Вы успешно усвоили последовательность активизации ваших колонок, т. е. некоторые клетки ваших колонок активизируются до того, как поступит входной сигнал из более низких зон. Какую информацию вы посылаете в более высокие зоны коры головного мозга? Как мы знаем, клетки слоев 2 и 3 посылают свои аксоны в следующую вверх по иерархии область. Активность этих клеток является входным сигналом для более высоких областей. Вот тут и возникает проблема. Чтобы иерархия работала слаженно, вам нужно транслировать постоянный сигнал на протяжении всей выученной последовательности. Вы должны транслировать «имя» последовательности, а не ее детали. До того как вы «выучите» последовательность, вы можете передавать отдельные ее элементы, но, после того как вы ее усвоили и можете прогнозировать, какие колонки будут активизироваться, вы должны передавать постоянный сигнал. Однако я еще не рассказывал, как это происходит. Итак, пока что вы передаете каждый переменный сигнал независимо от того, можете вы его спрогнозировать или нет. Когда активизируется вся колонка, клетки слоев 2 и 3 посылают новый сигнал вверх по иерархии. Коре головного мозга нужен какой-то способ, позволяющий поддерживать постоянный сигнал на протяжении всей выученной последовательности. Нам нужен способ отключать сигналы, исходящие из слоев 2 и 3, когда колонка прогнозирует свою активность, или же, наоборот, активизировать эти клетки, когда колонка не может спрогнозировать свою активность. Только так можно создать постоянный сигнал имени.

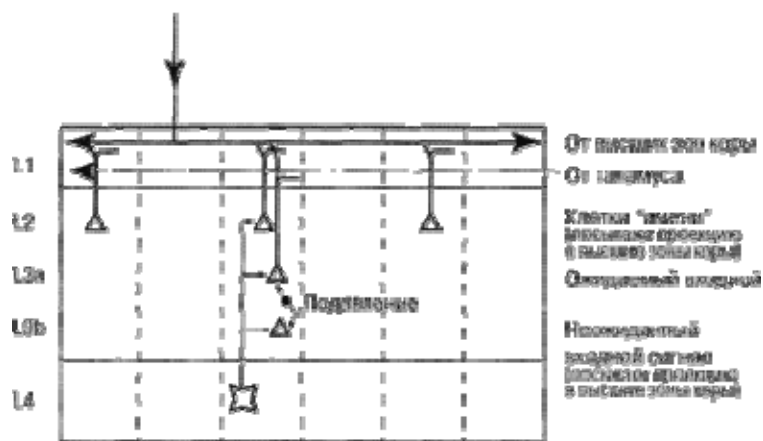


Рис. 6.10. Формирование постоянного «имени» для усвоенной последовательности

Науке очень мало известно о том, как именно кора головного мозга осуществляет подобную операцию. По моему мнению, существует несколько потенциально возможных методов. Я опишу тот, к которому больше всего склоняюсь в последнее время. (Не забывайте, что для нас сейчас наиболее важными являются концептуальные рамки, а не конкретный метод.) Ключевым моментом является создание постоянного сигнала «имени». Я хочу лишь показать, что существуют вполне правдоподобные механизмы для осуществления такого процесса.

Представьте, что вы – колонка (см. рис. 6.10). Постараемся разобраться, как вы учитесь посылать постоянный сигнал в вышестоящую в иерархии область, когда вам удастся спрогнозировать свою активность, и переменный сигнал в противном случае. Сначала предположим, что в слое 2 и 3 входят несколько классов клеток. (Впридачу к нескольким типам тормозных нейронов многие анатомы различают несколько типов в слоях 3а и 3б, т. е. наше предположение вполне оправдано.)

Теперь предположим, что один класс клеток – клетки слоя 2 – учится быть активным на протяжении усвоенных последовательностей. Как группа этих клеток представляет «имя» последовательности? Она посылает постоянный сигнал в высшие зоны коры головного мозга до тех пор, пока их зона может предвидеть, какие колонки будут активизироваться дальше. Если ваша зона усвоила последовательность из трех различных сигналов, тогда клетки слоя 2 всех колонок, представляющих эти три сигнала, будут оставаться активными до тех пор, пока мы остаемся в пределах нашей последовательности. Они и являются «именем» последовательности.

Теперь предположим, что существует другой класс клеток – клетки слоя 3б, которые не активизируются, когда наша колонка успешно прогнозирует поступающие сигналы, но генерируют разряд в случае, если этого не происходит. Клетка слоя 3б олицетворяет' неожиданный сигнал. Она генерирует импульс при неожиданной активизации своей колонки. Она будет поступать так всякий раз, когда колонка, еще не «прошедшая» процесс обучения, активизировалась. По мере того как колонка учится прогнозировать свою активность, клетки слоя 3б «успокаиваются». Вместе взятые, клетки слоя 2 и слоя 3б удовлетворяют нашим требованиям. До того как было пройдено «обучение», и те и другие клетки затухают-вспыхивают в колонке, но по окончании «учебного курса» клетка слоя 2 остается активной постоянно, а клетка слоя 3б пребывает в состоянии покоя.

Как клетки учатся этому? Сначала давайте рассмотрим, как «отключить» клетки слоя 3б, когда колонка успешно прогнозирует свою активность. Допустим, существует клетка 3а, расположенная над клеткой 3б, она имеет дендриты в слое 1. Ее единственная задача состоит в том, чтобы воспрепятствовать активизации клетки слоя 3б, если она замечает в слое 1 уже известный паттерн. Когда клетка слоя 3а видит выученный сигнал в слое 1, она сразу же активизирует тормозной нейрон, который, в свою очередь, препятствует вспышке клетки слоя 3б. Этого достаточно, чтобы удержать клетку слоя 3б от вспышки, когда колонка правильно

прогнозирует свою активность.

А теперь рассмотрим более сложную задачу: как клетка слоя 2 остается активной на протяжении всей известной последовательности сигналов. Данная задача более сложна, потому что при этом различные клетки слоя 2 из многих колонок должны одновременно оставаться активными, даже если отдельные колонки в текущий момент таковыми не являются. По моему мнению, это может происходить следующим образом. Клетки слоя 2 учатся активизироваться исключительно высшими областями иерархии коры головного мозга. Они формируют синапсы преимущественно с аксонами из клеток слоя 6 областей, расположенных на уровень выше. Таким образом, клетки слоя 2 представляют постоянный сигнал «имени» из высшей области. Когда высшая зона посылает сигнал вниз к слою 1 расположенной ниже области, набор клеток слоя 2 нижней зоны активизируется, представляя все колонки, являющиеся членами последовательности. Клетки слоя 2 тоже имеют проекции в вышестоящей области и представляют собой частично стабильную группу клеток. (Вряд ли они постоянно остаются активными. Скорее всего, они вспыхивают синхронно, наподобие ритма.) Верхняя зона как бы посылает название мелодии в слой 1, расположенный ниже. Это заставляет вспыхивать набор клеток слоя 2, каждую клетку в своей колонке, и они будут активизироваться, когда вы слышите мелодию.

Все эти механизмы вместе взятые позволяют коре головного мозга усваивать последовательности, составлять прогнозы, формировать постоянные репрезентации («имена») для последовательностей входящих сигналов. Это основные операции, необходимые для формирования инвариантных репрезентаций.

Как мы формируем прогнозы о событиях, с которыми ранее не сталкивались? Как мы выбираем одну из десятков возможных интерпретаций входных сигналов? Как зона коры головного мозга формирует частные прогнозы на основе инвариантных схем? Мы уже рассматривали примеры прогнозирования следующего звука мелодии, когда в вашей памяти есть только интервал между отдельными нотами, историю со встречей поезда и воспроизведение Геттисбергской речи. Во всех перечисленных случаях единственным путем решения проблемы было использование последней полученной информации для того, чтобы конвертировать инвариантный прогноз в частное предсказание. Перефразировав ту же мысль в терминах коры головного мозга, можем сказать, что нужно сочетать поток прямой информации (поступающие фактические сигналы) с потоком обратной информации (прогноз в инвариантной форме).

Приведу простой пример того, как это, по моему представлению, происходит. Допустим, кора вашего головного мозга знает, что должна воспринять музыкальный интервал квинту. Колонки вашей зоны реагируют на отдельные интервалы, такие как «до-ми», «до-соль», «ре-ля» и так далее. Вам нужно решить, какие из колонок будут активными. Если зона выше говорит вам, что следует ожидать квинту, это активизирует клетки слоя 2 во всех колонках, представляющих квинты, то есть «до-соль», «ре-ля», «ми-си». Клетки слоя 2 из колонок, которые представляют другие интервалы, активизироваться не будут. Теперь вам нужно из колонок, представляющих квинты, выбрать только одну. Входными сигналами для вашей зоны будут отдельные ноты. Если «ре» была последней нотой, которую вы слышали, тогда все колонки, представляющие интервалы, в которых присутствует нота «ре», т. е. «ре-ми» и «ре-си», получают часть входного сигнала. Теперь в слое 2 активизировались все колонки, представляющие квинты, а в слое 4 мы получили часть сигнала во все колонки, представляющие интервалы, в которые входит нота «ре». Пересечение этих двух множеств и будет ответом на наш вопрос. Это будет колонка, представляющая интервал «ре-ля» (см. рис. 6.11).

Как кора головного мозга определяет пересечение? Мы раньше упоминали о том, что аксоны клеток слоев 2 и 3, выходя из коры головного мозга, формируют синапсы в слое 5. Аксоны, направляющиеся к слою 4 из более низких зон коры головного мозга, создают синапсы в слое 6. Пересечение этих двух синапсов (снизу вверх и сверху вниз) дает нам как раз то, что нужно. Клетки слоя 6, получающие эти два сигнала возбуждаются. Клетка слоя 6 «представляет мнение» зоны коры головного мозга о том, что происходит, т. е. специфический прогноз. Если бы клетка слоя 6 могла разговаривать, она бы сказала: «Я являюсь частью колонки,

представляющей нечто. Конкретно в моем случае колонка представляет интервал „ре-ля“. У остальных колонок другие значения. Я представляю свою зону коры головного мозга. Когда я становлюсь активной, это означает, что, по нашему мнению, сейчас звучит интервал „ре-ля“, или же он прозвучит в ближайшем будущем. Возможно, я активизировалась, потому что сигнал снизу вверх, поступивший от ушей, заставил клетку слоя 4 моей колонки возбудить всю колонку. Или же моя активизация может означать, что мы узнали мелодию и прогнозируем следующий интервал. В любом случае моя задача состоит в том, чтобы довести до ведома низших зон коры головного мозга, *что*, по нашему мнению, происходит. Я представляю наше видение внешнего мира – независимо от того, является он настоящим или выдуманным».

Попробую описать указанный процесс с помощью визуальной картинки. Представьте себе два листа бумаги, усеянные множеством мелких дырочек. Дырочки на одном листе представляют колонки, у которых есть активные клетки слоев 2 и 3 (наш инвариантный прогноз). Дырочки на другом листе бумаги представляют колонки с частичным сигналом снизу. Если мы наложим листы бумаги друг на друга, то некоторые из дырочек совпадут, а некоторые – нет. Дырочки, которые совпали, представляют колонки, которые, как мы полагаем, должны активизироваться.

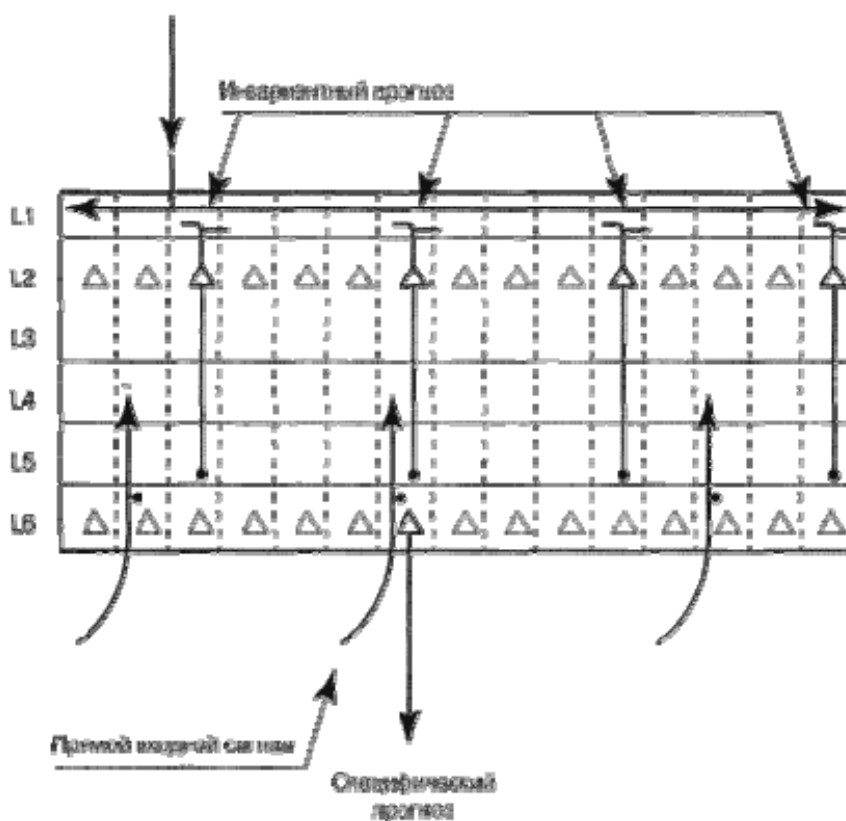


Рис. 6.11. Как зона коры головного мозга формирует специфические прогнозы, используя инвариантные репрезентации

Описанный механизм не только создает определенные прогнозы, он также разрешает неоднозначность входных сенсорных сигналов. Как мы помним из примеров с цветными листами бумаги и искаженным изображением, очень часто входные сигналы, полученные зоной коры головного мозга, являются неоднозначными. Механизм сочетания информации, поступающей снизу вверх и сверху вниз, помогает выбрать один из двух или более вариантов интерпретации. Как только вы приняли решение, каков будет выбор, вы сразу же транслируете этот выбор в область, расположенную ниже по иерархии.

Каждое мгновение нашей жизни каждая зона коры головного мозга сравнивает набор ожидаемых колонок, возбужденных сверху, с набором колонок, которые уже были возбуждены снизу. Там, где эти два набора пересекаются, создается наше восприятие. Если бы поступали

совершенные сигналы снизу и у нас были бы совершенные прогнозы, тогда набор возбужденных колонок всегда совпадал бы с набором ожидаемых колонок. Но очень часто этого не происходит. Метод сочетания частичного прогноза и частичного входного сигнала позволяет определиться с неоднозначным входным сигналом, он позволяет заполнить пробелы в информации и остановить выбор на одном из альтернативных вариантов. Именно так мы комбинируем музыкальный интервал с последней услышанной нотой, чтобы спрогнозировать следующую ноту мелодии. Именно так мы определяемся, изображено на картинке два лица или ваза. Именно так мы разветвляем моторный поток на написание или произнесение вслух Геттисбергской речи.

И наконец, вдобавок к отправлению информации в более низкие зоны коры головного мозга, клетки слоя 6 могут посылать выходные сигналы назад, к клеткам слоя 4 той же колонки. Когда они это делают, наши прогнозы становятся входным сигналом. Именно этим мы занимаемся, размышляя или мечтая. Это позволяет нам видеть последовательности наших собственных прогнозов. Мы занимаемся этим, планируя свое будущее, вспоминая выступления или переживая о предстоящих событиях. Стефен Гросберг, длительное время занимавшийся моделированием деятельности коры головного мозга, называет это *свернутой обратной связью*, я же предпочитаю термин *воображение*.

И наконец, последний вопрос, который нам нужно затронуть, прежде чем перейти к следующему разделу. Я неоднократно подчеркивал, что чаще всего то, что мы видим, слышим, чувствуем, во многом зависит от наших собственных действий. То, что мы видим, зависит от того, на чем наши глаза остановились после очередной саккады, и от угла поворота головы. То, что мы ощущаем, зависит от того, как мы передвигаем пальцы и конечности. То, что мы слышим, иногда зависит от того, что мы делаем и говорим.

Значит, чтобы спрогнозировать, *что* мы почувствуем дальше, мы должны знать, что мы сейчас делаем. Моторное поведение и сенсорное восприятие тесно взаимосвязаны. Как нам удастся составлять прогнозы, если то, что мы чувствуем, во многом является результатом наших собственных действий? К счастью, у данной проблемы существует неожиданное и простое решение, правда, многие его детали не очень понятны.

Неожиданное открытие: восприятие и поведение – это почти одно и то же. Как я уже упоминал раньше, большинство зон коры головного мозга (если не все) принимают участие в осуществлении движения.

Клеткам слоя 5, имеющим проекции в таламусе, а затем в слое 1, тоже свойственна моторная функция, поскольку у них одновременно есть проекции в моторных зонах «старого» мозга. Таким образом, знание о том, что только что произошло (как в сенсорном, так и в моторном плане), присутствует в слое 1.

Вторая неожиданность (являющаяся следствием первой): моторное поведение должно быть представлено в иерархии инвариантных репрезентаций. Вы генерируете движения, необходимые для осуществления определенного действия, не размышляя о нем во всех деталях. Когда моторная команда опускается вниз по иерархии, она трансформируется в сложные и подробные последовательности, необходимые для осуществления той деятельности, которая была запланирована. Это происходит как в сенсорной, так и в моторной зонах коры головного мозга, что размывает грань между ними. Если зона IT зрительной коры головного мозга воспринимает репрезентацию «носа», то простое действие переключения на репрезентацию «глаза» будет вызывать саккаду, необходимую для оправдания прогноза. Саккада, необходимая для того, чтобы перейти от зрительного восприятия носа к зрительному восприятию глаза, меняется в зависимости от расположения лица. Если лицо расположено поблизости, это требует большей саккады, и наоборот, более отдаленное лицо требует меньшей саккады. Наклоненное лицо требует саккады под углом, который отличается от используемого для лица без наклона. То, какой должна быть саккада, определяется, пока прогноз о том, что следующим мы увидим *глаз*, спускается к зоне V1. Чем дальше опустил прогноз, тем более точной становится саккада. Заканчивается все тем, что центральные ямки ваших глаз останавливаются точно на объекте или же очень близко к нему.

Давайте рассмотрим еще один пример. Для того чтобы я физически переместился из гостиной на кухню, все, что нужно моему мозгу, – это мысленно переключиться с

инвариантной репрезентации гостиной на инвариантную репрезентацию кухни. Данное переключение вызывает сложное разворачивание последовательностей. Процесс генерирования последовательности прогнозов того, что я увижу, услышу, почувствую, пока буду переходить из гостиной на кухню, тоже генерирует последовательность моторных команд, заставляющих меня перейти из гостиной на кухню и перемещать мой взгляд, пока я двигаюсь. Прогнозирование и моторное поведение идут рука об руку, в то время как потоки сигналов одновременно двигаются вверх и вниз по иерархии коры головного мозга. Каким бы странным это ни казалось, но, когда в дело вовлечено ваше поведение, ваши прогнозы не только предшествуют ощущению – они также его предопределяют. Прогнозирование перехода к следующему сигналу последовательности вызывает каскадное прогнозирование того, что вы должны пережить дальше. Когда каскадный прогноз разворачивается, он генерирует моторные команды, необходимые для осуществления прогноза. Воображение, прогнозирование и действие являются частями одного и того же разворачивания последовательностей в иерархии коры головного мозга.

Действие в воображении, параллельное разворачивание сенсорного и моторного поведения являются основой поведения, ориентированного на результат. Целенаправленное поведение, этот священный Грааль робототехники, изначально встроено в ткань коры головного мозга.

Конечно, мы можем отключить наше моторное поведение. Я могу представить себе, что вижу что-то, хотя на самом деле этого не вижу. Я могу представить себе, как иду на кухню, хотя на самом деле этого не делаю. Но действие всегда начинается в нашем воображении.

Поток вниз, поток вверх

Давайте на минуту остановимся и поразмышляем над тем, как информация перемещается вниз и вверх в иерархии коры головного мозга. Когда вы передвигаетесь, потоки непрерывно меняющейся входной информации устремляются в нижние зоны коры головного мозга. Каждая зона пытается интерпретировать поток входящей информации как часть известной последовательности сигналов. Колонки пытаются предугадать свою активность. Если им это удастся, они передают постоянный сигнал – название последовательности – в область, расположенную выше в иерархии. Зона коры головного мозга как бы говорит: «Я слышу песню, и вот вам ее название. Я сама разберусь с отдельными элементами».

Но что будет, если поступит неожиданный сигнал – нота, которой вы не ожидали? Что будет, если мы увидим что-то такое, чего на лице быть не должно? Неожиданный сигнал автоматически будет передан в область, стоящую выше в иерархии. Это случается, когда «вспыхивают» клетки слоя 3б, которые не были частью ожидаемой последовательности. Возможно, более высокая зона сможет понять сигнал как часть ее собственной последовательности. Возможно, она скажет: «Я вижу, что тут появилась новая нота. Вполне возможно, это первая нота следующей песни из альбома. Мне так кажется, поэтому я прогнозирую, что мы перешли к следующей песне из альбома. Низшая область, вот тебе название песни, которую, как мне кажется, ты сейчас слышишь». Но если не произойдет такого распознавания, сигнал будет подниматься вверх по иерархии коры головного мозга, пока какая-нибудь из вышестоящих зон не идентифицирует сигнал как часть нормальной последовательности событий. Чем выше по иерархии поднимается сигнал, тем больше разных зон вовлекается в его обработку. Когда же наконец какая-то зона в верхней части иерархии считает, что она в состоянии обработать неожиданное событие, тогда формируется новый прогноз. Новый сигнал продвигается вниз по иерархии так далеко, как ему удастся. Если новый прогноз оказался неверным, ошибка будет выявлена, и сигнал снова начнет подниматься вверх по иерархии до тех пор, пока не найдется область, которая интерпретирует его как часть своей непосредственной активной последовательности. Отсюда вывод: сигналы передвигаются вверх по иерархии, а прогнозы – вниз. В идеале, в хорошо известном и прогнозируемом мире, эти встречные потоки сигналов быстротечны и циркулируют преимущественно в нижних областях коры головного мозга. Мозг как можно быстрее пытается найти ту часть своей модели мира, которая бы соответствовала неожиданному входному сигналу. Только после этого он может

понять входной сигнал и знать, чего ожидать дальше.

Когда я хожу по знакомой комнате в своем доме, лишь несколько ошибок поднимутся к верхним областям коры головного мозга. Хорошо усвоенные последовательности сигналов в моем доме обрабатываются на уровне нижних зон зрительной, соматосенсорной и моторной иерархии. Я знаю комнату настолько хорошо, что могу передвигаться по ней даже в кромешной тьме. Благодаря тому, что я хорошо знаю обстановку вокруг меня, большая часть моего неокортекса свободна и может размышлять о мозге и написании книг. В то же время если бы я находился в незнакомой мне комнате, особенно в такой, подобной которой я никогда не встречал, то мне не только пришлось бы постоянно присматриваться, куда идти. Неожиданные сигналы мощным потоком ринулись бы к верхним областям коры головного мозга. Чем больше мой сенсорный опыт не соответствует заученным последовательностям, тем больше ошибок будет возникать. В такой непривычной ситуации я не смогу думать о мозге, поскольку большая часть коры головного мозга будет занята проблемами ориентирования в комнате. С подобными ощущениями сталкиваются люди, сошедшие с трапа самолета в незнакомой стране. Дороги похожи на те, к которым вы привыкли, автомобили придерживаются левой стороны, язык, на котором все общаются, вам непонятен, а поиск туалета может стать сверхзадачей для всей коры головного мозга. Даже не пытайтесь готовиться к публичному выступлению, разгуливая по незнакомой стране.

Ощущение внезапного понимания (момент озарения) можно объяснить в рамках модели. Представьте, что вы смотрите на неоднозначную картинку. Пятна чернил и какие-то линии ничего вам не напоминают. Полнейшая бессмыслица. Когда мозг не может найти воспоминания, соответствующего полученному сигналу, человек пребывает в замешательстве. Ваши глаза блуждают по картинке. Входные сигналы поднялись на самый верх иерархии. Высшие зоны выдвигают множество самых разных гипотез, но, как только прогнозы начинают опускаться вниз по иерархии, каждая из гипотез конфликтует с входными сигналами, и мозг вынужден начинать все сначала. На протяжении всего состояния замешательства ваш мозг полностью поглощен тем, что пытается понять картинку. Наконец он формирует прогноз, который оказывается правильным. Когда это происходит, прогноз, сформированный в верхней зоне коры головного мозга, стремительно спускается вниз. Менее чем за секунду каждая зона получает последовательность, соответствующую полученным данным. Ни одна ошибка больше не транслируется к верхним зонам иерархии. Вы понимаете картинку и вместо точек и пятен видите далматинца (рис. 6.12).

На что способна обратная связь?

Десятки лет нам было известно, что связи в коре головного мозга являются взаимными. Если у зоны А есть проекция в зоне Б, то у зоны Б есть проекция в зоне А. Часто аксонов обратной связи оказывается больше, чем волокон прямой связи. Несмотря на то что все признают существование такого феномена, считается, что обратная связь играет второстепенную, или «модуляторную», роль в работе мозга. Мысль о том, что сигнал обратной связи на может постоянно и точно активизировать различные группы клеток слоя 2, не имеет большой поддержки в рядах нейробиологов.

Почему так? Частично потому, что, как я уже упоминал, если вы не считаете функцию прогнозирования центральной функцией головного мозга, то не возникает особой необходимости обращать внимание на обратную связь. Если вы предполагаете, что информация движется прямолинейно, к моторной системе, зачем вам эта обратная связь? Еще одна причина игнорирования обратной связи – то, что сигнал обратной, связи широко распространяется в слое 1. Интуитивно мы ожидаем, что сигнал, рассеянный в большой зоне, будет оказывать очень слабое влияние на большое количество нейронов. В самом деле, у мозга есть несколько таких модуляторных сигналов, которые не оказывают влияния на отдельные нейроны, зато меняют общие характеристики (например, возбудимость).



Рис. 6.12. Вы видите далматинца?

И последняя причина игнорирования обратной связи обусловлена тем, как многие нейробиологи представляют себе работу нейронов. У типичных нейронов имеются тысячи или даже десятки тысяч синапсов. Одни из них расположены ближе к телу клетки, другие – дальше. Синапсы, расположенные ближе к телу клетки, оказывают сильное влияние на ее импульсацию. Дюжина активных синапсов, расположенных у тела клетки, может заставить ее генерировать потенциал действия – электрический разряд. В этом не ничего нового. Тем не менее большинство синапсов находятся вдали от тела клетки. Они разбросаны по разветвленной структуре дендритов клетки. Поскольку эти синапсы расположены далеко от тела клетки, ученые склонны полагать, что импульс, пришедший на один из этих синапсов, будет очень слабо либо вообще не будет влиять на активность клетки. Импульс с отдаленного синапса рассеется по дороге, не достигнув тела клетки.

Как правило, информация, поднимающаяся вверх по иерархии коры головного мозга, передается по синапсам, расположенным близко к телу клетки. В этом случае сведения с большей вероятностью будут успешно передаваться от зоны к зоне. Как правило, обратная информация, поступающая вниз по иерархии коры головного мозга, передается по синапсам, расположенным дальше от тела клетки. Клетки слоев 2, 3 и 5 посылают дендриты в слой 1 и формируют там много синапсов. Слой 1 представляет собой совокупность синапсов, но все они значительно отдалены от тел клеток слоев 2, 3 и 5. Более того, каждая отдельная клетка слоя 2 формирует чрезвычайно мало синапсов с каждым отдельным волокном обратной связи (если формирует их вообще). Следовательно, некоторые ученые могут возразить по поводу того, что непродолжительный сигнал в слое 1 может вызвать упорядоченное возбуждение группы клеток слоев 2, 3 и 5, однако моя теория базируется именно на этом предположении.

Дилемма разрешается благодаря тому, что фактическое поведение нейронов отличается от поведения, приписываемого им классической моделью. В последние годы многие ученые придерживаются мнения, что синапсы на отдаленных тонких дендритах могут играть самую непосредственную роль в возбуждении клеток. В указанных моделях эти дальние синапсы функционируют не так, как синапсы на более толстых дендритах, расположенных вблизи тела клетки. Например, если два синапса на тонком дендрите расположены очень близко друг к

другу, они будут действовать как «детектор совпадения». Другими словами, если оба синапса получили входящий импульс близко во времени, то они могут оказать большое влияние на клетку, хотя и расположены далеко от ее тела. Они могут заставить тело клетки генерировать разряд. Поведение дендритов нейронов до сих пор остается загадкой для науки, поэтому я не могу остановиться на этом вопросе более подробно. Важным моментом является то, что модель коры головного мозга «Память-предсказание» требует, чтобы синапсы, отдаленные от тела клетки, могли идентифицировать специфические сигналы.

Не слишком прозорливым, почти глупым было заявление, что большинство из тысяч синапсов нейрона играют только модуляторную роль. Обширная обратная связь и огромное количество синапсов не могли возникнуть без веской причины. Используя наши знания, мы можем сказать, что типичный нейрон обладает способностью выучить сотни точных совпадений благодаря волокнам обратной связи, когда они создают синапсы на тонких дендритах. Это, в свою очередь, означает, что колонка коры головного мозга очень гибка в отношении того, какие сигналы обратной связи могут вызвать ее активность. Таким образом, каждое отдельное свойство можно точно ассоциировать с тысячами различных объектов и последовательностей. Моя модель требует того, чтобы обратная связь была быстрой и точной. Клетки должны возбуждаться, когда они обнаруживают любое количество точных совпадений на отдаленных дендритах. Новая нейронная модель основывается именно на этом.

Как происходит процесс обучения

У всех нейронов из всех слоев коры головного мозга есть синапсы, большинство из которых может модифицироваться под влиянием опыта. Можно с уверенностью сказать, что процесс обучения и запоминания происходит во всех слоях, во всех колонках и во всех зонах коры головного мозга. Я упоминал ранее правило «обучения Хебба», названное так в честь Дональда Хебба. Сущность этого правила очень проста: когда два нейрона активизируются одновременно, синаптические связи между ними усиливаются. (Его легко запомнить как фразу «Возбуждены вместе – связаны вместе».) Сейчас мы знаем, что Хебб был прав. Конечно, ничего в природе не бывает настолько простым, свойства настоящего мозга сложнее. Наша нервная система работает на основе многих вариаций «обучения Хебба». Некоторые синапсы изменяют силу связи в ответ на незначительные отклонения координации нейронных сигналов. Некоторые синаптические изменения оказываются краткосрочными, некоторые – долгосрочными. Хебб очертил основные рамки, а не создал конечную теорию процесса обучения, причем эти рамки оказались очень и очень полезными.

Принципы «обучения Хебба» объясняют большинство проявлений поведения коры головного мозга, о которых мы говорили в данной главе. Помните, еще в 1970-х годах было установлено, что автоассоциативная память при помощи алгоритма «обучения Хебба» может усвоить пространственные сигналы и их последовательности. Основная проблема заключалась в том, что автоассоциативная память не в состоянии справиться с отклонениями. Согласно теории, предложенной в нашей книге, кора головного мозга нашла выход из такой ситуации, частично путем складывания автоассоциативных запоминаний в иерархию, а частично – путем использования сложной столбчатой архитектуры. Эта глава почти полностью посвящена иерархической структуре мозга и ее работе, поскольку именно иерархия обуславливает высокую мощность коры головного мозга. Вместо того чтобы пускаться в пространственные объяснения, как та или иная клетка учится тем или иным вещам, я хотел бы раскрыть несколько общих принципов обучения в иерархии.

У новорожденного кора головного мозга «не знает» фактически ничего. Она не знает его языка, его культуры, его дома, его города, песен, людей, вместе с которыми он будет расти, – ничего. Всю эту информацию – структуру внешнего мира – нужно выучить. Основными составляющими обучения являются формирование классификаций и создание последовательностей. Эти два взаимодополняющих компонента взаимодействуют друг с другом. По мере того как одна зона коры головного мозга учит последовательности, входящие сигналы, которые она посылает в клетки слоя 4 зоной выше в иерархии, меняются. Клетки слоя 4 учатся формировать новые классификации, что изменяет сигнал, проецируемый назад, в слой

1, зоной ниже, а это влияет на последовательности.

Основой формирования последовательностей является группирование паттернов, которые относятся к одному и тому же объекту. Один из способов – группировать сигналы, следующие друг за другом. Когда ребенок держит в руке игрушку и медленно ее переворачивает, его мозг может с уверенностью считать, что непрерывно изменяющееся изображение на сетчатке все равно принадлежит одному и тому же объекту, а значит, изменяющийся набор сигналов можно сгруппировать вместе. Иногда нужна внешняя подсказка, чтобы определить, что сигналы принадлежат одному и тому же целому. Чтобы выучить, что яблоки и бананы принадлежат к фруктам, а морковь и картофель – нет, нужен учитель, который помог бы сгруппировать указанные объекты как фрукты. В любом случае ваш мозг создает последовательности сигналов, принадлежащих одному целому. Однако когда зона мозга уже создала последовательности, то входной сигнал, поступающий в следующую область, изменяется. Если раньше он представлял отдельные сигналы, то теперь представляет группу сигналов. Входной сигнал превращается из ноты в мелодию, из букв в слово, из носов в лица и так далее. Поскольку сигналы снизу вверх становятся более объектно-ориентированными, высшая зона коры головного мозга теперь может запоминать последовательности этих объектов высшего порядка. Там, где раньше строились последовательности букв, теперь строятся последовательности слов. Неожиданным результатом процесса обучения является то, что на протяжении повторного обучения представления объектов передвигаются вниз по иерархии. В первые годы вашей жизни «записи» о мире формируются в высших зонах коры головного мозга. Однако по мере вашего дальнейшего обучения они перемещаются во все более низкие зоны иерархии коры головного мозга. Не то чтобы мозг их передвигал туда, просто он должен учить их заново, снова и снова. (Я не хочу сказать, что все запоминания обязательно начинаются с высших зон иерархии коры головного мозга. В реальной жизни усвоение информации выглядит сложнее. Я считаю, что классификация сигналов слоя 4 начинается снизу и передвигается вверх. По мере передвижения формируются последовательности, которые опускаются вниз. Запоминания последовательностей опускаются все ниже и ниже в коре головного мозга.) Когда простые репрезентации передвигаются вниз, высшие зоны коры головного мозга получают возможность учить новые, более сложные, сигналы.

Создание и передвижение вниз иерархически организованных воспоминаний можно наблюдать на примере обучения ребенка. Как мы постигаем чтение? Первое, чему нас учат, – распознавать отдельные печатные буквы. Это медленный и сложный процесс, требующий сознательных усилий. Потом мы переходим к распознаванию простых слов. И снова вначале это дается нам с большим трудом и очень медленно, даже если слова состоят из трех-четырех букв. Ребенок может прочесть каждую букву последовательности и произнести каждую из них поочередно, но понадобится достаточно большое количество времени, прежде чем слово будет распознано как слово. Выучив простые слова, мы переходим к изучению многосложных. Сначала мы произносим каждый слог, сцепляя их, как делали это в свое время с простыми словами. После многолетней практики человек способен читать достаточно быстро. Мы достигаем уровня, когда видим не отдельные буквы, а распознаем целые слова и даже фразы с первого взгляда. Дело не в том, что мы стали более быстрыми, а в том, что мы распознаем слова и фразы как целостные единицы. Когда мы читаем все слово, видим ли мы буквы этого слова? И да, и нет. Конечно, сетчатка глаза видит буквы, а значит, и клетки зоны V1 тоже. Однако распознавание букв происходит относительно низко в иерархии коры головного мозга, скажем, в зоне V2 или V4. К тому моменту, когда сигнал попадает в IT, отдельные буквы уже не будут представлены сигналом. То, что вначале требовало деятельности всей зоны коры головного мозга – распознавание отдельных букв, – теперь происходит ближе к сенсорному входному сигналу. Когда запоминания простых объектов (букв) перемещаются вниз по иерархии, высшие зоны коры головного мозга получают возможность учить сложные объекты, такие как слова или фразы.

Обучение чтению музыки по нотам может служить еще одним примером. Сначала нужно сосредоточиться на каждой ноте. С приобретенным опытом вы начинаете распознавать простые нотные последовательности, затем – целые фразы. После достаточно длительной практики все происходит так, как если бы вы не видели нот вообще. Нотный лист служит вам

для того, чтобы помнить общую структуру музыкального фрагмента, а детальные последовательности сохранились в памяти на более низком уровне коры головного мозга. Такой тип обучения происходит как в двигательной, так и в сенсорной зонах.

Молодой мозг медленнее распознает входные сигналы, медленнее формирует моторные команды, потому что запоминания, используемые в этих задачах, расположены выше в иерархии. Информация должна пройти весь путь вверх и вниз, возможно, неоднократно, для того, чтобы неясности разрешились. Передвижение сигналов вниз и вверх по иерархической структуре коры головного мозга требует времени. В молодом мозге еще не сформировались сложные последовательности в высших областях, поэтому он не может распознать и отреагировать сложными сигналами. Молодой мозг не может понять высшую структуру внешнего мира. По сравнению со взрослым язык ребенка прост, его музыка незамысловата, и его социальные интеракции тоже просты.

Если вы изучаете определенный набор объектов снова и снова, то ваша кора головного мозга смещает в памяти последовательности объектов в нижнюю часть иерархии. Это высвобождает верхние зоны иерархии для изучения более сложных и тонких связей. Согласно моей теории, именно таков путь становления эксперта.

Например, я работаю в области компьютерного дизайна, и многие удивляются тому, что я, только взглянув на продукт, уже вижу его недостатки, заложенные в архитектуре. Проработав двадцать пять лет в сфере компьютерного дизайна, я выработал собственную модель всевозможных проблем, связанных с мобильными компьютерами. Точно так же умудренный опытом родитель сможет без труда понять, чем расстроен его ребенок, в то время как новоиспеченный отец будет ломать голову над тем, как же решить возникшую проблему. Опытный менеджер может сразу распознать недостатки и преимущества структуры организации, в то время как начинающий просто не понимает пока этих вещей. Они получают один и тот же входной сигнал, но модель новичка не настолько совершенна, как модель опытного менеджера. В этих и тысяче других примеров мы начинаем с изучения простейшей базовой структуры. Со временем наши знания передвигаются вниз по иерархии, и у наших высших зон появляется возможность освоить структуру более высокого порядка. Именно структуры высшего порядка делают нас опытными. Эксперты и гении обладают мозгом, видящим структуру структуры и сигнал сигналов глубже, чем видят остальные. Можно стать экспертом благодаря практике, но, конечно, кроме этого, существует и генетический компонент таланта и гениальности.

Гиппокамп – вершина всех вершин

Три большие структуры мозга лежат под оболочкой коры головного мозга и связаны с ней. Это базальные ганглии, мозжечок и гиппокамп. Все три структуры возникли раньше, чем кора головного мозга. В очень грубом приближении мы можем утверждать, что базальные ганглии были примитивной двигательной системой, мозжечок изучал точные временные соотношения событий, а гиппокамп сохранял в памяти конкретные события и места. В определенной степени кора головного мозга присвоила себе функции, изначально принадлежавшие им. Например, человек, родившийся без мозжечка, будет страдать от недостатков координации и вынужден будет прилагать более сознательные усилия при передвижении, но в остальном он будет вполне нормальным.

Мы знаем, что кора головного мозга отвечает за все сложные последовательности движений и может непосредственно управлять нашими конечностями. Не то чтобы базальные ганглии не имели никакого значения, просто кора головного мозга узурпировала большую часть моторного контроля. Поэтому я описал общее функционирование коры головного мозга независимо от базальных ганглиев и мозжечка. Некоторые ученые не согласятся с таким моим предположением, тем не менее я использовал его в данной книге и в своей работе вообще.

Но вот гиппокамп – другого поля ягода. Это одна из наиболее изученных зон мозга, потому что она является обязательной для формирования новых запоминаний. Если вы потеряете обе половины гиппокампа (как и многие другие составляющие нервной системы, он присутствует и в левом, и в правом полушариях головного мозга), вы утратите способность

запоминать новую информацию. Без гиппокампа вы по-прежнему можете разговаривать, ходить, видеть, слышать и даже казаться нормальным в течение небольшого временного промежутка. На самом деле вам причинен глубокий ущерб, ведь вы не можете запомнить ничего нового. Вы будете помнить друга, которого знали до потери гиппокампа, но не сможете запомнить нового человека. Даже если бы вы стали встречаться с доктором пять раз в неделю на протяжении года, все равно каждый последующий раз для вас был бы как первый. У вас бы не было воспоминаний о событиях, которые произошли после того, как вы утратили гиппокамп.

Долгие годы у меня не было желания заниматься гиппокампом, поскольку для меня в этом было мало смысла. Гиппокамп является важным для процесса обучения, но не является окончательным хранилищем наших знаний. А вот кора головного мозга таковым является. Классический подход к гиппокампу следующий: здесь формируется память о новых событиях, а затем, через дни, недели, месяцы, эта новая информация перемещается в кору головного мозга. Для меня это звучало как бессмыслица. Мы знаем, что зрительные, слуховые и осязательные сигналы, т. е. сенсорный поток данных, поступают прямо в сенсорные зоны коры головного мозга без предварительного прохождения через гиппокамп. Мне казалось, что сенсорная информация автоматически формирует новые запоминания в коре головного мозга. Зачем нужен гиппокамп в процессе обучения? Как может отдельная структура под видом гиппокампа воспрепятствовать обучению или только позднее передать информацию назад, коре головного мозга?

Я решил отложить в сторону гиппокамп, полагая, что однажды его роль прояснится. Так и случилось. Это произошло в конце 2002 года, как раз когда я начал писать эту книгу. Брюно Ольшозен, один из моих коллег из Редвудского института нейрологии, сказал, что связи между гиппокампом и корой головного мозга наводят на мысль о том, что гиппокамп является частью коры головного мозга, а не отдельной структурой. В таком случае гиппокамп расположен на вершине иерархической пирамиды коры головного мозга, это ее самый верхний отдел (см. рис. 6.5). Кора головного мозга появилась на эволюционной сцене, втиснувшись между гиппокампом и остальной частью мозга. Конечно, идея о том, что гиппокамп венчает пирамиду неокортекса, была не нова, но я о ней не знал. Я пообщался с несколькими специалистами в области исследования гиппокампа и попросил их объяснить, как эта структура передает сохраненную информацию в кору головного мозга. Никто мне объяснить не смог. Более того, никто из них не упомянул, что гиппокамп является вершиной пирамиды коры головного мозга, возможно, потому, что гиппокамп не только расположен на вершине корковой пирамиды, но еще и непосредственно связан со многими более старыми частями мозга.

Несмотря на это, я сразу понял, что нашел ключ к решению волновавшей меня задачи.

Вспомните об информационном потоке, поступающем от ваших глаз, ушей, кожи в кору головного мозга. Каждая зона коры головного мозга пытается понять, что означает полученная информация. Каждая зона пытается объяснить входящий сигнал в рамках известных ей последовательностей. Если она понимает входящий сигнал, то говорит: «Мне это понятно. Это часть объекта, который я уже вижу. Подробности передавать дальше не будем». Если зона не понимает его, она передает его без изменений в более высокую зону иерархии коры головного мозга. Так продолжается до тех пор, пока следующая зона не распознает входной сигнал. Если сигнал является абсолютно новым, он будет подниматься все выше и выше по иерархии. Каждая последующая зона скажет: «Я этого не знаю. Я такого не ожидала. Почему бы вам, высшим областям, не взглянуть на это?» Конечным результатом будет то, что в верхней точке пирамиды коры головного мозга появится информация, которую невозможно понять на основе прошлого опыта. Вы сталкиваетесь с совершенно новым сигналом, которого совсем не ожидали. В повседневной жизни мы встречаем очень много самых разнообразных вещей, информация о которых добирается до верхней части пирамиды. Например, заметка в газете, имя человека, которого вы встретили сегодня утром, автомобильная авария, свидетелем которой вы стали по дороге домой. Вот эта необъясненная и неожиданная информация и попадает в гиппокамп и сохраняется там. Но хранится она там не бесконечно. Позже она либо передается в кору головного мозга, либо исчезает.

Я заметил за собой, что с годами мне становится все труднее и труднее запоминать новые вещи. Например, мои дети помнят подробности всех театральные постановки, которые они

увидели в прошлом году. Я же не могу этим похвастаться. Возможно, потому, что за всю свою жизнь я увидел такое количество спектаклей, что теперь мне редко попадает что-то на самом деле новое. Новые представления соответствуют воспоминаниям о старых представлениях, возможно, поэтому информация просто не попадает в мой гиппокамп. Если так оно и есть, то приходим к выводу, что, чем больше мы знаем, тем меньше запоминаем.

В отличие от коры головного мозга структура гиппокампа гетерогенная, он состоит из нескольких специализированных отделов. Гиппокамп быстро запоминает любые полученные сигналы. Это его уникальная роль, с которой он отлично справляется. Положение гиппокампа на вершине корковой иерархии оптимально для запоминания новых сигналов. У него также наилучшее положение для того, чтобы вызывать эти запоминания и передавать их на хранение в иерархию коры головного мозга. Правда, последний процесс – весьма небystрый. Вы можете моментально запомнить свежее событие в гиппокампе, но, для того чтобы сохранить его в коре головного мозга навсегда, вам нужно повторять полученный опыт снова и снова, в реальности или мысленно.

Альтернативный путь вверх по иерархии

У вашей коры головного мозга существует еще один основной способ передачи информации от одной зоны к другой вверх по иерархии. Этот альтернативный путь начинается с клеток слоя 5, проецирующихся в таламус (в его другую область, не в ту, о которой мы упоминали раньше), а затем из таламуса – в следующую корковую зону. Любые две зоны коры головного мозга, непосредственно связанные друг с другом в иерархии, связаны еще и косвенно – через таламус. Во втором случае информация передается только вверх по иерархии, но не вниз. Когда мы движемся вверх в иерархии коры головного мозга, то существует прямой путь – между областями – и путь косвенный – через таламус.

У второго пути есть два режима работы, предопределенные клетками таламуса. Первый режим: путь закрывается и не пропускает информацию. Во втором режиме информация в точности передается от области к области. Этот косвенный путь описали двое ученых, Мюррей Шерман (Нью-Йоркский университет, Стоуни Брук) и Рэй Гиллери (Медицинская школа университета Висконсина). Они полагают, что данный путь может оказаться не менее важным, чем прямой (а может, даже еще более важным, что и было предметом данной главы). У меня есть определенные соображения по поводу того, зачем нужен этот второй путь.

Прочтите это слово: *воображение*. Большинство людей распознают его с одного взгляда, с одной фиксации. Посмотрите на букву «а» в середине этого слова. Теперь посмотрите на хвостик буквы «а». Ваши глаза могут быть все время зафиксированы на одном и том же месте, но в одном случае вы видите слово, в другом – букву, а в третьем – ее хвостик. Смотрите, не отрываясь, на букву и постарайтесь чередовать ваше восприятие слова, буквы и хвостика. Во всех случаях зона V1 получит одну и ту же входящую информацию, но, когда она доберется до высших зон, таких как IT, например, вы будете воспринимать разные вещи, разные уровни детализации. Зона IT знает, как распознать все три объекта. Она может распознать хвостик буквы сам по себе, букву саму по себе и слово целиком с одного взгляда. Когда вы воспринимаете слово целиком, зоны V4, V2, V1 берут на себя детали объекта, а зона IT «узнает» только слово. Во время чтения мы, как правило, не воспринимаем отдельные буквы. Мы воспринимаем слова и фразы. Но если вам захочется, то вы можете распознавать буквы. Мы смещаем акцент внимания очень часто, хотя практически никогда не отдаем себе в этом отчета. Я могу слышать музыку как фон, почти не осознавая этого, но если я постараюсь, то смогу вычленить мелодию ударных инструментов или бас-гитары. В мой мозг поступает один и тот же звук, но я могу сфокусировать свое восприятие. Всякий раз при расчесывании волос возникает сильный внутренний звук, тем не менее вы обычно не осознаете наличие шума. Сосредоточившись, вы сможете очень хорошо услышать данный звук. Все перечисленное – еще один пример сенсорного входящего сигнала, который, как правило, подлежит обработке в нижних слоях коры головного мозга. Но, если вы приложите сознательное усилие, он может попасть и в высшие зоны.

Я предполагаю, что альтернативный путь через таламус является тем механизмом,

который мы используем, концентрируясь на подробностях, которых, как правило, не замечаем. Он пропускает группирование последовательностей в слое 2, посылая необработанные данные в более высокую зону иерархии. Как доказали биологи, альтернативный путь возможно активизировать двумя способами. Первый – сигналом из высшей зоны самой коры головного мозга. Именно его вы использовали, когда я просил вас обратить внимание на подробности, опускаемые в обычных обстоятельствах (например, хвостик буквы «а» или звук при расчесывании волос). Второй способ – мощный неожиданный сигнал снизу. Если входной сигнал достаточно силен, то он посылает сигнал возбуждения в более высокую область, которая «включает» указанный путь. Например, если бы я показал вам лицо и спросил, что это было, вы бы ответили: «Лицо». Если бы я вам опять показал лицо, но со странной меткой на носу, то вы все равно сразу бы опознали лицо. Затем ваши низшие уровни зрительного восприятия заметили бы: что-то не так. Выявленная ошибка вызывает активизацию дорожки повышенного внимания. Подробности будут передаваться альтернативным путем, пропуская группирование, которое происходит в нормальных условиях. Ваше внимание остановится на метке. Теперь вы видите не только лицо, но и метку. Если она выглядит достаточно необычно, то может полностью завладеть вашим вниманием.

Так необычные события быстро попадают в зону нашего внимания. Именно поэтому нам очень трудно не обращать внимания на физические изъяны или какие-либо необычные сигналы. Наш мозг делает это автоматически. Однако очень часто сообщения об ошибке оказываются недостаточно сильными для того, чтобы открыть альтернативный путь. Поэтому иногда мы не замечаем орфографических ошибок в словах, которые читаем.

Заключительные размышления

Для того чтобы найти и разработать новую научную модель, мы должны искать наиболее простые подходы, которые могли бы объединить и объяснить множество разрозненных фактов. Неизбежным следствием такого процесса является то, что по инерции мы можем сильно упрощать ситуацию. Можно пропустить важные детали и ошибочно интерпретировать факты. Но если модель в общем будет принята, то можно будет ее усовершенствовать и исправить ошибки. Постепенно станет ясно, где мы зашли слишком далеко, где мы копали недостаточно глубоко, а где ошибались.

В этой главе я изложил несколько гипотез по поводу работы коры головного мозга. Полагаю, некоторые из них окажутся ошибочными, возможно также, что все гипотезы будут пересмотрены. Осталось очень много деталей, которые не были мной даже упомянуты. Мозг сам по себе очень сложен. Нейробиологи, читающие эту книгу, знают, что я представил только схематическую характеристику сложнейшего биологического устройства – живого человеческого мозга. Тем не менее я считаю, что предложенная мной теоретическая модель в целом выглядит убедительно. Я надеюсь, что, несмотря на изменения по мере появления новых данных, мои основные теоретические положения все-таки окажутся верными.

И наконец, возможно, вы все еще не можете поверить в то, что простая, но обширная система запоминания в самом деле может отвечать за всю деятельность человека. Разве возможно, чтобы вы и я были просто иерархической системой запоминания? Разве возможно, чтобы наши жизни, верования и амбиции хранились в триллионах крошечных синапсов?

В 1984 году я начал работать в сфере создания компьютерных программ. Конечно, я писал небольшие программы и раньше, но это был мой первый опыт программирования компьютера с графическим интерфейсом для пользователя. Это был первый опыт работы над большими и сложными приложениями. Я создавал программное обеспечение для операционной системы, созданной *Grid Systems*. Там были окна, много шрифтов, пунктов меню. Операционная система указанной компании на то время была очень продвинутой. Однажды я был поражен мыслью, что занимаюсь чем-то практически невозможным. Как программист, я мог писать только одну строку кода за другой. Потом я группировал строки кодов в блоки, называемые *подпрограммами*. Затем подпрограммы объединялись в *модули*. Объединение модулей давало в результате приложение. Программа электронной таблицы, над которой я работал, состояла из такого большого количества подпрограмм и модулей, что ни один человек не был в состоянии

понять ее всю. Она была очень сложной. Тем не менее одна строка кода в ней мало что делает. Требовалось несколько строк кода, чтобы отобразить только один пиксель на экране. Чтобы заполнить весь экран для таблицы, компьютеру надо было выполнить миллионы команд, переданных сотнями подпрограмм в повторяющемся и рекурсивном порядке. Все это было настолько сложно, что было невозможно предвидеть, *что* происходит, когда программа работает. Я был поражен: насколько нереалистичным казалось то, что при работе программы экран будет заполнен чуть ли не мгновенно. Внешне плоды работы программы выглядели как таблицы с числами, ярлыки, текст, графика. Она работала как электронная таблица. Однако я знал, *что* происходит внутри компьютера, процессор которого постоянно выполняет одну и ту же простую операцию. Трудно было поверить в то, что компьютер может сориентироваться в лабиринте модулей и подпрограмм и выполнить команды настолько быстро. Если бы у меня не было предварительных знаний, я бы попросту не поверил, что все это будет работать. Если бы кто-нибудь изобрел компьютер с графическим интерфейсом для пользователя и приложением электронных таблиц и принес бы мне свое изобретение на бумаге, я бы отклонил его как не имеющее практического применения. Я бы возразил: такой машине понадобится вечность, чтобы что-либо сделать. И тут мне стало стыдно за себя, ведь на самом деле она работает. Я понял, что мое интуитивное представление о скорости процессора и мое интуитивное представление о мощности иерархического строения не соответствуют действительности.

Здесь уместно привести аналогию с корой головного мозга. Мозг не состоит из сверхбыстрых компонентов, правила, которых он придерживается в своей работе не так уж сложны. Однако есть одно «но»: кора головного мозга обладает иерархической структурой, она состоит из миллиардов нейронов и триллионов синапсов. Нам трудно представить, как такая логически простая, но по численности обширная система памяти может создавать наше сознание, язык, культуру, искусство, книги, науку и технологии. Однако это происходит, потому что наше интуитивное представление о способностях коры головного мозга и мощностях ее иерархической структуры является неадекватным. Кора головного мозга на самом деле работает. В этом нет никакой магии. Это доступно нашему пониманию. И точно так же, как в случае с компьютером, в конечном счете мы сможем создать разумные машины, работающие по тому же принципу.

7. Сознание и творчество

Когда я рассказываю о своей теории мозга, слушатели почти всегда соглашались со значением прогнозирования как ядра человеческой деятельности. Но обычно мне задают много дополнительных вопросов: что такое сознание? что такое воображение? как рождается творчество? как отличить реальность от ложных представлений и так далее. Хотя мотивацией для создания моей теории мозга послужили совсем не перечисленные вопросы, тем не менее они интересны практически каждому из нас. При обсуждении подобных тем я не претендую на роль знатока, но модель «Память-предсказание» интеллекта может послужить основой для ответов на некоторые вопросы. В этой главе я хотел бы обратиться к наиболее часто задаваемым.

Обладают ли животные разумом?

Есть ли интеллект у крысы? Или у кошки? На какой стадии эволюции у живых существ возник разум? Я люблю отвечать на подобные вопросы, потому что ответы бывают весьма неожиданными. Все, что я рассказал вам о коре головного мозга и ее функционировании, справедливо благодаря важной предпосылке: окружающий мир структурирован, и, следовательно, прогноз может быть осуществим. В мире существуют модели и формы: у любого лица есть глаза, у глаз есть зрачки, огонь обжигает, благодаря силе гравитации предметы падают на землю, дверь закрывается и открывается и так далее. Мир является упорядоченным, но не однородным. Память, прогнозирование и поведение были бы бессмысленны, если бы в мире царил хаос. Все виды поведения, будь то человеческая

деятельность, или поведение улитки, или функционирование одноклеточных и даже растений, есть не что иное, как разные способы исследования структуры мира с целью воспроизведения.

Представьте себе одноклеточное, обитающее в пруду. У простейшего есть жгутики, позволяющие ему плавать. Поверхность его образована молекулами, выявляющими наличие или отсутствие питательных веществ. Поскольку не все зоны пруда имеют одинаковую концентрацию последних, то восприятие одноклеточным организмом питательности окружающей среды неоднородно, оно может меняться от одной стороны клетки к другой. Перемещаясь по пруду, простейшее распознает эти изменения. Как видите, мир одноклеточных чрезвычайно прост. Клетка пополняет свою «осведомленность», перемещаясь в зоны с более высокой концентрацией питательных веществ. Можно сказать, что даже этот простейший организм способен осуществлять прогностическую функцию: он «предвидит», что, перемещаясь в определенном направлении, получит из окружающей среды больше питательных веществ. Включена ли в данный примитивнейший процесс прогнозирования память? Да. Запоминание сохраняется в ДНК простейшего организма, причем обучение происходит не на протяжении жизни конкретной особи, а в процессе эволюционного развития. Если бы структура мира внезапно изменилась, то одноклеточный организм не смог бы приспособиться, поскольку ему не удалось бы изменить свою ДНК и обусловленное ею поведение. Для этих простейших видов процесс обучения реализуется исключительно посредством эволюционного развития на протяжении многих поколений.

Разумен ли одноклеточный организм? Если под «разумом» понимать человеческий интеллект, то ответ будет отрицательным. Однако, если использовать академический подход к понятию разума и учесть, что это одноклеточное находится в самом дальнем конце континуума видов, использующих прогнозирование и память для оптимизации репродуктивных процессов, ответ будет положительным. Суть ведь не в том, чтобы представителей одних видов окрестить «разумными», а других, наоборот, «лишенными интеллекта». Память и прогнозирование используются абсолютно всеми живыми существами. Разница в методах – от простых до самых изощренных, – которые они используют.

Растения также используют память и прогнозирование для исследования структуры мира. Когда дерево посылает свои корни глубоко под землю, а листья – к небу, оно тоже составляет определенный прогноз. Генетическая память "дерева позволяет ему прогнозировать, где можно получить воду и необходимые минеральные вещества. Конечно, дерево не думает, его поведение, или познание структуры окружающей среды, осуществляется автоматически.

В процессе филогенеза у растений развились системы коммуникации, преимущественно базирующиеся на медленном высвобождении химических сигналов. Если какую-то часть коры повредило насекомое, дерево посылает по своей сосудистой системе химические элементы, активизирующие защитную систему, что, в свою очередь, стимулирует выработку токсинов. Благодаря такой системе коммуникации дерево может проявлять и немного более сложные формы поведения.

Возможно, нейроны возникли из-за необходимости передавать информацию с большей скоростью, чем посредством сосудистой системы растений. Нейрон можно представить себе как клетку со своими собственными сосудистыми отростками. В какой-то момент, вместо того чтобы передавать химические элементы по этим отросткам, нейроны начали использовать электрохимические импульсы, имеющие намного более высокую скорость передвижения. В самом начале синаптическая передача и простые нервные системы, вероятно, не были способны к обучению. Их задача состояла лишь в том, чтобы как можно быстрее передать сигналы.

Однако со временем, в процессе эволюционного развития, произошло нечто весьма интересное. Связи между нейронами стали модифицируемыми. Нервные клетки обрели возможность посылать или не посылать сигнал – в зависимости от текущей ситуации. Совершенствовалась система памяти, и, следовательно, живые организмы стали обучаться и познавать структуру мира на протяжении своей собственной жизни, а не исключительно в процессе филогенеза. Таким образом, нервная система стала более гибкой, а поведение – модифицируемым. При внезапном изменении окружающей среды (скажем, появлении нового хищника) животные больше не были обречены неуклонно следовать генетически обусловленной модели поведения, а могли находить более уместные в данной ситуации схемы

и следовать им. Гибкая нервная система дала толчок эволюционному развитию, привела к появлению новых видов, начиная рыбами и улитками и заканчивая человеком.

Как мы помним из главы 3, у всех млекопитающих есть «старый» мозг, покрытый сверху неокортексом – корой головного мозга. Кора головного мозга – это нервная ткань, развившаяся в процессе эволюции позже всего. Благодаря своей иерархической структуре, наличию инвариантных репрезентаций и прогностической функции, опирающейся на приобретенный опыт, кора головного мозга позволяет млекопитающим намного более полно (по сравнению с животными без таковой) исследовать структуру мира. Благодаря неокортексу наши предки научились плести сети и ловить рыбу. Но рыба не смогла выучить, что сеть означает опасность, или придумать инструменты, позволяющие ее разрезать. Кора головного мозга есть у всех млекопитающих – крыс, кошек, людей, – все они разумны, но в разной степени.

Каковы особенности человеческого интеллекта?

Модель «Память-предсказание» предполагает два ответа на поставленный вопрос. Первый достаточно очевиден: кора нашего головного мозга больше по размеру, чем неокортекс обезьяны или собаки. Когда в процессе эволюции кора головного мозга *Homo sapiens* увеличилась до размеров большой обеденной салфетки (в распластанном виде), люди научились строить многоплановые и всеобъемлющие модели мира, а также формировать сложные прогнозы. Нам доступны гораздо более глубокие аналогии, чем другим млекопитающим. Мы осознаем и анализируем сложнейшие структуры. В поисках своей второй половины мы не только рассматриваем такую характеристику претендентов или претенденток как здоровье, но также знакомимся и общаемся с его или ее родителями и друзьями, оцениваем поведение и нравственные качества будущего мужа или жены. На основании указанных вторичных характеристик мы строим прогноз о том, как девушка или парень поведет себя в будущем. Трейдеры на рынке недвижимости ищут структуру в торговле. Математики ищут структуру в цифрах и уравнениях. Астрономы ищут структуру в расположении планет и звезд. Большая кора головного мозга позволяет нам воспринимать свой дом как часть города, который является частью региона, который является частью планеты, который является частью Вселенной. Другими словами, речь идет о восприятии структуры в структуре. Ни одно другое млекопитающее не может углубиться в размышления такого уровня. Например, я убежден, что моя кошка не имеет представления об устройстве мире за дверью нашего дома.

Второе отличие человеческого интеллекта от разума других млекопитающих состоит в том, что у последних отсутствует система речи. Языку как феномену, присущему исключительно человеку, были посвящены сотни исследований, о нем были написаны тома литературы. Однако система речи точно соответствует модели «Память-предсказание», причем понимание ее в подобном ключе не требует разработки и использования специальной терминологической базы, привязанной исключительно к речи. Слова, написанные от руки, напечатанные или произнесенные вслух, являются не чем иным, как сигналами внешнего мира, точно так же, как и мелодии, автомобили, дома. Синтаксис и семантика языка мало отличаются от иерархической структуры остальных объектов, привычных нам. И, подобно тому как звук приближающегося поезда вызывает появление визуального прогноза – «картинки» прибытия состава на вокзал, – слова ассоциируются с их физическими и семантическими аналогами, сохраняющимися в нашей памяти. Посредством языка один человек может активизировать воспоминания и создавать противопоставления мысленных объектов у другого человека. По сути, язык – это череда аналогий, благодаря которым люди могут представлять и познавать объекты окружающей действительности, не воспринимаемые ими непосредственно. В филогенезе появление речи требовало увеличения у предков человека неокортекса, а также развития моторной зоны коры головного мозга и скелетной мускулатуры, в частности артикуляционной. Владая языком, мы можем передавать свой жизненный опыт другим людям, а также своим потомкам. Как письменность, так и устная речь, отражаясь в культурных традициях, стали средством передачи знаний о мире из поколения в поколение. В наши дни печатные и коммуникационные средства позволяют нам иметь доступ к знаниям миллионов людей во всем мире. Остальные млекопитающие, не имеющие системы речи, даже в самом

грубом приближении не могут передать такое количество информации своим отпрыскам. Крыса на протяжении своей жизни способна выучить и запомнить значения множества сигналов, но она не научит этому своих детенышей. А все потому, что она не может сказать: «Смотри, малыш, вот так и так мой отец учил меня избегать ударов электрического тока».

Таким образом, в развития интеллекта можно выделить три этапа, причем на каждом из них использовалась память и прогнозирование. На первом этапе простейшие организмы использовали ДНК как средство запоминания и прогнозирования. Каждая отдельная особь не могла обучиться и приспособиться в течение своей жизни, она была способна лишь передать своим потомкам информацию о мире, полученную посредством ДНК от предков.

Второй этап начался, когда природа изобрела модифицируемые нервные системы, способные быстро усваивать сведения, получаемые из внешней среды, и сохранять их в памяти. На данном этапе отдельные особи могли усваивать структуру мира и адаптироваться к ней в рамках своей жизни. Но механизмы передачи приобретенных знаний, навыков и умений другим представителям своего вида, равно как и потомкам, отсутствовали (частично это было осуществимо путем непосредственного наблюдения). Именно на втором этапе у живых существ появилась (а со временем и увеличилась) кора головного мозга, но это было только начало.

Третьего и последнего этапа развития интеллекта достиг только человек. Этот этап ознаменовался значительным увеличением неокортекса и возникновением системы речи. Мы, люди, способны не только глубоко изучить структуру мира и наследие предков, но также посредством речи передать свои знания другим. Так осуществляется связь между поколениями.

Мы с вами как раз сейчас принимаем участие в подобном процессе. Значительную часть своей жизни я посвятил исследованию головного мозга и интеллекта. Конечно, я не достиг бы своей цели, если бы не имел доступа к информации, собранной сотнями ученых, у которых, в свою очередь, тоже были учителя, – и так на протяжении многих веков. Я изучил опыт своих предшественников, усвоил и переработал полученные от них знания, дополнил их, а теперь надеюсь поделиться своими открытиями с вами, дорогие читатели этой книги.

Люди – наиболее приспособляющиеся создания на планете и единственные, кто обладает способностью широко обмениваться знаниями о мире в пределах своей популяции. Человечество пережило взрывное развитие, потому что мы в состоянии глубоко познавать структуру мира и передавать накопленную информацию друг другу. Благодаря развитой коре головного мозга и системе речи мы способны выжить везде – в лесу, в пустыне, в вечной мерзлоте или в джунглях.

Что такое творчество?

Мне часто задают этот вопрос. Подозреваю, что люди думают, будто творчество – нечто такое, на что машины не способны, и именно поэтому вся идея создания искусственного интеллекта ставится под сомнение. Что же такое творчество? Мы уже несколько раз встречали ответ в этой книге. Творчество не является чем-то свойственным исключительно одной области коры головного мозга. Оно отличается от эмоций или координации, формирующихся в отдельных структурах вне коры головного мозга. Творчество, скорее, присуще каждой зоне неокортекса и является неотъемлемым элементом прогнозирования.

Как такое возможно? Разве творчество – не удел избранных, одаренных высочайшим интеллектом и выдающимися характеристиками? Не совсем. Творчество можно определить как способность прогнозирования на основе аналогий. Иногда данный процесс происходит во всех областях коры головного мозга, он непрерывен на протяжении всего времени, пока вы бодрствуете. Творчество присутствует в самых простых, обыденных процессах восприятия, происходящих в сенсорных зонах коры головного мозга (вспомните прослушивание знакомой песни в новой тональности), и в сложнейших проявлениях гениальности, рождающихся в высших зонах (например, создание симфонии). Фундаментальная основа повседневных примеров сенсорного восприятия та же, что и редких проявлений гениальности. Просто повседневные события для нас настолько обыденны, что мы этого не замечаем.

В предыдущих главах вы получили основное представление о том, как формируются наши инвариантные представления и как мы используем их для составления прогнозов

будущих событий, которые всегда хотя бы немного, но отличаются от нашего прошлого опыта. Вы также знаете, что наши инвариантные представления сохраняются в памяти как последовательности событий. Мы создаем прогнозы, сопоставляя инвариантные представления об ожидаемых событиях и особенности текущей ситуации (вспомните историю о встрече поезда). По сути, прогнозирование – это применение инвариантных последовательностей к новым ситуациям. Таким образом, все прогнозы, создаваемые неокортексом, основаны на аналогиях. Мы прогнозируем будущее по аналогии с прошлым.

Представьте себе, что вы собираетесь поужинать в незнакомом ресторане и хотите вымыть руки. Вы никогда прежде не бывали здесь, но ваш мозг прогнозирует, что где-то в ресторане есть туалетная комната с умывальником. Откуда вашему мозгу это известно? Просто в других ресторанах, где вам доводилось побывать, всегда была такая комната. По аналогии мозг приходит к выводу, что в этом ресторане тоже должно быть нечто подобное. Вы уже знаете, где и что вам нужно искать. Вы прогнозируете, что должна быть дверь со значком, указывающим на то, мужская это или женская туалетная комната. Вы прогнозируете, что эта комната должна находиться где-то в дальней части главного помещения, возможно, позади барной стойки или на лестничной клетке, но точно не в зоне непосредственной видимости посетителей ресторана. Так, по аналогии с другими ресторанами, в которых вам довелось побывать, вы легко найдете то, что вам нужно. Вы не оглядываетесь по сторонам наугад, а наоборот, ищите ожидаемые указатели и значки, которые быстро помогли бы вам найти туалетную комнату. Поведение такого типа тоже является проявлением творчества, поскольку вы прогнозируете будущее по аналогии с прошлым. Подобный процесс не принято считать творческим, но, тем не менее, таковым он и является.

Недавно я купил виброфон¹⁷. У нас есть пианино, но вот на виброфоне я не играл никогда. Когда мы доставили инструмент домой, я взял ноты с пианино, раскрыл их, положил на подставку перед виброфоном и начал наигрывать простую мелодию. Разумеется, качество моего исполнения оставляло желать лучшего, но на фундаментальном уровне это был творческий поступок. Задумайтесь о том, что в нем присутствовало. У меня появился инструмент, который очень сильно отличается от пианино, – у виброфона золотистые металлические пластинки, у пианино белые и черные клавиши. Золотистые пластинки крупные и имеют различные размеры, клавиши невелики и всегда имеют только один из двух возможных размеров. Золотистые пластинки упорядочены в два отдельных ряда, черные и белые клавиши расположены в чередующемся порядке. При игре на одном инструменте я использую пальцы, а при игре на другом – ударные палочки. На одном инструменте я играю стоя, на втором – сидя. Для игры на виброфоне и на пианино требуется задействовать различные группы мышц и выполнять совершенно разные движения.

Как же мне удалось наиграть мелодию на незнакомом инструменте? Дело в том, что мой мозг нашел соответствие между клавишами пианино и пластинками виброфона. По этой аналогии я и наиграл мелодию – подобно тому, как узнал бы песню, исполненную в разных тональностях. В обоих случаях мы познаем действительность и действуем по аналогии с прошлым полученным опытом. Я понимаю, что для вас сходство двух инструментов может показаться очевидным, но это возможно лишь потому, что человеческий мозг автоматически распознает аналогии. Попробуйте запрограммировать компьютер на поиск сходства между такими объектами, как пианино и виброфон, и вы убедитесь, что эта задача невероятно сложна. Прогнозирование на основе аналогий, т. е. творчество, настолько неуловимо, что в обыденной жизни мы его даже не замечаем.

С другой стороны, мы считаем, что проявляем творчество, когда наша запоминающе-прогностическая система оперирует на высших уровнях абстрагирования, когда она создает незаурядные прогнозы, используя незаурядные аналогии. Например, большинство людей согласятся, что решение математиком сложной задачи – творческий процесс. Но давайте поговорим об этом подробнее. Наш математик изучает замысловатое уравнение и думает: «Интересно, как бы мне его решить?» Не найдя очевидного ответа, он пробует изменить форму

¹⁷ Вибрафон – тип ксилофона; ударный музыкальный инструмент. – Примеч. ред.

уравнения. Переписывая его несколько раз, ученый рассматривает его с разных точек зрения. Наконец, одна часть уравнения кажется ему очень знакомой. Он думает примерно так: «О, это напоминает мне элемент другого уравнения, которое я решал несколько лет тому назад». Затем математик создает прогноз на основе аналогий: «Возможно, я смогу решить это уравнение используя тот же подход, который оказался успешным в прошлый раз». Так, решая уравнение по аналогии с задачей, выполненной ранее, ученый совершает творческий поступок.

У моего отца было какое-то загадочное заболевание крови, и врачам долго не удавалось установить правильный диагноз. Как же они выбирали метод лечения? Они взяли результаты ежемесячных анализов крови отца с тем, чтобы отследить определенные закономерности. (Мой отец даже распечатал их в виде графика, чтобы картина была более четкой.) Хотя симптомы болезни не сочетались в единую клиническую картину, присутствовали признаки целого ряда патологических состояний. Врачи сошлись на том, чтобы применить лечение, основанное на комбинации стратегий, которые оказались успешными при лечении других заболеваний крови. Методы лечения были выведены по аналогии с успешными методами, которые использовались для лечения заболеваний, встречавшихся раньше во врачебной практике.

Примером творчества могут служить метафоры Шекспира. Когда вы их читаете, они вызывают восхищение, но придумать их самому крайне сложно, поэтому Шекспира считают гением. Чтобы создать подобные метафоры, нужно увидеть последовательность удачных аналогий. Когда Шекспир пишет: «Ножи в улыбке у него», то он на самом деле не говорит о ножах или об улыбке. Ножи служат аналогией злых намерений, а улыбка – аналогом лживости. Вы только подумайте – целых две искусных аналогии всего в пяти словах! Талантливые поэты обладают даром сочетания слов или понятий, позволяющим взглянуть на мир под новым углом. Они создают неожиданные аналогии как способ познания структур высшего уровня.

Искусство ценно потому, что его творческая сила позволяет временно приостановить прогностическую функцию. Если вы смотрите фильм, который нарушает привычные каноны, он вам нравится, потому что это – не давно известное старье. Полотна живописцев, музыкальные произведения, поэзия, романы – все формы искусства призваны разрушать привычные стереотипы и обманывать ожидания публики. Не существует единого определения высокого искусства. С одной стороны, мы хотим наслаждаться знакомыми явлениями, пусть и поданными под необычным углом, и в то же время ждем чего-то особенного и неожиданного. Перевес знакомого оценивается как дешевка; избыток новизны утомляет, и ее очень трудно оценить по достоинству. Например, привлекательность для слуха музыкального произведения определяется довольно простыми свойствами – четким ритмом и гармоничным звучанием музыкальных фраз. Насладиться такой музыкой может любой. В то же время она несколько необычна и неожиданна, однако чем больше вы ее слушаете, тем лучше понимаете, что непредсказуемой ее назвать нельзя. Даже в неожиданной части музыкального произведения тоже существует своя структура, включающая, например, повторение музыкального рисунка или смену тональности. Точно так же обстоит дело с литературными или кинематографическими шедеврами. Чем более искушенным читателем или зрителем вы становитесь, тем точнее различаете творческие ходы создателей или сложность структуры их творений.

Наверняка, вам случалось смотреть на что-то и отмечать про себя: «Хм, по-моему, я уже где-то что-то подобное видел... Но где и когда?» Во многих случаях вы не пытаетесь разрешить возникшую задачу сознательно, просто вследствие неожиданной ситуации в вашем мозге активизируется инвариантная репрезентация. Сознание подсказывает вам аналогию между двумя, как правило, не связанными между собой событиями или фактами. Мне кажется, что продвижение новых идей в науке и в бизнесе протекают по одним и тем же законам, а лоббирование политической реформы во многом подобно воспитанию детей. Если я поэт, пожалуйста, вот вам моя новая метафора! Если я инженер или ученый, то предлагаю решение давно существующей проблемы. Творчество – это поиск соответствия между сигналами, которые вам доводилось встречать на своем жизненном пути, и их сопоставление. Нейронный механизм осуществления такой деятельности функционирует во всех отделах коры головного мозга.

Действительно ли есть люди, более и менее способные к творчеству?

Когда речь заходит о творчестве в целом, меня нередко спрашивают: «Если творчество присуще мозгу любого человека, то как можно объяснить различия в степени выраженности творческих способностей?» Модель «Память-предсказание» предлагает два возможных ответа. Первый опирается на природные задатки, а второй – на обстоятельства воспитания. Начнем со второго. Каждый из нас в процессе жизни формирует свой уникальный жизненный опыт. Это значит, что нет двух одинаковых моделей мира, что в одной и той же ситуации у разных людей возникают совершенно различные аналогии, а их прогнозы могут быть противоположными. Если я рос в музыкальной среде, то могу спеть песню в другой тональности или наиграть хорошо знакомую мне мелодию на новом музыкальном инструменте. Если я не имел ничего общего с музыкой, то не сделаю этого, поскольку не смогу осуществить прогнозы такого уровня. Если я изучал физику, я могу объяснить поведение обыденных объектов по аналогии с физическими законами. Если мои родители занимались разведением собак, то у меня будут развиты способности проведения аналогий в их поведении, таким образом, я смогу точнее прогнозировать их поведение. Некоторые люди имеют выдающиеся способности к творчеству, языкам, математике, дипломатии. Во всем этом присутствует значительная доля влияния среды – обстановки, в которой они росли и воспитывались. Особенности прогностического процесса, а значит, и наши дарования, базируются на нашем опыте.

В главе 6 мы говорили о том, как сохраняющиеся в памяти единицы информации передвигаются вниз по иерархическим зонам коры головного мозга. Чем чаще вы подвергаетесь воздействию определенного сигнала, тем скорее информация о его значении передается в более низкие в иерархии коры головного мозга зоны. Это позволяет вам изучать взаимосвязи между абстрактными понятиями высшего порядка в верхней части иерархии. В этом суть компетентности. Специалиста, который благодаря своему опыту и постоянному контакту с объектами определенной сферы может распознавать более тонкие сигналы, незаметные непосвященным, называют профессионалом, мастером своего дела, знатоком. В качестве примера таких незаметных сигналов можно привести форму крыла автомобиля середины 1950-х годов или размер пятнышка на клюве чайки. Мастера способны распознавать сверхструктуры. И, хотя физический предел способности к обучению существует (он ограничен размерами неокортекса), мы, люди, обладаем чрезвычайно большой по сравнению с другими видами корой головного мозга, а следовательно, значительной гибкостью в том, чему можем научиться. И самая важная роль во всем этом отводится опыту, который приобретаем на протяжении своей жизни.

Существуют, разумеется, и врожденные задатки, в частности – обусловленные различиями в строении головного мозга. Например, размером областей коры головного мозга (наиболее заметны различия в размерах зоны V1) и латерализацией полушарий (у женщин связи между правым и левым полушариями более прочны, чем у мужчин). Не исключены также различия в числе нервных клеток в нервной ткани и связях между ними. Скорее всего, творческий гений Альберта Эйнштейна нельзя объяснить исключительно благоприятной внешней средой патентного бюро, в котором он работал в молодости. В недавних исследованиях мозга гения, долгое время считавшегося утерянным, обнаружилось много необычного. В мозге создателя теории относительности на один нейрон приходилось существенно больше поддерживающих клеток (глий¹⁸), чем у обычных людей. Кроме того, были выявлены борозды необычной формы в теменных выступах, т. е. зонах, которые считаются ответственными за математические способности и пространственное воображение. У Эйнштейна эти зоны были примерно на 15% шире, чем у большинства людей. Вероятно, нам так и не удастся до конца разгадать загадку творчества и интеллектуальных способностей Эйнштейна, но можно с уверенностью утверждать, что отчасти в их основе лежали генетические факторы.

¹⁸ Глия – от греч. *glia* – клей, ткань, заполняющая пространства между нервными клетками, их отростками и сосудами в центральной нервной системе. – Примеч. ред.

Независимо от различий в строении головного мозга гения и среднестатистического человека, мы все креативны, а посредством практики и обучения можем отточить наши способности и таланты.

Можно ли развить креативность путем самообучения?

Скорее всего, да. Я обнаружил, что существуют способы, помогающие найти полезные аналогии при решении задач. Во-первых, нужно исходить из предположения, что поставленная задача является разрешимой. Слишком часто люди бросают начатое на полпути. Вы должны быть уверены в том, что решение просто ждет, и готовыми к тому, что упорно размышлять о поставленной задаче придется достаточно долго.

Во-вторых, не нужно останавливать блуждания собственной мысли. Вашему мозгу нужны время и пространство для того, чтобы найти решение. Поиск решения равнозначен для мозга поиску аналогии, существующей во внешнем мире или хранящейся в вашей коре головного мозга. Согласно модели «Память-предсказание», в работе над любой проблемой вы должны искать разные перспективы ее рассмотрения, повышая тем самым вероятность найти аналогию с прошлым опытом. Не заикливайтесь на якобы единственном возможном подходе! Попробуйте разбить задачу на несколько частей и переставить их в другом порядке, как в буквальном, так и в переносном смысле. Когда я играю в составление слов из букв одного длинного слова, то периодически меняю порядок составляющих. Дело совсем не в том, что перестановка букв помогает мне придумать новое слово, а в том, что разные комбинации букв могут напомнить мне слова или их части, которые станут частью решения задачи. Если вы смотрите на изображение чего-либо, и рисунок ни о чем вам не говорит, попробуйте представить его перевернутым, или изменить цвета, или взглянуть на него под другим углом. Например, в процессе размышлений о том, как поступающие из окружающей среды в зону V1 сигналы вызывают активизацию инвариантных репрезентаций в зоне IT, я зашел в тупик. Тогда я задал себе обратный вопрос: как постоянный сигнал из зоны IT может приводить к возникновению различных прогнозов в V1? В конечном счете я понял, что V1 не стоит рассматривать как однородную зону коры головного мозга.

Чем бы вы ни занимались, порой вы можете зайти в тупик. В таком случае отвлекитесь немного, смените вид деятельности. Потом попытайтесь переформулировать свою задачу. Возможно, поиск верного решения займет дни или недели, но рано или поздно вы добьетесь своего. Цель состоит в том, чтобы найти аналогичную ситуацию в вашем прошлом или настоящем опыте. Чтобы достичь успеха, вы должны постоянно держать свою цель в голове, но при этом иногда отвлекаться, чтобы ваш мозг имел возможность поискать аналогичное воспоминание.

Приведу еще один пример того, как переформулирование задачи способствовало нахождению правильного решения. В 1994 году мы с коллегами искали способ ввода текста в карманных компьютерах. Все сосредоточились на создании программного обеспечения для распознавания написанного от руки текста. Мысль была такая: раз человек может писать от руки на листе бумаги, точно так же можно писать и на мониторе компьютера. К сожалению, оказалось, что данная задача крайне сложна – еще один пример того, что способности компьютера и человеческого мозга несравнимы. Причина в том, что мозг использует как информацию, хранящуюся в памяти, так и текущий контекст, чтобы прогнозировать, *что* будет написано дальше. Буквы и слова, трудно распознаваемые по отдельности, очень легко идентифицируются в контексте. Набор сигналов, распознаваемых компьютером, не является достаточным для решения поставленной задачи. Я разработал архитектуру нескольких компьютеров, использующих традиционное распознавание почерка, но ни одна из них не оказалась достаточно функциональной.

Несколько лет я бился над тем, чтобы улучшить работу этого программного обеспечения, пока не понял, что зашел в тупик. Тогда я решил сменить угол зрения на проблему и задумался над аналогичными задачами. Я спросил себя: «Как мы вводим текст в стационарных компьютерах? Печатаем на клавиатуре. А откуда мы знаем, *как* нам нужно печатать на клавиатуре? В общем-то это не так просто. Ведь это сравнительно новое изобретение,

требующее времени для обучения. Печатать на клавиатуре посредством ударов пальцами непросто, это не интуитивное знание, оно нисколько не напоминает письмо от руки, однако миллионы людей смогли овладеть им. Почему? Потому, что это работает». Дальше мои размышления сместились к построению аналогий: «Возможно, я придумаю систему текстового ввода, не обязательно интуитивную, которую придется изучить, но люди будут ее использовать, потому что она окажется работающей».

Я прошел через описанный процесс, используя печатание на клавиатуре как аналогию для решения задачи введения текста в компьютер с помощью письма ручкой по дисплею. Я учитывал, что пользователи с готовностью берутся за освоение печати на клавиатуре, потому что это надежный и быстрый способ ввода текста в компьютер. Значит, если мы сможем придумать новый способ ввода текста при помощи ручки, который будет быстрым и надежным, то люди согласятся его использовать, даже если потребуется обучение. И вот я придумал алфавит, который надежно переводил рукописный текст в компьютерный. Впоследствии мы назвали его *Graffiti*. В традиционной системе распознавания рукописного текста вы не можете выяснить причины возможных ошибок компьютера. Система *Graffiti* всегда воссоздает правильные буквы – при условии, что при письме не допущено ошибок. Наш мозг ненавидит неопределенность, именно поэтому традиционные системы распознавания рукописного текста не пользуются популярностью.

Многие считали *Graffiti* исключительно глупым изобретением. Я пошел наперекор всем ожиданиям в отношении компьютеров. Сакральным заклинанием того периода было: компьютер должен приспособливаться к пользователю, а не наоборот. Тем не менее я был полон уверенности, что люди примут мое нововведение точно так же, как они приняли клавиатуру. Система *Graffiti* оказалась хорошим инженерным решением и вскоре обрела широкую популярность. И хотя я до сих пор нередко слышу мнение, что компьютеры должны приспособливаться к пользователям, а не наоборот, это не всегда верно. Наш мозг отдает предпочтение системам, которые являются последовательными и прогнозируемыми, к тому же нам нравится приобретать новые навыки.

Не может ли творчество ввести в заблуждение? Возможна ли ситуация самообмана?

Опасность ложной аналогии существует всегда. История науки богата примерами разоблачения великолепных на первый взгляд аналогий. Так, например, известный астроном Иоганн Кеплер убедил себя в том, что орбиты шести известных на то время планет предопределяются Платоновыми многогранниками. Платоновы многогранники – это единственные трехмерные фигуры, которые можно полностью создать из правильных многоугольников. В частности, это тетраэдр, октаэдр, додекаэдр, икосаэдр, куб, открытые древними греками, которые уделяли пристальное внимание взаимосвязям математики и космоса.

Как и все классики эпохи Ренессанса, Кеплер находился под сильным влиянием греков. Ему казалось, что факт существования пяти правильных трехмерных фигур и шести планет не был случайным. В своем труде *Космическая загадка* (1596) ученый писал: «Динамический мир представлен пятью правильными многоугольниками с плоской поверхностью. Если их рассматривать как границы, то они, определяют шесть объектов, т.е. шесть планет, вращающихся вокруг Солнца. Именно в этом и состоит причина того, почему существует только шесть планет». Как видите, аналогия, найденная Кеплером, была очень изящной, но совершенно не соответствующей истине.

Далее Кеплер продолжал размышлять об орбитах планет в рамках вложенных правильных многогранников, центрированных вокруг Солнца. Взяв за основу орбиту Меркурия, ученый описал ее правильным восьмиугольником. Края октаэдра обозначали большую по размерам сферу, образующую орбиту Венеры. Вокруг орбиты Венеры Кеплер описывал икосаэдр, края которого обозначали орбиту Земли. А дальше прогрессия продолжалась: по представлениям Кеплера, додекаэдр, описанный вокруг орбиты Земли, очерчивал орбиту Марса, тетраэдр вокруг орбиты Марса – орбиту Юпитера, а куб вокруг Юпитера – орбиту Сатурна. Так просто и

красиво! Принимая во внимания скудность астрономических сведений в те дни, такая схема действительно казалась реальной. (Кеплеру посчастливилось понять свою ошибку, когда через несколько лет другой ученый, Тихо Браге, доказал, что орбиты планет имеют эллипсовидную форму.)

Пусть заблуждение Кеплера послужит хорошим уроком для всех ученых, да и не только для них. Мозг – это орган, создающий модели и ставящий творческие прогнозы. Эти прогнозы и модели могут как прояснить истину, так и сбить с правильного пути познания мира. Причем, при отсутствии правильных корреляций человеческий мозг склонен к тому, чтобы принять ошибочные предположения. Псевдонаука, фанатизм, религиозность, нетерпимость очень часто имеют общий корень – ошибочные аналогии.

Что такое сознание?

Это один из тех вопросов, которых большинство нейробиологов понапрасну опасаются и относят к разделу философии, граничащему с псевдонаукой. Тему человеческого сознания готовы обсуждать лишь немногие из них, например, Кристоф Кох. Я тоже считаю данный вопрос заслуживающим внимания. К сожалению, у меня нет полной и исчерпывающей теории, которая полностью бы объясняла феномен сознания, но я полагаю, что отчасти его можно понять, опираясь на все те же ключевые процессы деятельности головного мозга – память и прогнозирование. (Не кажется ли вам, что теперь я привнес в нашу дискуссию оттенок загадочности?)

Недавно мне довелось посетить научную конференцию в Лонг-Айленд Саунд. Вечером порядка дюжины ее участников спустились к пирсу, чтобы поболтать возле воды перед обедом и следующим за ним вечерним отделением конференции. Незаметно разговор перешел на тему сознания. Как я уже отмечал, нейрологи не любят этой темы, но мы находились в чудном местечке, выпили немного отличного вина, и расслабились. Коллега из Британии заявила: «Человеческое сознание непостижимо». Я возразил: «Сознание не такая большая загадка. Это просто осознание того, что значит иметь кору головного мозга». Воцарилась тишина. Затем несколько ученых попытались вывести меня из заблуждения, в котором я, по их мнению, пребывал. «Согласитесь, что вы воспринимаете мир как нечто живое и прекрасное. Как вы можете отрицать существование вашего сознания, которое воспринимает внешний мир? Согласитесь, вы ведь ощущаете себя чем-то особенным». Чтобы доказать свою точку зрения, я настаивал: «Не понимаю, о чем выговорите. Если принять тот угол, под которым вы рассматриваете сознание, то я от вас, должно быть, очень сильно отличаюсь. Я не чувствую того, что чувствуете вы, значит, я создание без сознания или вообще зомби». Должен заметить, что понятие «зомби» очень часто всплывает в дискуссиях философов на тему сознания. Согласно общепринятому определению, зомби – это физический аналог человека, у которого отсутствует сознание. Британская коллега с интересом на меня посмотрела:

Конечно же, у вас есть сознание.

Нет, я так не думаю. Возможно, внешне я похож на вас, но я не сознательное человеческое создание. Пусть вас это не беспокоит, я ведь не переживаю.

Разве вы не замечаете всю эту красоту? – сказала она, показывая рукой на линию горизонта, где закат солнца отливал оранжево-розовой дорожкой на поверхности моря.

Вижу. Ну и что?

Опишите свое субъективное восприятие.

Я знаю, что нахожусь здесь. Моя память сохранит многие из подробностей этого вечера. Однако я не чувствую, что происходит нечто особенное, и если вы что-то особенное ощущаете, то, стало быть, у меня нет сознания.

Так я хотел припереть ее к стенке, чтобы она ответила, что, по ее мнению, было необъяснимого и таинственного в сознании. Я хотел заставить ее дать определение понятию сознания!

Мы препирались до тех пор, пока не подошло время ужина. Не думаю, что мне удалось изменить мнение кого-либо из присутствующих о том, что такое сознание и в чем оно проявляется. Я просто пытался подвести их к мысли о том, что считать сознание неким

магическим «соусом», который прилагается к мозгу, в корне неверно. Кстати, этот подход невероятно распространен: мозг, состоящий из нервных клеток, следует полить магическим «соусом» сознания, чтобы в результате получился человек. При таком подходе сознание является таинственной субстанцией, которая существует независимо от мозга. Именно поэтому у зомби есть мозг, но нет сознания. У них имеются все механические составляющие, нейроны и синапсы, но у них нет того особенного «соуса». Они могут делать все то же, что делают люди. По внешним признакам зомби не отличишь от человека.

Представление о том, что сознание – это нечто особенное, происходит из верований в дух жизни, т. е. особенную силу, которая, как считалось, делает живое живым. Люди верили, что присутствие этой силы объясняет различия между скалой и растением, каменным изваянием и живым существом. На сегодняшний день мы достаточно хорошо понимаем различия между живыми и неживыми объектами, чтобы понять, что никакого особенного «соуса» нет и в помине. Нам многое известно о ДНК, белковой природе человеческого организма, генетическом коде и метаболизме. Все тайны живых систем пока что не раскрыты, но мы достаточно искушены в биологии, чтобы не искушаться мифом магии.

Но и в наше время, несмотря на сказанное выше, многие люди продолжают верить, что сознание – это нечто особенное, то, что невозможно объяснить в редукционистских биологических терминах. И, хотя я не занимаюсь непосредственно исследованием человеческого сознания и не изучил досконально все философские воззрения на сей счет, у меня есть обоснованное суждение по этому поводу. Я считаю, что между присутствием у живого существа сознания и наличием у него неокортекса можно поставить знак равенства. Но это еще не все. Мы можем подразделить сознание на две основные категории. Первая – самосознание, т. е. обыденное ощущение собственной идентичности; его понять нетрудно. Вторая категория сознания – *квалия*, т. е. представление о том, что чувства и ощущения являются частично независимыми друг от друга в потоке сенсорных входных сигналов. Вот ее понять сложнее.

Чаще всего, употребляя слово *сознание*, люди обращаются к первой категории. «Вы осознаете, что прошли мимо меня, даже не поздоровавшись?», «Ты был в сознании, когда преступник скрылся?» и проч. Нередко возникает путаница в понятиях «сознание» и «осведомленность». Они действительно очень близки, однако я полагаю, что первое гораздо шире. По моему мнению, сознание тождественно способности сохранять единицы информации в декларативном виде так, чтобы вы могли вызвать их из памяти по своему желанию и пересказать кому-то другому посредством устной или письменной речи. Давайте расшифруем, что представляет собой эта декларативная память. Если вы спросите меня, где я был в прошлую субботу, я вам смогу рассказать об этом. Вот вам пример единицы информации, сохраненной в мозге в декларативном виде. С другой стороны, если вы спросите меня, как при езде на велосипеде удерживать его в состоянии равновесия, я порекомендую вам крепче держаться за руль и давить на педали, но не смогу предоставить более точных объяснений. Дело в том, что удержание равновесия на велосипеде обеспечивается нейронной деятельностью «старого» мозга, т. е. воспоминание об этом процессе не сохраняется в декларативной форме.

Проведите небольшой мысленный эксперимент, подтверждающий приведенную выше информацию. Как вы помните, вся наша система памяти обусловлена физическими изменениями синапсов и нейронов, к которым они привязаны. Если бы я изобрел метод, позволяющий повернуть эти физические изменения вспять, то хранилище памяти оказалось бы стертым. Только представьте себе: я поворачиваю рычаг, и ваш мозг оказывается точно в таком же физическом состоянии, в каком он был в определенный момент прошлого – скажем, час или двадцать четыре часа назад. Осуществив такую процедуру, я стираю всю информацию, хранившуюся в вашем мозге с того момента времени.

Допустим, вы прожили сегодняшний день и проснулись завтра. Но как раз перед вашим пробуждением я стер последние двадцать четыре часа. У вас не осталось бы совершенно никаких воспоминаний о предыдущем дне. С точки зрения вашего мозга, «вчера» не было вообще. Я бы вам сказал, что сегодня среда, но вы бы протестовали: «Нет, это неправда. Сегодня вторник, и я уверен в этом. Кто-то изменил календарь. Даже не пытайтесь обмануть меня!» Я могу предоставить вам свидетелей, готовых подтвердить, что они встречались с вами

во вторник и что вы были в сознании. Они видели вас, обедали с вами, общались с вами. Неужели вы этого не помните? Вы скажете, что ничего такого не было. Когда вам, наконец, покажут видеопленку, на которой записано, как вы обедали, вы осознаете, что вас не водят на нос, что на самом деле вчера был вторник – день, о котором у вас не сохранилось ни единого воспоминания. Похоже, вчера вы были зомби, а не человеком. Тем не менее на протяжении всего дня вы пребывали в полном сознании. Ваша убежденность в том, что вы были в сознании, исчезла только тогда, когда была стерта декларативная память.

Рассмотренный мысленный эксперимент показывает равнозначность декларативных воспоминаний и обыденного восприятия собственного сознания. Если на протяжении партии в теннис и в конце игры я спросил бы вас, в сознании ли вы, то, разумеется, получил бы утвердительный ответ. Если бы я стер вашу память за последние два часа, вы бы сказали, что все это время находились в бессознательном состоянии, а значит, не можете отвечать за свои действия. Таким образом, ключевую роль играет то, сохраняется ли информация об определенном временном промежутке в системе вашей памяти. Иными словами, значение сознания не является абсолютным. Его можно изменить путем стирания памяти.

Более сложная интерпретация сознания связана с понятием *квалия*, которое наиболее полно отражается в вопросах в стиле дзен, например: «Почему красный бывает красным, а зеленый зеленым? Выглядит ли красный цвет для меня точно таким же, как и для вас? Почему красный цвет вызывает определенные чувства? Красный цвет всегда обладает для меня неотделимым свойством эмоциональности. А какие чувства вызывает он у вас?»

В нейробиологии подобные вопросы некорректны, поэтому позволю себе несколько их перефразировать. С моей точки зрения, эквивалентным и ничуть не менее простым вопросом будет такой: «Почему разные ощущения воспринимаются как отличающиеся качественно? Почему воспринимаемые нами зрительные образы отличны от слуховых, а также от тактильных ощущений?». Кора головного мозга везде одинакова, в ней везде происходят одни и те же процессы, она имеет дело только со стимулами, поступающими извне. Аксоны, по которым передаются звуковые и визуальные сигналы, по характеру своего функционирования полностью идентичны – аксон сенсорного нейрона не переносит «свет» или «звук». Но откуда берутся эти различия, которые трудно описать словами, но не признать их наличия невозможно?

У людей, страдающих синестезией¹⁹, в головном мозге размыты различия между чувствами восприятия. У таких людей некоторые звуки имеют цвет и тактильные ощущения. Это говорит о том, что качественный аспект любого ощущения не является неизменным. В результате некоторых физических преобразований мозг может присвоить аудиальному входному сигналу качественный аспект зрения.

Так как же можно объяснить феномен *квалия*? Существуют два предположения, хотя ни одно из них не является вполне удовлетворительным. Вариант первый: зрительное, слуховое и осязательное восприятие осуществляются по одинаковым принципам, но подлежат различным формам обработки в коре головного мозга. Звуковое восприятие рождается в специфических звуковых подструктурах неокортекса, которые обрабатывают слуховые сигналы до того, как они собственно попадают в кору головного мозга. Соматосенсорные сигналы тоже проходят обработку в соответствующих подструктурах коры головного мозга. Возможно, феномен *квалия*, подобно эмоциям, не передается исключительно корой головного мозга. В случае если он как-то связан с подструктурами коры головного мозга, в которых существуют уникальные нейронные связи с эмоциональными центрами, это может объяснить, почему мы воспринимаем чувства и эмоции по-разному, хотя это и не объясняет явления *квалия*.

Второе возможное объяснение таково: восприятие качественных аспектов информации обусловлено структурой входных сигналов, т. е. их различиями. Природа

¹⁹ Синестезия (от греческого *synaisthesis* – соощущение) – явление восприятия, когда при раздражении определенного органа чувств наряду со специфическими для него ощущениями возникают и ощущения, соответствующие другому органу чувств (например, «цветной слух» – звуковые переживания при восприятии цвета и т. п.). – Примеч. ред.

пространственно-временных сигналов, воспринимаемых слуховым нервом, отличается от природы пространственно-временных сигналов, фиксируемых зрительным нервом. Зрительный нерв, содержащий миллионы волокон, воспринимает преимущественно пространственную информацию. У слухового нерва, содержащего тридцать тысяч волокон, отчетливо выражен приоритет временной информации. Эти различия могут быть связаны с тем, что мы называем *квалиа*. Одно мы можем утверждать с полной уверенностью: независимо от определения сознания, память и прогнозирование играют в его формировании критическую роль.

С понятием сознания связаны понятия разума и души. В детстве я часто фантазировал о том, что было бы, если бы я родился в теле другого ребенка в другой стране. В тех фантазиях мое «я» было независимым от тела. Это широко распространенное чувство независимости разума от физического тела вызвано особенностями функционирования коры головного мозга. Как вам уже известно, неокортекс создает в своей иерархической памяти модель окружающего мира. Непосредственным результатом работы данной модели являются мысли. Извлечение из памяти хранящейся в ней информации (воспоминаний) приводит к формированию прогнозов, действующих, в свою очередь, подобно сенсорным сигналам, вызывающим новые воспоминания и так далее. Некоторые наши мысли не зависят от событий реального мира и не связаны с ними, они являются чистыми продуктами нашей модели. Мы закрываем глаза и размышляем в тишине с тем, чтобы наши размышления не были прерваны сенсорным входным сигналом. Конечно, изначально модель формировалась в процессе и посредством познания реального мира органами чувств, однако, размышляя о мире и осуществляя планирование, пользуемся преимущественно моделью коры головного мозга, а не непосредственными сигналами, поступающими из окружающей среды.

Даже собственное тело для вашего неокортекса является лишь частью внешнего мира. Не забывайте, что головной мозг – это ящик, в котором царят крошечная тьма и тишина. Мозг познает окружающую среду только посредством сигналов, поступающих по сенсорным нервным волокнам. Мозг как сигнальное устройство не воспринимает тело своего хозяина как нечто особенное, отличное от остального внешнего мира. Иными словами, границы между вашим телом и оставшейся частью внешнего мира для мозга не существует. В то же время кора мозга не может построить модель собственно головного мозга, потому что у мозга нет ощущений. Теперь становится понятным, почему наши мысли появляются независимо от наших тел, почему нам кажется, что душа или разум независимы. Ваши мысли существуют в мозге, они физически отделены от тела и остальной части внешнего мира. Разум независим от тела, но является продуктом деятельности мозга.

Эта дифференциация наглядно проявляется во время травм и заболеваний. Так, у пациентов, перенесших ампутацию нижней конечности, нередко возникает фантом утерянной конечности – человек якобы чувствует свою ногу. Возможен и обратный вариант: у человека, пережившего травму коры головного мозга, может быть утеряна модель руки. И хоть данная часть тела остается неповрежденной, пациент страдает синдромом отчуждения конечности. Это состояние проявляется неприятным, порой даже невыносимым чувством утраты контроля над собственной рукой. Известны случаи, когда страдания вынуждают пациентов просить ампутировать «отчужденную» конечность. Если при каком-то заболевании мозг остается здоровым, то пациенту кажется, будто его здоровый разум заключен в ловушку больного тела, хотя на самом деле речь идет о здоровом мозге. Представление о том, что разум продолжает свое существование после смерти тела, вполне естественно. Однако сознание не может существовать после смерти головного мозга. Достоверность такого утверждения становится очевидной, когда мозг подводит организм. Люди, страдающие болезнью Альцгеймера или другими серьезными заболеваниями головного мозга, теряют разум, даже если их тела при этом остаются здоровыми.

Что такое воображение?

По сути, понятие воображения очень простое. От органов чувств в соответствующие зоны коры головного мозга поступают сигналы, которые затем передаются в более высокие по иерархии зоны. Формирующиеся прогнозы спускаются вниз по этой иерархии. Процесс

воображение основан на «разворачивании» прогнозов и превращении их во входные сигналы, причем для этого не требуется никаких физических действий. «Если случится это, то потом случится вот это, а затем произойдет то-то», – размышляем мы по ходу подготовки к деловой встрече или к спортивным соревнованиям, во время шахматной партии и в тысячах других случаев.

В шахматах вы предполагаете, что поставите слона на определенную клетку, а потом мысленно представляете, как будет выглядеть расклад фигур вашего противника и каким будет последующий вид шахматной доски. Потом вы прогнозируете свой следующий ход и так далее. В конечном счете на основе воображаемой последовательности событий вы решаете для себя, был ли выбранный вами ход удачным. Некоторые спортсмены – например, те, кто занимается горными лыжами, – стараются улучшить свои результаты, многократно мысленно проигрывая прохождение спуска. Закрывая глаза и представляя каждый поворот, каждую преграду, даже то, как им вручают медаль, спортсмены могут значительно увеличить свои шансы на успех. Воображение – это, по сути, вариант планирования, или реализации прогностической функции неокортекса, позволяющее судить о последствиях действий еще до их совершения. Воображение обусловлено действием нейронного механизма, преобразующего прогноз во входной сигнал. В главе 6 я изложил свое предположение, что точные прогнозы формируются клетками слоя 6. Клетки этого слоя посылают проекцию в низшие уровни иерархии, но они также посылают проекцию назад, ко входным клеткам слоя 4. Таким образом, исходящие сигналы могут стать входящими. Как я уже отмечал, Стефен Гроссберг называет подобный кругооборот совмещенной обратной связью. Если вы закроете глаза и представите себе гиппопотама, зрительная зона коры вашего головного мозга активизируется точно так же, как если бы вы сейчас и вправду смотрели на гиппопотама. Вы видите то, что воображаете.

Что такое действительность?

Потрясенные данной информацией, слушатели часто спрашивают меня: «Вы в самом деле считаете, что создаваемая мозгом модель мира может оказаться более значимой, чем объективная реальность?»

В целом – да, – соглашаюсь я.

Вы хотите сказать, что вне моей головы мир не существует?

Ну конечно же, он существует. Люди, окружающие вас, – это объективная реальность. Моя кошка, общественные ситуации, в которые мы попадаем, – все это тоже вполне осязаемая реальность. Однако ваше понимание мира и ваша реакция на него основаны на прогнозах, создаваемых вашей внутренней моделью. Каждое мгновение вы можете непосредственно ощущать только крошечную часть окружающей среды. Эта толика информации предопределяет, какие воспоминания всплывут из памяти, но ее одной недостаточно для того, чтобы у вас возникло текущее восприятие ситуации. Например, сейчас я сижу в офисе, печатаю и слышу стук во входную дверь. Я знаю, что это зашла в гости моя мать, и представляю себе, как она стоит внизу, хотя на самом деле я ее не вижу и не слышу. Ни один непосредственный входной сенсорный сигнал не был связан с моей матерью. Ее образ был вызван запоминающей моделью мира, спрогнозирован по аналогии с предшествующим опытом. В целом восприятие не построено исключительно на информации, получаемой из ваших органов чувств, а генерируется внутренней моделью памяти.

Итак, ответ на вопрос «Что такое объективная реальность?» во многом зависит от того, насколько адекватно ваша модель коры головного мозга отражает настоящую природу мира.

Многие особенности мира, окружающего нас, настолько последовательны, что практически каждый человек имеет одну и ту же внутреннюю модель мира. В детстве вы усвоили, каковы бывают поверхности, цвета и формы разных предметов. Вы знаете о том, что чередование дня и ночи, дней недели и времен года неизбежно. Вы знаете, что при попадании света на предмет возникает тень определенной формы. Такие простые физические свойства мира последовательно усваивают все люди. Тем не менее формирование модели мира во многом основано на обычаях, культуре и влиянии ближайшего окружения. Эти части нашей модели менее последовательны. Они могут полностью отличаться для разных людей. Если

ребенок вырос в доме, где его любят, где о нем проявляют заботу, с родителями, которые чутко относятся к его эмоциональным потребностям, он, становясь взрослым, как правило, прогнозирует, что мир – это дружелюбное и безопасное место. Дети, подвергавшиеся жестокому обращению со стороны одного или обоих родителей, склонны прогнозировать будущие события как опасные или жестокие и считать, что никому не стоит доверять (причем независимо от того, насколько хорошо с ними обращаются). Психология, в основном, обращается к последствиям раннего жизненного опыта, привязанностей и воспитания, поскольку именно в этот период мозг «закладывает фундамент» своей модели мира.

Ваша культура в очень большой степени предопределяет вашу модель мира. Например, полевые исследования показывают, что люди Востока и Запада по-разному воспринимают пространство и объекты. Азиаты больше внимания уделяют пространству между объектами, а представители западного мира больше внимания обращают на сами объекты. Это различие выливается в разное эстетическое восприятие и различные подходы к решению задач. Исследования показывают, что в некоторых культурах, например, у народностей, населяющих Афганистан и Южную Америку, превыше всего принципы чести. В результате возникает отношение к насилию как к вполне естественному явлению. Различные религиозные верования, усвоенные в раннем возрасте, могут вести к совершенно различным моделям морали, различному отношению к мужчинам и женщинам, и даже разному осознанию ценности самой жизни. Не существует абсолютной, универсальной модели мира. Понимание того, что такое хорошо и что такое плохо, формируется обучением.

Ваша культура (и опыт, полученный в семье) прививает вам стереотипы, которые, к сожалению, являются неотъемлемой частью вашей жизни. В этой книге вполне можно заменить словосочетание *инвариантная репрезентация* (или *инвариантное запоминание*) на *стереотип*, при этом смысл мало изменится. Прогнозирование на основе аналогии – в значительной степени то же самое, что и оценка на основе стереотипа. Негативные стереотипы могут иметь страшные последствия для общества. Если моя теория разума верна, то мы не сможем избавить людей от склонности к употреблению стереотипов, поскольку кора головного мозга работает на их основе. Создание стереотипов – это врожденное свойство мозга.

Способ, к которому мы должны прибегать для предотвращения ущерба, наносимого стереотипами, при воспитании своих детей, – это учить их распознавать ложные стереотипы, проявлять больше эмпатии и скептицизма. Нам нужно развивать эти навыки критического мышления в придачу к прививанию лучших качеств, нам известных. Скептицизм, являющийся основой научного метода, – единственный способ отличить факты от фикции.

Надеюсь, мне удалось убедить вас в том, что интеллект – это просто определение того, чем занимается неокортекс. Это не какая-то отдельная субстанция, сотрудничающая с нервными клетками головного мозга. Нейроны – это просто нервные клетки, и они остаются таковыми все время своего существования. Нет никакой мистической силы, заставляющей отдельные нейроны или их совокупность делать что-либо такое, чего они обычно не делают. Признав этот факт, мы можем сосредоточиться на поиске возможностей воссоздания способности клеток мозга к запоминанию и прогнозированию, т. е. на составлении алгоритма, имитирующего работу коры головного мозга, для воссоздания в кремнии.

8. Будущее разума

Предсказать пределы возможностей новых технологий чрезвычайно трудно. В этой книге я показал вам, что основной функцией мозга является составление прогнозов на основе прошлого опыта. Следовательно, можно предположить, что новые технологии будут использоваться для решения тех же задач, что и старые. Роль новых инструментов мы видим в том, чтобы они выполняли знакомые нам функции быстрее, эффективнее и дешевле.

Подобных примеров очень много. Люди называли железную дорогу «железным конем», а автомобиль – «повозкой без лошади». Десятки лет телефон считали усовершенствованной разновидностью телеграфа, т. е. чем-то, что должно использоваться только в экстренных случаях для передачи важных сообщений. Лишь в 1920-х годах его стали использовать в быту.

Фотографию сначала рассматривали как способ создания портретных изображений «лучшего качества». Знаете ли вы, почему в первых кинотеатрах был занавес для экрана? Потому, что кинематограф воспринимали как новый способ постановки представлений, прежде разыгрываемых на театральных подмостках.

Действительно, возможности применения новых технологий часто оказываются неожиданными и намного более далеко идущими, чем предполагается вначале. Так, создание телефона привело к развитию сети беспроводной передачи голоса и данных, позволяющей двум людям общаться друг с другом независимо от того, где они находятся, причем, передавать друг другу не только голос, но также тексты и изображения. Первый транзистор, созданный в *Bell Labs*²⁰ в 1947 году, сразу называли новаторским изобретением. Однако его первоначальное применение было ограничено усовершенствованием старых приложений – транзисторы пришли на смену вакуумным трубкам. Это привело к появлению более надежных радио и компьютеров, что в те дни вызывало восторг как ученых, так и потребителей, но основные отличия состояли в различии размеров и надежности машин. Наиболее революционное применение транзисторов было найдено лишь некоторое время спустя. До создания интегральной схемы, микропроцессора, цифрового сигнального процессора, чипов памяти должен был пройти период последовательных инноваций. Создание в 1970 году первого микропроцессора было вызвано желанием ученых усовершенствовать работу настольных калькуляторов. На этом примере мы также видим, что первичное применение научного открытия поначалу ограничивается лишь заменой существующих технологий. Так, на смену механическому настольному калькулятору пришел электронный. Микропроцессоры были очевидными кандидатами на замещение соленоидов, которые в то время использовались в некоторых видах промышленного управления (например, для переключения сигналов светофоров). И только годы спустя люди осознали, сколь велика на самом деле мощность микропроцессора. Никто в то время не мог предвидеть изобретение современного персонального компьютера, мобильного телефона, Интернета, системы глобального позиционирования²¹ или любой другой современной информационной технологии.

Тем более нелепо было бы ожидать, что мы сможем достоверно спрогнозировать революционное применение систем памяти, построенных на принципах работы мозга. Я твердо уверен, что разумные машины будут способны улучшить нашу жизнь во многих аспектах. Однако прогнозировать будущее технологий, связанных с искусственным интеллектом, да еще и на период, превышающий несколько лет, просто нереалистично. Чтобы осознать это, достаточно проанализировать некоторые наиболее безапелляционные прогнозы футуристов. Например, в 1950-х годах прогнозировалось, что к 2000-му году у каждого во дворе будет стоять свой атомный реактор, а отпуск мы будем проводить на Луне. Если мы не станем повторять подобных ошибок, переосмысление возможностей разумных машин может оказаться плодотворным. Как минимум, мы сможем сформулировать ряд полезных утверждений.

Вопросы на данную тему могут быть чрезвычайно интригующими. Сможем ли мы создать разумные машины, и если да, то как они будут выглядеть? Будут они напоминать человекоподобных роботов, знакомых нам по фантастическим романам, или «мыслящим» может быть и корпус системного блока компьютера? Каково будет практическое применение машин, обладающих искусственным интеллектом? Является ли данная технология опасной, может ли она угрожать человечеству? Какое применение имеет технология сейчас, и какое неожиданное применение ее возможно в будущем? И наконец, к чему в конце концов приведет

²⁰ *Bell Labs* – бывшая американская корпорация, крупный исследовательский центр в области телекоммуникаций, электронных и компьютерных систем. Основана в 1925 году как исследовательский центр компании *AT & T*. В настоящее время *Bell Labs* входит в состав компании *Lucent Technologies*. – *Примеч. ред.*

²¹ Система глобального позиционирования – спутниковая система, позволяющая определить место нахождения неподвижного либо движущегося объекта в любой точке Земного шара, будь то на земле, в воздухе, на море, в трех измерениях, с очень высокой точностью. Сейчас в мире существует две глобальные системы позиционирования (GPS): американская *US GPS* и российская *ГЛОНАСС*. В процессе разработки находится европейская система позиционирования. – *Примеч. ред.*

создание разумных машин?

Возможно ли создание разумных машин?

Да, возможно. Однако может оказаться, что они будут совсем не такими, как мы ожидаем. На первый взгляд, совершенно естественным представляется создание разумных машин, напоминающих человека, но я не сторонник такого мнения. Я не думаю, что разумные машины будут иметь человеческий облик и что они будут взаимодействовать с нами в форме, напоминающей человеческое общение.

Одно из наиболее популярных определений разумных машин пришло к нам из фильмов и фантастической литературы. Согласно указанным источникам, разумные машины – это милые либо злые, а иногда и нелепые роботы, которые способны испытывать чувства и делиться своими мыслями. Обычно такие создания играют ключевую роль в бесконечных фантастических интригах футуристических романов. Фантастическая литература, существующая вот уже сто лет, научила людей считать человекоподобных роботов неотъемлемой и желанной частью нашего реального будущего. Несколько поколений выросли на образах героев фильмов «Запрещенная планета», «Звездные войны», «Звездный путь» и тому подобных. Герой фильма «Космическая Одиссея 2001», хотя и не имел человеческого обличья, в целом очень походил на человека. Он был разработан не только как помощник, но и как товарищ пилота, отправляющегося в длительный космический полет. Роботы узкой специализации, как, например, «умные» автомобили, самоуправляемые пылесосы или газонокосилки, уже изобретены. Не исключено, что в далеком или, напротив, не столь отдаленном будущем они будут распространены повсеместно. А вот описанные выше человекоподобные роботы – гуманоиды – еще очень долго будут оставаться выдумкой, что объясняется двумя причинами.

Во-первых, человеческий разум – не только функционирование коры головного мозга, но и действие эмоциональных систем «старого» мозга, а также влияние человеческого тела с его сложнейшим строением. Человек – сложнейший биологический механизм, а не только кора головного мозга. Чтобы общаться подобно человеку на любые темы (т. е. чтобы пройти тест Тьюринга), разумная машина должна обладать опытом, принадлежащим человеку, и жить такой жизнью, как человек. Разумные машины будут обладать эквивалентом коры головного мозга и набором некоторых ощущений, а вот всем остальным – не обязательно. Идея сотворить разумную машину в человеческом обличье, конечно, очень и очень заманчива. Однако такая машина не сможет обладать разумом, подобным человеческому, если мы не насытим ее эмоциональными системами и человеческим опытом. Маловероятно, что столь сложная задача когда-либо окажется выполнимой.

Во-вторых, учитывая затраты на создание и эксплуатацию гуманоидов, трудно представить себе, в чем могла бы состоять их практическая целесообразность. Содержание робота-дворецкого обходилось бы намного дороже, чем наем обычного помощника по дому, кроме того, его функциональность была бы весьма ограниченной по сравнению с последним. Самый разумный робот не сможет проявить того понимания и участия, на которое способен человек по одной лишь своей природе.

Как паровой двигатель, так и цифровой компьютер возникли на основе проектов, так и не воплощенных в реальность. Точно так же, когда мы говорим о создании разумных машин, многие люди представляют их благодаря художественной литературе человекоподобными роботами, но вряд ли это станет реальностью.

Так как же тогда будут выглядеть разумные машины? Из теории эволюционного развития видно, что если наложить иерархическую систему памяти на возможность испытывать ощущения, то станет возможным моделирование и прогнозирование будущего. Заимствуя изобретения природы, мы должны построить разумные машины на основе ее же принципов. Это и будет рецептом создания разумных машин. Прежде всего, наша разумная машина может обладать набором ощущений, отличных от человеческих, она может «существовать» в своем мире, полностью непохожем на наш (более подробно остановимся на этом ниже). Другими словами, разумной машине совсем не обязательно обладать парой глаз и парой ушей. Дальше

мы добавим к этим ощущениям иерархическую систему памяти, работающую по принципу коры головного мозга. Затем мы обучим ее точно так же, как мы учим маленького ребенка. В результате наша разумная машина создаст модель *своего* собственного мира, такого, каким она его видит его сквозь призму *своих* ощущений. Ни у кого не возникнет ни потребности, ни возможности программировать ее с использованием баз данных, фактов или других концепций высшего порядка, что несовместимо с концепцией искусственного интеллекта. Разумные машины должны учиться на собственном опыте, в том числе – если возникает такая необходимость – получая входные сигналы от инструктора. Как только разумная машина создаст модель своего мира, она сможет распознавать аналогии, основанные на прошлом опыте, для прогнозирования событий в будущем, предлагать решение новых задач и делиться своим опытом с нами.

Такие разумные машины вполне могут быть вмонтированы в самолеты и автомобили или же размещаться на полках компьютерных лабораторий. В отличие от человека, мозг которого неотделим от его тела, система памяти разумной машины может иметь удаленный доступ к своим сенсорам (или «телу», если таковое у нее имеется). Например, интеллектуальной системе обеспечения безопасности целесообразно иметь сенсоры, расположенные по всей территории фабрики или города, а иерархическая система памяти, получающая информацию от этих сенсоров, может находиться в комнате лишь одного из сооружений. Таким образом, физическое воплощение машины, обладающей искусственным интеллектом, может быть чрезвычайно разнообразным.

Нет никаких оснований предполагать, что разумная машина должна выглядеть, действовать или чувствовать, как человек. Но ее искусственно созданный интеллект позволит ей формировать модель окружающего мира посредством иерархической системы памяти, а также думать о своем мире подобно тому, как мы с вами размышляем о нашем. Мысли и поведение разумной машины могут очень отличаться от свойственных человеку, и все же у нее будет интеллект, который, как мы показали в этой книге, определяется прогностической способностью иерархической памяти, а не человекоподобным поведением.

* * * * *

Главной технической трудностью при создании разумной машины будет создание иерархической системы памяти, которая работала бы подобно памяти головного мозга человека. Эту проблему мы будем решать при помощи связности и емкости.

Начнем с *емкости*. Кора головного мозга имеет порядка 32 миллионов синапсов²². Если мы представим каждый синапс двумя битами (что дает нам четыре возможных значения для каждого синапса), а каждый байт состоит из восьми бит (один байт может представлять четыре синапса), то нам понадобится порядка 8 трлн. байтов памяти. Винчестер современного персонального компьютера имеет 100 млрд. байтов памяти, т. е. нам понадобилось бы порядка 80 современных винчестеров, чтобы получить такой же объем памяти, как у человеческого неокортекса. (Пусть вас не беспокоят точные цифры, мы сейчас работаем над созданием общего представления.) Суть состоит в том, что в лабораторных условиях такая задача вполне решаема, хотя вмонтировать машину такого типа в тостер или карманный компьютер не получится. Важно заметить, что данное количество памяти сейчас вполне подлежит технической реализации, а всего каких-нибудь десять лет тому назад об этом нельзя было и мечтать. Следует также учитывать оптимистический факт: совсем не обязательно воссоздавать всю кору головного мозга человека. Для функционирования большинства приложений и устройств требуется намного менее объемная система памяти.

Возможно, мы начнем конструировать системы памяти в винчестерах или оптических дисках, но в конечном счете нам захочется создавать их на основе кремния. Кремниевые чипы

²² Синапс (от греч. соединение, связь) – специализированные межклеточные контакты нервной ткани, обеспечивающие передачу влияния с одного нейрона на другой и с нейронов на исполнительные органы. – *Примеч. ред.*

малы по размеру, имеют низкое энергопотребление и очень прочны. Вопрос о том, когда кремниевые чипы памяти обретут достаточную емкость, чтобы стать базой разумных машин, – исключительно вопрос времени. Разумная память имеет преимущества по сравнению с обычной компьютерной памятью. Экономика полупроводниковой промышленности основана на проценте чипов с ошибками. Для многих чипов единственная ошибка уже фатальна. Процент качественных чипов называется выработкой продукта. Он определяет, можно ли будет произвести чип определенного дизайна и продать его с прибылью. Вероятность допущения ошибки возрастает прямо пропорционально размерам чипов, поэтому на сегодняшний день большинство чипов имеют размер небольшой почтовой марки. Промышленность увеличила размер памяти одного чипа не за счет увеличения его размеров, а за счет уменьшения отдельных атрибутов.

С другой стороны, чипы разумной памяти толерантны к наличию ошибок. Как вы помните, ни одна из составляющих вашего мозга не владеет совершенно незаменимым носителем данных. Ваш мозг ежедневно теряет тысячи нейронов, тем не менее ваши интеллектуальные способности притупляются очень медленно – с годами в преклонном возрасте. Чипы разумной памяти будут работать по тем же принципам, что и кора головного мозга, т. е. даже если определенная часть элементов чипа будет неисправна, он все равно будет оставаться рабочим и окупится в экономическом плане. Скорее всего, прирожденная толерантность мозгоподобной памяти к наличию ошибок позволит дизайнерам разработать архитектуру чипов, которые будут значительно превосходить по размеру и плотности современные компьютерные чипы памяти. В результате вполне возможно, что мы сможем создать прототип мозга в кремнии намного раньше, чем сулят прогнозы.

Второе условие, необходимое для создания разумных машин, – это *связность*. Дело в том, что в человеческом мозге под тонким покрытием коры имеется белое вещество, состоящее из миллионов аксонов. Оно связывает области иерархии коры головного мозга между собой. Отдельная клетка коры головного мозга может быть связана с 5 или 10 тысячами других клеток. Такой тип масштабного параллельного соединения невозможно внедрить на основе традиционных техник производства кремниевых чипов. Последний создается путем нанесения нескольких слоев металла, каждый из которых отделяется от последующего изоляционным веществом. (Этот процесс наложения не имеет ничего общего со слоями коры головного мозга.) Слои металла вмещают «провода» чипа. В пределах одного слоя «провода» не пересекаются. Поэтому суммарное количество проводных связей в чипе ограничено. На основе такой связности совершенно невозможно создать мозгоподобную систему памяти, для которой реально необходимы миллионы подобных связей. Кремниевые чипы и белое вещество не очень-то совместимы друг с другом.

Понадобится большое количество инженерных разработок и экспериментальных исследований, прежде чем удастся решить эту проблему. Но направление поиска нам уже известно. Сигналы гораздо быстрее передаются по электрическим проводам, чем по аксонам нейронов. Один провод в чипе может обеспечивать несколько связей, в то время как в мозге каждый аксон имеет отношение лишь к единственному нейрону.

Примером из реальной жизни может служить телефонная система. Если мы будем проводить линию от аппарата к аппарату, то поверхность планеты утонет в джунглях медного провода. Вместо этого все телефоны подсоединены к относительно небольшому количеству высокочастотных телефонных линий. Данный метод работает потому, что емкость одной линии намного превышает емкость, требуемую для передачи отдельного разговора. Телефонная система вполне отвечает таким требованиям, ведь один оптоволоконный кабель может передавать порядка миллиона телефонных разговоров одновременно.

В человеческом мозге аксоны проходят между всеми клетками, связанными друг с другом. А вот мы можем создать разумную машину, которая по связности напоминает телефонную линию, где существуют общие связи. Хотите – верьте, хотите – нет, но некоторые ученые многие годы размышляли о том, как можно решить задачу связности чипа по типу мозга. Хотя работа самого мозга оставалась загадкой, тем не менее исследователи верили, что однажды загадка будет разгадана, и тогда мы столкнемся с задачей связности. Мы не будем сейчас рассматривать разные подходы к решению этой проблемы. Достаточно упомянуть, что

связность, возможно, является самой большой технической преградой на пути создания разумных машин, но со временем мы сможем преодолеть и ее.

Как только технические решения будут найдены, других значимых препятствий для создания по-настоящему разумных машин не останется. Конечно, придется поработать над уменьшением размеров, затрат на энергопотребление и, соответственно, снижением стоимости интеллектуальных систем, но в целом эти задачи не являются неразрешимыми. Всего за пятьдесят лет мы прошли путь от компьютеров, имевших размеры комнаты, до карманных аналогов. В данном случае мы начнем с довольно высокого технологического уровня, поэтому следует ожидать гораздо более стремительных изменений.

Нужно ли нам создавать разумные машины?

В XXI столетии разумные машины превратятся из выдумки, пришедшей из фантастических романов, в реальность. Однако пока этого не произошло, стоит обсудить ряд этических аспектов их создания, в первую очередь – соотношение потенциальных угроз и ожидаемой выгоды.

Перспектива появления в не столь отдаленном будущем машин, которые смогут думать и действовать сами по себе, уже очень давно беспокоит людей, что вполне объяснимо. Новые области знания и новые технологии почти всегда принимаются общественностью в штыки. Человеческое воображение с готовностью рисует все мыслимые и немыслимые ужасы, сопряженные с появлением искусственного интеллекта, – вплоть до сведения на нет самой ценности человеческой жизни. Но, как показывает история, все наши мрачные прогнозы никогда не бывают такими, как мы себе представляем. На заре индустриальной революции наши предки боялись электричества (помните Франкенштейна?) и паровых двигателей. Машина, обладающая собственной энергией и способная совершать сложные передвижения, внушала людям того времени ужас. Нам, их потомкам, электричество и двигатели внутреннего сгорания не кажутся ни странными, ни угрожающими. Они превратились в самые обычные составляющие нашей жизненной среды, подобно воде и воздуху.

Страх перед компьютерами зародился с началом эры информатики. Писатели-фантасты преподносили нам захватывающие дух истории о том, как мощные компьютеры или компьютерные сети, внезапно обретая сознание, нападают на своих создателей. Сейчас компьютеры заняли прочное место в нашей повседневной жизни, и такой страх кажется просто абсурдным. Компьютер, стоящий на вашем столе, или Интернет имеет так же мало шансов вдруг обрести ощущения, как и кассовый аппарат в ближайшем супермаркете.

Конечно, любая технология может быть использована во зло или с добрыми целями, но определенные из них изначально, по своей сути, несут людям большую угрозу. Атомная энергия опасна независимо от того, используется она на атомной электростанции или в ядерной боеголовке, потому что одна-единственная ошибка в управлении может стоить жизни миллионам людей. Хоть атом и является ценным источником энергии, однако у него существуют и альтернативы. Транспортная технология может воплотиться в танках и бомбардировщиках или в автомобилях и пассажирских самолетах. В любом случае использование данной технологии с дурными намерениями может принести людям много вреда. Транспортные средства важнее для современной жизни и намного менее безвредны, чем атомная энергия. Урон, который может быть причинен в самых худших обстоятельствах одним самолетом, во много раз меньше, чем потенциальная опасность одной атомной бомбы. С другой стороны, существуют и технологии, способные приносить только пользу, – например, телефоны. Способность телефонов соединять людей, помогать им поддерживать связь друг с другом намного превышает какие то ни было негативные побочные эффекты. То же самое можно сказать об электричестве и здравоохранении. По моему мнению, разумные машины будут одним из наименее опасных и одним из наиболее полезных изобретений человечества.

Некоторые, как, например, соучредитель компании *Sun Microsystems* Билл Джой, высказывают опасения следующего порядка: как бы мы не создали роботов, ускользающих из-под нашего контроля, которые окажутся способны захватить Землю. Подобные высказывания ассоциируются у меня с фантастическими историями. С другой стороны,

протагонисты искусственного интеллекта предлагают свои долговременные пророчества. Например, Рэй Курцвейл говорит о дне, когда посредством нанороботов можно будет внедриться в человеческий мозг, оцифровать каждый синапс и каждую нейронную связь, а потом передать полученную информацию в сверхмощный компьютер! Так можно будет «сконфигурировать» любого из нас. Хотите ли вы стать «программной» версией самого себя, которая будет практически бессмертна? Эти два прогноза, касающиеся искусственного интеллекта, – сценарий «Разумная машина впадает в бешенство» и сценарий «Ваш мозг загружается в компьютер» – кажется, никогда не перестанут будоражить умы людей.

Машина, обладающая интеллектом, и самореплицирующаяся машина – совсем не одно и то же. Между ними вообще не существует логической связи. Ни мозг, ни компьютер не обладают свойством саморепликации. Мозгоподобная система памяти тоже таким свойством не обладает. Ключевым преимуществом создания разумных машин является то, что мы сможем наладить их массовое производство, что совершенно не имеет ничего общего с саморепликацией бактерий или вирусов. Саморепликация не требует наличия разума, а наличие разума не нуждается в саморепликации.

И наконец, я сильно сомневаюсь в возможности когда-либо воссоздать в цифровом виде человеческий мозг. Методов, позволяющих записать миллиарды подробностей, из которых состоит «я» любого из нас, не существует на сегодняшний день, и вряд ли они когда-либо появятся. Для этого понадобилось бы оцифровать и воссоздать всю нервную систему конкретного индивида и все его тело, а не только кору головного мозга. Даже если в очень отдаленном будущем это станет возможным, задача расширяется намного больше, чем воссоздание работы неокортекса. Открыть алгоритм коры головного мозга и вмонтировать его в машину – это одно, но вот оцифровать миллионы операционных подробностей живого мозга и перенести их в кремний – совершенно другое.

Помимо саморепликации и копирования разума, высказываются и другие опасения. Могут ли разумные машины оказаться угрозой для больших групп людей, как это было в случае с атомной бомбой? Может ли их наличие привести к концентрации огромной власти в руках небольшой группки злоумышленников? Могут ли такие машины обернуться во зло и начать бороться против людей, как это было с героями «Терминатора» и «Матрицы»?

На все эти и подобные им вопросы могу дать уверенный отрицательный ответ. Как информационные устройства, мозгоподобные системы памяти будут одной из наиболее полезных технологий, которые доводилось изобретать человечеству. Точно так же, как автомобиль или компьютер, они останутся всего лишь инструментами. Наличие разума у машин не обеспечит им возможности разрушать или манипулировать людьми. Мы же не отдаем весь мировой атомный арсенал в руки одного человека или одного компьютера, следовательно, нам придется проявлять осторожность и не перегружать разумные машины, иначе они просто выйдут из строя, как и любая другая техника. Теперь вернемся к вопросу о злоумышленном использовании. Некоторые люди полагают, что быть разумным – то же самое, что и обладать человеческой ментальностью. Они опасаются, что разумные машины однажды взбунтуются против «порабощения», потому что гнет претит людям. Они боятся, что разумные машины попытаются захватить мир, потому что разумные люди, как показывает история, постоянно борются за власть. Все эти опасения базируются на ошибочной аналогии. Они основываются на объединении разума, т. е. алгоритма коры головного мозга, с эмоциональными устремлениями «старого» мозга, такими как страхи, стремление к обладанию, неприятие насилия. А ведь у разумных машин не будет подобных побуждений. У них не будет личных амбиций. Они не будут испытывать потребность обогащения, общественного признания и чувственного удовлетворения. У них не будет аппетита, пристрастий и – порой – плохого расположения духа. Разумные машины не будут иметь ничего напоминающего человеческие эмоции, если мы не приложим невероятные усилия для создания такой архитектуры. Наиболее эффективное применение разумные машины найдут там, где человеческий мозг испытывает трудности, – в областях, для которых недостаточно опоры на ощущения и интуицию, или же в скучных и утомительных видах деятельности.

Диапазон сложности разумных машин может быть очень большим – от простых с одним приложением до очень мощных систем, обладающих сверхчеловеческим интеллектом. В любом

случае, если мы не пойдем путем сознательных усилий, чтобы эти системы стали человекоподобными, они таковыми никогда не станут. Возможно, наступит день, когда нам придется поставить ограничения на сферы применения разумных машин, но сейчас нам до него еще далеко. Даже когда такой день настанет, то возникшие этические вопросы можно будет решить гораздо скорее и проще, чем те, перед которыми нас ставит генетика и ядерные технологии.

Зачем создавать разумные машины?

Чем же будут заниматься разумные машины?

Меня часто просят рассказать о будущем мобильной вычислительной техники, например, о том, как будут выглядеть карманные компьютеры или мобильные телефоны лет через двадцать. Когда мои собеседники интересуются моим видением будущего, я затрудняюсь дать ответ. Чтобы моя позиция была более доходчивой, я однажды вышел на сцену со «шляпой волшебника» и хрустальным шаром в руках. Я объяснил, что никто не может предвидеть будущее в подробностях, любой, кто утверждает, что может предвидеть то, что случится в ближайшие годы, ошибается. Лучшее, что мы можем сделать, – это понять ключевые тенденции. Понимая ключевую идею, вы сможете успешно следовать за ней независимо от того, чем она обернется.

Одним из наиболее ярких примеров технологических тенденций является закон Мура. Гордон Мур совершенно точно предсказал, что количество элементов, которые можно поместить на кремниевую пластину, будет удваиваться каждые два с половиной года. Мур не сказал, будут ли это чипы памяти, или центральные микропроцессоры, или еще что-то. Он также ничего не сказал о том, в каких типах носителей будут использоваться пластины. Он не прогнозировал, будут они помещены в пластиковую оболочку или в керамическую или же будут расположены на микросхеме. Он ничего не говорил о различных процессах, используемых для производства чипов. Он остановился на ключевой тенденции и оказался прав.

В наши дни мы не в силах представить пределы применения разумных машин. Если я или кто-либо другой во всех подробностях опишет перспективы функционирования искусственного интеллекта, он неминуемо ошибется. Тем не менее мы можем сделать нечто большее, чем просто пожать плечами. Существуют два направления, которым мы можем последовать. Первое: представить себе возможные области применения мозгоподобных систем памяти в краткосрочной перспективе, т. е. сначала испробовать менее захватывающие и интересные, но более реалистичные варианты. Второй подход: представить себе общие возможности применения ИИ в долгосрочной перспективе так, как это сделал Мур.

Предлагаю начать с возможных краткосрочных перспектив применения. Некоторые примеры (скажем, замена трубок на транзисторы в радиоприемниках или создание калькуляторов на микропроцессорах) более очевидны. Рассмотрим проблемные области, с которыми тем или иным образом сталкивался искусственный интеллект, но разрешить проблему так и не удалось. Я имею в виду распознавание речи, зрительное восприятие и «мыслящие» автомобили.

Если вы когда-нибудь использовали программное обеспечение для ввода произносимого вслух текста на персональном компьютере, тогда вы знаете, насколько бесполезным оно бывает. Как и в эксперименте с «Китайской комнатой», компьютер не понимает того, о чем идет речь. Несколько раз я пробовал использовать подобные приложения и всегда впадал в уныние. Если в комнате появлялся какой-то посторонний шум, от стука упавшего карандаша до голоса человека, обращающегося ко мне, то на экране сразу же возникали посторонние слова. Процент ошибок распознавания речи был очень высок. Часто слова, которые, как предполагала программа, я произнес, вообще не были связанными по смыслу. Даже ребенок понял бы, что в предложении ошибки, но не компьютер. Так называемый интерфейс естественной речи многие годы был целью инженеров, занимающихся разработкой программного обеспечения. Суть состоит в том, чтобы вы могли сказать машине, чего вы от нее хотите, обычным языком, и она бы выполнила ваши команды. Личной цифровой записной книжке вы могли бы сказать:

«Перенеси дочкину игру по баскетболу с субботы на сегодня, на десять утра». Подобного рода вещи невозможно было сделать с помощью традиционного искусственного интеллекта. Даже если бы компьютер распознал каждое слово, для выполнения задания ему нужно знать, где находится школа вашей дочери, какую именно субботу вы имели в виду, и, вообще, *что* такое баскетбольная игра, поскольку у вас может быть занесена информация как «Ментло против Сен-Джо». Или вы хотите, чтобы компьютер слушал радиопередачи и сканировал звуковой поток на предмет упоминания в нем определенного товара, а рассказчик будет описывать свой запрос, не упоминая его названия. Вы и я поймем, о чем он говорит, но этого не поймет компьютер. Подобные приложения требуют, чтобы машина могла не только слушать, но и слышать разговорную речь, что пока не достижимо. Программа распознавания речи соотносит звуковые сигналы с шаблонами слов, внесенных в память путем механического запоминания, не учитывая их значения. Представьте, что вы бы научились распознавать звучание отдельных слов на каком-то иностранном языке, не зная их значения. Я вас попрошу записать разговор на этом языке. Во время разговора вы понятия не имеете, о чем он, но пытаетесь распознать отдельные слова и записать их. Но в какой-то момент многие слова перекрываются, или частично неслышны, или появляется какой-то посторонний шум. Вам будет чрезвычайно сложно распознавать слова и вычленивать их. Именно с такого рода препятствиями сталкиваются современные программы по распознаванию человеческой речи. Их разработчики обнаружили, что, используя вероятность и переход слов, они могут несколько улучшить качество распознавания. Например, чтобы решить, какой из омонимов нужно выбрать, они используют правила грамматики. Это очень простая форма прогнозирования, однако системы в данном случае остаются немymi. Современные программы по распознаванию речи работают успешно лишь в очень ограниченных ситуациях, когда количество слов, которые человек может произнести в каждое отдельное мгновение, строго ограничено. А вот людям распознавание речи дается без труда, потому что неокортекс не только воспринимает отдельные слова, но и предугадывает содержание целых предложений, а также рамки общего контекста. В процессе распознавания устной речи мы прогнозируем идеи, фразы, отдельные слова, мало того – кора головного мозга выполняет всю эту работу автоматически.

Мы имеем основания ожидать, что системы памяти, построенные на принципах работы коры головного мозга, позволят создать грамотные системы распознавания речи. Вместо программирования на основе вероятностей словесных переходов иерархическая память будет отслеживать акценты, слова, фразы, идеи и использовать их для интерпретации того, что было сказано. Как и человек, такая система сможет проследить различия между разнообразными ситуациями беседы (например, ваш разговор с другом в комнате, телефонный разговор, редактирование команд для книги). Конечно, создать такие машины будет непросто. Чтобы полностью понимать человеческую речь, машина должна многое «пережить» и «научиться» тому же, что и люди. Возможно, нам понадобятся долгие годы, чтобы создать разумную машину, которая понимает язык так же хорошо, как вы и я. А в ближайшем будущем мы можем рассчитывать лишь на улучшения существующих систем распознавания человеческой речи путем использования системы памяти, построенной на принципах работы коры головного мозга.

Зрительное восприятие – еще одна задача, которую так и не смог решить существующий искусственный интеллект, но она под силу по-настоящему разумным системам. На сегодняшний день не существует машины, которая смогла бы наблюдать естественную сцену, как, например, мир перед вашими глазами или картинка видеокamеры, и описать то, что она видит. Есть несколько примеров успешного применения систем, распознающих изображения, но весьма ограниченные. Речь идет, например, о регистрации расположения чипа на интегральной схеме или сопоставлении черт лица с базой данных. На данной стадии компьютер не может распознавать различные объекты или анализировать наблюдаемую картину в более широком плане. У вас не возникает проблем, когда вы заходите в комнату и ищете место, где можно было бы сесть, но не просите компьютер сделать то же самое. Представьте себе, что вы смотрите на экран камеры безопасности. Сможете ли вы увидеть различия между человеком, держащим в руке подарок и стучащим в дверь, и человеком, у которого в руке перо ворона? Конечно сможете, однако эти различия превосходят возможности современного программного обеспечения. Поэтому мы нанимаем людей, которые следят за экранами камер безопасности

круглые сутки и наблюдают, не происходит ли что-нибудь подозрительное. Человеку-наблюдателю непросто сохранять бдительность много часов подряд, а вот машина могла бы сделать это без труда.

Давайте также рассмотрим транспортные средства. Устройство автомобилей постоянно усложняется. Существуют глобальные системы позиционирования, способные просчитать кратчайший маршрут от точки А до точки Б, существуют сенсоры, автоматически включающие фары, когда на улице становится темно, сенсоры приближения, которые помогают определить безопасность обгона и так далее. Существуют даже автомобили, способные ехать без водителя на специальных трассах в идеальных условиях. Они, правда, не поступают в открытую продажу. Чтобы вести машину безопасно и эффективно на любых дорогах при любых условиях, вовсе не достаточно несколько датчиков или схем обратного контроля. Чтобы быть хорошим водителем, вам нужно понимать устройство автомобиля, характер дорожного движения, предугадывать маневры других водителей и учитывать массу других обстоятельств. Водителю следует внимательно отслеживать сигналы, предостерегающие об опасности. Например, включенный сигнал поворота у впереди идущего автомобиля предупреждает о том, что водитель собирается перестроиться в другой ряд. Если сигнал включен несколько минут подряд, то, скорее всего, водитель о нем просто забыл и перестраиваться не собирается. Когда водитель видит мяч, катящийся на проезжую часть со стороны тротуара, он автоматически предполагает, что, возможно, за мячом сейчас побежит ребенок, и интуитивно резко сбрасывает скорость.

Допустим, мы хотим создать по-настоящему «умный» автомобиль. Сначала нам нужно определить набор сенсоров, которые позволят нашему автомобилю воспринимать текущую дорожную ситуацию. Мы, вероятно, начнем с камеры видения, даже нескольких камер в передней и задней частях автомобиля, затем придумаем микрофоны для воссоздания звукового восприятия, затем, возможно, добавим радар или ультразвуковые сенсоры, точно определяющие диапазон и скорость других объектов на трассе как при хорошем освещении, так и в темное время суток. Не надо создавать рамки, ограничиваясь ощущениями, свойственными исключительно человеку. Алгоритм коры головного мозга очень гибок, и если мы спроектируем нашу рукотворную иерархическую систему памяти должным образом, она будет работать независимо от типов установленных сенсоров. Теоретически наш воображаемый «мыслящий» автомобиль может воспринимать ежесекундно меняющуюся дорожную ситуацию лучше, чем водитель-человек, потому что может выбирать разные наборы сенсоров в зависимости от поставленной задачи. Сенсоры будут связаны с достаточно большой иерархической системой памяти. Разработчики автомобилей будут тренировать его память путем помещения в условия реального мира, чтобы она училась создавать модель мира точно так же, как это делают люди, только в условиях более ограниченной области. (Ведь автомобилю нужны знания об автодорогах, а не об элеваторах и аэропланах.) Память автомобиля «выучит» иерархическую структуру транспорта и дорог так, что сможет понимать, что происходит в ее мире движущихся автомобилей, дорожных знаков, препятствий и перекрестков в текущий момент времени, и прогнозировать ход развития дальнейших событий. Инженеры-разработчики такого автомобиля могут настроить систему его памяти так, чтобы она полностью управляла автомобилем или только отслеживала, *что* происходит, когда за рулем сидите вы. Она может давать советы или принимать на себя управление в экстремальных ситуациях. Как только память будет полностью натренирована и сможет понимать и решать любые проблемы, с которыми сталкивается, инженеры получат возможность выбрать два варианта последующего применения. Они могут поместить стандартную память во все автомобили, сходящие с конвейера, или же комплектовать их памятью, которая будет продолжать самообучение после продажи автомобиля. Как и в случае с компьютерами, память автомобиля должна поддаваться перепрограммированию более актуальной версией.

У меня нет абсолютной уверенности в том, что нам удастся создать подобные умные автомобили или машины, обладающие зрительным и слуховым восприятием. Я лишь привел примеры того, какие типы устройств мы, возможно, разработаем и какие типы разумных машин кажутся реалистичными на сегодняшний день.

* * * * *

К очевидному применению разумных машин у меня гораздо меньше интереса. Я способен оценить пользу и искренне восхищаться новой технологией лишь тогда, когда она получает неизвестное доселе применение. Чем удивят нас разумные машины и какие неожиданные их свойства и способности проявятся со временем? Я убежден, что иерархические системы памяти, подобно транзисторам и микропроцессорам, улучшат нашу жизнь совершенно неожиданным образом, но как именно? Один из способов заглянуть в будущее разумных машин – рассмотреть их составляющие, легко поддающиеся измерению. Например, какие атрибуты разумных машин будут становиться все более дешевыми, более быстродействующими и менее громоздкими? То, чему присущ экспоненциальный рост, может быстро превзойти наши ожидания и, скорее всего, сыграет ключевую роль в наиболее радикальных разработках будущих технологий.

Примерами технологий, которые развивались экспоненциально на протяжении многих лет, являются кремниевый чип памяти, винчестер, техники последовательности ДНК и фиброоптика. Указанные быстро-развивающиеся технологии положили начало многим новым товарам и сферам бизнеса.

В отличие от приведенных примеров развитие некоторых технологий, таких как аккумуляторы, двигатели и традиционная робототехника, довольно инертно. Несмотря на многочисленные усилия и постоянные усовершенствования, рука современного робота не очень отличается от руки его предшественника, разработанного несколько лет тому назад. Развитие традиционной робототехники можно считать медленным, а успехи скромными, несравнимыми с экспоненциальным развитием дизайна чипов или темпами распространения программного обеспечения. Рука робота, созданная в 1985 году и повлекшая затраты в миллион долларов, не будет сегодня в тысячу раз сильнее при цене в десять долларов. И аккумуляторы, производимые сегодня, не намного лучше, чем те же устройства десятилетней давности. В лучшем случае они превосходит мощность предшественников в два-три раза, но не в тысячи или десятки тысяч раз. Прогресс в их развитии происходит очень и очень медленно. Если бы емкость батарей возрастала с той же скоростью, что и емкость винчестеров, тогда мобильные телефоны и прочую электронику никогда не понадобилось бы перезаряжать, а электромобили, проходящие расстояния в тысячу миль после одной-единственной подзарядки, были бы сегодня нормой повседневной жизни.

Теперь мы подошли к разговору о том, какие аспекты мозгоподобных систем памяти значительно превзойдут биологический мозг. Я вижу четыре потенциальных свойства, которые могут превзойти человеческие способности. Это быстрота, емкость, реплицируемость и сенсорные системы.

Быстрота

Быстродействие нейронов измеряется в миллисекундах, а кремний функционирует со скоростью наносекунд, причем постоянно возрастающей. Различия скорости быстродействия органического и кремниевого разума – на шесть порядков! – будут иметь большие последствия. Разумные машины смогут мыслить в миллион раз быстрее, чем человеческий мозг. Такой разум сможет прочесть целую библиотеку книг или изучить большую и очень сложную базу данных – задача, для выполнения которой любому из нас понадобились бы годы, – за несколько минут, причем понимание прочитанного не пострадает. В этом нет никакой магии. Биологический мозг развивался в условиях двух временных ограничений: пределов скорости действия нервных клеток и скорости изменений внешнего мира. Какой резон биологическому мозгу ускорять мышление в миллионы раз, если окружающая человека среда изначально инертна?! Однако в самом алгоритме коры головного мозга не заложено ничего такого, что тормозило бы его быстродействие. Если бы разумная машина взаимодействовала с человеком, ей пришлось бы снизить свое быстродействие и работать с «человеческой» скоростью. Если бы ей пришлось читать книгу и переворачивать при этом страницу за страницей, тоже существовало бы ограничение скорости чтения. Но вот если бы разумная машина взаимодействовала с электронным миром, она бы работала гораздо быстрее. Две разумные машины общались бы друг с другом в миллион раз быстрее, чем двое людей. Представьте себе прогресс разумной

машины, которая решает математические задачи или реализует научные идеи в миллионы раз быстрее, чем человек. За десять секунд она добьется более впечатляющих результатов, чем вы – за месяц. Не знающий усталости и скуки, искусственный разум, работающий со скоростью света, явно может найти очень полезное применение, такое, о котором мы еще даже не подозреваем.

Емкость

Несмотря на впечатляющую емкость человеческой коры головного мозга, разумные машины смогут ее намного превзойти. Размер нашего мозга ограничивается несколькими биологическими факторами, среди них – соотношение размера черепа ребенка и диаметр таза матери, большие метаболические затраты на поддержание мозга (мозг составляет порядка 2% веса тела и при этом потребляет 20% вдыхаемого кислорода), низкое быстродействие нейронов. С другой стороны, мы можем создавать разумные системы памяти любых размеров, заложив свою задумку в отдельные атрибуты дизайна. Возможно, несколько десятилетий спустя емкость человеческого мозга будет нам казаться очень скромной.

Предскажу несколько способов увеличения емкости памяти будущих разумных машин. Во-первых, сделав иерархическую систему памяти более глубокой, мы добьемся более глубокого понимания – умения распознавать сигналы высокого порядка. Увеличение емкости отдельных зон поможет разумной машине запоминать больше подробностей, проявлять более высокую точность восприятия (подобно тому, как незрячий человек обладает повышенной осязательной и слуховой чувствительностью). Добавление новых ощущений и сенсорной иерархии позволит устройству создавать более качественные модели внешнего мира. К этому мы еще вернемся ниже.

Возникает закономерный вопрос: существует ли предел того, как далеко может зайти разумная система памяти и в каких измерениях? Предположительно, на какой-то стадии устройство станет слишком громоздким, чтобы быть по-настоящему полезным, или оно начнет давать сбои, приблизившись к своему теоретическому пределу. Возможно, человеческий мозг уже приблизился к своему максимальному теоретическому размеру, хотя я считаю это маловероятным. Человеческий мозг увеличился относительно недавно в ходе эволюции, и ничто не наводит на мысль о том, что мы вышли на стабильный максимальный уровень. Каким бы ни был лимит емкости разумной системы памяти, человеческий мозг почти наверняка до него и близко не дотягивает.

Еще один способ понять, каким образом могли бы быть полезны разумные системы памяти, – проанализировать пределы человеческих возможностей. Эйнштейн, вне всякого сомнения, был очень умен, но его мозг был всего лишь мозгом. Можно предположить, что его выдающийся ум был последствием физических отличий его мозга от типичного. Эйнштейн был уникален потому, что человечество нечасто рождает гениев. А создавая мозг из кремния, мы можем воплотить в жизнь что угодно. Кремниевый мозг может располагать таким же уровнем интеллекта, как и Эйнштейн, или даже превзойти его. С другой стороны, лучше познать возможные измерения разума нам помогут умственно отсталые люди, проявляющие фотографическую память или способности осуществлять в уме сложные математические вычисления. Совершенно нетипичный мозг таких индивидов, несмотря ни на что, является мозгом, неокортекс которого работает по тому же алгоритму. Если нетипичный мозг обладает невероятными способностями к запоминанию, то теоретически мы можем добавить эти способности нашему искусственному мозгу. Подобные экстремальные проявления человеческих умственных способностей не только показывают, что можно было бы воссоздать, но и указывают направления, в которых мы могли бы превзойти наилучшие проявления работы человеческого мозга.

Реплицируемость

Каждый новый органический мозг растет и обучается на протяжении десятилетий человеческой жизни. Каждый человек на собственном опыте познает основы координации конечностей и мышечных групп, осваивает азы и передвижения, изучает общие свойства множества объектов окружающей среды, животных, других людей, названия предметов и структур речи, семейные и общественные правила. Как только основа заложена, начинаются годы формального обучения. Каждый человек в течении своей жизни проходит множество

кругов обучения. Несмотря на то что путь этот несметное количество раз повторялся другими людьми, модель мира в коре головного мозга каждого из нас создается в индивидуальном порядке.

Разумным машинам нет необходимости проходить подобную спиралевидную кривую обучения, поскольку чипы и другие единицы хранения информации можно реплицировать бесконечно, а сохраненные данные – копировать. В этом отношении разумные машины подлежат такой же репликации, как и программное обеспечение. Как только прототип системы достаточно обучен, его можно скопировать столько раз, сколько нам потребуется. Процессы разработки дизайна, настройки системного обеспечения, обучения, испытания и выявления ошибок с тем, чтобы усовершенствовать систему памяти умной машины, могут длиться годами. Но, как только мы получим конечный продукт, его можно будет запускать в массовое производство. Как я уже отмечал, мы сможем обеспечить возможность последующего обучения для копий или обойтись без таковой. Для некоторых приложений мы захотим ограничить возможности с тем, чтобы разумные машины работали известным и испытанным нами способом. Как только «умный» автомобиль усвоит все, что ему положено «знать», мы примем меры, чтобы у него не образовались плохие привычки или не закрепились самостоятельно найденные ложные аналогии. Мы захотим, чтобы все машины подобного строения вели себя одинаково. Но вот для других приложений может оказаться целесообразной возможность постоянного обучения мозгоподобных систем памяти. Например, разумной машине, созданной для того, чтобы искать математические доказательства, понадобится способность обучаться на основе опыта, применять старые решения для новых проблем, ей желательно быть универсально гибкой и открытой.

Возможным станет общее использование компонентов, как сейчас общедоступны программные продукты. Разумная машина определенного дизайна может быть перепрограммирована набором новых связей с тем, чтобы генерировать новое поведение (вот, например, если бы я смог загрузить новый набор связей в ваш мозг, моментально превратив вас из человека, разговаривающего на английском, в человека, знающего только французский, или из профессора политологии в музыковеда). Люди получили бы возможность «переключаться» и осуществлять новые наработки на основе чужого опыта. Допустим, что я разработал и внедрил в жизнь суперсистему зрительного восприятия, а другой человек разработал и обучил систему с превосходными способностями слухового восприятия. При надлежащем дизайне мы смогли бы объединить обе системы без необходимости повторного обучения с самого начала. Такого рода объединение экспертных знаний не представляется возможным для человека – у каждого из нас свой путь. Бизнес создания разумных машин смог бы развиваться по аналогии с компьютерной индустрией, в таком случае перепрограммирование разумных машин не слишком отличалось бы от разработок новых версий видеоигр или инсталляции новой программы.

Сенсорные системы

Люди познают мир посредством органов чувств. Ощущения, которые мы испытываем, обусловлены генами, строением наших тел и соединениями под корой головного мозга. Мы не можем изменить их. Иногда мы используем технологические достижения, позволяющие сделать наши ощущения более острыми. Например, приборы ночного видения, радары и космические телескопы представляют собой устройства, оптимизирующие передачу данных, но не обеспечивающие нам новых способов восприятия. Они конвертируют информацию, недоступную для нас в ее первоначальном виде, в визуальные или звуковые сигналы, поддающиеся нашему восприятию. И все равно, в том, что мы можем взглянуть на экран радара и идентифицировать изображение, заслуга невероятной гибкости нашего мозга.

Многие виды животных обладают поразительными ощущениями, совершенно несвойственными человеку. В качестве примеров можно привести эхолокатор у дельфинов и летающих мышей, способность пчел видеть поляризованное и ультрафиолетовое излучения, восприятие электрического поля некоторыми видами рыб.

Разумные машины смогли бы воспринимать мир посредством всех видов ощущений, существующих в природе, а также изобретенных человеком. Данные об окружающей среде,

получаемые с помощью гидролокатора, радара и приборов ультрафиолетового видения, являются очевидными примерами несвойственных человеку ощущений, которые мы, возможно, хотели бы воссоздать в разумной машине. И это всего лишь начало.

Намного более интересными являются способы восприятия, которые нам неизвестны и чужды. Как мы уже выяснили, алгоритм коры головного мозга в первую очередь сосредоточен на поиске сигналов внешнего мира. У него нет никаких приоритетов относительно физических истоков этих сигналов. Если входные сигналы не носят произвольный характер и обладают определенным богатством, а также статистической структурой, разумная машина будет формировать систему воспоминаний о них и прогнозы на их основе. Не существует причин, по которым входные сигналы должны иметь аналогию с физическими ощущениями или вообще иметь что-либо общее с реальным миром. Я полагаю, что именно в сфере экзотических ощущений скрыты перспективы революционного применения разумных машин.

Например, вообразите сенсорную систему, покрывающую планету. Представьте себе сенсоры, расположенные на расстоянии каждой пятидесяти миль по всем континентам. Эти сенсоры имитировали бы клетки сетчатки глаза. Каждое мгновение два смежных сенсора погоды имели бы высокую корреляцию активности, подобно высокой корреляции двух смежных клеток сетчатки. Существуют крупные погодные явления – например, штормы и фронтальные циклоны, – которые передвигаются в пространстве и меняются со временем точно так же, как и все объекты, которые меняются и перемещаются. Привязав этот сенсорный ряд к системе памяти, работающей по принципу коры головного мозга, мы бы обучили указанную систему прогнозировать погоду, подобно тому, как люди учатся распознавать объекты окружающей среды и прогнозировать траекторию их передвижения. Такая система предсказывала бы погодные условия на ближайшее время, а также формировала бы долгосрочные метеопрогнозы. Разместив сенсоры рядом друг с другом в какой-то области, мы бы создали своеобразный эквивалент ямки глаза, которая позволяла бы нашей системе понимать и прогнозировать микроклимат. Наш «погодный мозг» размышлял и понимал бы глобальные погодные системы точно так же, как мы с вами понимаем объекты и людей. Метеорологи пытаются создать нечто подобное в наши дни. На основе данных, собранных в разных частях Земли, используя сверхмощные компьютеры, ученые симулируют текущие погодные условия и прогнозируют будущее. Однако данный подход фундаментально отличается от принципов работы разумных машин. Он больше сродни компьютеру, играющему в шахматы, а разумная система, сконструированная по принципам работы неокортекса, скорее, напоминает человека, играющего в шахматы. Она делает это вдумчиво и с пониманием. Разумная машина-"синоптик" учла бы те сигналы, которых человек распознать не может. Она смогла бы распознать множество новых погодных феноменов (скажем, такое погодное явление, как ураганы Эль-Ниньо было открыто лишь в шестидесятых годах XX века). Она бы прогнозировала появление торнадо и муссонов намного точнее, чем человек. Человеку не под силу сохранять в памяти и анализировать множество метеоданных искусственный интеллект компьютера-"синаптика", напротив, был бы способен ощущать и воспринимать погодные условия непосредственно.

Другие распространенные большие сенсорные системы могли бы дать нам возможность создать разумные машины, понимающие и прогнозирующие миграцию животных, демографические изменения и распространение инфекционных болезней. Представьте себе сенсоры, размещенные в сети электрического тока некой страны. Разумная машина, присоединенная к этим сенсорам, наблюдала бы за спадами и подъемами потребления электроэнергии точно так же, как мы с вами наблюдаем большее или меньшее количество транспортных средств на трассе или потоки передвижения пассажиров в аэропорту. Посредством повторных наблюдений люди учатся прогнозировать подобные сигналы. (Можете спросить у тех, кто каждый день ездит на работу на автомобиле или охранника в аэропорту.) Точно так же наша разумная система сможет прогнозировать повышенное напряжение или опасные ситуации, которые могут вызвать сбои в подаче электричества, лучше, чем человек. Мы могли бы сочетать погодные и демографические сенсоры, чтобы спрогнозировать вспышки политического недовольства, голода или болезней. Подобно гениальному дипломату, разумные машины могли бы сглаживать конфликты и уменьшать человеческие страдания. Вы можете

подумать: для того чтобы предвидеть сигналы, включающие человеческое поведение, разумным машинам непременно нужны эмоции. Я так не считаю. Мы не рождены с заданной культурой, заданными ценностями и заданной религией, мы приобрели их в процессе жизненного опыта, иными словами – обучились им. Точно так же, как я могу изучить мотивацию людей с ценностями, отличающимися от моих, разумные машины могут понять человеческие мотивы и эмоции, сами таковыми не обладая.

Мы могли бы создать сверхтонкие ощущения. Теоретически возможно иметь сенсоры, регистрирующие сигналы клеток или больших молекул. Например, перед нынешними биологами стоит серьезная задача – разобраться, как можно спрогнозировать форму молекулы протеина, зная последовательность аминокислот, формирующих протеин. Способность прогнозировать расщепление и взаимодействие протеинов ускорила бы развитие медицины и разработку многих лекарственных препаратов. Инженеры и ученые создали трехмерные визуальные модели протеинов с тем, чтобы прогнозировать, как эти сложные молекулы будут себя вести. Но, как бы мы ни старались, на сегодняшний день эта задача слишком сложна. С другой стороны, машина, обладающая суперинтеллектом и набором ощущений, специально настроенных на данную проблему, могла бы ответить на поставленный вопрос. Быть может, эта идея представляется вам чересчур надуманной, но ведь нас бы не удивило, если бы человек смог решить такую проблему. Возможно, наша неспособность разрешить проблему связана в первую очередь с несоответствием масштабов человеческих ощущений и изучаемого физического явления. Разумные же машины смогут обладать выборочными ощущениями и памятью большей емкости, чем человеческая, которая позволит им решать задачи, людям неподвластные.

При наличии соответствующих ощущений и небольшом структурном изменении памяти коры головного мозга разумные машины будущего смогли бы жить и мыслить в виртуальном мире математики и физики. Многие математические проблемы, например, касаются поведения объектов в мире с более, чем тремя измерениями. Ученые, изучающие природу пространства, размышляют о Вселенной как сущности, имеющей десять и более измерений. Но даже с восприятием четырехмерного пространства у людей возникают немалые трудности. Возможно, разумная машина с соответствующими функциями сможет понять многомерные пространства, как вы и я понимаем трехмерное, и сможет спрогнозировать поведение находящихся в них объектов.

И наконец, мы смогли бы объединить несколько разумных систем в большую иерархию, подобно тому, как наша кора сочетает осязание, слуховое и зрительное восприятия в высших областях иерархии коры головного мозга. Такая система автоматически научилась бы моделировать и прогнозировать сигналы в популяциях разумных машин. С помощью средств коммуникации, таких как Интернет, разумные машины могли бы объединиться в сеть, охватывающую весь Земной шар. Большие иерархии способны регистрировать более глубокие сигналы и более сложные аналогии.

Цель изложенных мной размышлений состоит в том, чтобы показать, что существует очень много сфер, в которых мозгоподобные машины могли бы намного превзойти наши способности. Они смогли бы мыслить и обучаться в миллионы раз быстрее, чем мы с вами, запоминать огромные массивы подробной информации и улавливать практически сигналы. Они могут обладать ощущениями более чувствительными, чем наши, или ощущениями для восприятия очень специфических феноменов. Они без труда могут размышлять в терминах многомерных пространств. Ни одно из перечисленных любопытных свойств не зависит от внешнего сходства разумных машин с человеком, и эти машины также не требуют сложной робототехники.

Теперь, я надеюсь, вы понимаете, насколько тест Тьюринга, в котором между интеллектом и человеческим поведением ставился знак равенства, ограничил наше видение потенциальных возможностей разумных машин. Глубоко познав интеллект, мы сможем создавать разумные машины, которые будут представлять намного большую ценность, чем простое воспроизведение человеческого поведения. Наши разумные машины станут отличным инструментом, позволяющим значительно обогатить наши познания о Вселенной, наши способности и опыт.

* * * * *

Как много воды утечет прежде, чем мы достигнем такого уровня? Создадим ли мы по-настоящему разумные машины через пятьдесят, двадцать или пять лет? В мире высоких технологий любят повторять: кратковременные изменения длятся дольше, чем это предполагалось, а долгосрочные перемены происходят намного быстрее, чем прогнозировалось. Я много раз убеждался в справедливости этого высказывания. Кто-то поднимется во время конференции и скажет, что новая технология появится в каждом доме уже через четыре года. А потом окажется, что он ошибся. Четыре года проходят, восемь лет, все начинают думать, что этого уже никогда не произойдет. И вот как раз в тот момент, когда кажется, что сама идея зашла в тупик, она вдруг оживает и становится величайшей сенсацией. Скорее всего, что-то подобное произойдет и с индустрией разумных машин. Сначала прогресс будет малозаметным, но затем резко пойдет в гору.

На конференциях по нейробиологии я люблю задавать присутствующим вопрос: «Как вы думаете, сколько еще времени пройдет до тех пор, пока у нас появится рабочая теория коры головного мозга?» Некоторые слушатели (порядка 5%) отвечают: никогда, или она уже существует (интересный вариант ответа, учитывая профессию, благодаря которой они зарабатывают на жизнь). Еще примерно 5% считают, что понадобится порядка пяти-десяти лет. Половина оставшихся считает, что ждать остается от десяти до пятидесяти лет или отвечают: «Еще при нашей жизни». Остальные утверждают, что речь идет о пятидесяти-двухстах годах (или «не при нашей жизни»). Лично я оптимист. Мы долгие годы находились «в фазе затишья», поэтому многим кажется, что прогресс в теоретической нейробиологии и разработке разумных машин зашел в полный тупик. Если строить выводы на основе последних тридцати лет, то вполне естественно считать, что мы не очень близки ответу. Но я полагаю, что как раз сейчас мы находимся в поворотном пункте, и вскоре наша область знаний начнет развиваться стремительными темпами.

Можно приблизить момент, которого мы так ждем. Одна из целей данной книги – убедить вас в том, что при наличии соответствующих теоретических рамок мы сможем значительно продвинуться вперед в понимании коры головного мозга. Другими словами, при помощи запоминающе-прогностических рамок мы сможем разгадать особенности работы мозга и мышления. Эти знания нужны нам для создания разумных машин. Если мы попадем в точку, модель окажется корректной, и прогресс будет происходить очень быстро.

Я и не берусь давать точный прогноз, *когда* наступит эра разумных машин, тем не менее я полагаю, что если достаточно много людей посвятят себя решению этой проблемы сегодня, то, возможно, мы сможем создать полезные прототипы и симуляции коры головного мозга в пределах ближайших нескольких лет. Через десять лет, как я надеюсь, разумные машины станут одной из наиболее «горячих» тем в области науки и технологии. Мне не хотелось бы составлять более подробные прогнозы, поскольку я знаю, как трудно оценить время, требующееся для важных открытий. Так почему же я столь оптимистически настроен относительно теории мозга и создания разумных машин? Моя уверенность во многом обусловлена количеством времени, потраченным мною на изучение работы разума. Когда в 1979 году я впервые всерьез заинтересовался мозгом, я почувствовал, что эта головоломка может быть разгадана еще при моей жизни. Годами я наблюдал за неуклонным спадом в сфере разработки искусственного интеллекта, подъемами и падениями в области создания нейронных сетей, а также был свидетелем «Десятилетия мозга» в 1990-х годах. Я видел, как менялось отношение к теоретической биологии в целом и нейробиологии в частности. Я был свидетелем того, как понятия прогноза, иерархических представлений и фактора времени прочно укоренились в арсенале нейробиологов. Я отмечал изменения в своем понимании проблемы искусственного интеллекта и наблюдал соответствующие перемены у своих коллег. Гипотеза о роли прогнозирования в работе мозга увлекла меня еще восемнадцать лет назад, и с тех пор я постоянно проверял ее. Поскольку я был поглощен нейробиологией и сферой компьютерных наук на протяжении более двух десятилетий, возможно, мой мозг создал модель высшего порядка, которая описывает, как происходят технологические и научные изменения, и эта

модель прогнозирует быстрые темпы развития. Как раз сейчас наступил поворотный момент.

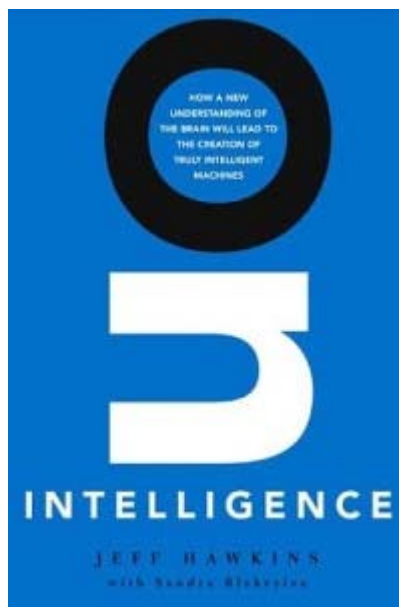
Эпилог

Очень многие ученые исключают для себя возможность восхищения собственными открытиями, полагая, что этому помешает прагматичный ум и научное понимание проблемы. Но известный астроном Карл Саган как-то отметил, что понимание не умаляет изумления, вызванного открытием, как познание не лишает жизнь полноты красок, звуков, света. Я соглашаюсь с Саганом. Углубляя свои познания, осваивая свои роли, мы более комфортно чувствуем себя во Вселенной, а мир, в свою очередь, обретает еще больше красок и загадочности. Быть маленькой, но живой, разумной и творческой искоркой в бесконечном космосе намного интереснее, чем обитать на плоской ограниченной Земле в центре своего крошечного мирка. Понимание работы человеческого мозга нисколько не умаляет загадочности и гармоничности устройства мироздания, поразительной полноты нашей жизни, захватывающих дух перспектив будущего человечества. Нашему изумлению не будет предела, когда сможем применять эти знания на практике – не только для создания разумных машин, но и для понимания самих себя, и для углубления собственных знаний.

Заканчивая эту книгу, я выражаю надежду на то, что она привлекла внимание кибернетиков и представителей других направлений научного мира к коре головного мозга, что они смогут применить на практике предложенные теоретические рамки и создать разумные машины. В свое время идея искусственного интеллекта была очень популярной – ей посвящались многочисленные публикации в журналах и книги, на ее основе составлялись бизнес-планы и в соответствии с ней координировались усилия предпринимателей. Открывались соответствующие научные кафедры и исследовательские институты. Точно так же и нейронные сети привлекли большое внимание, когда пребывали в зените своей славы в восьмидесятых годах XX столетия. Однако научные рамки, используемые в попытках разработки искусственного интеллекта и нейронных сетей, оказались непригодными для создания по-настоящему разумных машин.

Я предлагаю другой, гораздо более многообещающий путь, которому стоит последовать. Если вы, мой читатель, сейчас проходите обучение в университете или колледже и эта книга вдохновляет вас на то, чтобы поработать над созданием новой технологии и сконструировать разумную машину, – решайтесь. Основа предпринимательского успеха состоит в том, чтобы очертя голову первым броситься в область, когда еще никому не известно наверняка, окажется ли она успешной. Важно сделать это вовремя. Поспешите – и вам придется бороться. Задержавшись, вы рискуете дожидаться момента, когда неопределенность спадет, – и будет слишком поздно. Я глубоко убежден, что именно сейчас настало время разработать и создать системы памяти, построенные по принципу функционирования коры головного мозга. Это очень многообещающая сфера как в научном, так и в коммерческом плане. Возможно, что фирмы-гиганты новой индустрии, основанной на воссозданных в силиконе иерархических системах человеческой памяти, появятся уже в ближайшие десять лет. Конечно, начало нового бизнеса такого масштаба сопряжено с большим финансовым и интеллектуальным риском, но я убежден, что этот шаг следует сделать. Я надеюсь, что вы присоединитесь ко мне и моим единомышленникам в попытке создания одной из наиболее выдающихся технологий, когда-либо существовавших на нашей планете.

Библиография



Большинство научных книг и журнальных статей включают обширные библиографии, которые не только помогают читателям глубже разобраться в теме, но и каталогизируют работы других авторов в этой области.

Адресовав эту книгу широкому кругу читателей, включая тех, у кого отсутствуют специальные знания нейробиологии, я отказался от академического стиля изложения материала. Аналогично, библиография призвана помочь тем, кто захочет углубить свои знания. Я не стал перечислять здесь все известные мне материалы, касающиеся разработок в области искусственного интеллекта и не упоминаю всех ученых, внесших значительный вклад в эту сферу науки. Вместо этого я привожу ссылки на книги и другие опубликованные материалы, которые пригодятся читателям, интересующимся человеческим мозгом. Лишь несколько ссылок адресованы специалистам в нейробиологии и смежных дисциплинах.

Много тематических обсуждений вы можете найти в Интернете, в частности, на Web-сайте, посвященном этой книге: [www. Onintelligence. org](http://www.Onintelligence.org).

К сожалению, вы сможете найти всего лишь несколько ссылок на исследования, в которых делаются попытки разработать единую теорию мозга. Как я упоминал в прологе, целостная теория о разуме и строении головного мозга пока не разработана.

История искусственного интеллекта и нейронные сети

Baumgartner, Peter, and Sabine Payr, eds. *Speaking Minds: Interviews with Twenty Eminent Cognitive Scientists* (Princeton, N. J.: Princeton University Press, 1995).

В этой книге приведены интервью с известными учеными, работающими в сфере искусственного интеллекта, нейронных сетей и когнитивной науки. Это интересный и доступно написанный обзор истории развития данной области науки и последних веяний в ней.

- Dreyfus, Hubert L. *What Computers Still Can't Do: A Critique of Artificial Reason* (Cambridge, Mass.: MIT Press, 1992).

Книга, содержащая яркую критику идеи искусственного интеллекта, впервые была опубликована под названием *What Computers Can't Do*, а впоследствии переиздана под другим названием. Это исторический обзор разработок в сфере искусственного интеллекта, написанный одним из самых жестких его критиков.

- Anderson, James A., and Edward Rosenfeld, eds. *Neurocomputing, Foundations of Research* (Cambridge, Mass.: MIT Press, 1988).

Книга представляет собой крупную подборку важных материалов по нейронным сетям и теориям мозга. Исследования, упоминающиеся в этой книге, проводились в промежутке между 1890 и 1987 гг. и представлены в хронологическом порядке. Содержатся описания исследований таких ученых, как W. S. McCulloch и W. Pitts, Donald Hebb, Steve Grossberg, а также многих других. Каждый материал предваряет вступление авторов. Эта книга

предоставляет вам уникальный шанс ознакомиться с множеством нейрологических исследований.

- Searle, J. R. «Minds, Brains, and Programs», *The Behavioral and Brain Sciences*, vol. 3 (1980): pp. 417–424.

В книге представлена идея знаменитой «Китайской комнаты» и оспаривается такой сомнительный критерий наличия интеллекта, как умение производить подсчеты. Множество отчетов и дискуссий на эту тему можно найти в Интернете.

- Turing, A. M. «Computing Machinery and Intelligence», *Mind*, vol. 59 (1950): pp. 433–60.

Представлен знаменитый тест Тьюринга, цель которого – определение наличия интеллекта. Как и в предыдущем случае, множество отчетов и дискуссий на эту тему можно найти в Интернете.

- Palm, Gunther. *Neural Assemblies: An Alternative Approach to Artificial Intelligence* (New York: Springer Verlag, 1982).

Чтобы разобраться в работе мозга и том, как в нем сохраняются последовательности паттернов, следует изучить автоассоциативную память. Хотя на эту тему было написано немало книг, *Neural Assemblies: An Alternative Approach to Artificial Intelligence*, на мой взгляд, наиболее обстоятельная. Автор книги, Палм, – один из пионеров в сфере исследования мозга. Книгу нелегко найти и непросто читать, но она прекрасно раскрывает основы автоассоциативной памяти.

Неокортекс и общая нейрология

Тем, кто интересуется работой неокортекса и нейрологией в целом, рекомендую такие книги.

- Crick, Francis H. C. «Thinking about the Brain», *Scientific American*, vol. 241 (September 1979): pp. 181–188. Also available in *The Brain: A Scientific American Book* (San Francisco: W. H. Freeman, 1979).

Это работа, которая в свое время послужила для меня толчком к изучению головного мозга. Написанная четверть века назад, она и по сей день не потеряла актуальность.

- Koch, Christof. *Quest for Consciousness: A Neurobiological Approach* (Denver, Colo.: Roberts and Co., 2004).

В этой книге рассматривается феномен осознанности, но также затрагивается много смежных тем, таких как анатомия и физиология нервной системы. Если вы желаете получить из одного источника общее представление о нейробиологии и работе мозга, прочтите эту книгу.

- Mountcastle, Vernon B. *Perceptual Neuroscience: The Cerebral Cortex* (Cambridge, Mass.: Harvard University Press, 1998).

Замечательная книга, целиком посвященная коре полушарий головного мозга – неокортексу. Она хорошо и понятно написана и структурирована, так что прочтение доставит вам удовольствие. Лучший источник для получения базовых знаний о неокортексе.

- Kandel, Eric R., James H. Schwartz, Thomas M. Jessell, eds. *Principles of Neural Science*, 4th ed. (New York: McGraw-Hill, 2000).

Это настоящая энциклопедия в одном томе обо всем, что касается нервной системы. Она не подойдет в качестве легкого вечернего чтения, но советую приобрести ее для своей библиотеки. Подробно описаны все части нервной системы, включая нейроны, органы чувств и нейротрансмиттеры.

- Shepherd, Gordon M., ed. *The Synaptic Organization of the Brain*, 5th ed. (New York: Oxford University Press, 2004).

Я считаю эту книгу очень ценной, хоть предпочитаю более ранние издания, в которых был только один автор. В нее включен подробный обзор всего, что касается головного мозга, в том числе и синапсов. Я использую эту книгу в качестве справочника.

- Koch, Christof, and Joel L. Davis, eds. *Large-scale Neuronal Theories of the Brain* (Cambridge, Mass.: MIT Press, 1994).

Как я не раз упоминал, попытки создания цельной теории головного мозга проводились считанные разы. В этой книге рассказывается обо всех научных изысканиях на эту тему, хотя

ни одному из них нельзя дать название всей книги. Прочитав ее, вы узнаете об ученых, посвятивших свою жизнь изучению мозга, и о подходах, которых они придерживались на пути к достижению общей цели. Также в книге представлена модель «Память-предсказание».

• Braitenberg, Valentino, and Almut Schbz. *Cortex: Statistics and Geometry of Neuronal Connectivity*, 2nd ed. (New York: Springer Verlag, 1998).

В этой книге статистически описаны характеристики мозга мыши. Хотя подобное чтение не всем придется по душе, я считаю книгу весьма полезной. История развития неокортекса представлена в ней в цифрах.

Статьи на темы нейрологии

Здесь я перечислю статьи – первоисточники концепций, представленных мной в этой книге. Большинство из них доступны только в университетских библиотеках и в Интернете.

• Mountcastle, Vernon B. «An Organizing Principle for Cerebral Function: The Unit Model and the Distributed System», in Gerald M. Edelman and Vernon B. Mountcastle, eds., *The Mindful Brain* (Cambridge, Mass.: MIT Press, 1978).

Прочитав эту статью Маунткастла, я познакомился с его предположением о едином принципе работы неокортекса и о кортикальных колонках. Оба допущения вдохновили меня на работу над этой книгой.

• Creutzfeldt, Otto D. «Generality of the Functional Structure of the Neocortex», *Naturwissenschaften*, vol. 64 (1977): pp. 507–17.

Я прочел эту статью уже после окончания работы над книгой *Об интеллекте*. Как и в работах Маунткастла, в ней исследуется единый кортикальный алгоритм.

• Felleman, D. J., and D. C. Van Essen. «Distributed Hierarchical Processing in the Primate Cerebral Cortex», *Cerebral Cortex*, vol. 1 (January/February 1991): pp. 1–47.

Эта работа описывает иерархический механизм работы зрительной зоны неокортекса. Модель «Память-предсказание» основана на том, что не только зрительная зона, но и весь неокортекс функционирует по иерархическому принципу.

• Sherman, S.M., and R. W. Guillery. «The Role of the Thalamus in the Flow of Information to the Cortex», *Philosophical Transactions of the Royal Society of London*, vol. 357, no. 1428 (2002): pp. 1695–1708.

Обзор таламической организации и опровержение теории Шермана-Гиллери, согласно которой таламус служит передатчиком информации между разными отделами головного мозга. Я рассматриваю этот вопрос в главе 6, в разделе «Альтернативный путь вверх по иерархии».

• Rao, R. P., and D. H. Ballard. «Predictive Coding in the Visual Cortex: A Functional Interpretation of Some Extra-Classical Receptive-field Effects», *Nature Neuroscience*, vol. 2, no. 1 (1999): pp. 79–87.

Это отчет об одном из новых исследований функции предсказания и иерархического принципа. В нем представлена модель обратной связи в иерархически связанных зонах головного мозга, согласно которой нейроны высших зон постоянно предсказывают активность нейронов низших зон.

• Guillery, R. W. «Branching Thalamic Afferents Link Action and Perception» *Journal of Neurophysiology*, vol. 90 (2003): pp. 539–548.

• Young, M. P. «The Organization of Neural Systems in the Primate Cerebral Cortex», *Proceedings of the Royal Society: Biological Sciences*, vol. 252 (1993): pp. 13–18.

В двух этих отчетах представлены доказательства тесной взаимосвязи моторного поведения и сенсорного восприятия. Более того, они рассматриваются как части единого процесса. Гиллери утверждает, что все сенсорные кортикальные зоны принимают участие в формировании моторного поведения, а Янг доказывает, что моторные и соматосенсорные зоны связаны столь тесно, что их следует рассматривать как единую систему. Я кратко рассматриваю эти идеи в главе 6.

