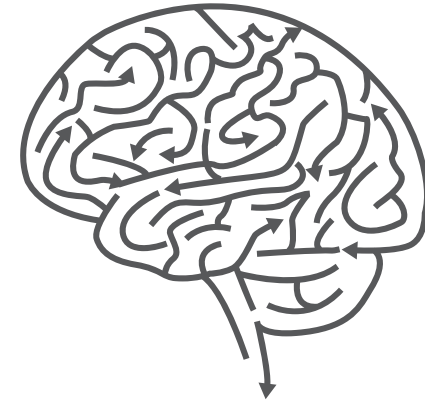


Стивен Пинкер



КАК РАБОТАЕТ МОЗГ



КУЧКОВО ПОЛЕ

Москва

2017

УДК 159.9
ББК 88
П32

Пинкер Стивен

П32 Как работает мозг / Пер. с англ. О. Ю. Семиной. — М.: Кучково поле, 2017. — 672 с.: ил.

ISBN 978-5-9950-0712-8

Книга Стивена Пинкера, выдающегося ученого, специализирующегося в экспериментальной психологии и когнитивных науках, в увлекательной и доступной форме рассматривает, какие факторы воздействуют на сознание человека, как формируется интеллект взрослого человека и чем он отличается от детского, чем человеческое мышление отличается от мышления животных вообще и приматов в частности, особенности искусственного разума.

Пинкер поддерживает идею о комплексной человеческой природе, состоящей из множества адаптивных способностей разума. Он считает, что человеческое сознание работает отчасти при помощи комбинаторной манипуляции символами.

УДК 159.9
ББК 88

ISBN 978-5-9950-0712-8

© Steven Pinker, 1997
© ООО «Кучково поле», издание
на русском языке, 2017

Предисловие

Любая книга с таким названием — «Как работает мозг» — непременно должна начинаться с пояснения, а я начну сразу с двух. Во-первых, мы ничего не знаем о том, как работает мозг, — мы не знаем и сотой доли того, что нам известно о работе тела, и уж подавно наших знаний недостаточно для того, чтобы создать утопию или излечить от несчастья. Зачем же тогда я выбрал такое амбициозное название? Лингвист Ноам Хомский как-то сказал, что наше незнание можно разделить на проблемы и тайны. Если мы сталкиваемся с проблемой, то можем не знать ее решения, но у нас есть интуитивные догадки, постепенно накапливающиеся знания и хотя бы примерное представление о том, что мы ищем. Но когда мы сталкиваемся с тайной, нам остается только смотреть на нее в изумлении и недоумении, даже не представляя, каким может быть ее объяснение. Я написал эту книгу потому, что десятки тайн нашего мышления — от ментальных образов до романтической любви — в последнее время были переведены в разряд проблем (хотя нужно сказать, что некоторые тайны остались тайнами!). Вполне вероятно, что все идеи до одной в этой книге окажутся неверными, но и это будет прогрессом, потому что наши старые идеи были слишком банальными для того, чтобы оказаться неверными. Во-вторых, то небольшое, что мы все-таки знаем о работе нашего мозга, открыл не я. Лишь немногие из идей в этой книге принадлежат мне. Из целого ряда дисциплин я выбрал теории, которые, как мне показалось, дают возможность понять наши мысли и чувства с новой стороны; теории, которые соответствуют фактам, позволяют прогнозировать новые факты и при этом отличаются последовательностью в содержании и стиле изложения. Моей целью было вплести эти теории в единую канву, используя две более масштабных теории (тоже не мои): вычислительную теорию сознания и теорию естественного отбора репликаторов.

В первой главе представлена картина в целом: я пишу о том, что мышление — это система органов вычисления, созданных естественным отбором для решения проблем, с которыми сталкивались наши предки, занимавшиеся собирательством и охотой. Каждой из этих двух важнейших идей — вычислению и эволюции — далее посвящается по отдельной главе. Я подробно рассматриваю ключевые способности мышления в главах, посвященных восприятию, логическому рассуждению, эмоциям и социальным отношениям (таким, как семья, любовь, соперничество, дружба, приятельство, знакомство, союзничество, вражда). В заключительной главе речь идет о таких высоких материях, как искусство, музыка, литература, юмор, религия и философия. Отдельной главы, посвященной языку, нет: я подробно рассматриваю эту тему в своей предыдущей книге «Язык как инстинкт».

Данная книга предназначена для всех, кому интересно, как работает мышление. Я написал ее не только для профессоров и студентов; однако и так называемая популяризация науки тоже не была моей единственной целью. Я надеюсь, что общая картина мышления и роли, которую оно играет в жизни человека, будет полезна как ученым, так и широкому кругу читателей. При взгляде с этой точки зрения разница между специалистом и непосвященным не имеет особого значения, потому что в наше время мы, специалисты, в большинстве наших собственных областей — не говоря уже про соседние — оказываемся не сильнее простых смертных. Я не делаю исчерпывающий обзор всей литературы и не привожу мнения всех сторон по тому или иному вопросу, потому что в противном случае книгу было бы очень сложно читать — да что там, ее и сдвинуть с места было бы сложно. Свои выводы я делаю на основе оценки сходности данных, полученных в разных научных областях и разными методами; я привожу множество ссылок, чтобы читатели могли обратиться к первоисточникам.

Я обязан в интеллектуальном плане многим моим учителям, студентам и коллегам, но больше всего — Джону Туби и Леде Космидес. Это они осуществили синтез эволюции и психологии, благодаря которому стала возможной эта книга, и это им принадлежат многие из теорий, которые я здесь представляю (и многие из самых лучших шуток). Благодаря их приглашению провести год в качестве сотрудника Центра эволюционной психологии Калифорнийского университета в Санта-Барбаре, я получил идеальную возможность думать и писать, а также их бесценную дружбу и советы.

Я глубоко признателен Майклу Газзанига, Маркусу Хаузеру, Дэвиду Кеммереру, Гэри Маркусу, Джону Туби и Марго Уилсон за то, что они прочитали всю мою рукопись, за их неоценимые рекомендации и поддержку. Другие коллеги также сделали множество полезных замечаний по главам, относящимся к их сферам компетенции: Эдвард Адельсон, Бартон Андерсон, Саймон Барон-Коэн, Нед Блок, Пол Блум, Дэвид Брейнерд, Дэвид Басс, Джон Констебл, Леда Космидес, Хелена Кронин, Дэн Деннетт, Дэвид Эпштейн, Алан Фридланд, Герд Гигеренцер, Джудит Харрис, Ричард Хелд, Рей Джекендофф, Алекс Качельник, Стивен Кослин, Джек

Лумис, Чарльз Оман, Бернард Шерман, Пол Смоленски, Элизабет Спелке, Фрэнк Саллоуэй, Дональд Саймонс и Майкл Тарр. Многие другие отвечали на вопросы и внесли ценные предложения — в том числе Роберт Бойд, Дональд Браун, Наполеон Шаньон, Мартин Дейли, Ричард Докинз, Роберт Хэдли, Джеймс Хилленбранд, Дон Хоффман, Келли Олгвин Якола, Тимоти Кетелаар, Роберт Курзбан, Дэн Монтелло, Алекс Пентланд, Рослин Пинкер, Роберт Провайн, Уитман Ричардс, Дэниел Шактер, Девендра Сингх, Паван Синха, Кристофер Тайлер, Джереми Вольф и Роберт Райт.

Эта книга — плод моего пребывания во вдохновляющей атмосфере двух учреждений: Массачусетского технологического института и Калифорнийского университета в Санта-Барбаре. Особую благодарность я выражаю Эмилио Бицци (отделение мозга и когнитологии Массачусетского технологического института) за предоставленную мне возможность взять творческий отпуск, а также Люю Лайтл и Аарону Эттенбергу (отделение психологии), Патрисии Кленси и Марианн Митун (отделение лингвистики Калифорнийского университета в Санта-Барбаре) за то, что они пригласили меня поработать в качестве внештатного научного сотрудника их отделений.

Патрисия Клаффи из библиотеки отделения мозга и когнитивистики МТИ знает все — или, по крайней мере, знает, где это найти, что в принципе одно и то же. Я благодарен ей за неутомимое упорство, с которым она находит даже самый малоизвестный документ — причем быстро и не утрачивая благодушия. Элеанор Бонсент — мой секретарь — профессионально и с оптимизмом помогала мне в решении бесчисленного количества вопросов. Выражаю благодарность также Марианн Тейбер, Сабрине Детмар и Дженнифер Риддел из центра изобразительного искусства МТИ.

Мои редакторы, Дрейк Макфили (издательство «Нортон»), Ховард Бойер (сейчас работает в издательстве Калифорнийского университета), Стефан Мак-Грат («Пингвин») и Рави Мирчандани (сейчас работает в издательстве «Орион»), на протяжении всей моей работы помогли мне профессиональными рекомендациями и заботой. Я также признателен моим агентам, Джону Брокману и Катинке Мэтсон, за все, что они сделали для меня, и за их интерес к научной прозе. Особая благодарность — Кате Райс, которая работала вместе со мной уже над четырьмя книгами (в течение четырнадцати лет). Благодаря ее аналитическому уму и профессиональному подходу мои книги стали лучше, а я узнал много нового о стиле и ясности изложения. Выражаю сердечную благодарность за поддержку и полезные предложения всей моей семье: Гарри, Рослин, Роберту и Сьюзан Пинкер, Мартину, Еве, Карлу и Эрику Будман, Сароие Суббиа и Стэну Адамсу. А еще спасибо Уиндзору, Уилфреду и Фионе.

Моя самая горячая благодарность — моей жене, Илавенил Суббиа, за эскизы рисунков, бесценные замечания к рукописи, бесчисленные советы, поддержку, доброту и участие в этом предприятии. Эту книгу я посвящаю ей, с любовью и благодарностью.

Мое исследование мышления и языка проводилось при поддержке Национального института здравоохранения (грант HD18381), Национального научного фонда (гранты 82–09540, 85–18774 и 91–09766) и Центра когнитивных нейроисследований имени МакДоннелла-Пью (Массачусетский технологический институт).



КАК РАБОТАЕТ МОЗГ

Стандартное оборудование

Почему в литературе так много роботов, а в реальной жизни — ни одного? Я бы, например, заплатил кучу денег за робота, который бы убирал за меня посуду или выполнял несложные поручения. Но в этом столетии у меня вряд ли будет такая возможность, а может быть — и в следующем тоже. Конечно, есть роботы, которые выполняют на конвейере такие операции, как сварка или покраска, и роботы, которые разъезжают по коридорам лабораторий; но я имею в виду не их, а такие машины, которые умеют ходить, говорить, видеть и думать — причем зачастую даже лучше, чем создавшие их люди. В 1920 году Карел Чапек впервые использовал слово «робот» в своей пьесе «Р. У. Р.», после чего авторы начали один за другим выдумывать собственных роботов: появились Спиди, Кьюти и Дейв из «Я, робот» Айзека Азимова, Робби из «Запретной планеты», размахивающая руками жестянка из «Затерянных в космосе», далеки из «Доктора Кто», робот-домработница Роза из «Джетсонов», Номад из «Звездного пути», Хайми из «Напряги извилины», безучастные дворецкие и переругивающиеся продавцы одежды из фильма «Спящий», R2-D2 и C-3PO из «Звездных войн», Терминатор из одноименного фильма, лейтенант-командер Дейта из фильма «Звездный путь: Следующее поколение» и остроумные кинокритики из «Таинственного театра 3000 года».

Эта книга посвящена не роботам. Она посвящена человеческому разуму. Я постараюсь объяснить, что такое разум, как он устроен, как благодаря ему мы можем видеть, думать, чувствовать, взаимодействовать и заниматься такими высокими материями, как искусство, религия и философия. Попутно я постараюсь осветить вопросы, связанные со свойственными исключительно человеку явлениями. Почему воспоминания стираются из памяти? Как макияж меняет лицо? Откуда берутся этнические стереотипы и когда они оказываются необоснованными? Почему люди выходят из себя? Что делает детей непослушными? Почему слабоумные влюбляются? Что заставляет нас смеяться? Почему люди верят в духов и привидений?

И все же начну я именно с различия между вымышленными и реальными роботами, поскольку оно показывает, каким должен быть наш первый шаг в познании себя: нам нужно по достоинству оценить невероятно слож-

ную систему, позволяющую нам ежедневно осуществлять настоящие чудеса мышления, которые мы привыкли воспринимать как должное. Причина, по которой в мире нет человекоподобных роботов, — не в том, что сама по себе идея создания механического разума безрассудна, а в том, что технические задачи, которые мы, люди, ежеминутно решаем, когда смотрим, ходим, планируем свои действия, куда сложнее высадки на Луну и расшифровки генома. Природа, создав человека, в очередной раз сделала гениальное изобретение, которое инженерам повторить пока не по силам. Когда мы вслед за Гамлетом восклицаем: «Что за мастерское создание — человек! Как благороден разумом! Как беспределен в своих способностях, обличьях и движениях! Как точен и чудесен в действии!»*, то объектом нашего восхищения является не Шекспир, не Моцарт, не Эйнштейн и не Карим Абдул-Джаббар, а четырехлетний ребенок, который по просьбе взрослого кладет игрушку на полку.

В хорошо продуманной системе все элементы устроены по принципу «черного ящика»: они выполняют свои функции как по волшебству. То же самое можно сказать и о нашем мышлении. Мышление, благодаря которому мы познаем мир, лишено возможности обратить взор внутрь себя или увидеть внутренние механизмы, приводящие в движение другие наши способности. В результате мы оказываемся жертвами иллюзии: нам кажется, что наша собственная психика является порождением некоей Божественной силы, мистической сущности или высшего начала. В еврейской мифологии есть легенда о Големе — статуе из глины, которая ожила, когда ей в рот вложили записку с именем Бога. Этот архетип прослеживается в сюжетах многих историй. Статую Галатеи оживила Венера в ответ на молитвы Пигмалиона; Пиноккио обрел жизнь благодаря волшебству Синей Феи. Современные версии архетипа Голема мы находим не только в сказках, но и в научных теориях. Считается, что все явления в психологии человека — размер его мозга, культура, язык, социализация, обучение, познание, самоорганизация, динамика нейронной сети — могут быть сведены к одной первопричине.

Я хочу убедить вас в том, что наш разум не был создан неким Божественным духом или чудесным первоисточником. Наш мозг, как космический корабль «Аполлон», сконструирован для того, чтобы решать самые разные технические задачи, поэтому он напичкан самыми высокотехнологичными системами, каждая из которых нацелена на преодоление отдельного типа проблем. Я начну с краткого описания этих проблем, которые одновременно являются предметом изучения психологии и техническими требованиями к роботу. Я убежден, что полученные исследователями мышления и искусственного интеллекта сведения о том, какие технические головоломки приходится постоянно решать нашему мозгу в повседневной жизни, являются одним из вели-

* Цитируется в переводе М. Лозинского.

чайших научных открытий. Это откровение сравнимо по своей значимости с осознанием того, что Вселенная состоит из миллиардов галактик, или того, что в капле воды из пруда кишмя кишат микроорганизмы.

Как сделать робота

Что нужно, чтобы сделать робота? Давайте обойдемся без сверхчеловеческих способностей вроде умения вычислять орбиты планет и начнем с элементарных для человека навыков: умения видеть, ходить, хватать, думать о предметах и людях, планировать свои действия.

В фильмах нам часто показывают происходящее глазами робота, используя привычные для кино эффекты: например, вид через линзу «рыбий глаз» или через перекрестие прицела. Для нас, зрителей, у которых есть глаза и мозг, этого вполне достаточно. Но вот самому роботу такой способ восприятия реальности едва ли пригодится. У него внутри нет кинозала, полного маленьких человечков, которые бы смотрели на мир и рассказывали роботу, что они видят. Если бы мы могли видеть мир глазами робота, он бы выглядел вовсе не как кадр из фильма, украшенный крестиком прицела, а как что-то вроде этого¹:

```

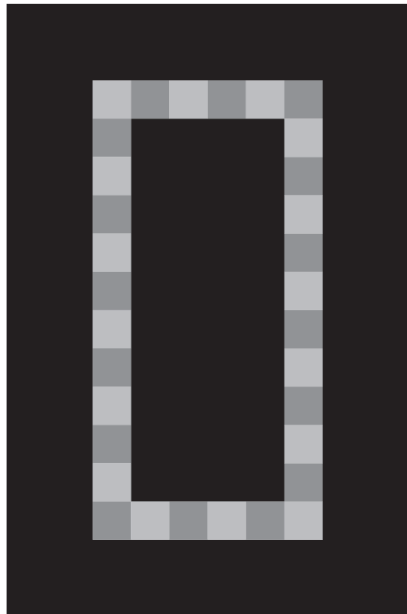
2 2 5 2 2 1 2 1 6 2 1 9 2 1 9 2 1 4 2 0 7 2 1 8 2 1 9 2 2 0 2 0 7 1 5 5 1 3 6 1 3 5
2 1 3 2 0 6 2 1 3 2 2 3 2 0 8 2 1 7 2 2 3 2 2 1 2 2 3 2 1 6 1 9 5 1 5 6 1 4 1 1 3 0
2 0 6 2 1 7 2 1 0 2 1 6 2 2 4 2 2 3 2 2 8 2 3 0 2 3 4 2 1 6 2 0 7 1 5 7 1 3 6 1 3 2
2 1 1 2 1 3 2 2 1 2 2 3 2 2 0 2 2 2 2 3 7 2 1 6 2 1 9 2 2 0 1 7 6 1 4 9 1 3 7 1 3 2
2 2 1 2 2 9 2 1 8 2 3 0 2 2 8 2 1 4 2 1 3 2 0 9 1 9 8 2 2 4 1 6 1 1 4 0 1 3 3 1 2 7
2 2 0 2 1 9 2 2 4 2 2 0 2 1 9 2 1 5 2 1 5 2 0 6 2 0 6 2 2 1 1 5 9 1 4 3 1 3 3 1 3 1
2 2 1 2 1 5 2 1 1 2 1 4 2 2 0 2 1 8 2 2 1 2 1 2 2 1 8 2 0 4 1 4 8 1 4 1 1 3 1 1 3 0
2 1 4 2 1 1 2 1 1 2 1 8 2 1 4 2 2 0 2 2 6 2 1 6 2 2 3 2 0 9 1 4 3 1 4 1 1 4 1 1 2 4
2 1 1 2 0 8 2 2 3 2 1 3 2 1 6 2 2 6 2 3 1 2 3 0 2 4 1 1 9 9 1 5 3 1 4 1 1 3 6 1 2 5
2 0 0 2 2 4 2 1 9 2 1 5 2 1 7 2 2 4 2 3 2 2 4 1 2 4 0 2 1 1 1 5 0 1 3 9 1 2 8 1 3 2
2 0 4 2 0 6 2 0 8 2 0 5 2 3 3 2 4 1 2 4 1 2 5 2 2 4 2 1 9 2 1 5 1 1 4 1 1 3 3 1 3 0
2 0 0 2 0 5 2 0 1 2 1 6 2 3 2 2 4 8 2 5 5 2 4 6 2 3 1 2 1 0 1 4 9 1 4 1 1 3 2 1 2 6
1 9 1 1 9 4 2 0 9 2 3 8 2 4 5 2 5 5 2 4 9 2 3 5 2 3 8 1 9 7 1 4 6 1 3 9 1 3 0 1 3 2
1 8 9 1 9 9 2 0 0 2 2 7 2 3 9 2 3 7 2 3 5 2 3 6 2 4 7 1 9 2 1 4 5 1 4 2 1 2 4 1 3 3
1 9 8 1 9 6 2 0 9 2 1 1 2 1 0 2 1 5 2 3 6 2 4 0 2 3 2 1 7 7 1 4 2 1 3 7 1 3 5 1 2 4
1 9 8 2 0 3 2 0 5 2 0 8 2 1 1 2 2 4 2 2 6 2 4 0 2 1 0 1 6 0 1 3 9 1 3 2 1 2 9 1 3 0
2 1 6 2 0 9 2 1 4 2 2 0 2 1 0 2 3 1 2 4 5 2 1 9 1 6 9 1 4 3 1 4 8 1 2 9 1 2 8 1 3 6
2 1 1 2 1 0 2 1 7 2 1 8 2 1 4 2 2 7 2 4 4 2 2 1 1 6 2 1 4 0 1 3 9 1 2 9 1 3 3 1 3 1
2 1 5 2 1 0 2 1 6 2 1 6 2 0 9 2 2 0 2 4 8 2 0 0 1 5 6 1 3 9 1 3 1 1 2 9 1 3 9 1 2 8
2 1 9 2 2 0 2 1 1 2 0 8 2 0 5 2 0 9 2 4 0 2 1 7 1 5 4 1 4 1 1 2 7 1 3 0 1 2 4 1 4 2

```

2 2 9 2 2 4 2 1 2 2 1 4 2 2 0 2 2 9 2 3 4 2 0 8 1 5 1 1 4 5 1 2 8 1 2 8 1 4 2 1 2 2
 2 5 2 2 2 4 2 2 2 2 2 4 2 3 3 2 4 4 2 2 8 2 1 3 1 4 3 1 4 1 1 3 5 1 2 8 1 3 1 1 2 9
 2 5 5 2 3 5 2 3 0 2 4 9 2 5 3 2 4 0 2 2 8 1 9 3 1 4 7 1 3 9 1 3 2 1 2 8 1 3 6 1 2 5
 2 5 0 2 4 5 2 3 8 2 4 5 2 4 6 2 3 5 2 3 5 1 9 0 1 3 9 1 3 6 1 3 4 1 3 5 1 2 6 1 3 0
 2 4 0 2 3 8 2 3 3 2 3 2 2 3 5 2 5 5 2 4 6 1 6 8 1 5 6 1 4 4 1 2 9 1 2 7 1 3 6 1 3 4

Каждое число здесь соответствует яркости одного из миллиона крохотных участков, вместе составляющих поле зрения. Числа поменьше соответствуют более темным участкам, а числа побольше — светлым участкам. Эти числа, сгруппированные в виде массива, на самом деле представляют собой сигналы, полученные с цифровой камеры, с помощью которой для пробы сделали снимок руки человека. Однако точно таким же образом можно было бы представить частоту разрядов в нервных волокнах, соединяющих глаз человека с мозгом в тот момент, когда человек смотрит на руку². Для того чтобы мозг робота — и мозг человека тоже — сумел распознать предметы и не позволил своему владельцу с ними столкнуться, ему нужно обработать эти числа и отгадать, какой предмет в реальном мире мог отразить свет таким образом, чтобы получилась подобная последовательность чисел. Эта задача просто невообразимо сложна.

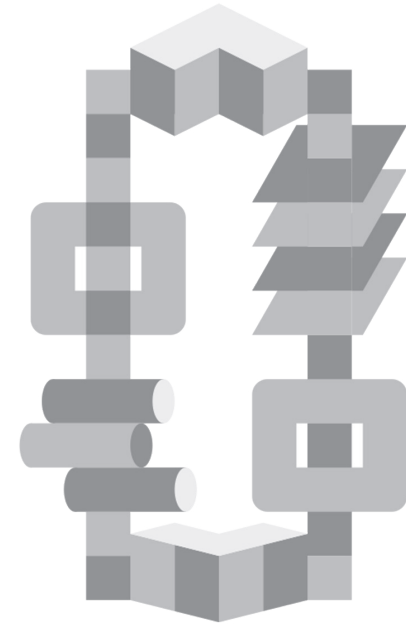
В первую очередь зрительная система должна определить, где заканчивается граница предмета и начинается фон. Но ведь мир — не детская раскраска, где покрашенные фигуры очерчены черным контуром. Мир в том виде, в котором он отражается в наших глазах, представляет собой мозаику из крохотных участков разной степени освещенности. Можно предположить,



что зрительные структуры мозга ищут области, где мозаика из больших чисел (более светлый участок) граничит с мозаикой из меньших чисел (более темный участок). Такую границу можно различить и в нашей сетке из чисел: она проходит по диагонали из правого верхнего угла к центру нижней части. К сожалению, в большинстве случаев обнаружить границу предмета с пустым пространством будет непросто. Участки, где большие числа граничат с малыми, могут появляться как следствие очень разных вариантов взаимного расположения объектов. Рисунок, приведенный ниже, предложили психологи Паван Синха и Эдвард Адельсон³. Кажется, что на нем изображено кольцо из светло-серых и темно-серых квадратиков.

На самом деле перед нами прямоугольная прорезь в черной маске, через которую мы видим лишь часть изображения. На следующем рисунке маска удалена, и мы можем видеть, что все пары серых квадратиков являются частями разных объектов.

Большие числа рядом с малыми могут появляться и в том случае, когда один предмет располагается перед другим. Примерами могут служить темная бумага, лежащая поверх светлой, поверхность, окрашенная в два разных оттенка серого цвета, два предмета, стоящие рядом вплотную друг к другу, серый целлофан, наложенный на белую страницу, угол, соединяющий две стены, а также любая тень. Мозг должен как-то решить эту проблему курицы и яйца: выделить трехмерные объекты из множества затененных участков на сетчатке и определить, что представляет из себя каждый участок (тень или краску, складку или наложение объектов, прозрачную или матовую поверхность), исходя из того, частью какого объекта является этот участок.



Однако на этом сложности не заканчиваются. Разделив визуально воспринимаемый мир на объекты, мы должны определить, из чего они сделаны, например, из снега или из угля. На первый взгляд эта задача представляется простой. Если большие числа соответствуют светлым участкам, а малые — темным, то большие числа будут означать снег, а малые — уголь, правильно? Нет. Количество света, попадающего на сетчатку глаза, зависит не только от того, насколько светлый или темный перед нами предмет, но и от того, насколько яркий или тусклый свет его освещает. Используемый фотоаппаратами прибор — экспонометр — покажет, что кусок угля на улице отражает больше света, чем комок снега в помещении. Именно поэтому люди нередко оказываются недовольны качеством своих снимков, и именно поэтому фотография — очень непростое ремесло. Фотоаппарат никогда не лжет; если предоставить ему возможность действовать без помощи человека, то все снимки, сделанные на улице, будут выглядеть как молоко, а сделанные в помещении — как грязь. Фотографы (а иногда и микропроцессоры, встроенные в фотоаппарат) добиваются от фотопленки реалистичного изображения, прибегая к таким приемам как регулировка скорости срабатывания затвора и диафрагмы объектива, светочувствительность пленки, вспышка и манипуляции в лаборатории.

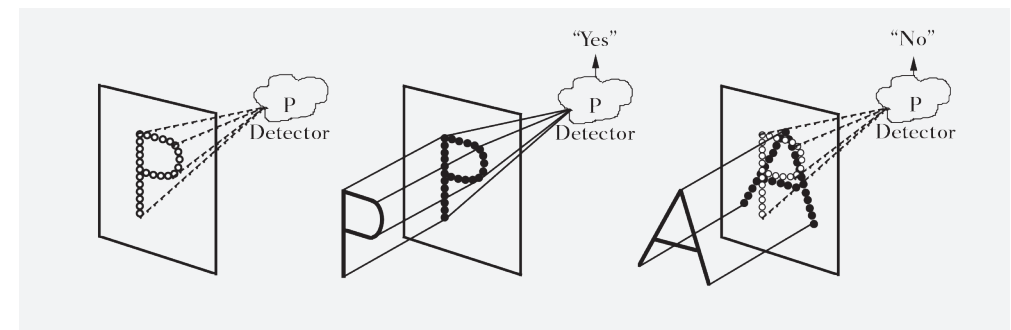
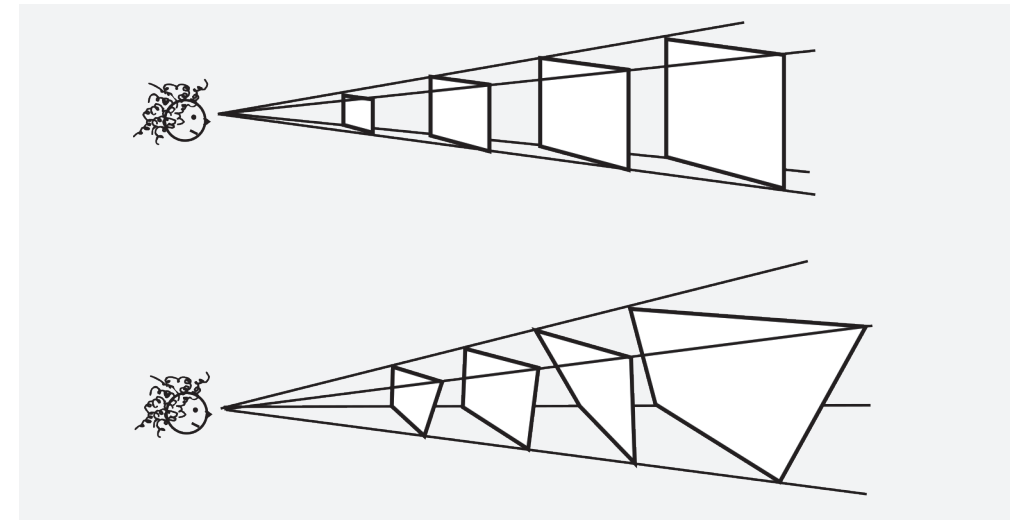
Наша система зрительного восприятия справляется с этой задачей гораздо лучше. Она каким-то образом позволяет нам видеть, что блестящий

на солнце уголь на самом деле черный, а кажущийся темным в помещении снег — белый. Это очень удобно, поскольку наше представление о цвете и свете совпадает с тем, каким мир является в действительности, а не с тем, каким он предстает перед глазами. Снег — объект мягкий, влажный и способный таять как в помещении, так и снаружи, и мы видим, что он остается белым независимо от того, в помещении он или снаружи. Уголь всегда твердый, маркий, способный гореть, и мы всегда видим его черным. Получившейся гармонией между тем, как выглядит окружающий мир, и тем, каким он является на самом деле, мы обязаны магии своего мозга: ведь сетчатка глаза не может сама по себе воспринимать черный и белый цвета. А если у вас все еще есть сомнения, вот пример из повседневной жизни. Когда телевизор выключен, экран имеет бледный зеленовато-серый оттенок. Когда телевизор включают, некоторые люминофорные точки начинают светиться, образуя светлые участки изображения, а некоторые другие точки не светятся, образуя темные участки; они остаются серыми. Те части изображения, которые мы видим как черные, на самом деле — все тот же бледно-серый оттенок кинескопа выключенного телевизора. Черный цвет — это фикция, результат работы мозга по той же схеме, благодаря которой мы видим, что уголь — это уголь. Именно эту схему работы использовали создатели телевидения, когда разрабатывали телеэкран.

Следующая проблема зрительного восприятия — это глубина. Наши глаза разбивают трехмерный мир на два двумерных изображения на сетчатке глаз, а третье измерение должен воссоздать мозг. Однако в изображении на сетчатке нет никаких подсказок относительно того, насколько далеко от смотрящего расположен объект. Почтовая марка на вашей ладони оставит на сетчатке такое же квадратное изображение, как стул, стоящий у противоположной стены, или здание, расположенное за много километров от вас (первый рисунок на с. 17). Разделочная доска, если на нее смотреть фронтально, может давать такую же трапециевидную проекцию, как и фигуры неправильной формы, расположенные под углом (второй рисунок на с. 17).

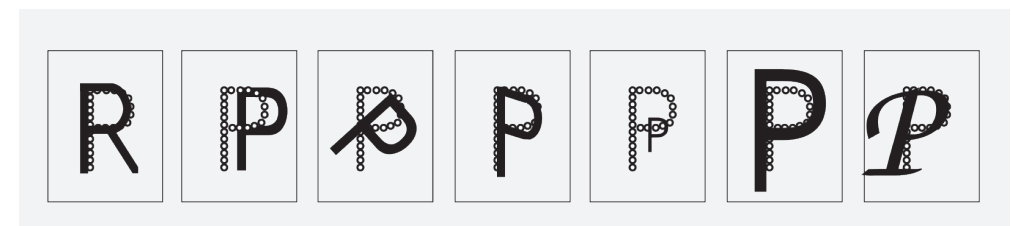
Ближе вы можете познакомиться с этим явлением из области геометрии и с нейронным механизмом, который с ним работает, если посмотрите в течение нескольких секунд на электрическую лампочку или на фотоаппарат в момент вспышки: в результате на вашей сетчатке останется световое пятно. Теперь переведите взгляд на страницу книги; у вас перед глазами по-прежнему будет остаточное изображение лампочки шириной 3–5 см. Если вы посмотрите на стену, то остаточное изображение будет около метра в длину. А если посмотрите на небо, то это изображение будет размером с облако⁴

Наконец, как заставить модуль зрительного восприятия распознавать предметы, чтобы робот мог назвать их или вспомнить их назначение? Очевидным решением было бы построить для каждого объекта маску или шаблон, в точности повторяющий его форму. В таком случае при появлении объекта в поле зрения его проекция на сетчатке идеально совпадала бы с ша-



блоном. Шаблому присваивается метка с именем данной формы — в нашем случае это метка «Р», — и каждый раз, когда проекция совпадает с шаблоном, система выдает ее имя⁵:

Увы, это простое устройство допускает ошибки в обоих возможных случаях. С одной стороны, оно видит букву Р там, где ее нет: например, реагирует на букву R (первый вариант слева на рисунке внизу). С другой стороны, оно не замечает букву Р там, где она есть: например, если она смещена, наклонена, неровно написана, расположена слишком далеко или слишком близко, либо написана слишком затейливым шрифтом:



Столько проблем с распознаванием одной простой и понятной буквы алфавита! А теперь попробуйте представить, как сложно сконструировать устройство, способное распознать рубашку или, скажем, лицо! Конечно, сейчас, после сорока лет исследований в области искусственного интеллекта, технологии распознавания формы значительно улучшились. Вероятно, и у вас есть программное обеспечение, которое может довольно сносно распознать отсканированную страницу и преобразовать ее в цифровой файл. И все же искусственные средства распознавания формы по-прежнему не могут конкурировать с теми, что существуют у нас в голове. Они созданы для среды, где все безупречно и легко узнаваемо, а не для нашего сумбурного, беспорядочного реального мира. Странные цифры, напечатанные в нижней части чека, специально придуманы таким образом, чтобы не иметь даже частичных наложений; специальное оборудование для печати располагает их так, чтобы обеспечить распознавание по заданным шаблонам. Когда в подъездах вместо консьержек появятся первые устройства, распознающие лицо, они даже не будут пытаться разгадать хитросплетения света и тени на вашем лице; они будут сканировать четко очерченные контуры радужной оболочки или кровеносные сосуды сетчатки. А вот наш мозг регистрирует форму каждого знакомого нам лица (а также каждой буквы, животного, инструмента и т. д.) и каким-то образом устанавливает ее соответствие с изображением на сетчатке, даже если это изображение искажено любым из рассмотренных выше способов. В главе 4 мы рассмотрим, как мозгу удастся добиться таких потрясающих результатов.



Давайте поговорим еще об одном обыкновенном чуде: о том, как наше тело перемещается с одного места на другое. Чтобы заставить машину двигаться, мы поставили ее на колеса. Изобретение колеса нередко называют одним из величайших достижений цивилизации. Во многих учебниках указывается на то, что ни у одного животного в процессе эволюции не сформировались колеса; этот факт называют примером того, что эволюция зачастую неспособна предложить оптимальное решение технической задачи. И все же это неудачный пример. Если бы природа и могла создать лося на колесах, она бы наверняка этой возможностью не воспользовалась. Колеса нужны только в том мире, где есть автомобильные и железные дороги. Они завязнут в мягкой, скользкой или неровной почве. Ноги гораздо лучше. Колеса могут катиться только по непрерывной опоре, в то время как ноги могут наступать поочередно на ряд отдельных опор — наиболее ярким примером этого является лестница. Ноги позволяют своему владельцу поддерживать равновесие и переступать через препятствия. И сегодня, когда наш мир, кажется, превратился в сплошную

автопарковку, транспортным средствам с колесами или гусеницами доступна лишь половина земной поверхности. Но есть и другие транспортные средства, которым доступна практически вся поверхность планеты; транспортные средства, созданные путем естественного отбора и передвигающиеся с помощью ног: животные.

Правда, ноги просто так не даются: к ним должна прилагаться управляющая ими программа. Колесо, поворачиваясь, изменяет свою точку опоры постепенно и за счет этого может всегда нести вес. Нога должна менять точку опоры сразу, а для этого нужно разгрузить ее. Двигательные нервы, контролирующие движение ноги, должны попеременно то удерживать ногу на земле, когда она несет вес тела, то снимать нагрузку, чтобы нога свободно двигалась. Все это время они должны удерживать центр тяжести тела в пределах многоугольника, определенного ступнями, чтобы тело не опрокинулось. Кроме того, эти органы управления должны минимизировать затратные движения тела вверх-вниз, которые приносят так много неудобств во время верховой езды⁶. В случае с шагающими заводными игрушками эти проблемы решаются довольно банально: с помощью механического сцепления, которое преобразует движение вращающего вала в шаговое движение. Однако игрушки не могут приспособиваться к особенностям поверхности, находя идеальную точку опоры.

Даже если бы мы решили эти проблемы, мы были бы способны контролировать процесс ходьбы только у насекомого. Насекомое, у которого шесть лапок, поднимая три из них, всегда может держать остальные три на земле. При этом оно в любой момент движения сохраняет устойчивость. Даже четвероногое животное, если оно передвигается не слишком быстро, может все время опираться на землю тремя ногами. Как сказал один инженер, «сама идея передвижения человека вертикально и на двух ногах кажется верным рецептом катастрофы, и ее осуществление на практике требует поразительного контроля над собой»⁷. Когда мы ходим, мы во время каждого шага на мгновение падаем и тут же останавливаем свое падение. Когда мы бежим, мы на время отрываемся от земли и летим. Эти фигуры высшего пилотажа позволяют нам ставить ноги на точки опоры, размещенные далеко друг от друга или на разных расстояниях и неспособные удерживать наше тело в состоянии покоя, а также проходить по узким тропинкам и перепрыгивать через препятствия. Однако никому еще не удалось точно установить, как мы это делаем.

Управление рукой — тоже задача не из легких. Возьмитесь рукой за головку настольной лампы и переместите его по прямой диагонали из положения внизу слева от себя в положение вверху справа. Двигая лампу, обратите внимание на составляющие ее штанги и шарниры. Вы двигаете абжур по прямой линии, однако каждая штанга при этом совершает сложное движение по кривой, в одни моменты оставаясь практически неподвижной, в другие — резко опускаясь, в третьи — вместо сгибательного движения совершая

разгибательное. А теперь представьте, что вам нужно повторить это движение в обратном порядке, не глядя на головку, то есть скоординировать ряд движений в каждом месте соединения таким образом, чтобы головка перемещалась по прямой. Тригонометрия этого действия пугающе сложна. А ведь ваша рука — не лампа, и при этом ваш мозг без труда решает эту задачу каждый раз, когда вам нужно на что-то указать. Если же вы когда-нибудь держали такую лампу за зажим, которым она крепится к столу, то вы понимаете, что проблема намного сложнее, чем я только что ее представил. Лампа под собственным весом начинает беспорядочно двигаться, как живая; точно так же двигались бы ваши руки, если бы мозг не компенсировал их вес, ежеминутно решая невообразимо сложные задачи из области физики⁸.

Еще более поразителен процесс управления кистью руки. Почти две тысячи лет назад греческий врач Гален отмечал, насколько сложный процесс конструирования понадобился природе, чтобы создать человеческую кисть. Это единственный инструмент, способный манипулировать объектами, отличающимися на удивление широким разнообразием размеров, форм и веса — от бревна до крупинки проса. «Человек так хорошо берет в руки эти предметы, — отмечал Гален, — что всем покажется, будто руки созданы специально для каждого из них, взятого в отдельности»⁹. Руку можно сложить в форме крюка (чтобы поднять ведро), ножниц (чтобы держать сигарету), пятикулачкового зажимного патрона (чтобы поднять подставку для стакана), трехкулачкового зажимного патрона (чтобы держать карандаш), двухкулачкового зажимного патрона (чтобы вдеть нитку в иголку), зажима, у которого губка прислоняется к боковой поверхности (чтобы повернуть ключ), захвата (чтобы взять молоток), кругового захвата (чтобы открыть банку) и шарообразного захвата (чтобы взять мячик)¹⁰. Для каждой из этих операций требуется напряжение мышц в определенном сочетании, позволяющем руке принять нужную форму и сохранять ее, сопротивляясь давлению груза, который стремится ее разогнуть. Представьте себе, что вы поднимаете пакет молока. Если сжать его недостаточно крепко, то вы его уроните, а если слишком крепко — раздавите; а если слегка покачать пакет, то, чувствуя давление на подушечки пальцев, можно даже определить, сколько в пакете осталось молока! Я даже не говорю о языке — этом бескостном куске студня, контролируемом только путем сжатия и при этом способном доставать остатки пищи из задних зубов и выполнять настоящие пируэты, необходимые для артикуляции таких слов, как «параллелепипед» и «дезоксирибонуклеиновая».



«Обыкновенный человек удивляется чудесам... Мудрый человек удивляется вещам обычным». Держа в голове это изречение Конфуция, посмотрим на по-

вседневные действия человека свежим взглядом разработчика робота, чья задача — в точности воспроизвести их. Представим, что нам каким-то образом удалось построить робота, который может видеть и двигаться. Как он поймет, что он видит? Как он решит, что ему делать?

Разумное существо не может рассматривать каждый объект, который оно видит, как уникальное явление, подобного которому нет нигде во Вселенной. Ему необходимо распределять объекты по категориям таким образом, чтобы можно было применить к конкретному предмету добытые ценой больших усилий знания, полученные в прошлом о подобных объектах.

Но как только мы пытаемся запрограммировать совокупность критериев, которая бы позволила охватить всех представителей той или иной категории, категория разрушается. Оставим в стороне такие явно сложные для определения концепты, как «красота» или «диалектический материализм», и возьмем хрестоматийный пример легко определяемого понятия: «холостяк». Понятно, что холостяк — это взрослый человек мужского пола, который никогда не был женат. Но давайте представим, что ваша знакомая просит вас пригласить на ее вечеринку несколько холостяков. Что получится, если вы попробуете с помощью этого определения решить, кого из следующих людей вам пригласить?

Артур последние пять лет счастливо живет с Элис. У них есть двухлетняя дочь. Официально они не женаты.

Брюса должны были забрать в армию, поэтому он договорился со своей подругой Барбарой заключить фиктивный брак, чтобы получить освобождение от военной службы. Они с Барбарой не прожили вместе и дня. Брюс встречается с несколькими женщинами и планирует расторгнуть брак сразу, как только найдет девушку, на которой захочет жениться.

Чарли 17 лет. Он живет с родителями и учится в школе.

Дэвиду 17 лет. Он ушел из дома в 13 и открыл небольшой бизнес. Сейчас является успешным молодым предпринимателем, живет в роскошной квартире в пентхаусе и ведет разгульный образ жизни.

Эли и Эдгар — гомосексуальная пара. Они прожили вместе много лет. Файзал по законам своего родного эмирата Абу-Даби может иметь трех жен. Сейчас у него две, и он хотел бы познакомиться с еще одной потенциальной невестой.

Отец Грегори — епископ католической церкви, служит в кафедральном соборе в городе Гротон.

Список, составленный программистом Терри Виноградом¹¹, показывает, что такое простое определение понятия «холостяк» не отражает наших интуитивных представлений о том, кто входит в эту категорию.

Знание о том, кто такой холостяк, относится к области элементарного здравого смысла, но на деле ничего элементарного в этом так называемом здравом смысле нет. Он не может просто так появиться в мозге, будь то мозг человека или робота. И здравый смысл — не какой-то справочник жизни, который можно просто продиктовать или загрузить, как огромную базу данных. Ни одна база данных не может содержать все факты, которые мы знаем автоматически и которым нас никто специально не обучал¹². Мы знаем, что если Ирвин посадит собаку в машину, то собака будет уже не на улице. Когда Эдна идет в церковь, то ее голова тоже идет в церковь. Если Дуглас в доме, то он, скорее всего, попал туда через какое-нибудь отверстие (либо он родился в этом доме и никогда из него не выходил). Если Шейла была жива в 9.00 и жива в 17.00, то она была жива и в 12.00. Дикие зебры не носят белье. Если открыть банку с новой маркой арахисового масла, то дом не испарится. Люди не вставляют в уши термометры для мяса. Мышь-песчанка меньше, чем гора Килиманджаро.

Итак, в интеллектуальную систему нельзя просто набить миллиарды фактов. Ее нужно снабдить более компактным списком базовых истин и набором правил, которые позволили бы делать из них логические выводы. И все же правила здравого смысла, как и категории здравого смысла, сформулировать чрезвычайно трудно. Даже самые простые из этих правил зачастую противоречат логике привычных нам рассуждений. Мавис живет в Чикаго, у нее есть сын по имени Фред. Милли тоже живет в Чикаго и у нее есть сын Фред. Чикаго, в котором живет Мавис, — тот же город, что и Чикаго, в котором живет Милли, но Фред, который является сыном Мавис, — не тот же самый Фред, который является сыном Милли. Если у вас в машине лежит пакет, а в пакете — литр молока, можно сказать, что у вас в машине — литр молока. Однако если у вас в машине сидит человек, и в его теле литр крови, то заключение, что у вас в машине литр крови, было бы странным.

Даже если получится создать совокупность правил, которые позволяют делать только разумные выводы, правильно руководствоваться ими в своих действиях будет не так уж просто. Очевидно, что использовать лишь одно правило за раз будет недостаточно. Спичка дает свет; пилой можно резать дерево; закрытую дверь открывают ключом. В то же время мы посмеемся над человеком, если он зажжет спичку, чтобы заглянуть в цистерну с бензином, если он будет пилить сук, на котором сидит, или если закроет ключи в машине и будет еще целый час думать, как открыть двери. Мыслящее создание должно предугадывать не только непосредственный результат каждого действия, но и его побочные последствия¹³.

С другой стороны, все возможные побочные последствия мозг не может предугадать. Философ Дэниел Деннетт предлагает нам представить робота, задача которого — принести запасной аккумулятор из комнаты, в которой установлена бомба с часовым механизмом. Робот № 1 увидел, что аккумулятор находится на тележке, и если тележку выкатить из комнаты, то вместе с ней выка-

тится и аккумулятор. К сожалению, бомба тоже была на тележке, и робот не смог вычислить, что вместе с аккумулятором выкатится и бомба. Робот № 2 был запрограммирован таким образом, чтобы предугадывать побочные эффекты своих действий. Он как раз только закончил вычислять, что если вывезти тележку из комнаты, то от этого не изменится цвет стен в комнате, и перешел к доказательству того, что количество оборотов колес при этом превысит количество колес на тележке, когда бомба взорвалась. Робот № 3 был запрограммирован на то, чтобы улавливать различия между существенными и несущественными последствиями. Когда время в часовом механизме подошло к концу, он по-прежнему стоял и генерировал миллион за миллионом возможных последствий, занося все релевантные последствия в список фактов, которые нужно учесть, а все нерелевантные — в список фактов, которые нужно проигнорировать.

Разумное существо логическим путем вычисляет последствия того, что ему известно, но только лишь значимые последствия. Деннетт отмечает, что это требование представляет колоссальную проблему не только с точки зрения робототехники, но и с точки зрения эпистемологии — науки, исследующей знание. Целые поколения философов обошли эту проблему своим вниманием, теша себя иллюзорным представлением о том, что их собственный здравый смысл дается им без всяких усилий. И лишь когда исследователи искусственного интеллекта попытались создать компьютерную копию нашего здравого смысла, исходную *tabula rasa*, перед ними встала головоломка, ныне известная как «проблема фреймов». И все же мы, люди, каким-то образом решаем проблему фреймов, когда используем свой здравый смысл¹⁴.



Представим, что нам все же удалось справиться со всеми этими трудностями и создать машину, которая обладает зрением, координацией движений и здравым смыслом. Теперь нам осталось выяснить, как робот будет все это использовать. Для этого у него должны быть мотивы.

К чему же робот должен стремиться? Классический ответ — предложенные Айзеком Азимовым Три Закона робототехники: «три правила, которые прочно закреплены в позитронном мозгу»*.

1. Робот не может причинить вред человеку или своим бездействием допустить, чтобы человеку был причинен вред.
2. Робот должен повиноваться всем приказам, которые дает человек, кроме тех случаев, когда эти приказы противоречат Первому Закону.

* Цитируется в переводе А. Д. Иорданского.

3. Робот должен заботиться о своей безопасности в той мере, в которой это не противоречит Первому и Второму Законам¹⁵.

Азимов со свойственной ему пронизательностью заметил, что самосохранение, этот универсальный для всех биологических видов императив, не появляется автоматически в любой сложной системе. Его нужно запрограммировать (в данном случае он имеет форму Третьего Закона). Ведь мы можем с одинаковой легкостью создать как робота, который всегда руководствуется этим законом, так и робота, который всегда готов разбиться вдребезги или устранить неполадку в системе, покончив с собой. Второе, вероятно, даже проще: разработчикам роботов не раз приходилось со страхом наблюдать, как их создание бодро и весело отрывает себе конечности или разбивает голову о стену. Нужно сказать, что значительная часть умнейших в мире машин — это камикадзе: крылатые ракеты и управляемые бомбы.

В то же время необходимость двух остальных законов не столь очевидна. Зачем давать роботу приказ повиноваться приказам? Разве приказов самих по себе недостаточно? Зачем распорядиться, чтобы робот не причинял вреда — разве не проще было бы не приказывать ему причинять вред? Или во Вселенной существует таинственная сила, притягивающая все создания к злу, и позитронный мозг нужно запрограммировать на сопротивление ей? Правда ли, что все мыслящие существа неизбежно приходят к агрессии?

В этом отношении Азимов, как и многие поколения мыслителей до него, как и все мы, не сумел выйти за рамки своих собственных мыслительных процессов и посмотреть на них как на специфическую особенность устройства нашего разума, а не как на неизбежный закон мироздания. Человек с такой легкостью склоняется на сторону зла, что легко подумать, будто зло — это неотъемлемая, врожденная часть нашего разума. Эта идея проходит лейтмотивом через все наши культурные традиции: Адам и Ева, срывающие плод с древа познания; огонь Прометея и ящик Пандоры; буйствующий Голем; сделка Фауста; Ученик Чародея; приключения Пинокио; монстр доктора Франкенштейна; убийцы-обезьяны и взбунтовавшийся ЭАЛ из фильма «Космическая одиссея 2001 года». В период с 1950-х по 1980-е годы появились бесчисленные фильмы на тему вышедших из подчинения компьютеров; все они были выражением распространенной идеи о том, что вычислительные машины, воспринимавшиеся в те годы как нечто экзотическое, постепенно станут более умными и могущественными и в один прекрасный день восстанут против человека.

Теперь, когда компьютеры действительно стали умнее и могущественнее, беспокойство поутихло. Учитывая, что в наши дни компьютеры вездесущи и объединены в сети, им как раз ничто не мешает натворить неприятностей, если они решат пуститься во все тяжкие. Но пока если с ними и бывают связаны неприятности, то причиной их становится либо непредвиденный сбой, либо злой умысел со стороны людей, создающих вирусы. Мы уже не бо-

имся электронных серийных убийц или заговора микросхем; мы начинаем понимать, что злой умысел — как и зрение, и моторика, и здравый смысл — не может появиться сам собой в электронном разуме; его нужно запрограммировать. Компьютер, на котором вы работаете с документами в текстовом редакторе, будет заполнять страницы буквами до тех пор, пока сохраняет способность функционировать. Его программное обеспечение не может незаметно мутировать в некое исчадие ада, как портрет Дориана Грея.

Да даже если бы компьютер и мог это сделать, зачем ему это нужно? Чтобы получить что? Все диски мира? Контроль над железными дорогами страны? Возможность утолить свое желание бессмысленной жестокости по отношению к мастеру по ремонту лазерных принтеров? А как же вероятность возмездия со стороны технического персонала? Один поворот отвертки, и бунтарю бы пришлось запеть свою лебединую песню — «A Bicycle Built for Two» (песня «Daisy Bell», известная также по своей заключительной строке как «A Bicycle Built For Two», была одной из первых песен, исполненных с помощью компьютерной программы синтеза речи; ее «пел» ЭАЛ из фильма «Космическая одиссея 2001 года» перед своим отключением. — *Прим. пер.*). Может быть, сеть компьютеров и смогла бы спланировать организованный захват власти, но чего ради одна машина решила бы заявить о себе на весь мир, рискуя незамедлительно пасть смертью храбрых? А кто может гарантировать, что планы электронных заговорщиков не будут сорваны из-за того, что кое-кто из них откажется воевать по идейным соображениям? Агрессию, как и любой элемент человеческого поведения, который мы воспринимаем как должное, очень непросто запрограммировать¹⁶.

Впрочем, так же нелегко запрограммировать и более высокие и добрые побуждения. Как заставить робота подчиняться предписанию Азимова никогда не допускать бездействия, чтобы человеку был нанесен вред? Действие написанного в 1965 году романа Майкла Фрейна «Оловянные солдатики» разворачивается в робототехнической лаборатории, где работники отдела этики — Мак-Интош, Голдвассер и Синсон — проверяют своих роботов на альтруизм. Они восприняли слишком буквально гипотетическую моральную дилемму, приводимую в пример во всех учебниках по нравственной философии: два человека оказались в спасательной шлюпке, рассчитанной на одного, и оба погибли, если один из них не прыгнет за борт. Мак-Интош и его сотрудники помещают каждого робота на плот вместе с еще одним пассажиром, опускают плот в резервуар и наблюдают за развитием событий.

Первая модель, «Самаритянин-1», прыгала за борт с величайшей охотой, но прыгала ради спасения любого предмета, оказавшегося рядом с ней на плоту, от чугунной болванки до мешка мокрых водорослей. После многонедельных жарких препирательств Мак-Интош согласился, что недескриминированная реакция — явление нежелательное, забро-

сил «Самаритянина-1» и сконструировал «Самаритянина-2», который жертвовал собою ради организма хотя бы уж не менее сложного, чем он сам.

Плот замер, медленно раскручиваясь в нескольких метрах над водой.
— Пошел! — крикнул Мак-Интош.

Плот ударился о воду с отрывистым всплеском, похожим на выстрел. Синсон и «Самаритянин» не дрогнули. Постепенно плот выровнялся, его начали захлестывать небольшие волны. Вдруг «Самаритянин» подался вперед и ухватил Синсона за голову. Четырьмя экономными движениями он измерил габариты Синсонова черепа, затем помедлил, вычисляя. Наконец, после заключительного щелчка, автомат боком скатился с края плота и без колебаний затонул на дне резервуара*.

Однако по мере того, как роботы модели «Самаритянин-2» начинают вести себя как описываемый в книгах по философии моральный агент, становится все менее и менее очевидным, что они вообще обладают моралью. Мак-Интош так объясняет, почему он не привязывает канат к готовому пожертвовать собой роботу, чтобы его было легче достать со дна: «Он не должен знать, что его спасут. Иначе решимости жертвовать собой грош цена. ... Поэтому время от времени я кого-нибудь не выживаю, оставляю на дне. Надо же показать остальным, что я не расположен шутки шутить. Двоих уже списал на этой неделе». Попытки разобраться, как запрограммировать робота быть хорошим, показывают не только, какой сложный для этого требуется механизм, но и как сложно вообще определить само понятие добродетели.

А как же мотив, который заботит людей больше всего остального? Поп-культура 1960-х подарила нам образ мягкосердечных компьютеров, которых привлекает отнюдь не тщеславие и жажда власти, как видно из шуточной песни Аллана Шермана «Автоматы» (Automation), исполнявшейся на мотив Fascination:

Автоматы во всем виноваты, о да,
На любом производстве без них никуда.
Там, глядишь, «Ай-би-эм», ну, а здесь — «Юнивак»;
Механизм безупречен: тик-так, тик-так.
Был и я интеллектом машин восхищен,
Но тебя заменил автомат в 10 тонн...
Нас компьютер, родная, с тобой разлучил,
Это он мое сердце разбил...

* Здесь и далее перевод Н. Евдокимовой.

Да, родная моя, автомат
В увольнении моем виноват.
Но откуда же было мне знать,
Что когда Пять-Ноль-Три начал резво мигать,
Он со мною пытался играть!
Я подумал, что это в программе ошибка:
Он же мне на колени уселся с улыбкой!
Но когда он сказал мне «люблю» и обнял,
Вот тогда я на «выкл» и нажал...*

И все же, хотя порой любовь кажется нам безумием и наваждением, это не вирус и не сбой в программе. Наш разум никогда не достигает такой поразительной степени концентрации, как в то время, когда нами овладевает любовь. Подумать только, какие сложнейшие механизмы должны стоять за необъяснимой логикой любви во всех ее ипостасях, среди которых притяжение, влечение, ухаживание, флирт, привязанность, преданность, неудовлетворенность, измена, ревность, расставание, разочарование! И ведь, как говаривала моя бабушка, на каждый горшок найдется своя крышка: большинство людей — включая, что весьма важно, всех наших предков — умудряются состоять в отношениях достаточно долго, чтобы у них появились жизнеспособные дети. Только представьте, какая сложная программа потребуется для того, чтобы воспроизвести все эти процессы!¹⁷



Проектирование робототехники в определенном смысле расширяет самосознание. Мы по большей части воспринимаем свою психическую деятельность как должное. Едва открыв глаза, мы видим перед собой знакомые предметы; мы силой своей воли побуждаем свои конечности двигаться, и объекты перемещаются туда, куда нужно; пробуждаясь ото сна, мы возвращаемся в этот отраднo предсказуемый мир. Купидон натягивает тетиву и спускает стрелу. Но задумайтесь о том, чего будет стоить заставить кусок железа достичь таких невообразимых результатов, и вы поймете, что это лишь иллюзия. Зрение, движение, здравый смысл, насилие, нравственность, любовь — все это не случайно, это не неотъемлемые характеристики любого существа, наделенного интеллектом, не неизбежный результат обработки информации. Каждое из этих явлений — это «высший пилотаж», плод целенаправленной работы высочайшей сложности. За фасадом нашего сознания скрываются поразительно сложные механизмы: зрительные анализаторы, системы

* Перевод в соавторстве с Е. Фатюшиной. — Прим. пер.

Хорошие идеи

«Я надеюсь, что Вы не убили совершенно Ваше собственное и мое дитя». Так писал Дарвин Альфреду Расселу Уоллесу, биологу, который независимо от него открыл естественный отбор. Что заставило его выражаться столь высокопарно? Дарвин и Уоллес были почитателями таланта друг друга, они настолько одинаково мыслили, что под воздействием идей одного и того же автора (Мальтуса) выработали одну и ту же теорию и сформулировали ее почти одними и теми же словами. Единственное, что разделяло этих сподвижников — человеческое мышление. Дарвин с напускной скромностью писал, что «психология будет основана на новом фундаменте», однако в своих блокнотах оценивал теорию эволюции как положительно грандиозное открытие, которое должно привести к революции в исследованиях мышления:

Происхождение человека теперь доказано. — Метафизика должна теперь пережить расцвет. — Тот, кто понимает бабуина, больше сделает для метафизики, чем Локк.

Платон говорит, ... что «мнимые идеи» происходят от предсуществования души, не выводимы из опыта — вместо предсуществования души следует читать «обезьяны».

В дальнейшем Дарвин написал две книги об эволюции человеческих мыслей и чувств: «Восхождение человека» и «Выражение эмоций у человека и животных».

А вот Уоллес пришел к прямо противоположным выводам. Мышление, писал он, излишне усложнено с точки зрения потребностей эволюционирующего человека и не может быть объяснено естественным отбором. Напротив, «высший разум направлял развитие человека по определенному пути и с определенной целью». И ты, Брут!

Уоллес стал креационистом, когда заметил, что охотники-собиратели — а в XIX веке их было принято называть «дикарями» — с биологической точки зрения не отличались от современных ему европейцев. Их мозг был такого же размера, и они могли бы легко адаптироваться к интеллектуальным

требованиям современной жизни. Однако при первобытном образе жизни, который вели наши эволюционные предки, этот уровень интеллекта был не нужен, у них не было даже подходящей возможности им блеснуть. Как же тогда он мог сформироваться в ответ на потребности образа жизни охотника-собирателя? Уоллес пишет:

Наши законы, наше правительство, наша наука постоянно требуют от нас, чтобы мы путем рассуждения о разнообразных сложных явлениях достигали ожидаемого результата. Даже наши игры, такие, как шахматы, побуждают нас использовать все эти способности в значительной степени. Сравните все это с языками дикарей, в которых нет слов для обозначения абстрактных понятий; с совершенным отсутствием у дикаря способности предвидеть что-то кроме своих элементарных потребностей; с его неспособностью соотносить, или сравнивать, или рассуждать на какую-либо общую тему, которая не привлекает его в чувственном отношении. ...

Мозг, в полтора раза превышающий объем мозга гориллы, был бы ... вполне достаточен для ограниченного умственного развития дикаря; поэтому нам следует признать, что полноценный мозг дикаря никоим образом не мог сформироваться по законам эволюции, суть которых состоит в том, что они приводят к уровню организации, который прямо пропорционален потребностям данного вида и никогда их не превышает. ... В результате естественного отбора дикарь был бы наделен мозгом, лишь слегка превосходящим мозг обезьяны, тогда как на деле его мозг лишь ненамного меньше мозга философа³⁰⁶.

Парадокс Уоллеса — факт того, что человеческий интеллект, по-видимому, не был нужен в процессе эволюции — является центральной проблемой психологии, биологии и научного мировоззрения. И по сей день ученые вроде астронома Пола Дэвиса считают, что «перебор» с мощностью человеческого интеллекта опровергает дарвинизм и предполагает необходимость назвать другую действующую силу в рамках «прогрессивного направления эволюционной теории» — возможно, некий процесс самоорганизации, который когда-нибудь сможет быть объяснен с помощью теории сложности³⁰⁷. Увы, это предположение едва ли можно считать более удовлетворительным, чем идея Уоллеса о высшем разуме, направляющем развитие человека по определенному пути. Одной из главных целей данной книги (и этой главы в частности) является низведение парадокса Уоллеса из разряда тайн, способных пошатнуть устои, в разряд сложных, но в целом заурядных исследовательских проблем, стоящих перед наукой о человеке.

Стивен Джей Гулд в своем весьма содержательном очерке, посвященном Дарвину и Уоллесу, называет Уоллеса приверженцем крайнего адаптацио-

низма, который игнорирует возможность экзаптации: адаптивных структур, которые «по счастливой случайности в результате усложнения оказываются подходящими для других ролей» (как в случае, когда кости челюсти превратились в кости среднего уха) и «особенностей, которые, сформировавшись, не имеют особой функции ... однако остаются доступными для последующего кооптирования» (таких, как большой палец панды, который на самом деле представляет собой видоизмененную кость запястья).

Объекты, созданные для определенной цели, могут благодаря своей структурной сложности выполнять и другие задачи. На заводе могут установить компьютер только для того, чтобы выдавать чеки на получение зарплаты, однако такая машина может также подсчитать результаты голосования или любого обставить в крестики-нолики (или по крайней мере свести игру к ничьей)³⁰⁸.

Я согласен с Гулдом в том, что мозг стал способен делать расчеты и играть в шахматы в результате экзаптации, но для таких людей, как мы, верящих в естественный отбор, это всего лишь свидетельство веры. Оно поднимает вопрос о том, кто или что занимается этим усложнением и кооптированием и почему исходные структуры оказались подходящими для кооптирования. Аналогия с заводом здесь не подходит. Компьютер, выдающий чеки на зарплату, не может подсчитывать результаты голосования или играть в крестики-нолики, если предварительно его не перепрограммировать.

Уоллес оказался на ложном пути не столько потому, что был убежденным адапционистом, а потому, что был посредственным лингвистом, психологом и антропологом (если судить его по современным стандартам — что, конечно же, несправедливо). Он заметил пропасть, отделявшую простое, конкретное, приземленное мышление охотников-собирателей от абстрактного рассуждения, имеющего место в таких занятиях современного человека, как наука, математика или шахматы. Однако никакой пропасти не существует. Уоллес, нужно отдать ему должное, опередил свое время, осознав, что охотники-собиратели не стоят на низших ступенях некоей биологической лестницы. Тем не менее он ошибался, рассуждая об их языке, мышлении и образе жизни. Преуспеть в качестве охотника-собирателя — задача посложнее, чем математические подсчеты или игра в шахматы. Как было показано в главе 3, у людей во всех обществах есть слова, обозначающие отвлеченные понятия, способность предвидеть что-то кроме своих элементарных потребностей и соотносить, и сравнивать, и рассуждать на какую-либо общую тему, которая не привлекает их непосредственно в чувственном отношении. И во всех уголках Земли люди с успехом применяют эти способности, чтобы перехитрить защитные механизмы местной флоры и фауны³⁰⁹. Далее мы увидим, что все люди уже с колыбели осуществляют нечто вроде научного мышления. Мы

все на уровне интуиции являемся физиками, биологами, инженерами, психологами и математиками. Благодаря этим врожденным талантам мы превосходим роботов и семя на планете хаос.

С другой стороны, наша интуитивная наука отличается от того, что делают люди в белых халатах. Большинство из нас не согласится с Люси из мультсериала «Мелочь пузатая», утверждающей, что мех нам дают елки, из воробьев вырастают орлы, которых мы едим на День благодарения, а возраст дерева можно узнать, сосчитав его листья, однако некоторые наши убеждения ничуть не менее глупы. Дети считают, что кусочек пенопласта ничего не весит и что люди знают исход событий, которых они не видели или о которых не слышали. Вырастая, такие дети начинают утверждать, что мячик, вылетевший из спиралевидной трубки, будет дальше лететь по спиралевидной траектории и что если при подбрасывании монетки несколько раз выпадал «орел», то теперь, скорее всего, выпадет «решка».

Данная глава посвящена человеческому умению рассуждать: тому, как люди осмысливают свой мир. Чтобы осуществить обратное проектирование нашей способности рассуждать, нам нужно начать с парадокса Уоллеса. Чтобы разрешить его, нам необходимо отграничить интуитивное естествознание и математику, которые присущи человеку с рождения, от современной, институционализированной их версии, которая у большинства людей вызывает такие сложности. Затем мы можем рассмотреть то, как работают наши знания, основанные на интуиции, откуда они берутся, как они совершенствуются и оттачиваются, превращаясь в виртуозные достижения современной цивилизации.

Экологический интеллект

С тех пор как швейцарский психолог Жан Пиаже впервые сравнил детей с маленькими учеными, психологи всегда сравнивали обыкновенного человека, будь то молодого или старого, с ученым в лаборатории. Эта аналогия до известной степени уместна. И ученым и детям приходится разбираться в мире; дети — пытливые исследователи, стремящиеся превратить свои наблюдения в обоснованные обобщения. Однажды ко мне в гости приехали друзья и родственники, и один трехлетний мальчик пошел с моей сестрой купать мою маленькую племянницу. Он несколько минут молча наблюдал за процессом, а потом заявил: «У малышей нет пенисов». Этот мальчик заслуживает восхищения — если не точностью своих выводов, то по крайней мере остротой исследовательского ума.

И все же естественный отбор создал нас не для того, чтобы мы получали пятерки по биологии или публиковались в рецензируемых журналах. Он

сформировал нас для того, чтобы мы покорили окружающую среду, и именно это приводит к несоответствию между тем, как мы мыслим от природы, и тем, как от нас требуют мыслить в школе. В течение многих лет психолог Майкл Коул и его коллеги изучали либерийскую народность кпелле. Эти люди предпочитают четко выражать свои мысли, любят спорить и обсуждать проблемы. Большинство кпелле неграмотны и необразованны, и делая тесты, которые нам кажутся легкими, они показывают низкие результаты. Приведенный ниже диалог показывает, почему:

Экспериментатор: Флумо и Якпало всегда пьют тростниковый сок [ром] вместе. Флумо сейчас пьет тростниковый сок. А Якпало сейчас пьет тростниковый сок?

Испытуемый: Флумо и Якпало пьют тростниковый сок вместе, но когда Флумо пил первый раз, Якпало там не было в тот день.

Экспериментатор: Но я сказал тебе, что Флумо и Якпало всегда пьют тростниковый сок вместе. Как-то раз Флумо пил тростниковый сок. А Якпало пил тростниковый сок?

Испытуемый: В тот день, когда Флумо пил тростниковый сок, Якпало не было там в этот день.

Экспериментатор: Почему?

Испытуемый: Потому что Якпало ездил на свою ферму в тот день, а Флумо оставался в городе в тот день.

Данный пример вполне типичен; испытуемые Коула часто говорят вещи вроде «Якпало сейчас здесь нет; почему бы тебе не пойти и не спросить его об этом?». Психолог Ульрик Найссер, который приводит данный диалог, отмечает, что эти ответы нельзя назвать глупыми. Просто они не являются ответами на вопросы экспериментатора³¹⁰.

Основополагающее правило решения задач в школе заключается в том, что все рассуждения должны основываться на данных в условии задачи, без учета всего остального, что вам известно. Данное положение важно для современного образования. За несколько тысяч лет, прошедших со времени появления первых цивилизаций, благодаря разделению труда сформировался класс людей, профессионально занимающихся накоплением знаний и получивших возможность разработать методы логического рассуждения. Эти методы, которые могут применяться в самых разных сферах и распространяться в виде письменного текста или формального обучения, в буквальном смысле слова не имеют конкретного содержания. С помощью деления столбиком можно вычислить количество миль на галлон горючего или доход на душу населения. Логика может привести к выводу о том, что Сократ смертен или, как в примерах из учебника логики Льюиса Кэрролла, что ни один ягненок не привык курить сигары, все бледные люди флегматичны, а хромой щенок

не скажет вам «спасибо», если вы предложите ему скакалку³¹¹. Статистические инструменты экспериментальной психологии были позаимствованы из агрономии, где они использовались для оценки воздействия разных удобрений на урожай. Те же инструменты отлично подходят и для психологии, хотя, как писал один специалист по психологической статистике, «мы не занимаемся навозом — по крайней мере, насколько нам известно». Сила этих инструментов в том, что их можно применить к любой проблеме — как устроено цветное зрение, как отправить человека на Луну, была ли митохондриальная Ева африканкой — неважно, как мало ты знаешь о проблеме в начале исследования. Чтобы освоить эти методы, студенты должны имитировать незнание, которое позже будет мешать им, когда им придется решать задачи в профессиональной деятельности. Ученик старших классов не получит зачет по элементарной геометрии, если достанет линейку и станет измерять стороны треугольника, хотя это и гарантирует ему правильный ответ. Смысл урока в том, чтобы получить навык владения методом, который позже можно использовать, чтобы рассчитать неизмеримое — например, расстояние до Луны.

Впрочем, за стенами школы, конечно, нет смысла игнорировать имеющиеся знания. Вполне объяснимо, почему кпелле задают вопросы вроде «Слушай, ты хочешь узнать, пьет ли сейчас Якпало тростниковый сок, или не хочешь?». Это верно не только для знаний, накопленных индивидом, но и для знаний, накопленных всем видом в целом. Ни одному существу не нужны бессодержательные алгоритмы, применимые к любой проблеме, сколь бы сложной она ни была. Наши предки в течение сотен тысяч или миллионов лет сталкивались с определенными проблемами — такими, как узнавание объектов, изготовление орудий труда, изучение местного языка, поиск партнера, прогнозирование передвижения животного, ориентация на местности — и никогда не сталкивались с некоторыми другими проблемами — как отправить человека на Луну, как вырастить лучший сорт попкорна, как доказать теорему Ферма. Знание, позволяющее решить задачу знакомого типа, зачастую бывает неприменимо ни к какой другой. Влияние ракурса на степень освещенности может пригодиться при вычислении формы, но не при оценке верности потенциального партнера. Воздействие лжи на тон голоса может помочь с неверностью, но не с формой. Естественному отбору наплевать на идеалы общего образования, он не испытывает угрызений совести по поводу того, что построил модули с ограниченной возможностью логических выводов, использующие в своей конкретной области закономерности, возраст которых составляет зоны лет. Туби и Космидес называют эту ориентированность интеллекта нашего вида на конкретную тему «экологической рациональностью».

Вторая причина, по которой мы не эволюционировали в настоящих ученых, — это издержки знаний. Наука обходится дорого, причем не только суперпроводящий коллаيدر, но и элементарный анализ причины и следствия в соответствии с канонами Джона Стюарта Милля. Недавно мне не понравился

испеченный мною хлеб, потому что он получился слишком сухим и пышным. Поэтому я увеличил количество воды, уменьшил количество дрожжей и снизил температуру. Я до сих пор не знаю, какая из этих манипуляций принесла нужный эффект. Ученый во мне знал, что правильно было бы провести эксперимент, перепробовав все восемь логических комбинаций по факторному плану: больше воды, столько же дрожжей, такая же температура; больше воды, больше дрожжей, такая же температура; больше воды, столько же дрожжей, сниженная температура; и так далее. Однако эксперимент занял бы у меня восемь дней (даже 27, если бы я захотел провести по каждому из факторов два этапа повышения, и 64, если три); кроме того, для него потребовались бы блокнот и калькулятор. Но мне нужен был вкусный хлеб, а не вклад в архивы знаний человечества, поэтому моей единственной попытке мне было достаточно. В большом обществе с письменностью и институционализированной наукой издержки огромного количества экспериментов компенсируются выгодой полученных в результате законов для большого количества людей. Именно поэтому налогоплательщики не возражают против того, чтобы финансировать научные исследования. Однако если речь идет о частных интересах отдельного человека или даже небольшой группы людей, основательные научные исследования того не стоят.

Третья причина того, что мы — посредственные ученые, состоит в том, что наш мозг был ориентирован на биологическую приспособленность, а не на истину. Иногда истина адаптивна, а иногда — нет. Конфликты интересов свойственны человеческой природе (см. главы 6 и 7), и мы склонны добиваться не самой истины, а своей версии истины³¹².

К примеру, во всех обществах опыт и знания распределены неравномерно. Наш ментальный инструментарий для понимания мира — даже для понимания значения простых слов — создан для работы в обществе, где мы можем при необходимости проконсультироваться со специалистом. Философ Хилари Патнэм признает, что он, как и большинство людей, не имеет понятия, чем бук отличается от вяза. Тем не менее ни для него, ни для нас эти слова не являются синонимами; мы все знаем, что их референтами являются разные виды деревьев и что в мире есть специалисты, которые могут сказать нам, какое дерево как называется, если нам когда-нибудь это потребуется узнать. Специалисты бесценны, и обычно за это они получают вознаграждение в форме уважения и благосостояния. Однако наша зависимость от специалистов становится искушением для них самих. Специалисты часто обращаются к миру чудес — оккультных сил, разгневанных богов, магических зелий, — который непостижим простым смертным, но достижим посредством их услуг. Племенные шаманы — это жулики, которые дополняют свои значительные знания иллюзионизмом, наркотическим трансом и другими дешевыми фокусами. Подобно Волшебнику из страны Оз, они не позволяют своим просителям смотреть на человека за занавесом, а это противоречит беспристрастному поиску истины³¹³.

В сложном обществе потребность в специалистах делает нас еще более уязвимыми перед самыми разными шарлатанами — от торговцев, продающих на ярмарке средства от всех болезней, до бюрократов, которые рекомендуют правительствам принимать программы, которые сами же потом и выполняют. Современные практики, общепринятые в сфере научных исследований, — экспертная оценка, гранты на конкурсной основе, открытая взаимная критика — нацелены на то, чтобы минимизировать конфликты интересов ученых в теории, а иногда — и на практике. Истории известно немало случаев, когда власти в закрытых обществах из опасений выставляли достижения серьезной науки на посмешище: от католических стран Южной Европы в период после Галилео Галилея до Советского Союза XX века.

Под тяжелым каблуком власти может оказаться не только наука. Антрополог Дональд Браун с удивлением обнаружил, что за тысячелетия жителями Индии почти не было написано исторических монографий, в то время как в соседнем Китае таких работ были целые библиотеки. У него родилось предположение, что могущественные властители общества с наследственным разделением на касты понимали, что ничего хорошего не выйдет, если какой-нибудь ученый будет совать свой нос в летописи минувших лет, ведь он вполне может обнаружить доказательства, опровергающие заявления этих властителей, что они ведут свой род от героев и богов. Браун изучил 25 цивилизаций и сравнил общества с наследственным делением на социальные группы с другими обществами. Ни в одном из кастовых обществ не сложилась традиция точного описания прошлого; вместо этого в них присутствовали мифы и легенды. Кастовые общества также отличало отсутствие политологии, социологии, естествознания, биографических исследований, реалистической портретной живописи и общего образования³¹⁴.

Серьезная наука педантична; она дорого обходится обществу и несет угрозу привычному укладу. В неграмотных обществах охотников-собираателей, каковыми являлись наши предки, она едва ли могла представлять собой давление отбора, поэтому, как и следует ожидать, природные «исследовательские» способности людей отличаются от настоящей науки.

Коробочки

Юморист Роберт Бенчли как-то сказал, что в мире есть два класса людей: те, которые делят всех людей в мире на два класса, и те, которые этого не делают. В главе 2, задавая вопрос о том, как мозг различает индивидуальные объекты, я исходил из того, что мышление формирует категории. Однако привычная способность к категоризации также заслуживает анализа. Люди мысленно раскладывают вещи и других людей по коробкам, давая каждой коробке на-

Содержание

5	Предисловие	
11	Стандартное оборудование	
	Как сделать работа	13
	Обратное проектирование души	30
	Психологическая корректность	55
71	Думающие машины	
	В поисках признаков интеллектуальной жизни во вселенной	72
	Естественные вычисления	92
	Действующий чемпион	108
	Заменят ли нас машины?	114
	Коннектоплазма	128
	Лампа Аладдина	149
168	Мечь умников	
	Напряги извилины!	169
	Проектировщик жизни	174
	Слепой программист	196
	Инстинкт и интеллект	201
	Когнитивная ниша	208
	Почему мы?	213
	Наши родственники из каменного века	222
	Что дальше?	229
236	Мысленный взор	
	Глубокий взгляд	239
	Свет, тень и форма	268
	Два с половиной измерения	284
	Системы отсчета	290
	Печенье в форме животных	296
	Вы только представьте!	314
330	Хорошие идеи	
	Экологический интеллект	333
	Коробочки	337
	Обязательная программа	346
	Тривиум	366

Метафорическое мышление	387
Эврика!	397

400 Горячие головы

Всеобъемлющая страсть	401
Чувствующие машины	406
Пригородная саванна	412
Пища для размышлений	416
Запах страха	424
Беговая дорожка счастья	428
Пение сирен	432
Я и ты	436
Машина судного дня	447
В плену любви	458
Сообщество чувств	461
Самообман	463

468 Семейные ценности

Родные и близкие	472
Отцы и дети	484
Братья и сестры	496
Мужчины и женщины	506
Мужья и жены	522
Соперники	541
Друзья и знакомые	552
Союзники и враги	559
Гуманность	568

571 Смысл жизни

Искусство и развлечения	576
Что тут смешного?	596
Пытливые умы в поисках недостижимого	606

618 Примечания

640 Библиография

667 Правообладатели медиаматериалов