



Десять уравнений, которые правят миром

и как их можете
использовать вы

Дэвид
Самптер

Эту книгу хорошо дополняют:

Модельное мышление

Скотт Пейдж

Как не ошибаться

Джордан Элленберг

Удовольствие от X

Стивен Строгац

Теория игр

Авинаш Диксит и Барри Нейлбраф

Голая статистика

Чарльз Уилан

David Sumpter

The Ten Equations that Rule the World

And How You Can Use Them Too

Penguin Books

Серия «Принятие решений»

Дэвид Самптер

Десять уравнений, которые правят миром

И как их можете использовать вы

Перевод с английского Евгения Поникарова

Москва
«Манн, Иванов и Фербер»
2022

УДК 519.2

ББК 22.17

С17

На русском языке публикуется впервые

Самптер, Дэвид

С17 Десять уравнений, которые правят миром. И как их можете использовать вы / Дэвид Самптер ; пер. с англ. Евгения Поникарова. — Москва : Манн, Иванов и Фербер, 2022. — 288 с. — (Принятие решений).

ISBN 978-5-00169-986-6

Если вы сомневались, что вам может пригодиться математика, эта книга развеет ваши сомнения.

Красота приведенных здесь 10 уравнений в том, что пронизывают все сферы жизни, будь то грамотные ставки, фильтрование значимой информации, точность прогнозов, степень влияния или эффективность рекламы.

Если научиться вычленять из происходящего данные и математические модели, то вы начнете видеть взаимосвязи, словно на рентгене. Более того, вы сможете управлять процессами, которые другим кажутся хаотичными.

В этом и есть смысл прикладной математики.

УДК 519.2

ББК 22.17

Все права защищены. Никакая часть данной книги не может быть воспроизведена в какой бы то ни было форме без письменного разрешения владельцев авторских прав.

This edition is published by arrangement with Aitken Alexander Associates Ltd.
and The Van Lear Agency LLC

ISBN 978-5-00169-986-6

© David Sumpter, 2020

© Издание на русском языке, перевод, оформление.
ООО «Манн, Иванов и Фербер», 2022

Оглавление

Список иллюстраций	8
Введение: «Десятка»	9
Глава 1. Уравнение ставок	15
Глава 2. Уравнение суждений	41
Глава 3. Уравнение уверенности	66
Глава 4. Уравнение умений	102
Глава 5. Уравнение влияния	129
Глава 6. Уравнение рынка	151
Глава 7. Уравнение рекламы	178
Глава 8. Уравнение вознаграждения	205
Глава 9. Уравнение обучения	227
Глава 10. Универсальное уравнение	246
Примечания	267
Благодарности	284
Об авторе	286

Список иллюстраций

Рис. 1. Иллюстрация, как логистическая регрессия дает оценки $\alpha = 1,16$ и $\beta = 1,25$	27
Рис. 2. Иллюстрация теоремы Байеса	49
Рис. 3. Нормальное распределение	72
Рис. 4. Как предположение о марковости может оценивать передачи в футболе	111
Рис. 5. Парадокс дружбы для четырех участников	142
Рис. 6. Как меняется со временем отношение к трем товарам	162
Рис. 7. Иллюстрация к вычислению корреляции между Кайли и косметикой	183
Рис. 8. Как отслеживающая переменная отслеживает вознаграждение	210
Рис. 9. Как обучается нейронная сеть	234
Рис. 10. Иллюстрация сортировки слиянием и алгоритма Дейкстры	251

Введение: «Десятка»

Существует ли секретная формула обретения богатства? Или счастья? Или популярности? Или уверенности в себе и здравомыслия?

Если вы просматриваете книги в местном магазине и выбрали эту или только что щелкнули по кнопке «Заглянуть» в онлайн-магазине, то знайте: есть множество заголовков, которые предлагают вам какую-нибудь формулу для успеха в жизни.

Мари Кондо советует убраться. Шерил Сэндберг — не бояться действовать. Джордан Питерсон призывает выпрямиться. Брене Браун — быть уязвимыми*. Вам говорят, что вам надо, пес раздери, успокоиться; прекратить делать вон ту фигню, забить на всё, не быть лохом и извлекать пользу из каждого чертова момента в жизни. Нужно рано вставать, заправлять постель, расчищать путь, меньше делать, больше запоминать, избавляться от хлама в голове, добиваться целей, возвращать силу воли, писать формулы счастья**, «поступать как женщина, думать как мужчина»***. Есть формулы любви, богатства, план успеха и пять (или восемь, или

* Мари Кондо — японская писательница, автор книги «Магическая уборка. Японское искусство наведения порядка дома и в жизни». Шерил Сэндберг — член совета директоров Facebook, автор книги «Не бойся действовать». Джордан Питерсон — канадский психолог, автор книги «12 правил жизни. Противоядие от хаоса», первое из которых гласит: «Выпрямись и расправь плечи». Брене Браун — американская писательница, автор лекции «Сила уязвимости», с которой она выступала на ежегодной конференции фонда TED (Technology, entertainment, design — технологии, развлечения, дизайн). Здесь и далее прим. пер., если не указано иное.

** Отсылка к названию книги египетского предпринимателя Мо Гавдата «Формула счастья» (Минск : Попурри, 2018).

*** Отсылка к названию книги американского комика Стива Харви «Поступай как женщина, думай как мужчина» (М. : Одри, 2011).

двенадцать) правил уверенности в себе. Есть даже, видимо, чудо-уравнение, которое обещает сделать невозможные цели неизбежными*.

Эти советы парадоксальны. Если бы все было так просто, если бы существовали несложные формулы для получения всего, чего мы хотим от жизни, то откуда же все эти книги и журналы, заполненные нравоучениями (зачастую противоречивыми)? Почему так много воодушевляющих телешоу и выступлений на конференциях TED с мотивационными монологами? Можно же просто сформулировать уравнения, дать несколько примеров, как они работают, и закрыть всю индустрию самосовершенствования и правильного мышления. Если здесь все так «математично», аксиоматично, почему не дать ответ прямо сейчас?

Количество предлагаемых решений жизненных проблем растет, и все труднее поверить, что есть одна формула для успеха (или даже несколько). Может, на самом деле нет простого лекарства от всех проблем, что ставит перед нами жизнь?

Я хочу, чтобы вы подумали о другой возможности — той, о которой повествует книга. Я расскажу историю о замечательном обществе людей, взломавших этот код. Они обнаружили несколько уравнений (десять), которые могут принести успех, популярность, богатство, уверенность в себе и здравомыслие. Они хранят секрет, пока остальные ищут ответы.

Это тайное общество было с нами столетиями. Его члены передавали свои знания из поколения в поколение. Они занимали высокие посты на государственной службе, в финансовой сфере, в академических кругах, а с недавнего времени — в технологических компаниях. Они живут среди нас, незримо, но активно консультируя нас, а иногда и контролируя. Они богаты, счастливы и уверены в себе. Они открыли секреты, которых так жаждут остальные.

В книге Дэна Брауна «Код да Винчи» криптограф Софи Невё при расследовании убийства своего деда обнаруживает какой-то математический код. Она знакомится с профессором Робертом

* Вероятно, отсылка к названию книги Аарона Росса и Джейсона Лемкина «От невозможного к неизбежному» (From Impossible To Inevitable).

Лэнгдоном, который открывает ей, что ее дед был главой тайного общества Приорат Сиона, где мир понимается через единственное число — золотое сечение ($\phi \approx 1,618$).

«Код да Винчи» — выдумка, но тайное общество, которое я изучал в рамках этой книги, имеет много общего с описанным Брауном. Его тайны зашифрованы в коде, который мало кто понимает до конца, а его члены общаются с помощью загадочных документов. Общество уходит корнями в христианство, его раздирают внутренние моральные сражения и конфликты. Но, как мы вскоре узнаем, оно во многом отличается от Приората Сиона. В частности, в нем нет ритуалов. Именно поэтому его гораздо труднее обнаружить, но его деятельность повсеместна. И оно невидимо для тех, кто снаружи.

Но откуда о нем знаю я? Ответ очень прост. Я в нем состою. Двадцать лет я участвовал в его работе, все более тяготея к его внутреннему кругу. Изучал работы этого общества и применял его уравнения на практике. На своем опыте ощутил успех, который может дать доступ к его коду. Работал в ведущих университетах мира и получил должность профессора прикладной математики за день до своего тридцать третьего дня рождения. Решал научные задачи в самых разных областях — от экологии и биологии до политологии и социологии. Консультировал людей из правительства, из области спорта, азартных игр, финансистов и специалистов по искусственному интеллекту. И я счастлив — отчасти благодаря своему успеху, но в основном, полагаю, из-за того, что секреты, которые узнал, сформировали мое мышление. Эти уравнения сделали меня лучше: сбалансировали мое мировоззрение и помогли понимать действия других.

Членство в клубе дало мне много знакомств с единомышленниками. С такими людьми, как Мариус и Ян — молодые профессиональные игроки, получающие выгоду на азиатских рынках ставок; Марк, микросекундные расчеты которого дают прибыль от небольших неэффективностей в ценах акций. Я работал с аналитиками футбольного клуба «Барселона», изучавшими, как Лионель Месси и компания контролируют поле. Я встречался с техническими экспертами, работающими в Google, Facebook, Snapchat и Cambridge

Analytica, которые управляют нашими соцсетями и создают будущий искусственный интеллект. Я лично наблюдал, как исследователи Моя Бёрселл, Николь Нисбетт и Виктория Спейсер используют уравнения, чтобы обнаружить дискриминацию, понять наши политические дискуссии и сделать мир лучше. Я учился у старшего поколения — например, у 94-летнего профессора Оксфорда сэра Дэвида Кокса, открывшего код, на котором построено это тайное общество.

Теперь я готов рассекретить это общество. Я буду называть его «Десятка» — по числу уравнений, которые должен знать каждый компетентный участник. Я готов открыть эти секретные уравнения.

Проблемы, которые решает «Десятка», включают повседневные вопросы. Нужно ли вам бросать работу (или разрывать отношения) и пробовать что-то другое? Почему вы ощущаете, что менее популярны, чем окружающие? Сколько усилий надо приложить, чтобы стать популярнее? Как лучше справляться с огромным потоком информации из соцсетей? Должны ли вы позволять своим детям плятиться в телефон по шесть часов в день? Сколько эпизодов какого-нибудь сериала Netflix вам нужно посмотреть, прежде чем пробовать что-то другое?

Возможно, это не те проблемы, которые, на ваш взгляд, должно решать тайное общество. Но дело вот в чем. Один и тот же небольшой набор формул может дать ответы на тривиальные и глубокие вопросы, рассказать о вас как о личности, а также об обществе в целом. Уравнение уверенности, введенное в главе 3, которое помогает решить, стоит ли вам бросать работу, также позволяет профессиональным игрокам узнать, когда у них есть преимущество на рынке ставок, и выявляет малозаметные расовые и гендерные предубеждения на работе. Уравнение вознаграждения, описанное в главе 8, показывает, как соцсети довели общество до критической точки и почему это не обязательно плохо. Если мы поймем, как интернет-гиганты используют данное и другие уравнения, чтобы вознаграждать нас, влиять на нас и классифицировать нас, мы сможем лучше сбалансировать свое пользование соцсетями, играми и рекламой, а также погруженность в них наших детей.

Мы знаем, что эти уравнения важны, поскольку принесли успех людям, которые их уже применяют. В главе 9 излагается история трех инженеров из Калифорнии, которые использовали уравнение обучения, чтобы увеличить на 2000% время, затрачиваемое людьми на просмотр YouTube. Уравнение ставок, уравнение влияния, уравнение рынка и уравнение рекламы изменили, соответственно, ставки, технологии, финансы и рекламу, принеся миллиарды долларов небольшому числу членов «Десятки».

По мере того как вы будете изучать уравнения в этой книге, начнут обретать смысл всё новые аспекты мира. Когда вы смотрите глазами участника «Десятки», крупные проблемы становятся мелкими, а мелкие — тривиальными.

Если вы просто ищете быстрые решения, то, конечно, есть одна загвоздка. Чтобы попасть в «Десятку», нужно овладеть новым способом мышления. Она просит разделить мир на три категории: *данные, модель и бессмыслица*.

Одна из причин, почему «Десятка» сегодня так сильна, в том, что *данных* сейчас больше, чем когда-либо: движения на фондовой бирже и рынках ставок, личная информация о том, что мы любим, покупаем и делаем, которую собирают Facebook и Instagram. Государственные организации владеют данными о том, где мы живем, как работаем, в какую школу ходят наши дети, сколько мы зарабатываем. В опросах общественного мнения собираются и обобщаются наши политические взгляды и позиции. В соцсетях, блогах и на новостных сайтах копятся новости и мнения. Регистрируется и сохраняется каждое движение звезд спорта на игровых полях.

Взрывной рост данных очевиден всем, но участники «Десятки» осознали важность математических *моделей* для их объяснения. Как и они, вы можете научиться строить модели, использовать уравнения для управления данными и их использования, чтобы получить выгоду — небольшое преимущество над другими.

Последняя категория, *бессмыслица*, — то, что нужно научиться распознавать. Вы поймете, что, как бы ни приятно было говорить чушь и как бы много времени мы этому ни посвящали, вам придется отказаться от этого, когда вы начнете мыслить как участник «Десятки». Вам нужно назвать бессмыслицу бессмыслицей,

когда вы с ней встретитесь, — и неважно, кто ее говорит. Я покажу вам, как игнорировать бессмыслицу и сосредоточиться на данных и моделях.

Это не просто очередная книга по самосовершенствованию. Это не Десять заповедей. Это не список того, что разрешено, а что запрещено. Здесь есть уравнения, но нет рецептов. Вы не можете просто заглянуть на страницу 181 и узнать точное количество эпизодов сериала Netflix, которые вам нужно посмотреть.

Правила и рецепты эксплуатируют наши страхи. Но эта книга не использует страхи, а объясняет, как эволюционировал код «Десятки» и как он определял последние 250 лет человеческой истории. Мы будем учиться у математиков, которые разработали этот код, и поймем философию, лежащую в основе их мышления. Обучение у «Десятки» подвергает сомнению многие наши повседневные принципы. Вас ждет переосмысление терминов (например, политкорректность), переоценка суждений о других людях и пересмотр стереотипов.

Это также рассказ о морали, поскольку с моей стороны было бы неправильно раскрыть так много секретов, не исследовав, как общество «Десятки» повлияло на мир. Если какая-то малая группа людей может управлять остальными, нам нужно знать, что определяет их выбор. Эта история заставила меня заново оценить себя и мои действия, задаться вопросом, можно ли считать «Десятку» добром или злом, и задуматься о моральных правилах, которые нам нужно установить для себя в будущем.

Передавая силу новому поколению, дядя Человека-паука говорит ему: «С большой силой приходит большая ответственность». Когда на карту поставлено так много, скрытые возможности «Десятки» требуют еще большей ответственности, чем костюм Человека-паука. Вы узнаете секреты, которые могут изменить вашу жизнь. И вам придется также задуматься о том, какое влияние они оказали на мир, в котором мы живем. Слишком долго этот код был доступен только немногим избранным.

Сейчас мы поговорим об этом — открыто и вместе.

Глава 1

Уравнение ставок

$$P\{\text{фаворит выигрывает}\} = \frac{1}{1 + \alpha x^\beta}.$$

Первое, что меня поразило в Яне и Мариусе, когда мы пожимали друг другу руки в холле гостиницы, — они были ненамного старше студентов, которым я преподавал в университете. И я надеялся узнать от них о мире азартных игр столько же, сколько они, по-видимому, рассчитывали узнать от меня о математике.

Мы поддерживали отношения в Сети, но увиделись впервые. Они прилетели пообщаться в рамках своего рода европейского турне, где встречались по очереди с прогнозистами и специалистами по ставкам на футбол — чтобы разработать собственную стратегию на следующий год. Мой родной город Уппсала в Швеции был последней их остановкой.

— Мы захватим ноутбуки в паб? — спросил меня Мариус, когда мы собирались уйти из гостиницы.

— Конечно, — ответил я.

Хотя это была просто встреча, «чтобы лучше узнать друг друга», а настоящая работа предполагалась на следующий день, все мы понимали, что даже для самых неформальных обсуждений нужно обрабатывать большие массивы данных. Ноутбуки всегда должны быть наготове.

Возможно, вы думаете, что для успешных футбольных ставок надо много знать, глубоко понимать эту игру, в том числе и форму всех спортсменов, иметь представление о травмах и, вероятно,

получить какую-то инсайдерскую информацию. Может быть, десять лет назад дела обстояли именно так. В те времена внимательное наблюдение за матчами, за языком тела футболистов и отслеживание того, как они действуют в ситуациях один на один, могло дать вам преимущество перед игроками, которые примитивно ставили на команду, игравшую на своем поле. Но теперь все не так.

Ян проявлял к футболу лишь поверхностный интерес и не планировал смотреть большую часть матчей, на которые мы собирались делать ставки на предстоящем чемпионате мира 2018 года. «Буду смотреть матчи сборной Германии», — сказал он с улыбкой.

Был вечер церемонии открытия, весь мир следил за событиями. Однако если не считать интереса Яна к своей сборной, ему было все равно, о чем речь — о Бундеслиге, высшей лиге чемпионата Норвегии или чемпионате мира; о теннисе или конном спорте. Любой турнир и любой вид спорта были для него и Мариуса очередной возможностью заработать деньги. Именно поиск таких возможностей и привел их ко мне.

Несколько месяцами ранее я опубликовал статью о моей собственной модели ставок¹. Это была необычная математическая модель. Перед началом сезона Премьер-лиги сезона 2015/16 я написал единственное уравнение и предположил, что оно может улучшить коэффициенты букмекеров для исходов матчей Премьер-лиги. Так и оказалось.

К маю 2018 года прибыль составила 1900%. Если бы вы в августе 2015 года вложили 100 фунтов в мою модель, то меньше чем за три года у вас было бы 2000 фунтов. Вам требовалось лишь бездумно делать те ставки, что предлагала моя модель.

Мое уравнение не имело ничего общего с тем, что происходило на поле. Оно никак не было связано с просмотром матчей, и ему однозначно было все равно, кто стал чемпионом мира. Эта математическая функция брала коэффициенты букмекеров, слегка подправляла их в соответствии с влиянием информации о прошлых матчах и предлагала выигрышные ставки.

Свое уравнение я совершенно не скрывал, и оно привлекло довольно большое внимание. Я опубликовал подробности в журнале «1843», принадлежащем The Economist Group, и рассказывал о нем

в интервью BBC, CNBC, в газетах и социальных сетях. И именно об этой модели меня спрашивали сейчас Ян и Мариус.

— Как вы считаете, почему у вас по-прежнему преимущество? — спросил Мариус.

Валюта азартных игр — информация. Если вы знаете то, чего не знают другие, и эта информация приносит деньги, вам не придет в голову делиться ею. Чтобы защитить преимущество, нужно соблюдать тайну. Если схему будут использовать и другие игроки, букмекеры скорректируют свои коэффициенты. Преимущество исчезнет. По крайней мере в теории. Но я поступил наоборот: рассказал всем о своем уравнении. Мариусу было интересно, почему, несмотря на огласку, моя модель по-прежнему работает.

Значительную часть ответа на вопрос Мариуса можно найти в ежедневно получаемых мной электронных письмах и личных сообщениях в твиттере, где меня спрашивают: «Как думаете, кто выиграет завтрашний матч? Много читал о вас и начал в вас верить»; «Намереваюсь накопить средства, чтобы начать бизнес. Ваши советы о футболе однозначно ведут меня в нужном направлении»; «Кто выиграет — Дания или Хорватия? Чувствую, что Дании удастся победить, но не совсем уверен»; «Как вы думаете, каков будет результат ближайшего матча сборной Англии? Ничья?». Такие просьбы не прекращаются.

Мне не особенно приятно об этом говорить, но причина отправки подобных сообщений и есть ответ на вопрос Мариуса. Как бы я ни демонстрировал ограничения своего подхода и как бы ни подчеркивал долговременность стратегии, основанной на статистике, публика в основном задавала вопросы: «Выиграет ли “Арсенал” на выходных?» или «Выйдет ли Египет из группы, если Салах не будет играть?»

И это еще цветочки. Люди, которые мне пишут, хотя бы ищут в интернете советы по математике и азартным играм. Гораздо больше тех, кто забавляется безо всяких исследований. Они играют на внутреннем чутье, ради развлечения, по пьяному делу, потому что не могут остановиться. И в целом их намного больше, чем тех, кто использует мой метод или нечто подобное.

— Моя модель продолжает выигрывать, потому что она предполагает ставки, которые многие не желают делать, — объяснил

я Мариусу. — Ставить на ничью в матче «Ливерпуль» — «Челси» или на то, что «Манчестер Сити» обыграет «Хаддерсфилд», где коэффициент очень мал, скучно и невесело. Чтобы получить какую-то прибыль, нужно время и терпение.

Первое электронное письмо Мариуса попало в один процент сообщений, отличившихся от общей массы. Он рассказал мне об автоматизированной системе, которую они с Яном разработали, чтобы получать выгоду на рынках ставок. Их идея заключалась в использовании того, что большинство букмекеров — «мягкие»: они предлагают коэффициенты, не всегда отражающие реальную вероятность победы той или иной команды.

Подавляющее большинство игроков (вероятно, все те, кто спрашивал у меня совета) используют мягких букмекеров. Ведущие конторы — Paddy Power, Ladbrokes и William Hill — мягкие, как и менее крупные организации вроде RedBet и 888sport. Они отдают предпочтение специальным предложениям, побуждающим людей играть, но уделяют меньше внимания определению тех коэффициентов, которые отражают вероятные исходы спортивных событий. А вот «резкие» букмекеры, например Pinnacle или Matchbook, определяют коэффициенты для предсказания результатов более точно; этими конторами, как правило, пользуются лишь 1% игроков*.

Идея Мариуса и Яна состояла в использовании резких букмекеров, чтобы забирать деньги у мягких. Их система отслеживала ставки у всех букмекеров и выискивала расхождения. Если один из мягких букмекеров давал более выгодные коэффициенты, чем резкие, их система предлагала сделать ставку у него. Такая стратегия не гарантировала победу, однако давала Яну и Мариусу важнейшее преимущество, потому что резкие букмекеры были точнее. В долгосрочной перспективе после сотен ставок они выиграли бы деньги у мягких.

* Разница между типами букмекеров не только в точности коэффициентов. Резкие быстрее реагируют на изменения ситуации корректировкой коэффициентов (часто это делается автоматически), ставки у них не ограничены. Мягкие меняют коэффициенты медленнее и вводят ограничения на ставки; они тоже имеют доступ к технологиям, но предпочитают охватить больше людей, которые в долгосрочной перспективе будут терять деньги. В России обычно говорят о рекреационных и профессиональных конторах. Первые (подавляющее большинство) рассчитаны на любителей азартных игр, вторые — на профессионалов, нацеленных на получение прибыли.

Однако у системы было одно ограничение: мягкие букмекеры блокируют победителей. Именно эти конторы решают, хотят ли они иметь с вами дело; и как только видят, что на счетах, скажем, Яна и Мариуса скапливается прибыль, отстраняют их от игры, отправляя сообщение: «Теперь ваша максимальная ставка — 2,5 фунта стерлингов».

Парни нашли способ обойти эту проблему. Они создали службу подписки. За ежемесячную плату подписчикам отправлялись прямые ссылки на выгодные коэффициенты у мягких букмекеров. Это означало, что Ян с Мариусом получали прибыль, даже если их самих заблокировали. Беспроигрышный вариант для всех, кроме букмекеров. Игроки могли получать советы, которые давали выигрыш в долгосрочной перспективе, а Ян с Мариусом брали свою долю.

Вот почему я сидел в пабе с этой парочкой. Они овладели искусством сбора данных и автоматического размещения ставок. Я написал уравнение, которое могло дополнительно увеличить их преимущество: моя модель Премьер-лиги способна была выигрывать не только у мягких букмекеров, но и у резких.

В тот момент я полагал, что обладаю преимуществом для предстоящего чемпионата мира. Однако для проверки моей гипотезы мне требовалось больше данных. Прежде чем я закончил рассказывать о своей идее, Ян открыл ноутбук и попытался поймать Wi-Fi.

— Уверен, что смогу найти коэффициенты для отборочного турнира и для восьми последних крупных международных соревнований, — сказал он. — У меня есть программа для веб-скрейпинга [автоматического просмотра сайтов и загрузки данных с них].

К концу ужина у нас был план, и мы определили данные, которые необходимы для его осуществления. Ян отправился обратно в гостиницу и поставил компьютер заниматься всю ночь веб-скрейпингом коэффициентов для прошлых турниров.

*

И Ян, и Мариус принадлежат к новому поколению профессиональных игроков. Они умеют программировать, знают, как добывать данные, и понимают математику. Игроки такого типа зачастую

меньше интересуются конкретным видом спорта и больше — числами, чем игроки старой школы. Но они точно так же заинтересованы в зарабатывании денег, и это у них получается гораздо лучше.

Обнаруженное мною преимущество при ставках привлекло ко мне внимание этой пары, и я оказался на периферии их игровой сети. Однако когда я задавал им вопросы о других проектах, по осторожным ответам было понятно, что о полноправном членстве в их клубе речи не было. Во всяком случае пока. Я был любителем — и они просто смеялись надо мной, когда я сказал, что планирую поставить 50 фунтов в разрабатываемой нами системе, — и информацию о других проектах выдавали только по мере необходимости.

Но один знакомый был более откровенен. Недавно он ушел из этой индустрии и был рад поделиться своим опытом, хотя и не пошел, чтобы я раскрыл его место работы и личность; так что назовем его Джеймсом.

Джеймс сказал мне:

— Если у вас есть реальное преимущество, то скорость вашего обогащения ограничена только скоростью, с которой вы способны делать ставки.

Чтобы понять смысл высказывания Джеймса, сначала представим традиционные инвестиции с доходностью 3%. Если вы вкладываете 1000 фунтов, то через год у вас будет 1030 — вы получите прибыль 30 фунтов.

Теперь представим игру на сумму 1000 фунтов с преимуществом в 3% перед букмекерами. Наверняка вы не захотите рисковать сразу всем капиталом при единственной ставке. Так что начнем со ставки в 10 фунтов — относительно умеренный риск. Вы не будете выигрывать каждый раз, но преимущество в 3% означает, что в среднем на одну ставку в 10 фунтов вы станете получать прибыль 30 пенсов, поэтому уровень доходности для капитала в 1000 фунтов составляет 0,03% при каждой ставке.

Чтобы получить прибыль 30 фунтов, вам придется сделать 100 ставок по 10 фунтов. Сто ставок в год, примерно две в неделю, — это больше, чем делает большинство из нас. И нас, любителей, отрезвляет мысль, что, даже если у вас есть какое-то преимущество,

вы как случайный игрок-любитель не можете ожидать слишком многоного от капиталовложения в 1000 фунтов.

Однако парни, с которыми работал Джеймс, не были случайными игроками. По всему миру нетрудно найти 100 матчей за один день. Ян недавно загрузил данные по 1085 различным лигам. Добавьте теннис, регби, скачки и все прочие виды спорта, и вы обнаружите массу возможностей для ставок.

А теперь представим, что Джеймс и его коллеги имеют какое-то преимущество и целый год ставят каждый день на сотню матчей. Примем также, что по мере роста своих доходов они увеличивают ставки пропорционально имеющемуся капиталу; то есть когда у них появилось 10 000 фунтов, они ставят по 100 фунтов, при капитале 100 000 фунтов — по 1000 фунтов и т. д. Сколько получат игроки к концу года, имея преимущество всего лишь в 3%? 1300, 3000, 13 000 или 310 000 фунтов? На самом деле к концу года у них будет 56 860 593,8 фунта. Почти 57 миллионов! Да, каждая ставка умножает их капитал всего на 1,0003, однако после 36 500 ставок вступает в игру мощь показательного распределения, и их прибыли резко взлетают².

Но на практике такой рост недостижим. Хотя резкие букмекерские конторы, которыми пользовались Джеймс с коллегами, разрешают ставить больше, чем мягкие, ограничения всё равно существуют. Джеймс сказал мне:

— Букмекерские компании в Лондоне выросли так быстро и стали такими масштабными, что они вынуждены теперь делать ставки через брокеров. Иначе, если станет широко известно, что они делают определенную ставку на конкретный матч, на этот рынок хлынут все остальные и их преимущество исчезнет.

Несмотря на такие ограничения, деньги по-прежнему текут в букмекерские компании, которыми управляют уравнения. Достаточно посмотреть на стильные интерьеры их офисов в Лондоне, чтобы убедиться в этом. Сотрудники одного из лидеров отрасли, Football Radar, начинают день с бесплатного завтрака, имеют доступ в роскошный спортзал, могут сделать перерыв и поиграть в теннис или PlayStation и располагают всем необходимым компьютерным оборудованием. Специалистам по обработке данных

и создателям программного обеспечения предлагают работать в свободное время, а компания утверждает, что гарантирует такую творческую среду, которая обычно ассоциируется с Google или Facebook.

В Лондоне базируются и два основных конкурента Football Radar — Smartodds и Starlizard. Они принадлежат соответственно Мэттю Бенхэму и Тони Блуму, карьера которых началась благодаря умению обращаться с числами. Бенхэм учился в Оксфорде, где начал работать в сфере азартных игр, основываясь на статистике, а Блум обладает опытом профессионального игрока в покер. В 2009 году оба они приобрели футбольные клубы из родных городов: Блум купил «Брайтон энд Хоув Альбион», а Бенхэм — «Брентфорд». Кроме того, второй решил добавить к своим активам и резкую букмекерскую контору Matchbook.

И Бенхэм, и Блум сумели использовать незначительное преимущество и с помощью больших данных получили колоссальные прибыли.

*

Преимущество, которое я предложил Яну и Мариусу для вероятности победы их фаворита в каком-нибудь матче чемпионата мира, основано на следующем уравнении:

$$P\{\text{фаворит выигрывает}\} = \frac{1}{1 + \alpha x^\beta} \quad (\text{Уравнение 1}),$$

где x — коэффициент букмекера на победу фаворита. Коэффициент здесь понимается в британском формате: если он составляет 3 к 2 или $x = 3/2$, это означает, что на каждые поставленные 2 фунта в случае победы чистый выигрыш составляет 3 фунта.

Разберемся, о чем на самом деле говорит уравнение 1. Начнем с левой стороны, где я написал: « $P\{\text{фаворит выигрывает}\}$ ». Ни одна математическая модель не предсказывает победу или поражение с абсолютной точностью. Они говорят о вероятности того, что выиграет фаворит, и эта вероятность — число от 0 до 100%. Оно определяет уровень уверенности, который я приписываю результату.

Эта вероятность зависит от того, что написано в правой части уравнения, куда входят три буквы: латинская x и греческие α и β . Одна студентка сказала мне, что математика казалась ей понятной, пока речь шла о латинских иксах и игреках, но стала трудной, когда начались греческие альфы и беты. Для математиков это звучит смешно, потому что x , α и β — только символы, они не делают науку проще или сложнее, так что я думаю, что студентка всего лишь шутила. Но она попала в точку: когда в уравнениях встречаются α и β , математика обычно сложнее.

Так что давайте начнем без них. Уравнение

$$P\{\text{фаворит выигрывает}\} = \frac{1}{1+x}$$

понять гораздо проще. Если, скажем, коэффициент был $3/2$ ($2,5$ в европейской системе или $+150$ в американской)*, вероятность того, что фаворит выиграет, равна

$$P\{\text{фаворит выигрывает}\} = \frac{1}{1 + \frac{3}{2}} = \frac{2}{2+3} = \frac{2}{5}.$$

По сути, это уравнение без α и β дает нам оценку букмекера для победы фаворита. Он считает, что шансы фаворита на победу в матче составляют $2/5$, или 40% . В остальных 60% случаев будет ничья или победит аутсайдер.

* В разных странах используются различные виды коэффициентов. Британский (или дробный) — отношение потенциальной чистой прибыли к сумме, которую надо поставить, чтобы ее получить. В нашем случае отношение $3/2$ означает, что мы ставим 2 фунта, чтобы в случае победы получить чистую прибыль 3 фунта. Европейский (десятичный) коэффициент, который используется и в России, — это число, на которое умножается ставка для определения потенциальной выплаты. В нашем случае коэффициент $2,5$ означает, что при ставке в 2 фунта мы получаем $2 \times 2,5 = 5$ фунтов, чистая прибыль снова составляет 3 фунта. Американский коэффициент — это сумма потенциальной чистой прибыли при ставке в 100 условных единиц (он может быть положительным или отрицательным). В нашем случае $+150$ означает, что на каждые 100 единиц ставки можно получить 150 единиц чистой прибыли. Для ставки в 2 фунта чистая прибыль равна 3 фунтам. Таким образом, все три числа, указанные автором, действительно определяют одно и то же.

Без α и β (точнее, при $\alpha = 1$ и $\beta = 1$) мое уравнение ставок относительно несложно понять. Однако без α и β оно не принесет денег. Почему? Поставим 1 фунт на фаворита. Если коэффициент букмекера верен, два раза из пяти вы выиграете 1,5 фунта, а три из пяти проиграете по 1 фунту. Поэтому в среднем вы выиграете

$$\frac{2}{5} \times 1,5 + \frac{3}{5} \times (-1) = \frac{3}{5} - \frac{3}{5} = 0.$$

Иными словами: после нескольких ставок вы почти ничего не выиграете. Нуль. Пшик. На деле всё еще хуже. Для начала я предположил, что коэффициенты букмекеров справедливы³. На самом деле нет. Букмекеры всегда подправляют их, чтобы ситуация складывалась в их пользу. И вместо того, чтобы предложить $3/2$, заявят, скажем, $7/5$. И если вы не знаете, что делаете, букмекеры всегда выиграют, а вы проиграете. При коэффициенте $7/5$ вы будете в среднем проигрывать 4 пенса на ставку в 1 фунт⁴.

Единственный способ обыграть букмекеров — рассмотреть эти числа, и именно такие данные компьютер Яна собирал после того, как мы посидели в пабе. Он скачал коэффициенты и результаты для всех матчей чемпионатов мира и Европы, включая отборочные игры, начиная с чемпионата мира в Германии в 2006 году. Утром, усевшись в моем офисе в университете, мы начали искать преимущество.

Сначала мы загрузили данные и посмотрели на них в таблице, подобной нижеприведенной.

Фаворит	Аутсайдер	Коэффициенты на победу фаворита, x	Вероятность букмекеров, что фаворит победит, $1/(1+x)$	Фаворит победил? («да» = 1, «нет» = 0)
Испания	Австралия	11/30	73%	1 (победил)
Англия	Уругвай	19/20	51%	0 (проиграл)
Швейцария	Гондурас	13/25	66%	1 (победил)
Италия	Коста-Рика	3/5	63%	0 (проиграл)
...				

Из таких прошлых результатов мы можем получить представление о том, насколько точны коэффициенты: для этого надо сравнить два последних столбца вышеприведенной таблицы. Например, в матче между Испанией и Австралией на чемпионате мира 2014 года коэффициенты дают вероятность 73%, что Испания выиграет, и она действительно победила. Это можно считать «хорошим» прогнозом. А вот Коста-Рика обыграла Италию, хотя коэффициенты давали 63% на победу итальянцев, — «плохой» прогноз.

Я пишу слова «хороший» и «плохой» в кавычках, поскольку нельзя сказать, хорош или плох прогноз, если нет альтернативы, с которой его можно сравнить. Вот здесь и появляются α и β . Их называют параметрами уравнения 1. Это величины, которые мы можем менять для тонкой настройки нашего уравнения, чтобы сделать его точнее. Мы не можем изменить итоговые коэффициенты для матча Испания — Австралия и определенно неспособны повлиять на результат этого матча сборных; но можем выбрать α и β так, чтобы получить более точный прогноз, чем у букмекеров.

Метод поиска наилучших параметров — логистическая регрессия. Чтобы описать, как она работает, сначала посмотрим, как можно улучшить наш прогноз на матч Испания — Австралия с помощью корректировки числа β . Если я приму $\beta = 1,2$ и оставлю $\alpha = 1$, получу

$$\frac{1}{1+\alpha x^\beta} = \frac{1}{1+\left(\frac{11}{30}\right)^{1,2}} = 77\%.$$

Поскольку результатом матча была победа Испании, прогноз на победу в 77% лучше, чем прогноз букмекеров, который давал 73%.

Но здесь есть проблема. Если я увеличу β , то повышу и прогнозируемую вероятность победы Англии над Уругваем — с 51 до 52%. Но Англия в том матче 2014 года уругвайцам проиграла. Чтобы справиться с этой проблемой, я могу увеличить другой параметр, назначив $\alpha = 1,1$ и оставив при этом $\beta = 1,2$. Теперь уравнение

предсказывает, что Испания обыграет Австралию с вероятностью 75%, а Англия обыграет Уругвай с вероятностью 49%. Изменив исходные значения $\alpha = 1$ и $\beta = 1$, мы улучшили прогноз на оба матча.

Я рассмотрел по одному изменению каждого из параметров α и β и сравнил результаты всего по двум матчам. Данные Яна включали 284 матча на всех чемпионатах мира и Европы с 2006 года. Потребовалось бы очень много времени, чтобы вручную менять значения параметров, подставлять их в уравнение и смотреть, улучшают они прогноз или нет. Однако мы можем использовать для вычислений компьютерный алгоритм; именно это и делает логистическая регрессия (см. рис. 1). Она меняет значения α и β так, чтобы дать прогнозы, которые максимально близки к реальным результатам матчей.

Я написал программу на языке Python, которая выполняет все эти вычисления. Запустил ее и смотрел, как она справляется со всеми этими расчетами. Через несколько секунд у меня был результат: наилучшие прогнозы получались при $\alpha = 1,16$ и $\beta = 1,25$.

Эти числа сразу привлекли мое внимание. Сам факт, что оба параметра $\alpha = 1,16$ и $\beta = 1,25$ превосходят 1, показывал сложную связь между коэффициентами и исходами матчей. Проще всего понять эту связь путем добавления к нашей таблице еще одной колонки и сравнения нашей модели логистической регрессии с прогнозами букмекеров.

Фаворит	Аутсайдер	Коэффициенты на победу фаворита, x	Вероятность букмекеров, что фаворит победит, $1/(1+x)$	Вероятность логистической регрессии, что фаворит победит $1/(1+1,16x^{1,25})$	Фаворит победил? («да» = 1, «нет» = 0)
Испания	Австралия	11/30	73%	75%	1 (победил)
Англия	Уругвай	19/20	51%	48%	0 (проиграл)
Швейцария	Гондурас	13/25	66%	66%	1 (победил)
Италия	Коста-Рика	3/5	63%	62%	0 (проиграл)
...					

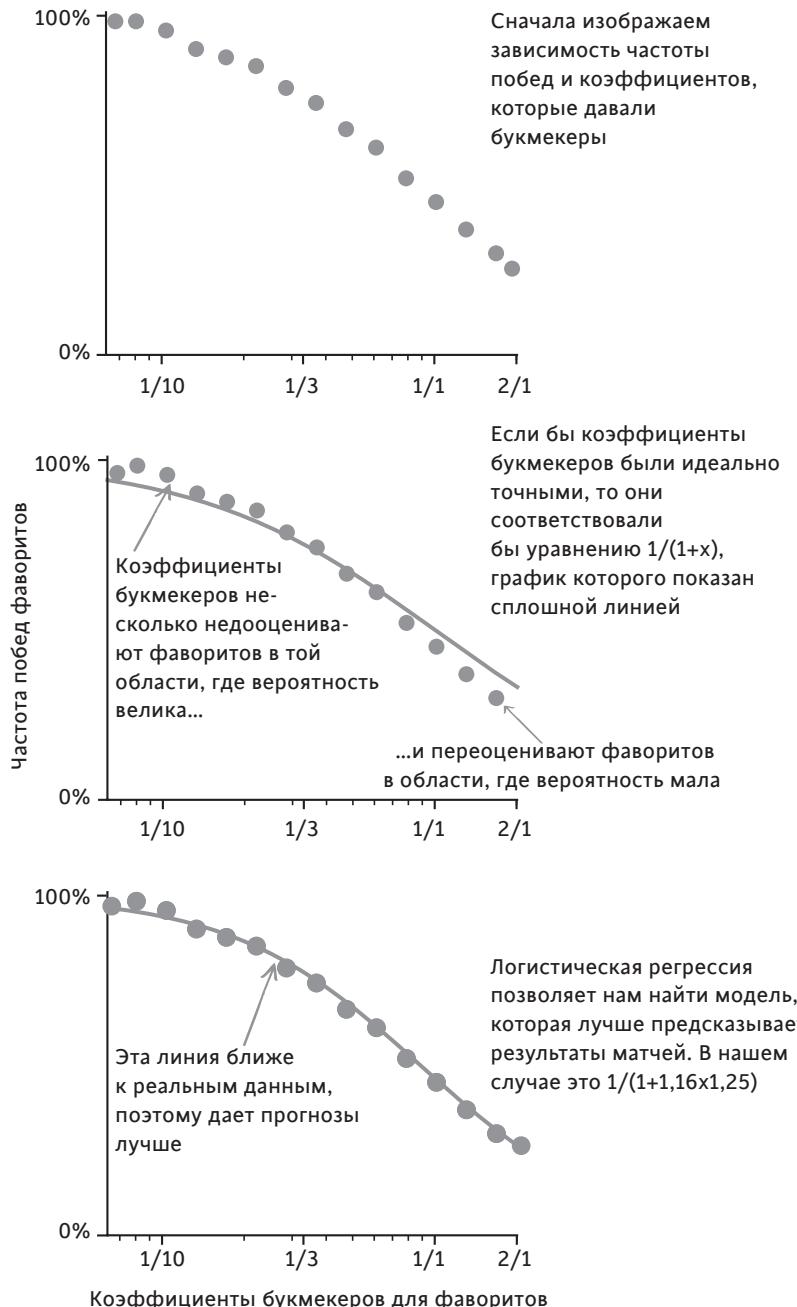


Рис. 1. Иллюстрация того, как логистическая регрессия дает оценки $\alpha = 1,16$ и $\beta = 1,25$

Здесь мы видим проявление известного опытным игрокам феномена с недооценкой записных фаворитов вроде Испании. Коэффициенты, которые букмекеры устанавливают для таких команд, как правило, занижены, поэтому на них стоит ставить. А более слабые фавориты, вроде Англии в 2014 году, бывают переоценены: их шансы на победу не так высоки, как предполагают коэффициенты. Хотя такие различия между прогнозами и моделью малы, мы с Яном и Мариусом знали, что их достаточно, чтобы получить прибыль.

Мы нашли небольшое преимущество для чемпионата мира. Еще не зная, будет ли это преимущество, замеченное на предыдущих чемпионатах, работать на новом турнире, мы были готовы рискнуть незначительной суммой. Чтобы реализовать систему ставок на основании моего уравнения, хватило времени до обеда. Мы нажали «Запуск» и привели систему в действие. В течение всего чемпионата мира наши ставки размещались автоматически.

После обеда вернулись ко мне домой. Мы с Мариусом уселись смотреть игру Уругвая с Египтом. Ян достал ноутбук и начал скачивать коэффициенты для тенниса.

*

Уравнение ставок — это не только один чемпионат мира и даже не только зарабатывание денег на букмекерах. Его настоящая сила в том, что оно заставляет нас смотреть в будущее с точки зрения вероятностей и исходов. Использование уравнения ставок означает следующее: надо отказаться от догадок и навсегда забыть идею, что результат футбольного матча, скачек, финансовой инвестиции, собеседования при приеме на работу или даже романтического свидания можно предсказать со стопроцентной уверенностью. Вы не можете знать наверняка, что произойдет.

Большинство из нас смутно осознают, что события в будущем во многом определяются случайностью. Когда прогноз погоды говорит, что завтрашний день будет солнечным с вероятностью 75%, не следует слишком сильно удивляться, если по дороге на работу вы попадете под ливень. Однако нахождение небольших преимуществ, скрытых в вероятностях, требует более глубокого понимания.

Если для вас важен конкретный результат, то подумайте, с какой вероятностью он реализуется, а с какой нет. Недавно я разговаривал с CEO одного очень успешного стартапа, который вырос за счет четырех этапов многомиллионных долларовых инвестиций и в котором работает сотня сотрудников, и он радостно признавал, что шансы на долговременную прибыль для него самого и его инвесторов по-прежнему всего лишь 1 из 10. Он самоотверженно и долго работал, но сознавал, что все может внезапно развалиться.

При поиске работы мечты или любви всей жизни шансы на успех при каком-нибудь конкретном заявлении на работу или свидании могут быть весьма малы. Меня часто удивляет, что люди, не прошедшие собеседования, ругают себя за то, что поступили неверно, а не учитывают, что, возможно, в этот день кто-то из других четырех кандидатов сделал все верно. Помните, что до появления на собеседовании ваши шансы составляли 20%. Пока вы не провалите примерно пять собеседований, нет особых причин беспокоиться о каком-то конкретном результате⁵.

Вводить числа в романтику труднее, но здесь применимы те же вероятностные принципы. Не ждите, что на вашем первом свидании в Tinder появится принц или принцесса, но если вы сидите в одиночестве после неудачного свидания номер 34, то полезно поразмыслить над своим подходом.

Определив соответствующие вероятности, подумайте, как они соотносятся с размером ваших инвестиций и потенциальных прибылей. Мой совет мыслить вероятностно не попытка призвать к кармическому спокойствию или внимательности. CEO с шансами на успех один из десяти располагал бизнес-идеей, которая потенциально могла дать результат, подобный Uber или Airbnb: создать компанию стоимостью 10 миллиардов долларов. Даже десятая часть от этой суммы — миллиард, и это огромная ожидаемая прибыль.

Вероятностное мышление поможет вам быть реалистом перед лицом шансов, которые часто обращаются против вас. В скачках и футболе наивные игроки нередко переоценивают маловероятные события, но в реальной жизни мы склонны их недооценивать. Мы по природе осторожны и избегаем рисков. Помните, что награда

после того, как вы получите действительно желательную работу или любимого человека, будет колоссальной. Это означает, что нужно быть готовым пойти на большой риск ради достижения цели.

*

Математика требует работы и упорства. Пять минут назад я закончил читать одну из самых примечательных работ в истории прикладной математики — статью, которая буквально стоит миллиард долларов. Я знал, что математика здесь важна, но, добравшись до уравнений, решил, что читать стало гораздо труднее. В первый раз я пропустил их, сказав себе, что вернусь к деталям позже, и перешел к интересным фрагментам.

Речь о статье Уильяма Бентера «Компьютерные системы прогнозирования и размещения ставок на скачках: отчет»⁶. Это своеобразный манифест, декларация о намерениях. И это работа человека, одержимого строгостью и верой в то, что он делает, который документировал свой план, прежде чем взялся его реализовывать, — чтобы показать всему миру, что он побеждает благодаря не удаче, а математической уверенности.

В конце 1980-х Уильям Бентер решил обыграть тотализатор скачек в Гонконге. До того как он начал свой проект, азартная игра с высокими ставками была уделом темных личностей, которые шлялись по ипподромам «Хэппи-Вэлли» и «Сатхинь» и по Гонконгскому жокей-клубу, пытаясь собрать инсайдерскую информацию у владельцев, персонала конюшен и тренеров. Они выясняли, завтракала лошадь или нет и была ли у нее дополнительная тайная тренировка. Они сходились с жокеями и расспрашивали их о стратегии в будущих скачках.

Будучи американцем, Бентер был в этом мире посторонним, однако он нашел другой способ получить инсайдерскую информацию — тот, который жучки упустили, хотя он прятался прямо здесь, в помещениях жокей-клуба. Бентер собрал копии справочников-ежегодников с результатами забегов и нанял двух женщин, чтобы вводить эти данные в компьютер. Затем было то, что журнал Bloomberg Businessweek назвал прорывом. Он взял коэффициенты ставок, также имевшиеся в жокей-клубе, и их оцифровал. Именно

они позволили Бентеру применить метод, аналогичный тому, что я показывал Яну и Мариусу: использовать уравнение ставок. Это был ключ к нахождению неточностей в предсказаниях игроков и прогнозистов*.

Бентер на этом не остановился. В основном уравнении, представленном выше, я ограничился выявлением ошибок в букмекерских коэффициентах для футбола. Теперь, при втором или третьем чтении его статьи, я начал понимать, как Бентер оказывался с прибылью за такой долгий промежуток времени. В своей модели я не рассматривал никаких дополнительных факторов, которые помогали бы мне прогнозировать результат матча. Но Бентер делал все возможное и невозможное. Его быстро разраставшиеся данные включали прошлые выступления, возраст лошади, вклад жокея, стартовый номер, местную погоду и множество других факторов. Каждый из них постепенно вносил свой вклад в уравнение ставок. По мере того как он включал все больше подробностей, точность его логистической регрессии и, соответственно, прогнозов увеличивалась. После пяти человеко-лет ввода данных модель была готова, и Бентер начал играть на скачках на ипподроме «Хэппи-Вэлли» с капиталом, заработанным в казино**.

За первые два месяца игры Бентер получил 50% прибыли от инвестиций, но еще через два месяца она исчезла. В течение двух следующих лет прибыли Бентера прыгали вверх и вниз — то взлетая до 100%, то падая почти до нуля. Примерно через два с половиной года модель действительно начала окупаться***. Прибыли поднялись до 200, 300, 400% и далее росли экспоненциально. Бентер сообщил Bloomberg Businessweek, что в сезоне 1990/91 выиграл 3 миллиона долларов⁷. То же издание оценивало, что за следующие два десятилетия Бентер и его немногочисленные конкуренты,

* Модель работала и без этих старых коэффициентов (по сути, вычисляла коэффициенты с нуля) и даже давала небольшую прибыль. Однако главную выгоду принесло именно сочетание коэффициентов и корректирующего их алгоритма.

** Бентер входил в состав команды, игроки которой отслеживали все вышедшие карты и за счет этого контролировали вероятность появления следующей карты. Команда получала крохотное преимущество, которого хватало, чтобы заработать. Однако в итоге Бентер с коллегами попали в черный список, и им было запрещено появляться в казино.

*** Модель стала всерьез окупаться как раз тогда, когда Бентер учел коэффициенты ставок.

использовавшие те же методы*, заработали на ипподромах Гонконга свыше миллиарда долларов**.

Самым примечательным в научной статье Бентера было не содержание, а то, что ее мало кто прочитал. За 25 лет после публикации в других научных журналах на нее сослались всего 92 раза. Для сравнения: когда я 15 лет назад написал статью, как муравьи рода *Temnothorax* выбирают новое жилье, на нее сослались 351 раз.

Игнорируется не только статья Бентера. Он ссылается на публикацию Рут Болтон и Рэндалла Чепмена 1986 года, называя ее «обязательным чтением» для его собственной работы⁸. Однако прошло почти 35 лет, а эта вдохновенная статья, показывающая, как можно получать деньги на американских ипподромах с помощью уравнения ставок, также цитировалась менее 100 раз.

У Бентера не имелось образования в области высшей математики***, но он был готов делать то, что требовалось. Его описывали как гения, но я так не считаю. Я знаком со многими нематематиками и негениями, которые настойчиво изучали те же статистические методы, что использовал Бентер. Обычно это не игроки. Это биологи, экономисты и социологи, которые используют статистику для проверки гипотез. Но они нашли время, чтобы понять математику.

Я никогда не улавливаю математические рассуждения при первом чтении. И я видел очень мало профессиональных математиков, которые могли бы прочитать и усвоить уравнения, не возвращаясь потом к ним детально. А именно в деталях и кроются секреты.

*

Самая большая угроза для любого тайного общества — раскрытие. Современная версия заговора иллюминатов, которая рисует технически подкованных правителей мира, требует, чтобы абсолютно

* Крупным игроком на гонконгских ипподромах был бывший партнер Бентера по играм в казино Аллан Вудс, с которым тот поссорился в первые годы работы над моделью. В итоге они использовали разные модели и ставили по отдельности (хотя часто привлекали одних и тех же людей в качестве помощников — такие ставки требовали участия множества людей).

** Сам Бентер соглашается с суммой около миллиарда, но говорит, что существенная часть денег ушла его партнерам в Гонконге и США.

*** Он покинул колледж ради игры в казино.

все участники общества молчали о его целях и методах. Если хотя бы один поделится кодом или разгласит планы общества, вся система окажется под угрозой.

Эта опасность раскрытия — главная причина, почему большинство ученых не верят в существование организаций, подобных иллюминатам. Контроль за всей человеческой деятельностью потребует большого общества и масштабной секретности. Риск того, что кто-нибудь расколется и все расскажет, крайне велик.

Однако когда мы погружаемся в уравнение ставок, то видим, что секрет «Десятки» скрыт у всех на виду. Только когда члены общества упорно учатся, он медленно открывается тем, кто его ищет. Этому коду учат во всех школах, а затем в университетах, но мы не понимаем, что мы изучаем. Члены общества лишь смутно осознают, что они участники этого обширного заговора. Они чувствуют, что им нечего раскрывать, нечего признавать и нечего прятать.

Когда молодой перспективный участник «Десятки» читает научную статью Бентера второй и третий раз, он заставляет себя понять ее правильно. Он начинает ощущать связь, простирающуюся на десятилетия и столетия. Бентер, несомненно, чувствовал ее, когда читал работу Рут Болтон и Рэндалла Чепмена. А те до него испытывали то же чувство, когда изучали статью Дэвида Кокса, который предложил в 1958 году логистическую регрессию, ставшую основой их работы. Связь, выкованная математикой, тянется в историческое прошлое, через Мориса Кендалла и Рональда Фишера, работавших между мировыми войнами, и до первых идей о вероятности, высказанных Абрахамом де Муавром и Томасом Байесом в Лондоне XVIII века.

По мере углубления в детали наш юный последователь видит, как перед ним шаг за шагом раскрываются все секреты. Бентер записал истоки своего успеха в коде уравнений. И теперь, спустя 25 лет, можно постепенно разобрать этот успех — один алгебраический символ за другим.

Именно математика, общий интерес к этому уравнению, объединяет нас в пропасти времени и пространства. Как и Бентер до

него, наш юный ученый начинает исследовать красоту размещения ставки не с помощью ощущений, а на основании статистической взаимосвязи данных.

*

Можно объяснить идею, которая стоит за разработанной нами с Яном и Мариусом стратегией ставок, даже без уравнений, одной фразой: мы обнаружили, что коэффициенты открытия для чемпионата мира (те, которые букмекеры предлагают задолго до начала матча) можно использовать для более точного прогнозирования результатов, чем коэффициенты закрытия (те, которые букмекеры предлагают прямо перед матчем).

Это наблюдение противоречит интуиции. Когда букмекеры устанавливают свои коэффициенты, до начала игры может произойти много неопределенных событий. Травма ведущего игрока (это случилось с Мохаммедом Салахом из сборной Египта), плохая форма целой команды (за несколько недель до чемпионата мира Франция сыграла вничью с командой США), смена тренера в последний момент (как произошло со сборной Испании). Теоретически для отражения всех таких событий коэффициенты должны меняться: если Испания внезапно уволит своего тренера, ее шансы на победу над Португалией упадут.

Коэффициенты действительно меняются, но при этом они скорее не отражают новую реальность, а завышаются. По мере приближения матча на рынке появляются игроки-любители, которые пытаются спрогнозировать результаты матчей, и коэффициенты букмекерских компаний меняются, чтобы отразить их ставки. Например, коэффициенты на победу Франции над Перу увеличились с 2/5 до 1/2 до их матча на групповой стадии турнира. Возможно, некоторые игроки полагали, что если Франция не смогла обыграть США в товарищеском матче, то Перу заработает одно очко, а то и три. Другие любители, несомненно, читали в газетах критику в адрес звездного полузащитника Поля Погба и начали сомневаться в его способностях привести сборную своей страны к успеху. Какими бы ни оказались причины, это был именно тот сценарий, который — как установила наша модель — давал

удачные ставки на предыдущих чемпионатах мира. Когда коэффициенты на записных фаворитов повышаются, выгодно их поддержать. Наша автоматизированная система обнаружила такое их изменение, активировала функцию ставки и поставила 50 фунтов на Францию. После матча мы получили 75 фунтов. Просто и эффективно.

Важный навык для специалистов по прикладной математике — объяснить логику в основе используемых моделей. Когда мы с Мариусом после запуска модели смотрели футбол, то обсуждали, почему по мере приближения к чемпионату мира коэффициенты становятся менее точными.

— Большинство наших стратегий основано на идее, что к матчу коэффициенты становятся гораздо точнее, — сказал он. — С чемпионатом мира должно быть что-то особенное.

— Большой объем ставок, — предположил я. — По телевизору сплошной футбол, интересно попробовать. Одни ставят деньги на свою страну из национальной гордости, другие — против какой-то страны.

Мариус согласился. Чемпионат мира дал футболу новую аудиторию, она не может устоять перед искушением поставить деньги на своих любимцев. По нашим прикидкам, верные английские фанаты считают, что было бы здорово заработать немного за счет французов. Мы подумали, как аргентинцы и немцы ставили на Швейцарию, игравшую свой первый матч с Бразилией. Когда на аутсайдеров хлынули деньги, букмекеры увеличили коэффициенты для фаворитов, и наша модель, пойдя против толпы, оказалась выгодной. Не каждый матч давал нам выигрыш (бразильцы неожиданно сыграли со швейцарцами вничью), но история показала, что ставка на серьезных фаворитов прямо перед началом игры, вероятнее всего, приносит прибыль.

Пристрастие любителей ставить на маловероятные события было только частью нашей модели. Уравнение давало более тонкие прогнозы: значения $\alpha = 1,16$ и $\beta = 1,25$ говорили, что в случае не совсем явного фаворита нам нужно, наоборот, поддержать аутсайдера — как в 2014 году, когда Англия проиграла Уругваю. Хорошим примером такого прогноза был матч между Колумбией

и Японией. За дни перед ним коэффициенты на победу Колумбии выросли с 7/10 до 8/9. Подставив эти числа в наше уравнение, можно прийти к выводу, что стоит поставить на Японию. Не потому, что у нее было больше шансов на победу в матче (фаворитом по-прежнему оставалась Колумбия), просто уравнение говорило, что теперь на Японию, коэффициенты для которой стали 26/5, ставить выгоднее. И мы оказались правы: Колумбия проиграла, а мы заработали 260 фунтов, поставив 50.

*

Сэру Дэвиду Коксу сейчас 95 лет, и он никогда не прекращал трудиться. За свою 80-летнюю карьеру Кокс написал 317 научных работ, и очень вероятно, что напишет еще. В своем офисе в Нраффилд-колледже в Оксфорде он продолжает писать комментарии и обзоры современной статистики, а также вносить новый вклад в эту область.

Я спросил его, каждый ли день он появляется в офисе.

— Не каждый, — ответил Дэвид. — В субботу и воскресенье не появляюсь.

Затем сделал паузу и поправился:

— Точнее, вероятность того, что я приду на работу в субботу или в воскресенье, отлична от нуля. Такое может произойти.

Сэр Дэвид Кокс любит точность. Его ответы на мои вопросы были осторожными и обдуманными; математик всегда оговаривал уровень уверенности в своей способности дать ответ.

Именно Кокс открыл уравнение ставок. Сам он, правда, никогда бы так не выразился, и это в любом случае не совсем верно. Точнее было бы сказать, что он разработал теорию логистической регрессии, которую я использовал для нахождения α и β , а Бентер — для определения факторов, влияющих на исход скачек⁹. Дэвид Кокс разработал статистический метод, благодаря которому уравнение ставок дает точные прогнозы.

Логистическая регрессия была продуктом послевоенной Британии. К финалу Второй мировой войны Дэвид закончил изучать математику в Кембридже, и его направили на работу в королевские BBC. Затем он перешел в текстильную промышленность, поскольку

в Великобритании начался процесс восстановления*. Кокс рассказал мне, что изначально его интересовала чистая математика, которую он изучал, но такая работа привлекла его внимание к новым задачам. «Текстильная промышленность была полна увлекательных математических проблем», — сказал он.

Сэр Дэвид признавал, что смутно помнит детали, но лучился энтузиазмом, говоря о тех временах. Он рассказал, как можно использовать тесты для различных характеристик материалов, чтобы определить вероятность их разрушения, и о проблемах, связанных с созданием более прочного и более однородного конечного продукта из грубо пряденой шерсти. Кроме того, в BBC он столкнулся с задачами, касавшимися частоты аварий и аэродинамики крыла. Это также дало ему много пищи для размышлений.

Именно из таких практических соображений у Дэвида Кокса зародился более общий, математический вопрос: каким образом лучше всего прогнозировать, как разные факторы (скорость ветра или напряжения в материале и другие) могут повлиять на что-то — например, на авиационную катастрофу или на то, порвется ли одеяло. Это вопрос того же типа, который Бентер задавал о лошадях: как зависит вероятность победы четвероногого от истории его прошлых выступлений и от погоды.

— Когда я создавал эту теорию [в середине 1950-х], самые серьезные разногласия в университетах касались анализа медицинских и психологических данных, предсказания того, как разные факторы связаны с медицинским результатом, — рассказывал Кокс. — Логистическая регрессия возникла в результате синтеза моего практического опыта и математического образования. Все известные мне проблемы медицины, психологии и промышленности можно было решить с помощью одного семейства математических функций.

Это семейство оказалось гораздо полезнее, чем предполагал даже сам учений. Начиная с промышленности 1950-х, когда логистическая регрессия играла важную роль в интерпретации результатов медицинских исследований, она успешно применяется

* В обоих случаях — в исследовательские организации.

к бесчисленным математическим задачам. Именно этот подход использует Facebook, чтобы определить, какую рекламу нам показать, а Spotify — чтобы рекомендовать нам музыку. Он же стал частью системы идентификации пешеходов в автомобилях без водителей. И конечно, он используется в азартных играх...

Я спросил у сэра Дэвида, знает ли он об успехах Бентера в применении логистической регрессии к скачкам. Он не слышал. Тогда я поведал ему, как логистическая регрессия принесла миллиард долларов, а также об оксфордском студенте Мэттью Бенхэме и его успехе в предсказании результатов футбольных матчей.

— Предпочту сказать, что вам не следует играть в азартные игры, — сказал он мне, а затем надолго задумался.

Потом Кокс начал вспоминать об одном из своих коллег в 1950-х. Он взял с меня обещание никогда не повторять эту историю, и я сдержу слово.

*

Ставки — это не прогнозирование будущего с определенностью. Это определение мелких различий между тем, как смотрите на мир вы и как смотрят на него другие. Если ваш взгляд чуть остree, а ваши параметры лучше объясняют данные, у вас есть преимущество. Не ждите, что оно появится сразу. Его нужно наращивать постепенно, методом проб и ошибок, улучшая оценки своих параметров. И не думайте, что будете постоянно побеждать. При регулярной игре у вас получится выигрывать лишь чуть чаще, чем проигрывать.

Иногда мы склонны сосредоточиваться на одной «грандиозной идее». Однако уравнение ставок говорит нам, что ключ к успеху — создать различные вариации нашей идеи. Представьте, что вы начинаете собственные уроки йоги или танцев. Пробуйте различные плей-листы с разными группами и отмечайте, какие композиции получают наилучший отклик. Проверяя множество мелких идей, мы сравниваем их между собой, как лошадей на ипподроме «Хэппи-Вэлли». В конце каждой гонки можем заново оценить победителей и неудачников и посмотреть на те свойства, которые привели к успеху или к неудаче.

Если вы начинаете проверять новую идею, вам нужно выполнять то, что именуется А/В-тестированием. Когда Netflix обновляет дизайн своего сайта, компания создает две или несколько версий (A, B, C и т. д.) и показывает их разным пользователям. Затем создатели смотрят, какой вариант наиболее привлекателен для людей. Это самое прямое применение уравнения ставок к «успеху» и «неудаче» параметров дизайна. Когда Netflix получает так много информации, она может быстро представить четкую картину, что работает, а что нет.

Вам не нужна логистическая регрессия, чтобы начать пользоваться уравнением ставок. Но как только вы поняли принцип настройки параметров для улучшения соответствия данным, рукой подать до изучения самого метода. Сэр Дэвид Кокс сказал мне, что, по его мнению, большинство людей могут и должны научиться использовать разработанный им метод. Он утверждал, будто для понимания того, что раскрывает логистическая регрессия в собранных вами данных, незачем вдаваться во все математические подробности ее работы.

*

Во время чемпионата мира я смотрел много матчей, но понятия не имел, принесет ли мне деньги тот или иной исход, потому что не следил за коэффициентами. Я просто наслаждался играми. Раз за разом Ян присыпал мне автоматически создаваемые таблицы со списками сделанных ставок и заработанных или проигранных денег. Во время первого круга групповой стадии мы проиграли, но затем начали выигрывать; и по ходу турнира наши прибыли стали превышать потери. К концу чемпионата мира я заработал примерно 200 фунтов при общем объеме ставок в 1400 фунтов, уровень доходности составил 14%.

После изучения очередной обновленной таблицы с нашими результатами я снова видел письма в своем почтовом ящике, которые по ходу чемпионата становились всё более отчаянными: «Пожалуйста, я знаю, что вы разбираетесь в футболе и правильном счете, не могли бы вы помочь мне, пожалуйста?»; «Я хотел бы следовать всем вашим прогнозам для правильных ставок, потому что я много

проиграл у букмекеров»; «Сегодня в стране разыгрывается футбольный джекпот с кучей денег. Помогая мне выиграть, вы поможете еще сотне людей за мной». Такие сообщения шли почти ежечасно.

Я не мог отделаться от мысли, что наша маленькая прибыль поступала из их маленьких карманов. Конечно, самая большая доля шла букмекерам, но те деньги, что заработали мы с Яном и Мариусом, когда-то принадлежали другим людям — возможно, тем, у кого их изначально было немного.

Тогда в моей голове зародилась идея: неравенство между теми, кто знает уравнения, и теми, кто их не знает, не ограничивается азартными играми. Статистические модели эра Дэвида Кокса работают во многих сферах современного общества. Математические методы привели к прогрессу и легли в основу технологий в самых разных областях — от шерстяной промышленности и проектирования летательных аппаратов до современной науки о данных и искусственного интеллекта. Этот прогресс контролируется очень небольшой группой — теми, кто знает уравнения. И часто именно люди, обладающие этими секретами, извлекали социальную и финансовую пользу из математики.

Дэвид Кокс — член «Десятки». Он не знает об этом, но он изобрел одно из уравнений и полностью понимает остальные девять. Его позиция в истории «Десятки» незыблема. Он заслуженный участник самого высокого уровня.

Бентер, Бенхэм и Блум тоже часть «Десятки». Может, они не знают уравнений на том формальном математическом уровне, на котором их понимает Кокс, но они разбираются в принципах и понимают, как применить их на практике. Ян и Мариус на пути вступления туда.

А я? Я знаю эти десять уравнений в чистом незамутненном виде — как их понимают ученые. Я также знаю их в практическом смысле, в котором их использовал Бентер. И хотя я этого не осознавал ранее, я понял, что «Десятка» определяет меня не только как работника, но и как личность.

Глава 2

Уравнение суждений

$$P\{M | D\} = \frac{P\{D | M\} \cdot P\{M\}}{P\{D | M\} \cdot P\{M\} + P\{D | \bar{M}\} \cdot P\{\bar{M}\}}$$

Мой друг Марк руководит командой финансовых трейдеров, и все они владеют математическими методами. Марк заметил, что у лучших из них есть общая черта — способность реагировать на новую информацию и обрабатывать ее. Когда происходят какие-то события, они быстро подстраиваются к новой реальности.

Эти люди рассуждают не в абсолютных характеристиках — «та компания получит прибыль в следующем квартале» или «тот стартап провалится», — а в терминах вероятностей: «компания получит прибыль с вероятностью 34%» или «для этого стартапа риск неудачи составляет 90%». Когда поступает новая информация — например, СЕО вынужден уйти в отставку или бета-версия, выпущенная стартапом, пользуется успехом, — они корректируют эти вероятности: 34% превращаются в 21%, а 90% — в 80%.

Аналогичные истории я слышал и от Джеймса, знакомого из индустрии азартных игр. Там в ходу варианты уравнения ставок, но при таком количестве денег на кону приходится принимать быстрые решения о том, годится ли их модель для предстоящих футбольных матчей. Что делать, если за час до игры стартовый состав команды меняется и предположения, лежащие в основе модели, становятся недействительными?

— Именно в такие моменты вы узнаете, кто действительно хороший специалист, — говорил Джеймс. — Он не реагирует резко. При одном изменении в стартовом составе ставка не меняется; при двух — четырех специалист оценивает разные возможности; при пяти или больше все ставки снимаются.

Чтобы научиться думать как эти аналитики, вам нужно ставить себя в эмоционально напряженную ситуацию. К примеру, на земле большинство из нас понимают, что полеты не опасны: вероятность попасть в авиакатастрофу со смертельным исходом не превосходит 1 на 10 миллионов. Но в воздухе все ощущается иначе.

Представьте, что вы опытный путешественник, уже летавший сотню раз. Однако этот рейс иной. При снижении самолет начинает грохотать и трястись — такой болтанки вы еще не ощущали. Женщина рядом хватает воздух ртом; мужчина, сидящий через проход, стискивает колени. Все вокруг явно напуганы. Что это? Может ли разворачиваться наихудший сценарий?

В подобных ситуациях математик глубоко вдохнет и сберет всю доступную информацию. Назовем катастрофой худший сценарий — крушение со смертельным (для вас) исходом. Обозначим его вероятность как $P\{\text{катастрофа}\}$. По статистике, $P\{\text{катастрофа}\} = 1/10\,000\,000$, или 1 на 10 миллионов¹⁰.

Чтобы понять, как события зависят друг от друга, обозначим $P\{\text{тряска}|\text{катастрофа}\}$ — вероятность, что самолет трястется, *при условии*, что произойдет катастрофа: «тряска» означает «самолет трястется», а вертикальная линия — «при условии». Сделаем разумное предположение, что $P\{\text{тряска}|\text{катастрофа}\} = 1$, то есть перед любой катастрофой самолет трястет.

Нам также необходимо знать $P\{\text{тряска} \mid \text{не катастрофа}\}$ — вероятность болтанки при безопасном приземлении. Здесь вам придется опираться на свои ощущения. Раз это самый страшный рейс из сотни ваших полетов, то лучшей оценкой будет $P\{\text{тряска}|\text{не катастрофа}\} = 1/100$.

Эти вероятности полезны, но вы желаете знать не их. Вам нужна величина $P\{\text{катастрофа}|\text{тряска}\}$, или вероятность того, что произойдет крушение, при условии, что самолет так трястется. Эту величину можно найти с помощью теоремы Байеса.

$$P\{\text{катастрофа} \mid \text{тряска}\} = \\ = \frac{P\{\text{тряска} \mid \text{катастрофа}\} \cdot P\{\text{катастрофа}\}}{P\{\text{тряска} \mid \text{катастрофа}\} \cdot P\{\text{катастрофа}\} + \\ + P\{\text{тряска} \mid \text{не катастрофа}\} \cdot P\{\text{не катастрофа}\}}.$$

Символ \cdot означает умножение. Вскоре я объясню, откуда появляется это уравнение, а пока просто примем его. Оно было рассмотрено преподобным Томасом Байесом в середине XVIII века и с тех пор используется математиками*. Подставив все нужные числа в наше уравнение, мы получаем:

$$P\{\text{катастрофа} \mid \text{тряска}\} = \\ = \frac{1 \cdot \frac{1}{10000000}}{1 \cdot \frac{1}{10000000} + \frac{1}{100} \cdot \frac{9999999}{10000000}} \approx 0,00001.$$

Даже если это самая сильная болтанка, которую вы когда-либо испытывали, шансы погибнуть составляют 0,00001. Вы благополучно приземлитесь с вероятностью 99,99999%.

То же рассуждение применимо к целому ряду различных, казалось бы, опасных ситуаций. Даже если во время купания на австралийском побережье вам кажется, что вы видите в воде нечто пугающее, вероятность того, что это акула, крохотная. Вы можете волноваться, когда ваши близкие возвращаются поздно домой, а вам не удается с ними связаться, но вероятнее всего, что они просто забыли зарядить телефон. Многое из того, что мы считаем новой информацией — тряска самолета, неясные фигуры в воде или отсутствие звонков, — не так уж страшно, если подходить к проблеме правильно.

* В своей работе «Очерки к решению проблемы доктрины шансов» Томас Байес рассмотрел одну конкретную задачу с шарами, бросаемыми на стол, и поставил вопрос: если нечто произошло или не произошло определенное количество раз, какова вероятность, что это произойдет в следующий раз? Формально это не было теоремой. Имя Байеса не связывалось с таким подходом до XX века.

Теорема Байеса позволяет вам верно оценивать важность информации и сохранять спокойствие, когда все вокруг паникуют.

*

Я смотрю на мир способом, который именую кинематографическим: часто (и один, и даже в компании) прокручиваю в голове фильмы о своем будущем. Это не один фильм или одно будущее; это много фильмов с разными поворотами сюжета и концовками. Объясню на примере самолета.

Когда я взлетаю и приземляюсь на самолете, то вижу катастрофу, которую описал выше. Если лечу с семьей, то представляю, как держу руки детей, говорю, что люблю их, чтобы они не беспокоились. Я воображаю, как мы держимся вместе, когда падаем навстречу смерти. Когда я лечу один, а вокруг только неизвестные, то смотрю другой фильм: наблюдаю целые годы, которые моя семья проведет без меня. Похороны проходят быстро, и я вижу, как моя жена в одиночку справляется с детьми и рассказывает им историю о нашей совместной жизни. Этот фильм неописуемо печален.

Подобные картины крутятся непрерывно и параллельно в какой-то области мозга выше левого глаза. По крайней мере, я так ощущаю. Большинство фильмов не настолько драматичны. У меня встреча с редактором книги, и в голове проигрывается дискуссия: я обдумываю, что ей скажу. Провожу какой-нибудь семинар и вижу, как излагаю материал, представляя сложные вопросы. Многие такие фильмы абстрактны: я прохожусь по научной статье, которую пишу; смотрю на структуру диссертации одного из моих аспирантов; работаю над какой-то математической задачей. Кинокартины такого рода не будут смотреться на большом экране. Они набиты цифрами, техническими терминами и научными ссылками. Мне они нравятся, но я очень специфическая аудитория.

Хочу убедиться, что у вас не сложилось впечатления, будто я воображаю себя всевидящим оракулом. Вовсе нет. Фильмы, которые я создаю, фрагментированы. Нет деталей, нет наполненности реальностью. И, что крайне важно, почти все они не соответствуют

действительности. Редактор книги ведет дискуссию в другом направлении, и я забываю все свои вопросы. В цепочке рассуждений в научной статье возникает дыра, и я не могу ее исправить. В первой строке моих математических выкладок обнаруживается ошибка в вычислениях, и результаты оказываются неверными.

Психологи изучали методы, с помощью которых люди смотрят на мир и конструируют повествования о будущем, однако научное описание процесса здесь не главное. Важно ваше представление о том, как *вы* видите будущее. В форме текстов, фильмов или компьютерных игр? Фотографий, звуков или запахов? Это абстрактное ощущение или вы визуализируете реальные события? Постарайтесь определить метод своих размышлений о вещах. Ваш взгляд на мир должен оставаться вашим личным — я не хочу его менять. Я бы не вынес, если бы кто-нибудь предложил мне отключить мои фильмы. Мой «кинематограф» — часть меня.

Математика участвует в моем мышлении, помогая мне организовать мою коллекцию фильмов. Авиакатастрофа — хороший пример. Когда прокручиваю фильм с катастрофой, я также оцениваю вероятность того, что такое событие произойдет, и считаю ее обнадеживающе низкой. Это не останавливает прокрутку фильма. Я по-прежнему боюсь летать или плавать в океане, но это помогает мне сосредоточиться. Вместо того чтобы бояться, думаю о том, как много значит для меня семья и почему мне нужно меньше путешествовать, а больше плавать в океане.

Научный термин для фильмов, которые я проигрываю в голове, — «модель». Крушение самолета — модель, нападение акулы — модель и план исследовательской работы — тоже. Модели могут быть чем угодно — от смутных мыслей до более формально определенных уравнений, подобных тому, что я создал для ставок на чемпионате мира. Первый шаг к математическому подходу к миру — понять, как мы используем модели.

*

У Эми начался новый курс в колледже, и она задается вопросом, с кем стоит дружить, а от кого держаться подальше. Она доверчивый человек, и ее «фильм» в голове — о людях, которые рады ей

и ведут себя мило. Но Эми не совсем наивна. Она уже знает, что не все люди приятны, и «стервозный» фильм в ее голове тоже есть. Не судите Эми за термины — в конце концов, она держит мысли при себе. Именно поэтому, когда Эми знакомится с Рэйчел — девушкой, сидящей рядом, — она полагает, будто шансы на то, что Рэйчел окажется стервой, довольно малы: скажем, примерно 1 из 20.

Я не думаю, что Эми при встрече с другими определяет вероятность «стервозности». Я определяю это число, чтобы вы лучше понимали ситуацию. Можете на секунду задуматься и решить, какая доля окружающих стервона. Надеюсь, она меньше 1/20, но вы можете выбрать свои числа.

В первое утро Рэйчел и Эми вместе разбирают лекцию. Эми плохо разбирается в деталях, потому что в ее школе не давали знаний, необходимых для понимания концепций, о которых говорит преподаватель. Рэйчел выглядит терпеливой, но Эми осознаёт, что та несколько раздосадована. Почему Эми не может учиться быстрее? Затем, после обеда, происходит ужасное. Эми сидит в одной из кабинок в туалете, рыщет по интернету с помощью смартфона и думает о своих делах. Она слышит, как входят Рэйчел и еще одна девушка.

— Эта новая девица тупа, — говорит Рэйчел. — Я пытала ее объяснить «культурную апоприацию», а она ничего об этом не знала. Она думала, что это когда белые люди учатся играть на бонго!*

Эми сидит неподвижно, не издает ни звука и ждет, пока девушки уйдут. Что она должна думать?

Большинство из нас на месте Эми огорчились бы, разозлились, а может, и то и другое. Но стоит ли? Да, Рэйчел явно поступила неправильно. У Эми первый день учебы, и лить грязь на человека нехорошо. Но вопрос в другом: должна ли Эми, несмотря на этот проступок Рэйчел, простить ее и дать ей еще один шанс?

Да. Да. Да. Должна! Обязана! Мы должны прощать такое. И не раз, а много. Нужно прощать людей за глупые комментарии, за то, что они обзываются за нашей спиной и не замечают нашего присутствия.

* Культурная апоприация — принятие одной культурой элементов другой. Многие считают это вредным, полагая, что происходит эксплуатация культуры, из которой совершается заимствование. Бонго — латиноамериканский ударный музыкальный инструмент, род барабана.

Но почему? Потому что мы хорошие? Потому что всегда позволяем наступать на себя? Потому что слабы и не можем за себя постоять?

Нет. Нет. Нет. Вовсе не поэтому. Нам нужно прощать их, потому что мы рациональны и верим в логику и разум. Мы хотим быть справедливыми. Мы знаем теорему преподобного Байеса. А второе уравнение говорит нам, что это — *единственно верное действие*.

И вот почему. Теорема Байеса — связь, которую нам нужно установить между моделью и данными. Она позволяет нам проверить, насколько хорошо наши картинки соотносятся с реальностью. В примере, который мы разбирали в начале главы, мы рассматривали вероятность $P\{\text{катастрофа}|\text{тряска}\}$ того, что самолет разобьется, при условии, что он попал в болтанку. Эми желает знать $P\{\text{стерва}|\text{грубость}\}$, и логика здесь та же.

Катастрофа и стерва — модели в наших головах. Это наши представления о мире, которые принимают форму мыслей или (в моем случае) фильмов. Тряска и грубость — данные, которые есть в нашем распоряжении. Это нечто осязаемое, то, что происходит, что мы можем ощущать. Значительная часть прикладной математики включает сопоставление моделей с данными, столкновение наших мечтаний с суровой реальностью.

Будем использовать букву M для модели и D для данных. Мы хотим знать сейчас вероятность того, что модель верна (Рэйчел — стерва), при условии истинности данных (грубый комментарий в туалете). Имеем:

$$P\{M|D\} = \frac{P\{D|M\} \cdot P\{M\}}{P\{D|M\} \cdot P\{M\} + P\{D|\bar{M}\} \cdot P\{\bar{M}\}} \quad (\text{Уравнение 2}).$$

Чтобы понять уравнение (формулу Байеса), лучше всего рассмотреть по отдельности компоненты правой части.

Числитель (часть над дробной чертой) — произведение двух вероятностей, $P\{M\}$ и $P\{D|M\}$. Множитель $P\{M\}$ — вероятность того, что модель истинна, до того, как произошло некое событие (статистическая вероятность крушения самолета или оценка Эми, что встреченный ею человек стервозен; в последнем случае — $1/20$).

Эту величину Эми знала прежде, чем отправилась в туалет. Второй множитель, $P\{\Delta|M\}$, касается того, что произошло в санузле. Это вероятность, что Рэйчел оскорбит Эми, если она в самом деле стерва, или — в общем случае — того, что мы наблюдаем некоторые конкретные данные, если наша модель верна. Трудно оценить это число количественно, поэтому будем считать его эквивалентом броска монетки: $P\{\Delta|M\} = 0,5$. Рэйчел не каждый раз при посещении туалета злословит о сокурсницах. Минимум 50% времени стервы тратят на разговоры о чем-то еще.

Мы перемножаем вероятности, чтобы найти вероятность того, что произошли одновременно оба события. Например, если я брошу две игральные кости и хочу найти вероятность выпадения двух шестерок, то я определяю вероятность $1/6$ для выпадения шестерки на первой кости, вероятность $1/6$ для шестерки на второй, а затем перемножаю их и получаю вероятность выпадения шестерок на обеих костях: $1/6 \cdot 1/6 = 1/36$. Тот же принцип применяется и здесь: числитель — вероятность того, что Рэйчел стерва и она отпустила стервозный комментарий при посещении туалета.

Итак, числитель описывает Рэйчел как стерву, но мы должны также рассмотреть альтернативную модель, в которой она хороший человек. Это делается в знаменателе дроби справа. Рэйчел может быть стервой, сделавшей стервозный комментарий (M), или хорошим человеком, допустившим ошибку (\bar{M}). Черта над буквой означает противоположность или дополнение. В нашем случае дополнение к «быть стервой» — «быть не стервой», «быть хорошим человеком». Обратите внимание, что первое слагаемое в знаменателе совпадает с числителем, $P\{\Delta|M\} \cdot P\{M\}$. Второе, $P\{\Delta|\bar{M}\} \cdot P\{\bar{M}\}$, — вероятность того, что Рэйчел сделала свое грязное замечание при условии, что она *не* стерва, умноженная на вероятность того, что она хороший человек. Мы охватили все возможные объяснения тех данных, что получила Эми, сидя в кабинке туалета. Теперь можно найти $P\{\bar{M}|\Delta\}$ — вероятность модели при условии этих данных.

Если Рэйчел не стерва, то она хороший человек, поэтому $P\{\bar{M}\} = 1 - P\{M\} = 0,95$. Теперь нужно определить вероятность, что хороший человек совершает ошибку и говорит гадость. Возможно,

у милой Рэйчел был неудачный день — у всех такое случается. Предположим, что один раз из десяти у хороших людей бывает плохой день и они говорят то, о чем потом жалеют. Иными словами, пусть $P\{\Delta|\bar{M}\}=0,1$ (см. рис. 2).

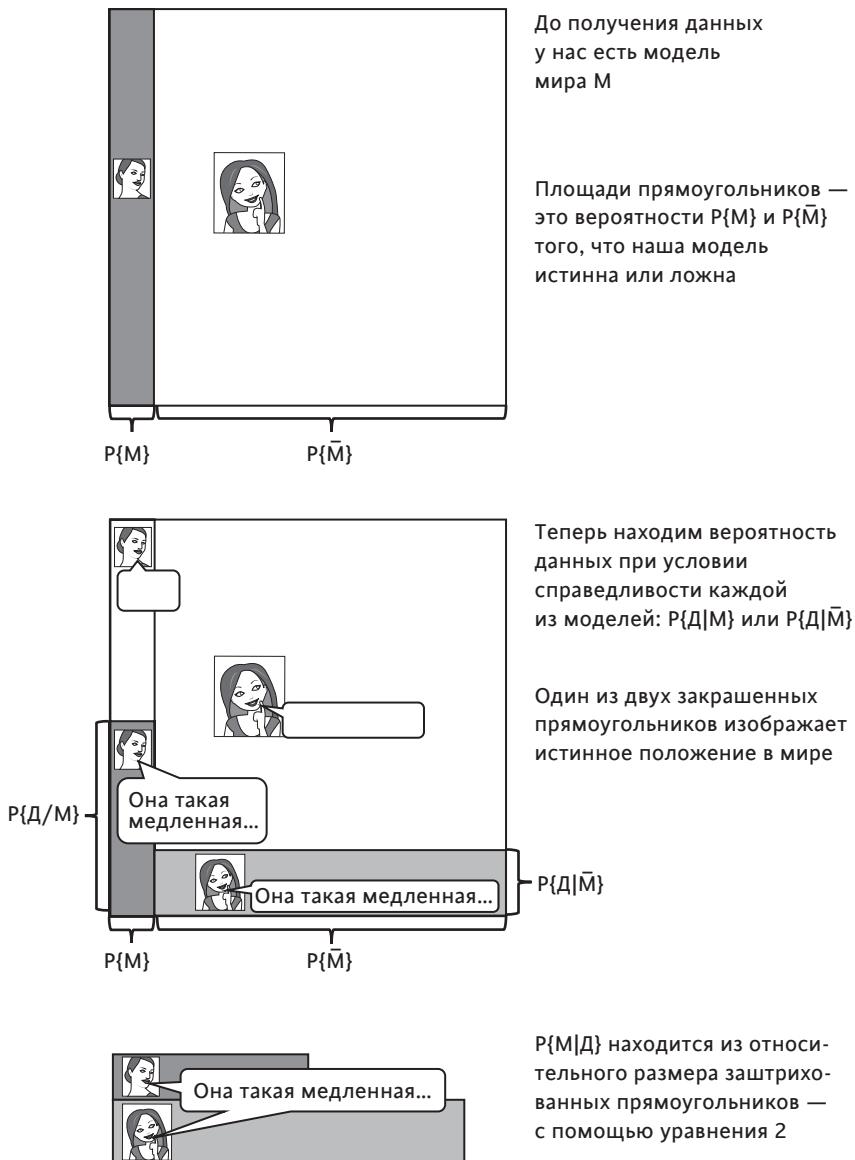


Рис. 2. Иллюстрация теоремы Байеса

Осталось только произвести подсчет — точно так же, как с крушением самолета, но с другими числами:

$$P\{M|D\} = \frac{0,5 \cdot 0,05}{0,5 \cdot 0,05 + 0,1 \cdot 0,95} \approx 0,21.$$

Вероятность того, что Рэйчел — стерва, примерно 1/5. Именно поэтому Эми стоит простить ее. С вероятностью 4/5 она хороший человек. Было бы нечестно судить девушку только по одному действию. Эми не следует упоминать, что она слышала Рэйчел, или допустить, чтобы эти слова влияли на их общение. Лучше ждать и смотреть, что произойдет завтра. С вероятностью 80% к концу года они будут вместе смеяться над этим случаем в туалете.

И еще один совет Эми, съежившейся за дверью туалета. Возможно, в то утро она не работала с максимальной отдачей. Может, в следующий раз ей нужно лучше стараться и уж точно не стоит сидеть в уборной и развлекаться с телефоном. Но помните, что Байес прощает прегрешения. Эми должна применять ту же формулу и к себе. Теорема Байеса предлагает ей медленно корректировать свое мнение о себе и не особо огорчаться из-за конкретных событий.

Вы продукт всех своих действий, а не результат одной или нескольких ошибок. Применяйте к себе то же рациональное прощение, которое Байес просит вас применять к другим.

*

Первый урок, который нужно извлечь из формулы Байеса (уравнения суждений), — не надо торопиться с выводами. Числа, которые я использовал в примере, влияют на результат, но не на саму логику. Вы можете спросить себя: какая доля людей, по-вашему, в целом хорошие? Как часто хорошие люди совершают ошибки? И как часто стервы поступают по-свински? Подставьте свои цифры в уравнение, и вы придетете к тому же заключению: нужно больше, чем один грязный комментарий, чтобы поставить на человеке клеймо «стерва» или «сволочь».

Иногда мой руководитель ведет себя как козел. Иногда моим студентам как будто не хватает сосредоточенности. Порой кто-то из коллег хочет приписать себе мою идею, заявляя, что придумал это первым. Иногда председатель комиссии, где я состою, кажется мне неорганизованным: он тратит мое время на бесполезный не-эффективный обмен электронными письмами. В таких ситуациях я использую уравнение суждений. Но я не вычисляю вероятность того, что каждый из моих коллег — скотина, рассеянный или некомпетентный. И не позволяю разовым событиям определять свои ощущения. Если я вижу, что некто, с кем я работаю, совершил ошибку, я жду развития ситуации. Вполне может оказаться, что неправ был я.

*

Мы не можем понять «Десятку» без раскрытия ее истории и философии. История «Десятки» — рассказ о небольшой группе людей, которые передавали секреты рационального мышления из поколения в поколение. Они ставили масштабные вопросы. Они хотели знать, как мыслить яснее и точнее, уметь оценивать истинность того, что говорят люди. Они даже задавались вопросом, что значит быть истинным или ложным. Это история о действительно важных вопросах: природе реальности и месте этих людей в ней.

Это также рассказ о религии, о том, что такое хорошо и что такое плохо, об этике, о добре и зле.

Наша первая остановка — 1761 год. Валлийский философ Ричард Прайс обнаружил в бумагах недавно умершего друга (того самого Томаса Байеса) эссе, полное математических символов и философских размышлений, и один вопрос звучал так: как на основе данных о предыдущих событиях оценить вероятность того, что подобное произойдет снова? Прайс опубликовал его со своим приложением, где просил читателя представить «человека, только что появившегося в этом мире, который заключает из своих наблюдений за порядком и ходом событий, какие силы и причины в нем действуют». Спрашивается, как такой человек должен рассуждать, увидев восход в первый раз, во второй и в третий. Что он должен сказать о вероятности того, что солнце встает каждый день?

Вывод примечателен. Ежедневный восход солнца не должен привести нашего «только появившегося» человека к выводу, будто солнце будет вставать всегда. Наоборот, его умоляют быть очень осторожным с этим событием — даже после сотни восходов и целой жизни восходов. Ничто не должно быть само собой разумеющимся.

«Только появившемуся» человеку предлагалось дать оценку вероятности ежедневного наступления восхода с помощью некоторого параметра θ . Перед первым восходом человек не имеет никаких априорных представлений о солнце и должен считать все значения параметра θ равновозможными. В этот момент одинаково вероятно, что солнце поднимается каждый день ($\theta = 1$), встает в половине дней ($\theta = 0,5$) или только один раз из ста ($\theta = 0,01$). Величина θ может принимать бесконечное число значений из интервала от 0 до 1 (все вероятности находятся в этом промежутке). Например, она может оказаться 0,8567, 0,1234792, 0,99999 и т. д. При этом число десятичных знаков любое, точность произвольна.

Далее человеку предлагается определить, какое значение тот сочтет минимальной правдоподобной вероятностью того, что солнце восходит ежедневно. Если человек думает, что шансы на восход превышают 50%, то $\theta > 0,5$. Если он считает, что они превышают 90%, то $\theta > 0,9$.

Теперь представим, что человек увидел 100 восходов подряд и пытается сделать из этого вывод о вероятности восхода в один день: он заявляет, что солнце поднимается чаще 99 раз из 100. Иными словами, он дает оценку $\theta > 0,99$. Выражение $P\{\theta > 0,99 \mid 100 \text{ восходов}\}$ определяет вероятность того, что он прав в своей оценке. Байес показал с помощью определенной разновидности уравнения 2, что $P\{\theta > 0,99 \mid 100 \text{ восходов}\} = 1 - 0,99^{100+1} \approx 63,8\%$. Соответственно, с вероятностью 36,2% наш человек ошибается и солнце встает реже, чем он полагает¹¹.

Если человек прожил 60 лет и видел восход солнца каждый день, он мог бы определенно быть уверен, что вероятность восхода солнца каждый день превышает 99%. Но если он желает быть уверенными, что вероятность восхода солнца превосходит 99,99%, мы бы посоветовали проявлять осторожность: $1 - 0,9999^{365 \times 60+1} \approx 88,8\%$.

Так что остается еще 11,2%, что он ошибается. Мы заставляем новоприбывшего обитателя мира определить свою модель, высказать минимальное возможное значение θ , а затем сообщаем ему вероятность, с которой он прав в своем предположении.

Ричард Прайс осознал, что формула Байеса связана со спорами о чудесах, которые происходили в XVIII веке. Как и Байес, Прайс был священником и интересовался, как новые научные открытия того времени могут сочетаться с чудесами, в которые он верил после чтения Библии.

Десятилетием ранее философ Дэвид Юм утверждал, что «никакое свидетельство недостаточно для установления чуда, кроме такого, ложность которого была бы большим чудом, нежели тот факт, который оно стремится установить»¹². Эти слова можно рассматривать как обращение к уравнению суждений. Он просит нас сравнивать модель M , что чудеса происходят, с альтернативной моделью \bar{M} , что они не происходят. Юм говорит, что, поскольку мы никогда раньше не были свидетелями чуда, $P\{\bar{M}\}$ близка к 1, а $P\{M\}$ очень мала, поэтому, чтобы убедить нас в обратном, понадобится очень серьезное чудо, у которого будет очень большая $P\{\Delta|M\}$ и маленькая $P\{\Delta|\bar{M}\}$. Аргументация Юма очень близка к моим рассуждениям о болтанке в начале этой главы: нужны очень веские доказательства, чтобы убедить нас, что в целом надежный самолет разобьется. Нам необходимы очень веские доказательства, чтобы убедить нас, что Иисус Христос воскрес.

Прайс считал, что рассуждения Юма «противоречат разуму»¹³, что Юм неправильно понял Байеса. Он объяснял, что Юму следовало быть более точным, когда он говорит о θ — вероятности чуда. Даже те, кто верит в чудеса, не считают, что они происходят каждый день. Конкретизируем: представьте, что Прайс предложил Юму дать какую-то оценку частоте чудес, а тот говорит, что они случаются реже одного раза в 10 миллионов дней (27 400 лет), что дает $\theta > 99,99999\%$. Предположим, Прайс верит, что $99,99999\% > \theta > 99,999\%$, то есть в то, что чудеса происходят реже, чем раз в 274 года, но чаще, чем раз в 27 400 лет. Теперь представьте, что 2000 лет подряд не происходило ни одного чуда. Вероятность того, что при таком условии прав Юм, составляет 7,04%¹⁴.

Вероятность, что прав Прайс, — 92,89%¹⁵. Даже если за несколько тысячелетий не произошло ни одного чуда, это недостаточно надежно подтверждает, что чудес не существует. Просто за время жизни одного человека не набирается достаточного объема данных, чтобы подтвердить заявление Юма, будто чудес не существует.

Ричард Прайс направил «Десятку» на путь христианской морали. Он верил в воскресение Христа и использовал рациональные аргументы, чтобы поставить под сомнение само сомнение. Прайс полагал, что логическое мышление поможет обнаружить истины о мире, скрытые от нашего повседневного опыта. Бог был одной из этих истин.

Двумя тысячелетиями ранее в своей аллегории греческий философ Платон описал находящихся в пещере людей в оковах, которые смотрят на тени и способны видеть только смутные проекции истинного, более логичного мира снаружи. Аллегорию Платона часто используют для объяснения силы математики, и Прайс воспринимал ее очень серьезно. Он считал, что мы открываем новые истины, когда признаём, что тени на стене пещеры не реальны. Наш повседневный опыт — путаное представление о какой-то большей истине. Размышляя четче об истинной форме мира — посредством моделей, не зависящих от данных, — мы можем яснее рассуждать о путанных ситуациях, о тенях нашей повседневной жизни.

«Десятка», которую предвидел Прайс, формировалась из его религиозных убеждений и метафизики Платона¹⁶. Он считал, что в математике есть мораль, что существует рациональный, правильный подход к жизни. Философ не только проповедовал такое представление, но и претворял его в жизнь. Он составил таблицы ожидаемой продолжительности жизни, которые почти целое столетие использовались для страховых выплат¹⁷. Он считал свою работу способом защиты бедных людей от неопределенности: в ней показывалось, что почти все страховые компании того времени не способны выполнить свои будущие обязательства и должны совершенствовать свою политику¹⁸. Прайс был ярым сторонником американской революции и близким другом Бенджамина Франклина; он полагал, что в Америке есть возможность создать систему, основанную на принципах свободы, где землей владели бы на равных,

а политическая власть справедливо распределялась среди всех людей¹⁹. Согласно видению Ричарда Прайса, США должны были стать страной, где религиозная рациональная «Десятка» могла бы в итоге процветать.

Современные практики «Десятки» редко говорят о морали, меньшинство из них верят в христианского Бога, однако многие унаследовали ценности Прайса: статистик, тщательно вычисляющий взносы автострахования для вашего тестя или свекра; государственный чиновник, планирующий наши пенсии или устанавливающий процентные ставки; научный сотрудник аппарата ООН, описывающий цели развития; климатолог, измеряющий вероятность различных повышений температуры на следующие двадцать лет; врач, балансирующий между рисками и стоимостью лечения. Все они используют байесовские рассуждения, чтобы создать более цивилизованное, справедливое и лучше организованное общество. Они помогают нам делиться рисками и непредопределенностями с другими — и в результате, когда с одним из нас происходит ужасное, но редкое событие, расходы покрывают другие люди.

Уравнение суждений помогает участникам общества действовать на благо всех. Правильные суждения, если смотреть глазами Прайса, требуют, чтобы все мы прощали других и были к ним внимательны. Оно говорит нам, что не стоит отказываться от чудес. Он предполагает, что как минимум одно из десяти уравнений ставит нас на путь праведности.

*

Аудитория затихла, ожидая начала мероприятия. Бьорн нервничает. Последние пять лет жизни он посвятил благородному делу открытия новых истин — научным исследованиям. Я был руководителем, направлявшим его к цели. Сейчас он стоит перед коллегами, комиссией, своей семьей и друзьями и собирается защищать свою диссертацию. Бьорна нервирует сочетание этой разнородной аудитории и сложной темы его работы. Одна из глав диссертации называется «Прошлая ночь в Швеции» и представляет собой исследование связей между насильственными преступлениями

и иммиграцией у него на родине. В другой главе он рассматривает, как «Шведские демократы» (популистская антииммиграционная партия) приобрели за последние десять лет известность в стране, знаменитой своей либеральной и социалистической политикой.

Для математиков из комиссии, сидящих в зале, это диссертация о статистических методах. Для второго научного руководителя, Ранжулы Бали Свайн, профессора экономики, которая работает в разных областях — от устойчивого развития до исследования того, как микрофинансирование помогает женщинам выбраться из бедности, — диссертация Бьорна нацелена на объяснение того, что происходит из-за смешения культур во всем мире. Семья Бьорна, Бломквисты, и его друзья хотят знать, что он выяснил об изменениях в Швеции. Их страна из однородной земли викингов превращается в мультикультурный плавильный котел из афганцев, эритреев, сирийцев, югославов и приехавшего британца.

Бьорн боится упасть с каната, на котором должен балансировать, чтобы осчастливить всех. Защита диссертации в Швеции начинается с выступления оппонента — человека, который должен прочитать работу, обсудить ее с кандидатом и представить исторический контекст в исследуемой области. Оппонент Бьорна — Иэн Вернон из Даремского университета.

Иэн рассказывает нам о принципах байесовского мышления. Мои примеры в этой главе относились к одной модели или использовали один параметр, но обычно ученые имеют дело с несколькими конкурирующими гипотезами. Задача Иэна — рассмотреть все эти альтернативные модели и присвоить им какие-то вероятности. Ни одна гипотеза не верна на 100%, но по мере накопления свидетельств одни становятся правдоподобнее других. Он демонстрирует разные примеры, начиная с поиска нефтяных месторождений. Нефтедобывающие компании используют алгоритм, запатентованный Иэном и его коллегами, для поиска месторождений, которые дают наилучшие долговременные перспективы. Затем он переходит к здравоохранению. Когда исследователи планируют действия против малярии или ВИЧ, они сначала проводят математическое моделирование для прогнозирования результатов своих действий. Фонд Билла и Мелинды Гейтс использует методы

Иэна для планирования своих программ по искоренению заболеваний.

Наконец Иэн переходит к одному из самых масштабных вопросов. Что происходило на первых стадиях существования Вселенной? Как впервые после Большого взрыва образовались галактики и какие модели объясняют размер и форму тех из них, что мы наблюдаем сегодня? Иэн смог уменьшить количество возможных моделей ранней Вселенной, найдя вероятные значения для семнадцати различных параметров, определяющих, как галактики расширяются в космическом пространстве²⁰. Презентация Иэна идеально ложится на аудиторию, демонстрируя мощь математических методов и широчайший спектр приложений. Семья и друзья Бьорна ахают, когда видят моделирование вращения и столкновения галактик, возможную модель начального этапа эволюции нашей Вселенной, параметры для которой были восстановлены с помощью формулы преподобного Байеса.

Приходит очередь Бьорна представить свою работу. Введение в масштабах Вселенной могло легко ошеломить и без того нервничающего аспиранта. Возможно, Бьорн обеспокоен сравнением его исследования всего одной страны в Скандинавии с огромным масштабом работ Иэна. Однако когда я смотрю на своего ученика, то вижу, что он спокоен и готов. А когда оглядываюсь на аудиторию, вижу гордость на их лицах. Бломквисты думают: *вот что можно сделать с помощью математики, которую изучал Бьорн. Вот чем овладел их Бьорн: математикой Вселенной.*

Социальные изменения так же сложны, как происхождение Вселенной, хотя и совершенно иначе. Бьорн показывает, как рост «Шведских демократов» можно объяснить преимущественно географическим положением. Эту партию поддерживают определенные регионы страны, особенно в самой южной области Сконе; однако поддержка есть и в центральной Даларне. Как ни странно, это не районы с самым высоким процентом иммигрантов. Возмущение появляется не потому, что в какую-то область приезжают иммигранты. Скорее, растет поддержка антииммиграционной политики в сельских областях — особенно там, где у людей низкий уровень образования.

После окончания презентации Иэн и комиссия задают Бьорну вопросы. Иэн и другие математики желают знать технические детали того, как Бьорн сравнивал модели и данные. Экономистка Линь Лерпольд, коллега Ранжулы и член комиссии, указывает на некоторые важные ограничения этого исследования. Бьорн не до конца разобрался в причинах антииммиграционных настроений. Он изучил закономерности изменений в местных сообществах, но не понял мышление их представителей. Чтобы ответить на вопросы Линь, нужны подробные беседы и анкеты.

Расспросы комиссии были дотошными, но справедливыми, а решение единогласным. Бьорн защитил диссертацию. Он присоединился к спецназу байесовских ученых.

*

Байесовское мышление изменило принципы науки в последние десятилетия. Оно идеально соответствует научному взгляду на мир. Экспериментаторы собирают данные (Δ), а теоретики разрабатывают гипотезы или модели (M). Формула Байеса объединяет оба эти компонента.

Рассмотрим следующую научную гипотезу: *использование мобильного телефона плохо влияет на психическое здоровье подростков*. В моей семье этот вопрос горячо обсуждается — два подростка (и, если честно, два взрослых) целый день приклеены к экранам. В годы моей юности родители беспокоились, где я и чем занимаюсь. У нас с женой другие проблемы: дети слишком много времени проводят дома, уставившись в телефоны. Чего бы мы только не отдали за старые добрые разборки «Почему не пришел домой вовремя», «С кем встречалась?»...

Доктор Кристина Картер, социолог и автор нескольких книг о воспитании детей и производительности, категорически против чрезмерного использования мобильных телефонов, говоря, что «время за экраном — вероятная причина волны депрессии, тревог и суицида среди подростков». Ее статья в журнале *Greater Good Magazine*, выпускающемся в Калифорнийском университете в Беркли, излагает аргументы в два этапа²¹. Во-первых, Картер ссылается на опрос родителей, почти половина которых считают, что их дети

тинейджеры зависимы от своих мобильных устройств, и 50% озабочены, что это негативно скажется на их психическом здоровье. Во-вторых, она ссылается на данные, полученные при исследовании 120 115 подростков в Великобритании, которые отвечали на 14 вопросов о том, как они чувствуют себя с точки зрения счастья и удовлетворенности жизнью, и об их социальной жизни. Это исследование показало, что при превышении порога в один час в день дети отличались пониженным психическим благополучием, что измерялось с помощью опросника. Иными словами, чем больше дети используют свои телефоны, тем они несчастнее.

Звучит весомо, не так ли? Должен признаться, когда я прочитал эту статью в первый раз, она меня убедила. Написана специалистом со степенью, опубликована в рецензируемом журнале, который связан с университетом с мировым именем, в статье есть данные качественно проведенного опроса, подтверждающие точку зрения автора. Но есть одна проблема, и серьезная.

Кристина Картер заполнила только верхнюю часть уравнения суждений. Ее первый шаг — описание родительских страхов — аналогичен вероятности модели $P\{M\}$: это вероятность того, что родители верят, будто время у экрана влияет на психическое состояние детей. Ее второй шаг — демонстрация, что имеющиеся данные согласуются с гипотезой встревоженных родителей, то есть рассматривается величина $P\{\Delta|M\}$, и она действительно оказывается большой. Но Картер не учла того, что другие модели тоже могли бы объяснить проблемы подростков. Она определила числитель дроби в уравнении 2, но пренебрегла знаменателем. Доктор не сообщает нам $P\{\Delta|\bar{M}\}$ для альтернативных гипотез, и мы не знаем $P\{M|\Delta\}$ — вероятность того, что мобильный телефон объясняет подростковую депрессию, то есть именно то, что мы хотим знать.

Кэндис Оджерс, профессор психологии из Калифорнийского университета в Ирвайне, заполнила пробелы, оставленные Картер. В комментарии, опубликованном в журнале *Nature*, она пришла к совершенно другому заключению²². Она начинает свою статью с признания проблемы. В США доля девочек в возрасте от 12 до 17, сообщавших о случаях депрессии, за период с 2005 по 2014 год

выросла с 13,3 до 17,3%; у мальчиков того же возраста тоже наблюдался рост, хотя несколько меньший. Нет сомнений, что использование мобильных телефонов за этот период увеличилось, здесь нам не нужны статистические данные. Оджерс также не оспаривает данные по британским подросткам, на которые ссылается Кристина Картер и которые указывают на рост случаев депрессии у тех, кто активно пользуется телефоном.

Однако Оджерс указывает, что депрессию у тинейджеров можно объяснить и другими причинами. Например, и отсутствие регулярного завтрака, и сон в разное время оказались втрое более важными факторами для прогнозирования психического состояния подростков, чем пользование телефонами²³. На языке теоремы Байеса завтрак и сон — альтернативные гипотезы, которые могут объяснить депрессию, и для них вероятность $P\{\Delta|\bar{M}\}$ велика. Если их учесть в знаменателе формулы Байеса, это перевесит числитель, и искомая вероятность $P\{M|\Delta\}$ того, что пользование мобильными телефонами связано с депрессией, становится меньше — не пренебрежимо малой, но достаточно маленькой, чтобы не считать ее важным фактором при объяснении психического здоровья подростков.

Более того, есть документально подтвержденная польза телефонов для тинейджеров. Во многих работах показано, что дети используют гаджеты, чтобы поддерживать друг друга и создавать долговременные социальные структуры. У большинства ребят среднего класса — группы, на которую ориентируются советчики по поводу экранного времени, — мобильные телефоны улучшают способность заводить настоящие, прочные дружеские отношения; и не только в интернете, но и в реальной жизни. Как говорит Кэндис Оджерс в своей статье, с проблемами сталкиваются дети из неблагополучных семей. Менее благополучные подростки скорее начнут драться из-за того, что происходило в соцсетях. Те, кого травили в жизни, позже с большей вероятностью станут жертвами и в интернете.

Мои дети общаются с детьми по всему миру и часто узнают в Сети что-то новое. Я слышал, как на прошлой неделе Элиза и Генри обсуждали барабаны бонго и культурную априориацию.

Элиза сказала:

— Это базовое уважение: если кто-то говорит, что его оскорбляет исполнение вами музыки его культуры, то лучше остановиться.

— Тогда Эминем — это культурная априориация? — возразил Генри.

Я точно не мог вести такие разговоры с сестрой, когда нам было тринадцать и пятнадцать. Я не уверен даже, что мы могли бы вести их сейчас. Дети, рожденные в 2000-х, имеют доступ к важным идеям и информации, которые находились вне рамок восприятия тех, кто рос в 1970-е, 1980-е и даже 1990-е.

*

Вернемся к Эми и Рэйчел, поскольку я пропустил кое-что важное.

Числа, которые я использовал в том примере (один человек из 20 стервозный; эти люди тратят на гадости половину времени; даже хорошие люди ведут себя так один день из десяти), не просто произвольны, но и субъективны — свои для каждого человека. В зависимости от вашего жизненного опыта вы будете доверять людям больше или меньше, чем Эми. Это как с авиакатастрофами: ужасная и объективная реальность. Выбор Эми, как смотреть на новых соучеников, или мой способ классифицировать коллег, полностью основан на нашем субъективном опыте общения с людьми, которых мы встречали ранее. Нет никаких объективных способов измерить «стервозность», «сущность» или «козлиность».

Да, числа в истории Эми субъективны. Однако дело вот в чем. Формула Байеса работает не только с объективными вероятностями, но и с субъективными. Она дает нам возможность рассуждать о числах, даже если они не абсолютно точные. Мы можем поменять их и получить другие результаты, но нельзя изменить логику, которую нам рекомендует байесовский подход.

Сделанные предположения называются *априорными*. В уравнении $2 P\{M\}$ — априорная вероятность того, что наша модель верна. Часто такие вероятности можно получить из субъективного опыта. Но не может быть субъективной $P\{M|\Delta\}$, вероятность истинности

нашей модели после того, как мы увидели какие-то данные. Пересчет вероятностей должен производиться с помощью формулы Байеса.

Многие полагают, что математика обязательно имеет дело с объективными вещами. Это не так. Это способ представления мира и рассуждений о нем; и иногда то, о чем мы спорим, известно только нам. В конце концов, никто другой, возможно, никогда не узнает, считает Эми Рэйчел стервой или нет. Этот процесс может быть навсегда скрыт в ее мозге.

Вспомните о моем кинематографическом представлении мира — фильмах, которые я проигрываю себе день за днем, порой очень личных. Это могут быть опасения по поводу того, что чувствует моя жена; мысли о будущем дочери; фантазии о том, как я веду мини-футбольную команду сына к победе в турнире; мечта о том, как в один прекрасный день я стану автором бестселлера. Мне не надо рассказывать вам о них, поскольку они принадлежат исключительно мне. Уравнение суждений не говорит нам, какие фильмы нужно иметь в коллекции или о чем мечтать. Оно просто говорит, как нужно рассуждать об этих мечтах, поскольку каждый из таких «фильмов» — модель мира. Уравнение суждений позволяет нам обновлять вероятности, которые мы связываем с каждой мечтой, но не говорит, какие мечты нужно иметь.

После защиты Бьорна Иэн Вернен за бокалом шампанского сказал мне: «Многие, даже математики и другие ученые, не осознают, что настоящая сила байесовского подхода в том, как он заставляет вас определить, что вы думали до эксперимента и после него. Байесовский анализ требует, чтобы вы разбили свои рассуждения на модели и по очереди искали свидетельства для них. Вы можете считать, что данные подтверждают ваше предположение, но нужно быть честным в отношении того, насколько сильно вы поддерживали свою гипотезу до этого эксперимента».

Я согласился. Иэн говорил в целом, думая при этом о защите Бьорна и использовании байесовской теории для объяснения подъема крайне правых сил в шведской политике. В этом проекте я проработал с Бьорном все детали, изучив факторы, которые заставляют людей голосовать за националистические партии. Теперь

пытался применить тот же подход к вопросу из собственной семейной жизни. Я не специалист по психическому здоровью или мобильным телефонам, но уравнение суждений дает мне способ интерпретировать результаты, полученные другими, — способ оценить относительную ценность аргументов, выдвинутых учеными. Я использовал теорему Байеса, чтобы проверить, соблюдали ли они критерии для правильного суждения. Посмотрели ли они на собственную и альтернативную модель мира? Кэндис Оджерс представила в своих рассуждениях все стороны; Кристина Картер — только одну.

Иногда я крайне разочарован, когда вижу, как некритично воспринимают люди советы так называемых экспертов по воспитанию, образу жизни или здоровью. Как неопытные игроки, которые просят у меня совета только о предстоящем серьезном матче, так и потребители таких советов смотрят не дальше последнего исследования. Они не понимают, что обеспечение здорового и сбалансированного образа жизни требует долгосрочной позиции, как успешная азартная игра — долгосрочной стратегии.

Но не только Кристина Картер ответственна за представление всех аспектов рассматриваемого вопроса. Вам может показаться странной такая моя позиция, поскольку я счел ее работу вводящей в заблуждение; но я также понимаю, что она отражает беспокойство многих родителей, включая меня. Данные, описанные ею, реальны, и она считает, что они подтверждают ее модель. Описывать поддержку альтернативных моделей не совсем ее дело.

Проверять достоверность ее модели — большей частью *наше* дело. Когда я читаю авторские колонки, то стараюсь убедиться, что их создатели — независимо от квалификации — разобрались во всех частях нашего уравнения. Мне как родителю оказалось не-трудно получить более полное представление о роли экранов в нашей жизни. Все статьи, что я использовал, есть в свободном доступе в интернете; и мне понадобилось несколько вечеров собственного времени у экрана, чтобы скачать и прочитать их. Как только разобрался в вопросе, обсудил результаты с детьми. Я сказал им, что и хороший сон, и правильный завтрак втрое важнее для их психического состояния, чем время в телефонах. Я рассказал, что

телефоны отчасти полезны, подчеркнув: это вовсе не подразумевает, что они должны проводить все вечера на диване за просмотром YouTube. Физические нагрузки и социальное общение необходимы, и уж точно не стоит брать телефоны в спальню. Мне кажется, Элиза и Генри всё поняли.

Можно услышать, как те же люди, которые некритично потребляют информацию о воспитании детей, выражают скептицизм, когда слышат более сбалансированную точку зрения от ученых, подобных Кэндису Оджерсу. Когда ученые излагают все стороны проблемы, такие люди считают, что это говорит о неуверенности специалистов. В научном сообществе активно обсуждаются такие темы, как изменение климата, достоинства разных диет или причины преступности. Но обсуждения и сравнение всех возможных гипотез вовсе не признак слабости или нерешительности участников дискуссии. Скорее, это признак силы и основательности. Это знак преимущества, того, что вы рассмотрели все возможности.

*

В мире полно людей, дающих советы. Как быть организованными на работе и дома. Как оставаться спокойными и сосредоточенными. Как стать лучше. Выберите идеальную работу. Идеального партнера. Идеальную жизнь. Десять главных вещей, которые нужно сразу сделать на новой работе. Десять вещей, которых нужно избегать. Десять главных уравнений.

Спокойствие, как у йогов. Самоосознание. Глубокое мышление и медленное дыхание. Тигры. Кошки и собаки. Популярная психология и эволюционное поведение. Будьте пещерным человеком, охотником-собирателем или греческим философом. Отключитесь. Подсоединитесь. Расслабьтесь. Зарядитесь. Выпрямитесь и никогда не лгите. Плюйте на всё, и вы всегда будете счастливы. Не откладывайте, делайте сейчас, и быстрее.

Всем таким советам не хватает структуризации. Важная информация перемешана с мнениями и ерундой. Уравнение суждений позволяет организовывать и оценивать. Оно превращает любой совет, желательный или нежелательный, в модель, которую

можно проверить, используя данные. Внимательно прислушивайтесь к мнению других, записывайте альтернативы, собирайте данные и выносите суждение. Корректируйте свое мнение по мере медленного накопления свидетельств за и против той или иной идеи. С помощью того же процесса судите о действиях других. Всегда давайте им второй и даже третий шанс — гарантируя, что ваши решения определяются фактами, а не эмоциями. Используя формулу Байеса, вы не только будете делать оптимальный выбор, но и обнаружите, что завоюете доверие других. Вы прославитесь своими здравыми суждениями.

Глава 3

Уравнение уверенности

$$h \cdot n \pm 1,96 \cdot \sigma \cdot \sqrt{n}$$

Не вся «Десятка» родилась из христианской морали. Если бы нас попросили отправиться в ту точку пространства и времени, где ей было положено начало, то это оказалось бы не смертное ложе Томаса Байеса. Мы очутились бы в Лондоне, но почти на тридцать лет раньше — на собрании друзей 12 ноября 1733 года, где слушали бы, как Абрахам де Муавр раскрывает секреты азартных игр.

Де Муавр был «нестандартным» математиком. Его изгнали из Франции как протестанта, а в Лондоне к нему относились с подозрением из-за его национальности. Его современники — Исаак Ньютон и Даниил Бернулли — стали профессорами в своих областях, а де Муавру приходилось зарабатывать другими способами. Он давал частные уроки для детей среднего класса в Лондоне (предполагается, хотя это и не доказано, что одним из его учеников был юный Томас Байес), занимался «консультациями». Его можно было найти в кофейне Old Slaughter's на Сент-Мартинс-Лейн, где он давал советы всем — от игроков и финансистов до самого сэра Исаака Ньютона.

Работа де Муавра, появившаяся в 1733 году, была более изощренной по сравнению с его ранними трудами. Он показал, как новая математика — анализ, недавно разработанный Исааком Ньютоном, — помогает определить уровень уверенности в долгосрочной прибыльности в играх, опирающихся на случай. В итоге представленное им уравнение станет основой того, как

специалисты в естественных и социальных науках определяют уровень доверия к их исследовательским результатам. Но чтобы понять, откуда взялось уравнение уверенности, нужно начать оттуда же, что и де Муавр. Нужно войти в сомнительный мир азартных игр.

*

Сегодня требуется всего несколько минут, чтобы открыть счет для ставок в онлайн-казино. Имя, адрес и (главное) данные карточки — и готово. Игры разные. Есть онлайн-покер, где вы играете против других, а казино получает свою долю. Есть игровые автоматы, похожие на те, что когда-то ставили в пабах; у них свои имена — например, Cleopatra's Tomb, Fruit vs Candy и Age of the Gods, а также торговые марки вроде Batman v Superman и Top Trumps Football Stars. Вы нажимаете на кнопку, вращаете колесо и, если боги выстраиваются в ряд или в ряду достаточно Бэтменов, выигрываете. Наконец, есть традиционные для казино игры — например, блекджек и рулетка, — которые транслируются в прямом эфире: элегантно одетые юноши раздают карты, а женщины в вечерних платьях с низким вырезом крутят колесо рулетки.

Я открыл счет на одном популярном сайте онлайн-игр. Положил 10 фунтов стерлингов и получил бонус в 10 фунтов за новый счет, что дало мне стартовый капитал в 20 фунтов. Я решил начать с «Века богов» — только по той причине, что там разрешались самые маленькие ставки. Соответственно, при цене 10 пенсов за вращение у меня было больше попыток.

Через двенадцать попыток у меня оказалось на 70 пенсов меньше и не было ощущения, что что-то происходит. Боги мне наскутили, и я отправился на Top Trumps крутить Криштиану Роналду, Месси и Неймара. Это было дороже, по 20 пенсов за вращение, но через шесть попыток я выиграл (1,5 фунта!) и почти вернулся к своему начальному капиталу. Я попробовал «Бэтмен против Супермена» и несколько других игр. Затем нашел кнопку автоматического запуска и теперь мог крутить постоянно без необходимости нажимать на кнопку. Впрочем, это была плохая идея. Спустя двести попыток у меня осталось 13 фунтов.

Мне показалось, что игровые автоматы не оптимальны по соотношению цены и качества, и я решил попробовать живое казино. За столом распоряжалась Керри, женщина двадцати с небольшим в черном платье. Она приветствовала меня и заговорила с другим клиентом. Это был странный опыт. Я мог набирать сообщения, а она отвечала.

— Как погода у вас? — спросил я.

— Хорошая, — ответила она, посмотрев прямо на меня. — Кажется, весна на подходе. Сейчас последние ставки. Удачи!

Она жила в Латвии и уже четыре раза бывала в Швеции. Немного поболтав, я спросил ее, были ли сегодня крупные выигрыши.

— Мы не видим, сколько вы ставите, — сообщила она.

Я чувствовал себя несколько глупо; старался ставить 1 фунт при каждом вращении колеса, так что она не думала, будто я дешевка.

Керри мне нравилась, но я ощущал, что мне нужно лучше осмотреться. Не знаю, как это выразить, но была какая-то причина, по которой она и большинство ее коллег-мужчин работали в зале с низкими ставками. Керри было слегка неловко и некомфортно в обтягивающем платье, и она не казалась сексуальной.

Залы с более высокими ставками выглядели иначе. Вырез на платьях был глубже, а улыбки завлекательнее. Перед каждым вращением Люси, работавшая в помещении, где я находился, осознанно смотрела в камеру, словно говоря, что мой выбор верный. Мне пришлось заставить себя вспомнить, что она смотрела не только на меня, но и на 163 других игроков со всего мира.

Люси отвечала на вопросы клиентов.

— Да, у меня есть партнер. Но все сложно, — говорила она одному.

— О, я люблю путешествовать, — сообщала другому. — Я бы с радостью поехала в Париж, Мадрид, Лондон...

Камера переключилась на вид сверху, давая возможность увидеть ее ноги перед тем, как колесо пришло во вращение.

Мне стало очень некомфортно. Приходилось напоминать себе, зачем я здесь в первую очередь. Вернулся в помещение с более низкими ставками — там работал Макс, вежливый молодой человек,

который давал статистические советы по выигрышным цветам и номерам. Похоже, на его рулетке хорошо шли большие числа.

Посмотрел на свой баланс. Я наугад ставил на красное и черное, особо не раздумывая, и с удивлением увидел, что после нескольких часов в казино у меня имелось 28 фунтов. Прибыль 8 фунтов за вечер. Дела шли хорошо.

*

Как узнать, почему мы выигрываем: потому что умело играем или потому что нам повезло? Я знал, что в онлайн-казино шансы против меня, пусть даже я после двух-трех часов и пополнил немного свой счет.

В других играх я не знаю, есть у меня преимущество или нет. Если я играю в покер с друзьями, то вижу, как моя стопка фишек растет и уменьшается. Но сколько времени пройдет, прежде чем я смогу сказать: я — лучший игрок? Если я установил стратегию спортивных ставок, как для чемпионата мира, — когда я узнаю, что она окупается?

Такие вопросы не ограничиваются азартными играми и спортом; они могут быть и политическими. Сколько избирателей нужно опросить, чтобы надежно оценить, кто выиграет президентские выборы в США? Они могут быть связаны с обществом: как узнать, проявляет ли компания расовую дискриминацию при найме людей? И даже личными: сколько времени вы должны отдать работе или отношениям с другим человеком, прежде чем решиться что-то поменять?

Удивительно, но существует формула, которая отвечает на все эти вопросы: назовем ее уравнением уверенности*. Вот оно:

$$h \cdot n \pm 1,96 \cdot \sigma \cdot \sqrt{n} \quad (\text{Уравнение 3}).$$

Понятие степени уверенности отражается центральным символом \pm (плюс-минус). Представьте, что вы спрашиваете меня, сколько чашек кофе в день я выпиваю. Я не знаю наверняка, поэтому могу

* Если быть точнее, это формула 95%-го доверительного интервала для числа успехов в схеме Бернулли.

сказать что-то вроде «четыре плюс-минус парочка», или 4 ± 2 . Это — доверительный интервал, удобное обозначение и среднего значения, и отклонения от него. Это не значит, что я никогда не могу выпить 7 чашек (или только одну), но я вполне уверен, что в большинстве случаев выпиваю от 2 до 6 чашек.

Уравнение 3 позволяет нам делать более точные утверждения о нашей уверенности в каком-либо событии. Представьте, что я прошу всех читателей этой книги запустить 400 раз колесо рулетки, ставя по 1 фунту на красное или черное. На рулетке 37 номеров: от 1 до 36, раскрашенные поочередно в красный и черный цвета, и особый зеленый 0 (зеро). Он обеспечивает преимущество казино. Например, если какой-то игрок ставит на красное, то вероятность попадания шарика на красный номер составляет $18/37$, и в случае этого события игрок возвращает свою ставку и получает еще такую же сумму. Вероятность потери денег (непопадания шарика на красный цвет) составляет $1 - 18/37 = 19/37$. Ожидаемый (средний) выигрыш игрока при ставке в 1 фунт составляет $1 \cdot 18/37 - 1 \cdot 19/37 = -1/37$; поэтому при каждом повороте колеса игрок в среднем проигрывает 2,7 пенса. В уравнении 3 средний проигрыш обозначен буквой h , и в нашем случае $h = -0,027$ (фунтов). За 400 попыток каждый из читателей проиграет в среднем $h \cdot n = 0,027 \cdot 400 = 10,8$ фунта.

Следующий шаг — определить степень отклонения от среднего. Не каждый читатель проиграет (или выиграет) одну и ту же сумму. Даже без арифметических подсчетов понятно, что при одном обороте рулетки можно наблюдать большую разницу в результатах: если я ставлю 1 фунт, то либо удвою свои деньги, либо потеряю их. Отклонение имеет такую же величину, как инвестируемая сумма, и гораздо больше, чем средняя потеря в 2,7 пенса.

Определим отклонение количественно. Для этого найдем средний квадрат разности между результатом одного вращения и средним значением. Среднее значение равно $-0,027$ фунта, и если мы выиграли фунт, то квадрат разности равен $(1 - (-0,027))^2 = 1,0547$, а если проиграли 1 фунт, то $(-1 - (-0,027))^2 = 0,9467$. Всего есть 18 удачных исходов и 19 неудачных, поэтому средний квадрат разности, который обозначают σ^2 , равен:

$$\sigma^2 = \frac{18}{37} \cdot 1,0547 + \frac{19}{37} \cdot 0,9467 = 0,9993.$$

Такой средний квадрат разности σ^2 называют дисперсией. У рулетки она очень близка к единице, но не равна ей. Если бы на рулетке было только 36 номеров, половина красных и половина черных, дисперсия была бы в точности единица.

Дисперсия увеличивается пропорционально количеству вращений колеса. Если я запускаю колесо рулетки дважды, она удваивается; если три раза — утраивается и т. д. Дисперсия при n попытках равна $n \cdot \sigma^2$.

Обратите внимание, что при вычислении дисперсии мы возводим разность в квадрат, поэтому ее размерность — фунты в квадрате, а не фунты. Чтобы получить снова фунты, можно извлечь из дисперсии квадратный корень и получить так называемое среднеквадратичное (стандартное) отклонение σ ; в нашем случае $\sigma = 0,9996$. Соответственно, за 400 оборотов мы получим среднеквадратичное отклонение

$$\sqrt{n \cdot \sigma^2} = \sigma \sqrt{n} = 0,996 \sqrt{400} = 0,9996 \cdot 20 \approx 19,99.$$

Теперь у нас есть большая часть компонентов для уравнения уверенности. Единственный элемент, который мы еще не объяснили, — число 1,96. Оно появляется из математической формулы, которая описывает кривую нормального (гауссовского) распределения; эта колоколообразная кривая используется для представления роста людей или их IQ. Вы можете вообразить нормальное распределение в виде колокола с точкой максимума в среднем значении (например, при запуске рулетки 400 раз средний выигрыш будет 10,8 фунта; если мы станем измерять рост мужчин в Великобритании, то среднее значение будет 175 сантиметров)²⁴. На рисунке 3 показана кривая нормального распределения для 400 запусков рулетки и ставок по 1 фунту.

Теперь представьте, что мы хотим найти интервал, который содержит 95% площади этой колоколообразной фигуры.

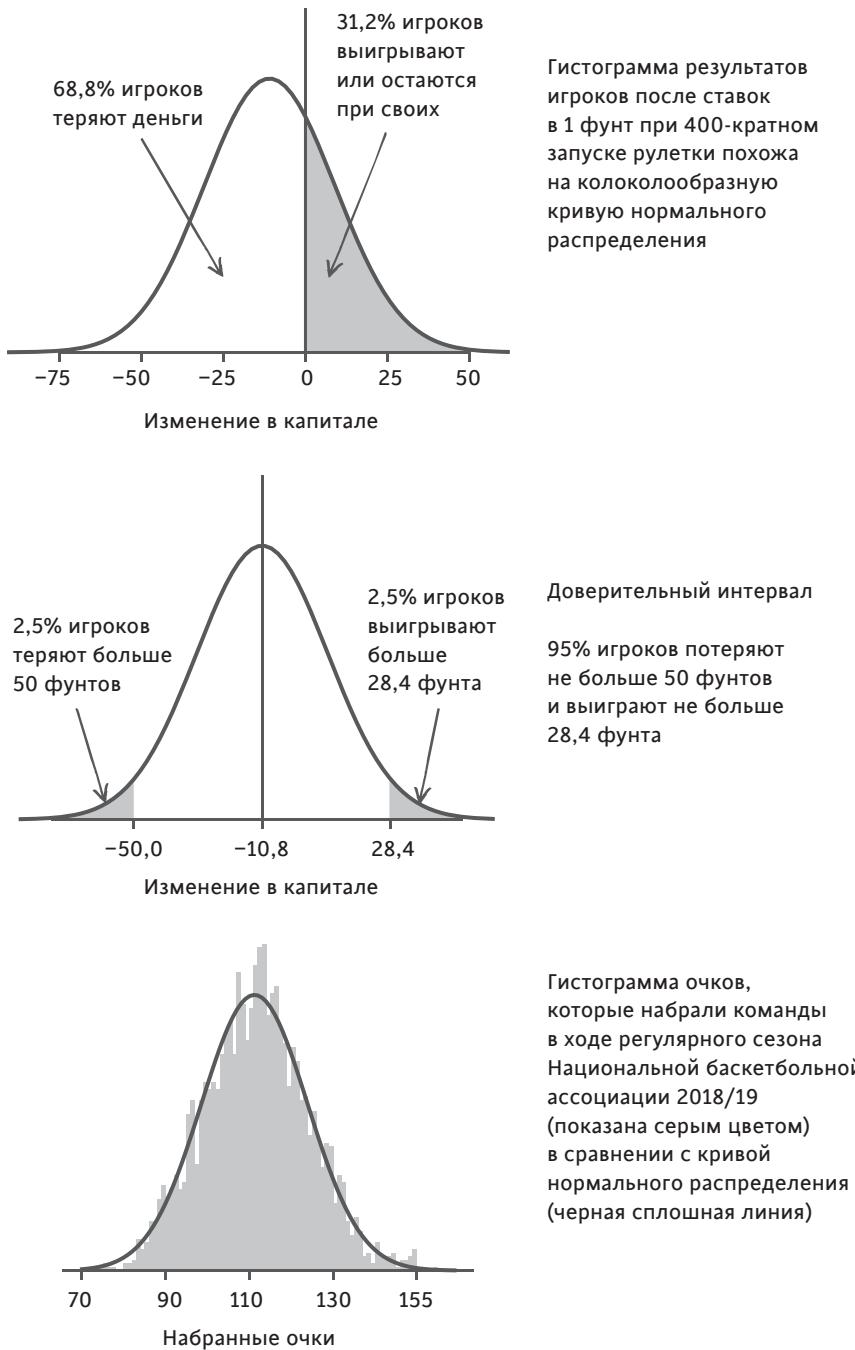


Рис. 3. Нормальное распределение

Для 400 запусков рулетки это интервал, куда попадут 95% прибылей или убытков читателей. Величина 1,96 берется именно отсюда. Чтобы интервал содержал 95% наблюдений, его граничные значения должны в 1,96 раза превосходить среднеквадратичное отклонение. Иными словами, в нашем случае 95%-й доверительный интервал для нашей прибыли после 400 запусков рулетки определяется уравнением 3:

$$h \cdot n \pm 1,96 \cdot \sigma \cdot \sqrt{n} = -0,027 \cdot 400 \pm 1,96 \cdot 0,9996 \cdot 20 = -10,8 \pm 39,2.$$

После 400 запусков рулетки читатель в среднем потеряет 10,8 фунта. Печально. С другой стороны, $\pm 39,2$ определяет довольно широкий доверительный интервал, поэтому некоторые читатели преуспеют. Получившие прибыль игроки будут в явном меньшинстве — их всего 31,2% от общего количества тех, кто крутил рулетку 400 раз. Я обращал на это внимание, когда ходил в казино или на скачки с небольшой группой людей. Обычно находится один человек, который выигрывает и остается в плюсе. Это ощущается как общая победа, особенно когда он покупает всем выпивку.

Итак, вот первый урок из уравнения уверенности. Победитель может считать, что у него была умная стратегия, а в реальности почти третья людей покидают казино победителями. Но случайность не должна их одурачивать. Они счастливчики, а не умельцы.

*

Я упустил важную деталь: сказал вам, что распределение результатов игры соответствует нормальному закону, но не объяснил почему. Объяснение восходит к работе Абрахама де Муавра 1733 года.

В своей первой книге «Доктрина шансов», посвященной азартным играм и опубликованной в 1718 году²⁵, де Муавр определял вероятность получения конкретных рук в карточных играх и выигрышных исходов при бросании костей — например, прихода двух тузов в пятикарточной руке или выпадения двух шестерок при бросании двух костей²⁶. Он вел читателя через вычисления, предлагая упражнения для улучшения понимания. Советы именно

такого рода просили у него игроки, искавшие его в кофейне Old Slaughter's.

В работе 1733 года де Муавр спрашивал своих читателей, как вычислить результат подбрасывания симметричной монеты 3600 раз. Для двух бросков монеты вероятность получить подряд два орла нетрудно найти прямым умножением: $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$. Вероятность получить три орла при пяти бросках можно найти, если выписать все возможные варианты, когда выпадает три орла (и, соответственно, две решки):

ГГГРР, ГГРГР, ГГРРГ, ГРГРГ, ГРГГР, ГРРГГ, РГРГГ, РГГРГ,
РГГГР, РРГГГ,

что дает нам 10 различных вариантов. Еще в 1653 году Блез Паскаль показал, что число способов взять k предметов из n (которое обозначается C_n^k) определяется формулой:

$$C_n^k = \frac{n!}{k(n-k)!}.$$

Выражение $k!$, которое называют факториалом, определяется так: $k! = k \cdot (k-1) \cdot (k-2) \cdot \dots \cdot 2 \cdot 1$. В нашем примере $n = 5$ (пять бросков монеты и их результатов), а $k = 3$ (три орла, которые должны оказаться среди этих результатов). Следовательно,

$$C_5^3 = \frac{5!}{3!(5-3)!} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1 \cdot 2 \cdot 1} = 10.$$

Получился тот же результат, что и при прямом переборе всех возможных вариантов. Поскольку вероятность выпадения орла на симметричной монете равна $1/2$, то вероятность получить k орлов при n бросаниях монеты равна

$$\frac{n!}{k!(n-k)!} \cdot \left(\frac{1}{2}\right)^n.$$

Для $n = 5$ и $k = 3$ получаем

$$\frac{5!}{3!(5-3)!} \cdot \left(\frac{1}{2}\right)^5 = \frac{10}{2 \cdot 2 \cdot 2 \cdot 2 \cdot 2} = 0,3125.$$

Следовательно, шансы на выпадение трех орлов при пяти бросках монеты равны 31,25%.

Де Муавр знал о таком распределении вероятностей (сейчас оно называется биномиальным распределением), но он также понимал, насколько непрактично применять этот способ, когда n — большое число. Чтобы решить аналогичную задачу для $n = 3600$ подбрасываний монеты, понадобится возвести двойку в степень 3600 и вычислить $3600 \cdot 3599 \cdot \dots \cdot 2 \cdot 1$. Попробуйте это посчитать. Такое невозможно осуществить вручную и трудно даже на компьютере.

Трюк, который проделал де Муавр, — отказ от непосредственного умножения и изучение математической формы биномиального распределения. Он вывел формулу для приближения факториалов больших чисел, а его друг, шотландский математик Джеймс Стирлинг, нашел точное значение константы в ней*, и де Муавр доказал, что при достаточно больших n вышеприведенное выражение для вероятности получить k орлов при n бросках монеты приблизительно равно

$$\frac{1}{\sqrt{2\pi} \sqrt{n/4}} \cdot \exp\left(-\frac{(k - n/2)^2}{n/2}\right).$$

На первый взгляд кажется, что это выражение сложнее, чем исходная формула вероятности для биномиального распределения, поскольку тут есть квадратные корни, константа $\pi = 3,14\dots$ и экспонента. Но здесь нет многочисленных умножений, необходимых для вычисления факториалов, и это главное в результате де Муавра.

* Формула по де Муавру выглядит как $n! \sim B \sqrt{n} \left(\frac{n}{e}\right)^n$, где константа B была вычислена приблизительно как 2,5074. Стирлинг нашел точное значение константы $B = \sqrt{2\pi}$. В итоге эта формула обычно называется формулой Стирлинга, хотя по справедливости ее нужно называть формулой Муавра или хотя бы формулой Муавра — Стирлинга.

Можно вычислять значения для 3600 или даже миллиона бросаний, просто подставляя нужные значения k и n . Теперь де Муавр мог для вычисления использовать таблицы логарифмов или логарифмическую линейку. Технологии XVIII века способны были вести расчеты для миллиона бросаний.

Де Муавр построил первый доверительный интервал для такого события. Он показал, что шансы получить при 3600 бросаниях меньше 1740 или больше 1860 орлов составляют примерно 21 к 1, то есть вероятность получить от 1740 до 1860 орлов примерно равна 95,4%²⁷.

В общем случае функция

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

с параметрами μ и σ^2 называется плотностью нормального (гауссовского) распределения, и это одна из самых важных функций в математике. Де Муавр, видимо, не осознавал всей важности своей формулы, и только в 1810-х Пьер-Симон, маркиз де Лаплас, понял весь ее потенциал. Лаплас изучал так называемую производящую функцию моментов, которая позволяет однозначно определить распределение через его моменты*. Производящие функции моментов позволили Лапласу исследовать, как меняется форма распределения при сложении множества случайных результатов (например, выпадения чисел на колесе рулетки или бросания костей). Лаплас продемонстрировал замечательную вещь: независимо от того, что суммировать, по мере увеличения числа слагаемых моменты суммы всегда становятся всё ближе к моментам нормального распределения.

Потребовалось несколько лет, чтобы справиться с некоторыми хитрыми исключениями в результате Лапласа (к некоторым из них

* Если имеется случайная величина X , то ее k -м начальным моментом называется число $\nu_k = E(X^k)$, где E — математическое ожидание. Аналогично k -м центральным моментом называется число $\mu_k = E((X - EX)^k)$. В частности, первый начальный момент X — среднее значение (математическое ожидание) случайной величины, второй — ее дисперсия, третий характеризует степень симметричности распределения, а четвертый — эксцесс, то есть показывает, насколько остра вершина распределения.

мы вернемся в главе 6). Над теми же вопросами работали в XX веке русский математик Александр Ляпунов и финский математик Ярл Вальдемар Линдеберг. Результат, доказанный Линдебергом в 1920 году, известен сегодня под названием «центральная предельная теорема»²⁸. Она говорит, что если мы складываем много независимых случайных величин (например, измерений) со средним значением h и среднеквадратичным отклонением σ , то распределение суммы этих величин будет близко к нормальному со средним значением $h \cdot n$ и среднеквадратичным отклонением $\sigma\sqrt{n}$.

Чтобы оценить масштабность этого результата, рассмотрим несколько примеров. Суммируя результаты бросания игральной кости, мы получим нормальное распределение. Суммируя результаты последовательных результатов в карточных играх, рулетке или онлайн-казино, — получим также нормальное распределение. И общее количество очков в играх сезона Национальной баскетбольной ассоциации имеет нормальное распределение (см. нижнюю диаграмму на рис. 3)²⁹. И урожайность имеет нормальное распределение³⁰. Скорость движения по автостраде тоже. Как и рост людей, их IQ и результаты личностных тестов*.

Везде, где результат обусловлен различными случайными факторами, можно найти нормальное распределение, поэтому уравнение 3 используют для построения доверительного интервала в любой области, где раз за разом происходит повторение одного типа действий или наблюдений.

*

В главе 1 я показал, как игрок с преимуществом в 3% может всего за один год превратить стартовый капитал в 1000 фунтов в сумму 57 миллионов. Если делать ставки и использовать выигрыши в дальнейшей игре, то капитал растет экспоненциально. И здесь я подошел к неизбежной ловушке для моего гипотетического игрока. Я назову этого человека Лизой. Как Лиза узнает, что ее преимущество составляет 3%?

* Если быть точнее, центральная предельная теорема говорит, что сумма случайных величин стремится к нормальному закону при числе слагаемых, стремящемся к бесконечности. При любом конечном числе слагаемых сумма лишь приблизительно является нормальной.

Нейт Сильвер, создатель и редактор сайта спортивных и политических прогнозов FiveThirtyEight, для объяснения таких ситуаций использует термины «сигнал» и «шум»³¹. В спортивных ставках значение средней прибыли (или потери) при одной ставке (величина h в уравнении 3) — это сигнал. Если у Лизы есть преимущество в 3%, то в среднем на каждую ставку в 1 фунт она выигрывает 3 пенса. Шум при ставке измеряется величиной среднеквадратичного (стандартного) отклонения σ . Как и в случае рулетки, шум в спортивных ставках гораздо больше, чем сигнал. Например, если Лиза ставит 1 фунт на какую-то команду с шансами $1/2$, то она либо проиграет 1 фунт, либо выиграет 50 пенсов. С помощью формулы на с. 71 можно показать, что стандартное отклонение в этом случае равно $0,71$ ³². Таким образом, шум, измеряемый стандартным отклонением ($\sigma = 0,71$), при одной ставке гораздо больше, чем сигнал ($h = 0,03$). Мы скажем, что отношение сигнала к шуму в нашем случае $h/\sigma = 0,03/0,71 \approx 1/24$.

Казино знает, что у него есть преимущество, поскольку поставило рулетки, его обеспечивающие, и отношение сигнала к шуму равно $1/37$. Лиза вынуждена при определении того, есть ли у нее преимущество, опираться на свой прошлый опыт. Именно здесь уравнение уверенности важнее всего для профессионального игрока. Если Лиза получила прибыль в h фунтов на ставку, а стандартное отклонение для ставки равно σ , то доверительный интервал для оценки преимущества h можно найти, поделив части уравнения 3 на n :

$$h \pm \frac{1,96 \cdot \sigma}{\sqrt{n}}.$$

Например, если Лиза сделала 100 ставок и получила в среднем 3 пенса на каждую, этот доверительный интервал выглядит так:

$$0,03 \pm \frac{1,96 \cdot 0,71}{\sqrt{100}} = 0,03 \pm 0,14.$$

Ее преимущество может составлять 17 пенсов ($0,03 + 0,14 = 0,17$), но оно иногда оборачивается потерей 11 пенсов в среднем. Все

возможные величины для преимущества между $-0,11$ и $+0,17$ лежат в этом 95%-м интервале*. Сделанные ею 100 ставок говорят слишком мало о том, работает ее стратегия или нет.

Поскольку доверительный интервал содержит 0, Лиза не может быть вполне уверена, что ее сигнал h положителен и ее игровая стратегия работает. Есть простое эмпирическое правило, которое она может использовать, чтобы узнать, сколько наблюдений нужно, чтобы надежно определить сигнал. Прежде всего округлим 1,96 до 2: для эмпирического правила разница между этими числами мала. Перепишем уравнение уверенности, чтобы определить условие, при котором доверительный интервал не включает ноль³³:

$$\frac{h}{\sigma} > \frac{2}{\sqrt{n}}.$$

Таким образом, если мы возьмем n наблюдений, то сможем обнаружить отношение сигнала к шуму, превосходящее $2/\sqrt{n}$.

Ниже дана таблица некоторых значений, чтобы вы представили, как работает это правило.

Число сделанных наблюдений (n)	16	36	64	100	400	1600	10 000
Обнаруженное отношение сигнала к шуму ($2/\sqrt{n}$)	1/2	1/3	1/4	1/5	1/10	1/20	1/50

Преимущество на ставках или в финансовой сфере, как правило, имеет отношение сигнала к шуму примерно 1/20 или даже 1/50; поэтому, чтобы его обнаружить, нужны тысячи и даже десятки тысяч наблюдений. Для отношения сигнала к шуму $h/\sigma = 1/24$, как на спортивных ставках у Лизы, потребуется $n > 2304$ наблюдений. Больше двух тысяч наблюдений — это очень много футбольных матчей. Если Лиза считает, что ее преимущество на рынке

* На самом деле ее преимущество может быть и больше 17, и меньше -11 пенсов, просто такие события выходят за рамки 95%-го интервала, то есть имеют в совокупности вероятность менее 5%. Мы считаем их маловероятными и пренебрегаем ими.

английской Премьер-лиги составляет 3%, то ей придется для уверенности прождать шесть сезонов.

За эти шесть лет другие игроки могут наткнуться на такое преимущество и нивелировать его. Обширные операции Мэттью Бенхэма и Тони Блаума постоянно направлены на поиск возможностей. Как только два этих больших Б оказываются в игре, букмекеры корректируют свои коэффициенты и преимущество исчезает. Лиза рискует, поскольку не осознает, что ее преимущество исчезло. Чтобы убедиться, что преимущество существует, требуется больше 1000 матчей. Чтобы понять, что оно исчезло, может потребоваться столько же больших убытков. Прибыль, которая экспоненциально росла, превратилась в такой же экспоненциальный ущерб.

Большинство инвесторов-любителей представляют, что им нужно отделить сигнал от шума, но мало кто понимает важность извлечения квадратного корня в полученном нами правиле. Например, для обнаружения сигнала вдвое меньшей мощности требуется вчетверо больше наблюдений, и, наоборот, увеличив число наблюдений с 400 до 1600, вы сможете заметить всего вдвое большее преимущество. Очень легко недооценить объем данных, необходимых для обнаружения крошечных преимуществ на рынках.

*

Я позвонил Яну в Берлин, чтобы спросить, как дела у него с Мариусом. Все шло хорошо — Мариус даже предупреждал Яна, что мне следует говорить. Однако Ян, как обычно, хотел беседовать о цифрах.

— Что я *не могу* тебе сказать, пока ты не поговоришь снова с Мариусом и не проверишь, что все в порядке? То, что наш оборот составляет 70 миллионов фунтов. За последний месяц мы сделали 50 000 ставок со средним преимуществом 1,5–2%.

По сравнению с этим те 50 фунтов, которые мы ставили на чемпионате мира, были мелочью. Когда я сказал Яну, что сейчас пишу про доверительные интервалы, он вспомнил ту игровую модель, что мы построили вместе.

— Да, мы на этом заработали, — сказал он. — Однако, если честно, это не то, что мы станем использовать в будущем.

Он оказался прав: наша модель для чемпионата мира была построена на 283 наблюдениях от предыдущих турниров. Теперь же Ян создал базу данных по 15 миллиардам позиций для ставок, охватывавших многие виды спорта за последние девять лет.

— Мы сосредоточиваемся на стратегиях, где у нас есть больше 10 тысяч подтверждающих наблюдений, — сказал он.

Это давало им уверенность, что их стратегия действительно обладает долгосрочным преимуществом.

Главное преимущество Яна и Мариуса основывалось на национальных различиях. Бразильцы ожидают больше голов в своих матчах, чем забивают на самом деле. Немцы — пессимисты и предпочитают ставить на скучные 0:0.

— Норвежцы отличаются точностью, — смеется Ян. — Идеально рациональные скандинавы.

Я вспомнил свой разговор с Мариусом — как раз рациональным норвежцем — во время чемпионата мира, когда мы обсуждали, как забраться в голову игрока. Он всегда считал, что важно иметь базовое объяснение для любой стратегии ставок. Теперь у него есть одно такое объяснение: применение национальных стереотипов.

*

Вы ищете гостиницу на сайте TripAdvisor. Вы желаете остановиться в месте, которое получило оценки в 4 звезды или больше, и скептически относитесь ко всему, что оценено в 3,5 звезды и ниже. Сигнал, который вы здесь ищете, — разница в ползвезды. Принципы присвоения звезд на сайте TripAdvisor различны: одни с энтузиазмом раздают пятерки, а несколько недовольных вечно ставят единицы. Однако в целом шум при оценке составляет около одной звезды: большинство людей ставят 3, 4 или 5 звездочек при среднем значении, слегка превышающем 4³⁴.

Мы можем ответить на вопрос, сколько обзоров нам нужно прочитать, чтобы надежно выявить отношение сигнала к шуму, равное 1/2: либо используя табличку на с. 79, либо решив уравнение $\frac{2}{\sqrt{n}} = \frac{1}{2}$, где 1/2 — отношение сигнала к шуму. Это дает нам $\sqrt{n}=4$,

что эквивалентно $n = 16$ оценкам на сайте. Вместо того чтобы искать среднее значение по всем сотням оценок для данной гостиницы, возьмите последние 16 и посчитайте среднее. Это даст вам одновременно и свежую, и надежную информацию.

Звездами можно оценивать не только гостиницы. Джесс не уверена в том, что правильно выбрала работу для карьеры. Она трудится в организации, занимающейся защитой прав человека. Это однозначно хорошее дело, однако руководительница просто несносна. Она целый день называет Джесс и предъявляет непомерные требования. Ее знакомый Стив шесть месяцев встречается с Анной. Их отношения непостоянны: сейчас горячие, а в следующую минуту уже холодные. Они ужасно ссорятся, но когда разборок нет, то все замечательно.

Уравнение уверенности дает определенные указания и для Джесс — сколько ей оставаться на работе, и для Стива — сколько ему продолжать отношения с Анной, прежде чем расстаться. Сначала им нужно установить подходящие интервалы времени. Стив и Джесс решают оценивать каждый день по шкале от 0 до 5 звезд. Затем они планируют регулярно встречаться и оценивать складывающиеся ситуации.

Вечером в пятницу первой недели у Стива происходит страшная ссора с Анной, потому что она отказывается пойти с ним в компанию друзей. Стив звонит Джесс и жалуется: в результате три дня в его неделе получают по 1 звезде. Джесс напоминает ему, что не надо делать выводы слишком быстро. В конце концов, $n = 7$, и сигнала в шуме пока не найти. У Джесс на работе все хорошо — в основном потому, что ее проблемная начальница была в командировке, так что все дни получили по 3 и 4 звезды.

Через месяц ($n = 30$) они встречаются за обедом и начинают лучше понимать, как идут дела. Стив провел с Анной несколько хороших недель. В прошлый уик-энд они ездили в Брайтон, плюс к этому было несколько приятных обедов, где они отлично провели время. У Стива были целые полосы 5-звездочных дней. У Джесс все наоборот: когда ее начальница вернулась, то постоянно сердилась, кричала и терпение по малейшему поводу. Дни у Джесс — это 1–2 звезды и даже найдется несколько 0-звездочных.

Прошло немногим более двух месяцев: $n = 64$, а $2/\sqrt{n}=1/4$. Теперь степень уверенности втрое выше, чем после первой недели. У Стива хорошие дни перевешивают плохие, но всё равно идут постоянные ссоры: оценки — по 3 и 4 звезды. Босс Джесс — реальная проблема, зато она работает над ценным проектом, на котором всегда хотела сосредоточиться. В лучшем случае она ставит 3 и 4 звезды, но в основном оценки — 1 и 2.

Хотя каждая неделя дает новые наблюдения, правило извлечения квадратного корня означает, что скорость получения Стивом и Джесс новой информации не пропорциональна времени, а ниже. Польза от новых наблюдений уменьшается, поэтому они решают ограничить свои обсуждения и установить срок в 100 дней (три месяца с небольшим), после чего с достаточной уверенностью окончательно решить вопрос о своем будущем.

Итак, этот день настал: $n = 100$, а $2/\sqrt{n}=1/5$. Они оглядываются не только на последние несколько недель, но и на всё, что происходило за это время. У Стива с Анной ссоры происходят реже. Они вместе стали ходить на уроки кулинарии, по вечерам с радостью готовят и приглашают к себе друзей. Жизнь хороша. Стив определяет доверительный интервал. У него получается средняя оценка $h = 4,3$ звезды, а стандартное отклонение $\sigma = 1,0$. Доверительный интервал для его отношений с Анной: $4,3 \pm 0,2$. Солидное среднее, уверенно выше четырех звезд. Стив решает прекратить жаловаться на Анну; он убежден, что нашел партнера на всю жизнь.

У Джесс ситуация не так благоприятна. Ее среднее $h = 2,1$ звезды. По-настоящему хороших дней было очень мало, поэтому стандартное отклонение меньше, чем у Стива: $\sigma = 0,5$. Ее доверительный интервал: $2,1 \pm 0,1$. По сути, у Джесс работа на две звезды. Она уже начала искать новую должность и в понедельник собирается подать заявление об уходе.

*

В 1964 году Мальcolm Икс сказал: «Неважно, насколько сильно белые уважают и признают меня; пока такое же уважение и признание не проявляется ко всему нашему народу, его для меня не существует».

Идея, выраженная в этих словах, рождается из математики. Опыт одного человека — будь то Мальcolm Икс или кто-то еще — дает нам очень мало информации. Это все равно что один раз дернуть ручку игрового автомата. Тот факт, что у Джесс был один хороший день на работе, ничего не говорит о ее карьере за долгий промежуток времени. Когда люди стали прислушиваться к Мальcolmу Иксу, это не значило ничего, пока они не прислушались к афроамериканцам в целом. Борьба цветных в США против дискриминации во всех формах, рассказанная через истории Мальcolmа Икса, Мартина Лютера Кинга и других, была и остается борьбой десятков миллионов людей.

Джоанна слышит о вакансии на своей работе. Вечером она встречает на вечеринке Джеймса и рассказывает ему об этом. Тот воодушевляется, говорит, что мечтал о такой работе, и в понедельник подает заявление о приеме. Через несколько дней Джеймс уже приступил к новой работе, а Джоанна вдруг у магазинчика с выпечкой сталкивается с Джамалом. Он спрашивает, как у нее дела, и Джоанна рассказывает, что у них только что начал работать Джеймс. Джамал воодушевляется, говорит, что мечтал о такой работе, и спрашивает Джоанну, нет ли других вакансий...

Джоанна белая. Джеймс тоже. Джамал — нет. Джоанна расистка? Нет. Она действовала бы точно так же, если бы первым встретила Джамала. Просто так случилось, что с Джеймсом она столкнулась раньше.

Однако можно еще поразмышлять о том, почему так вышло. Поскольку Джоанна и Джеймс принадлежат к одной социальной группе, они чаще встречаются и делятся информацией о возможностях. Их взаимная помощь может косвенно дискриминировать Джамала. Он не имеет доступа к тем же социальным возможностям, что Джеймс и Джоанна.

Здесь нужно быть осторожными. Мы не можем делать какие-либо выводы из истории Джоанны. У нас всего одно наблюдение, один случай ее взаимодействия с Джеймсом и Джамалом. По одному событию доверительный интервал не построить. Именно поэтому порой трудно обнаружить расовую дискриминацию. Каждый отдельный случай — всего одно наблюдение, из которого мы

узнаем очень мало. Единственный способ понять роль расы в обществе — изучить множество наблюдений и построить доверительный интервал.

*

Моа Бёрселл, исследователь и преподаватель кафедры социологии в Стокгольмском университете, два года писала резюме и отправляла заявки на работу в Швеции. Всего она подала заявления больше чем на две тысячи разных должностей в сферах информатики, бухгалтерии, преподавания, в качестве водителя и медсестры. Однако Моа вовсе не искала работу. Она проверяла предвзятость работодателей, которым писала.

Для каждого заявления Моа создавала два отдельных резюме и сопроводительных письма, в которых описывала близкую квалификацию и опыт работы. Затем случайным образом придумывала имена для каждой пары. Одно имя звучало по-шведски, например Йонас Сёдерстрём или Сара Андерссон; второе было не шведским — скажем, Камаль Ахмади или Фатима Ахмед, указывавшие на арабское мусульманское происхождение, или Мтугу Хандуле либо Уасила Балагве, говорившие об африканском не мусульманском происхождении. Схема эксперимента Моа была эквивалентна подбрасыванию монеты. Если работодатели будут беспристрастными, они с равной вероятностью позвонят человеку со шведским и именем.

Однако этого не происходило. Например, в одном исследовании с $n = 187$ заявками от шведских и арабских мужчин мужчинам с арабскими именами позвонили вдвое реже, чем шведам³⁵. Такие результаты нельзя объяснить случайностью. Мы можем увидеть это, построив доверительный интервал. Арабским мужчинам позвонили 43 раза, вероятность такого звонка (сигнал) $h = 43/187 = 23\%$. Чтобы оценить дисперсию, будем использовать 1 для человека, которому позвонили, и 0 для того, кому не звонили. Теперь, как и в случае рулетки, вычисляем среднеквадратичное расстояние между этими величинами и h и получаем стандартное отклонение $\sigma = 0,649$. Если мы подставим эти величины в уравнение 3, то получим доверительный интервал для звонков претендентам

с арабскими именами $43 \pm 17,3$, который содержит значительно меньшие числа, чем 79 звонков, полученных шведами.

Более того, Моа улучшила резюме арабских мужчин, добавив им от одного до трех лет опыта по сравнению со шведскими конкурентами. Это не помогало им найти работу. Более опытному арабу звонили только в 26 случаях — по сравнению с 69 звонками менее квалифицированным шведам; и это значение снова не попадает в доверительный интервал $26 \pm 15,9$.

«В моих результатах важнее всего то, — сказала она мне, — что их очень просто понять. С цифрами не поспоришь»*.

Когда Моа читает лекции по этой теме в Стокгольмском университете, она может увидеть реакцию студентов по их лицам. «Когда я смотрю на голубоглазых блондинов, они внимательно слушают. Они не думают, что это честно, но это их не трогает. Зато у людей с карими глазами, темными волосами и темной кожей я вижу другую реакцию, — продолжает она. — Это про них, их друзей, братьев и сестер. Для некоторых это означает окончательное признание. И это может быть облегчением. Они видят, что они не сумасшедшие. Их восприятие реальности подтверждается».

Эти учащиеся часто рассказывают ей о своем опыте, но другие по-прежнему молчат. «Знание о моем расследовании может травмировать, — сказала она мне. — Я вижу, что они расстроены. У них ощущение, словно им сказали: от вас мало толку, вам тут не место».

Моа аккуратно указывает, что из ее работы не следует, будто найти работу невозможно. Смысл исследования — выявить масштабы несправедливости; это не означает, что все в Швеции расисты. Оно показывает, что Камаль Ахмади и Йонас Сёдерстрём должны потратить разное количество времени, если хотят выиграть в этой лотерее вакансий.

* Никак не отрицая наличия расовых предрассудков в обществе, следует отметить, что предубеждения работодателя чаще экономические. Например, южане в шведском климате, вероятно, будут чаще болеть; мусульмане обязаны соблюдать пост длиной в месяц, который может отрицательно сказаться на их работоспособности. На работе скажется и низкий уровень владения шведским языком. Соответственно, иностранцу нужно лучшее резюме, которое перевесит опасения работодателя.

Когда реальный Камаль Ахмади подает заявление о работе в Швеции, он не знает наверняка, с каким игровым автоматом имеет дело. Если он подал заявление, а его не вызвали на собеседование, то он не может утверждать, что его дискриминируют. Точно так же реальный Йонас Сёдерстрём не может увидеть преимущество, которое отдает ему этот игровой автомат. Он имеет квалификацию для работы; он подал заявление; ему позвонили и пригласили на собеседование. С его точки зрения, ничего неправильного не произошло.

Я изложил Моа это соображение о Камале и Йонасе, и она сказала: «Это верно, но некоторые люди иногда сами ставят эксперименты. Приехавшие из-за границы рассказывали мне, как искали работу в местном супермаркете и им сообщали, что вакансия уже закрыта. Но когда они потом просили шведского приятеля позвонить по телефону в магазин и узнать, открыта ли еще вакансия, тому сообщали, что можно прийти на собеседование».

Моа и ее коллеги разослали уже свыше 10 000 резюме, чтобы проверить различные гипотезы о шведском рынке труда. Некоторые результаты удручают: дискриминация мужчин арабского происхождения сильнее всего проявляется в ситуациях с низкоквалифицированной работой. Другие результаты вдохновляют больше: дискриминация арабских женщин выражена слабее и исчезает, если у этих женщин больше опыта работы.

Подобные исследования повторялись по всему миру со схожими результатами³⁶. Работа Моа показывает пример структурного расизма, дискриминации, которую часто трудно увидеть на индивидуальном уровне, но легко обнаружить с помощью уравнения уверенности. Недавно ведущий медицинский журнал *The Lancet* опубликовал доверительные интервалы для измерения социального неравенства в США — от бедности, безработицы и лишения свободы до заболеваемости диабетом и сердечными недугами³⁷. Чернокожие американцы статистически отличались от белых по всем параметрам. Свалки токсичных отходов строятся близко к районам, выделенным по расовому признаку; правительство не может предотвратить попадание свинца в питьевую воду; мелкие оскорблении на расовой почве; пониженная зарплата за

одинаковую работу; целевой маркетинг сигарет и содержащих сахар продуктов; принудительная реконструкция и снос жилья; ограничения для избирателей; здравоохранение по сниженным стандартам вследствие явных или неявных предубеждений; исключение из социальных связей, которые могли бы помочь при поиске работы, — список можно продолжать. Психологическое и физическое здоровье отдельных афроамериканцев и коренных американцев ежедневно подвергается мелкой дискриминации, при этом никакого открытого расизма может и не быть.

Вернемся к Джоанне. Сталкивается ли она чаще с Джеймсами, чем с Джамалами? Чтобы это выяснить, она решает прибегнуть к уравнению уверенности. Она вспоминает всех, кому было бы интересно работать в издательстве, где она трудится: талантливых людей, затем собственных друзей, с которыми постоянно общается³⁸. Из 100 друзей Джоанны 93 белые; однако доля белых в населении США 72%: $93 - 72 = 21$. Ее дружеские отношения имеют перекос в расовом плане. Джоанна проверила свою привилегированность. Она пробудилась и осознала, что люди, с которыми она знакома, не отражают населения в целом и принадлежат к привилегированной группе, которая делится информацией о наличии работы в своей среде. Что Джоанне делать с этим — трудный вопрос.

Вот что я думаю. Не математический ответ на вопрос, а просто мои размышления. Джоанне не надо менять друзей. Она должна дружить с тем, с кем хочет. Но ей нужно подумать, что она может сделать в такой ситуации. Это просто. Она может написать Джамалу и другим семи приятелям из меньшинства, когда слышит о вакансии на работе, или просто поговорить с ними. У Джамала группа общения еще более перекошена: 85 из 100 черные — по сравнению с долей 12,6% в общем населении США и 25% в Нью-Йорке, где он живет. Одним быстрым сообщением Джоанна полностью меняет демографию людей, которые знают о вакансии.

Мои взгляды часто именуют политкорректностью. Я предпочитаю называть это статистической корректностью. Речь о статистическом осознании следующего факта: то, что мы ощущаем

индивидуально, часто не отражает мир в целом. Каждый из нас как личность должен сам решить, насколько статистически правильны наши жизни и что нам с этим делать.

*

Уравнение уверенности, возможно, и было создано для азартных игр, но изменили его именно естественные и в итоге социальные науки. Первым участником «Десятки», осознавшим научную мощь нормальной кривой, был Карл Фридрих Гаусс, который использовал ее свойства в 1809 году, чтобы оценить ошибки в определении местоположения астероида Церера. Сегодня нормальное распределение часто называют гауссовским, что несколько несправедливо, поскольку оно описано уже во втором издании (1738) работы де Муавра «Доктрина шансов»³⁹.

Статистика интегрируется в науку благодаря крупным достижениям XIX и начала XX века. После Второй мировой войны доверительные интервалы стали неотъемлемой частью научных работ, что заставило исследователей демонстрировать, почему их результаты не чистая случайность. Последняя моя научная статья содержала более пятидесяти расчетов различных доверительных интервалов. Существование бозона Хиггса было подтверждено только тогда, когда уровень статистической значимости достиг 5σ ; это означает, что вероятность получить такие результаты эксперимента при отсутствии бозона Хиггса составляет 1 к 3,5 миллиона.

Изначально прогресс «Десятки» в социальных науках шел медленнее, чем в естественных. До недавнего времени карикатурное изображение кафедры социологии могло изображать мужчин в потрепанной одежде, поклоняющихся умершим немецким мыслителям, и женщин с волосами пурпурного цвета, что появились в 1970-е, чтобы встрихнуть мир постмодернистскими идеями. Они спорят и дискутируют, но никогда не приходят к согласию. Они создают определения и рамки для обсуждения и спорят дальше. Сторонние наблюдатели понятия не имеют, о чем речь. Вплоть до начала нового тысячелетия эта карикатура во многом была правдива. Статистику и количественные методы применяли, однако способом изучения общества считались социологическая теория

и идеологическое обсуждение. За несколько быстро пролетевших лет «Десятка» разрушила этот старый мир. Внезапно с помощью Facebook и Instagram исследователи смогли измерить наши социальные связи. Они учили все блоги с мнениями и поняли наши методы коммуникации. Они смогли использовать государственные базы данных, чтобы определить факторы, которые заставляют нас переходить с одной работы на другую и менять место жительства. Благодаря доступности данных и статистических тестов, определивших доверительный уровень для каждого случая, раскрылась структура нашего общества.

Идеологические споры и теоретические рассуждения оказались вытесненными за пределы социальных наук. Теория ничего не стоит без данных, ее подтверждающих. Некоторые социологи старой гвардии присоединились к этой революции данных, другие же отстали, однако любой человек из университета не смог бы отрицать, что общественные науки навсегда изменились.

*

Не все заметили эту трансформацию социальных наук вследствие использования данных. Иногда я читаю сетевой журнал Quillette. Он гордится тем, что продолжает традиции научного диалога, восходящие к Ричарду Докинзу в 1980-е и 1990-е. Его заявленная цель — давать платформу для свободного мышления, даже для опасных идей; это означает, что он с радостью публикует мнения о гендере, расе и IQ, которые обязательно «политически корректны».

Статьи в Quillette регулярно нападают на работы в области социальных наук. Одна из любимых мишней — политика идентичности. Недавно я закончил читать одну статью, написанную ушедшим на пенсию профессором социологии, который заявил, что общественные науки превращаются в «сумятицу и бессмыслицу». Он возражал против книги Тукуфу Зубери и Эдуардо Бонилья-Сильвы «Белая логика, белые методы: расизм и идеология»⁴⁰. В ней исследовалось, насколько методы в социальных науках определяются «белой» культурой. Основываясь на своих сомнениях в утверждениях о «белых методах», профессор выступил

со встречным заявлением, что он не смог нигде в обществе найти систематических признаков расизма. Он предположил, что наблюдаемые нами различия лучше объясняются «способностями и интересами афроамериканцев»⁴¹.

Авторы многих других статей в Quillette вместо рассмотрения данных стараются разжечь дискуссию с учеными-социологами и левыми активистами. Меньше внимания обращено на цифры, больше — на войну идей. Как я покажу в главе 7, внутренняя разница между биологическими расами мала (фактически такого понятия, как биологическая раса, вообще не существует), но есть множество доказательств структурного расизма в США — например, в вышеупомянутой статье в The Lancet.

Я отправил по электронной почте автору публикации в Quillette копию статьи из The Lancet с предложением ознакомиться с ее содержанием. Мы обменялись несколькими дружескими электронными письмами. Оказалось, что в сфере изучения поведения животных у нас довольно много общих интересов.

Спустя несколько недель он прислал мне свой новый труд — нападение на саму идею структурного расизма. Среди прочего он заявлял, что доказать расизм невозможно, поскольку необходимо исключить слишком много других факторов. Профессор на пенсии, похоже, упустил из виду весь смысл сбора статистических данных — выявления дискриминации по большому числу повторяющихся наблюдений. Он повторил свои утверждения о значительной разнице в биологии рас.

В ответ на американский вариант исследования резюме, проведенного Моа Бёрселл, в котором продемонстрировали предубеждение против афроамериканских имен, профессор написал: «Это расизм? Мы не знаем предыдущего опыта нанимателя. Возможно, у него есть опыт неудачного найма темнокожих».

Это расизм? Да. Здесь не нужны доверительные интервалы. К моему крайнему удивлению, через несколько месяцев Quillette опубликовал эти его размышления. К счастью, фраза «прошлый опыт неудачного найма темнокожих» была убрана. Однако статья выдерживала тот же тон, бездоказательно отрицая факты из статьи в журнале The Lancet.

Не только Quillette использует такой подход к социальным наукам. Британское интернет-издание Spiked, сетевая реинкарнация бумажного журнала Living Marxism («Живой марксизм»), регулярно нападает на гендерную политику и идею существования структурного расизма. Раздел «Культурная волна» на сайте Reddit позволяет всем желающим присоединиться к дебатам. Те же понятия пронизывают The intellectual dark web («Темная сеть интеллектуалов») — самопровозглашенное движение за свободные идеи, высказывающееся на YouTube и в подкастах за право всех быть услышанными. Участники этой группы не просто пишут о гендере и расе, а оспаривают политкорректность; и обычно дискуссия быстро переходит к этим двум «табу», как они это называют.

Король «Темной сети интеллектуалов» — Джордан Питерсон. Как и Quillette, он ведет войну против того, что считает политкорректностью, захватившей общественные науки. Питерсон утверждает, что левая идеология заставила ученых сосредоточиться на вопросах пола и расовой идентичности. Он описывает университеты как места, где боятся сказать что-то не то. В итоге это оказывает негативное воздействие на общество в целом. Белые подвергаются несправедливым нападкам за их привилегированность, а женщинам предоставляются несправедливые преимущества при приеме на работу.

Когда я летел последний раз в бизнес-классе (иногда мне это приходится делать), два парня, сидевших за мной, весь полет обсуждали, насколько хорошо одет Питерсон и как он себя ведет в споре. Я хотел обернуться и возразить, но не смог точно сказать, что неверно в их разговоре. Он хорошо одет, может аргументированно спорить и даже плакать в нужные моменты во время интервью.

Я прочитал книгу Питерсона «12 правил жизни»⁴². Она мне понравилась. Книга полна интересных историй из его жизни. Несколько хороших советов о том, как быть человеком. Приятное название. Но это не современная социальная наука. Даже близко такого нет. Это привилегированный белый мужчина, крутящий свое колесо рулетки и рассказывающий, как ему повезло.

Современное научное сообщество сильно отличается от того, как его рисует Питерсон. Я работаю со многими социологами и не

думаю, что встречал хотя бы одного, кто боялся сказать то, что он думает*. Наоборот. Наличие противоречивых идей, размышления о различных моделях — важная часть нашей работы.

Уравнение уверенности сдерживает современных ученых. Если мы хотим проверить нашу модель, нам надо собрать данные. Социальные науки — это уже не анекдоты или абстрактные теории. Они занимаются обнаружением структурного расизма — путем создания и рассылки тысяч резюме или тщательного изучения литературы. Это тяжелая работа, а не вопрос о том, как хорошо смотреться в костюме и изображать тщательные раздумья перед ответом на вопрос.

Левые взгляды у Моа Бёрселл проявились еще в подростковом возрасте. «Многие мои лучшие друзья того времени [начала 1990-х] имели иностранное происхождение, — говорила она мне. — Когда мы выходили по вечерам, они боялись, а некоторых преследовали неонацисты. Приходилось бегать с ними от неприятностей. Этот опыт привел меня в политику».

Когда Моа говорит о годах своего становления, она открыта и эмоциональна — при этом бесстрастна, когда рассказывает о своих научных результатах. Она также говорила мне о том, как спустя много лет один молодой борец за равенство привел к ней в университет группу иммигрантов-тинейджеров. Он хотел, чтобы Моа рассказала подросткам о своих исследованиях в области поиска работы; однако Моа боялась, что это будет неправильно воспринято. И оказалась права: когда она объяснила свои результаты, дети разозлились. «Если у нас нет будущего, зачем вообще ходить в школу?» — спросили они.

Моа была глубоко потрясена и разочаровалась в себе. «Я знаю, что многие иммигранты чувствуют себя чужаками начиная со школы, — говорила она, — и все выглядело так, будто мы привели их в университет, чтобы сказать, что и на работе их будут дискриминировать». Сообщить людям о проблеме не всегда часть ее решения.

* В действительности сложно делать выводы о количестве социологов, которые боятся высказать свое мнение: именно потому, что они молчат.

Как и у всех социологов, у Моа есть идеалы, мечты и политические взгляды. Это ее модели мира. Когда мы находим мотивацию в наших убеждениях и опыте, в этом нет ничего ненаучного, если мы затем проверяем эти модели на соответствие данным. Когда я спросил Моа, как она начинала свою исследовательскую карьеру, она ответила: <[Социолог] Макс Вебер (если не ошибаюсь) как-то сказал — и я с ним согласна, — что вы должны выбирать тему для исследований сердцем, но затем подходить к ней максимально объективно».

Она продолжала: «Эксперименты с резюме заинтересовали меня, поскольку с результатами не споришь. Я изучала реальных людей, а не работала в лаборатории. В ходе этого эксперимента все контролируется, а результаты просты и понятны». Модель проверяется на данных. Моа сказала мне, что она была удивлена, когда не обнаружила гендерной дискриминации при оценке резюме работодателями, даже наоборот: женщинам звонили чаще в профессиях, связанных с компьютерами, где они представлены реже. Это шло вразрез с ее взглядами. «Но таковы факты, — заметила она. — С ними я тоже не могу спорить».

Один из вопросов, где Джордан Питерсон заметнее всего, — обсуждение гендерной разницы при оплате труда. Он правильно указывал: сам по себе тот факт, что в США на каждый доллар, выплачиваемый мужчинам, женщинам платят 77 центов, еще не свидетельствует о дискриминации. Он проводит полезное различие между равенством результатов и возможностей⁴³. Женщинам платят меньше, чем мужчинам, потому что они трудятся на более низкооплачиваемых работах — например, медсестрами. Вполне вероятно, что они могли бы найти более высокооплачиваемую работу, но они предпочли карьерный путь, отличный от мужского. Питерсон также утверждает, что женщины могут быть биологически менее приспособлены к определенным типам лучше оплачиваемой работы. В общем, по его мнению, мы не можем использовать разрыв в зарплате для заявления о наличии дискриминации; нам нужно проверять, были ли у женщин те же возможности, что и у мужчин.

Равенство возможностей — ровно то, что Моа проверяет своими экспериментами с резюме. Когда мусульмане подают заявления на

работу, им звонят реже, чем шведам, и поэтому они подвергаются дискриминации по возможностям. Точно так же результаты Моа показывают равенство в возможностях для шведок при рассылке резюме. В этом случае утверждение Питерсона об отсутствии дискриминации по возможностям справедливо.

Но равенство по результатам можно измерить одним числом (например, гендерный разрыв в оплате труда)*, а вот с равенством по возможностям так не получится. Есть много способов помешать женщинам полностью реализовать свой потенциал, и нужно исследовать много возможных препятствий.

К счастью, социологи стремятся определить их. В 2017 году Катрин Ауспург и ее коллеги опросили 1600 жителей Германии, предоставив им краткое описание возраста, пола, стажа и роли на работе гипотетической личности, а затем спросив, насколько указанная зарплата справедлива⁴⁴. Оказалось, респонденты склонны думать, что женщинам переплачивают, а мужчинам недоплачивают. В среднем они (как мужчины, так и женщины) считали, что женщинам нужно платить 92 цента на каждый доллар, выплачиваемый за ту же работу мужчинам. В то же время на прямой вопрос, следует ли мужчинам и женщинам платить одинаково, подавляющее большинство ответило утвердительно. Так что есть большая разница между тем, что мы говорим, и тем, как действуем на практике. Респонденты в этом эксперименте даже не осознавали, что фактически рекомендуют, чтобы женщинам платили меньше мужчин за одинаковую работу.

В одном исследовании 2012 года выявлено предвзятое отношение ученых к женщинам при оценке резюме на должность лаборанта⁴⁵. Снова оказалось, что и мужчины, и женщины при рассмотрении резюме от соискательниц-женщин считают их менее

* В этой сфере используется довольно сложная и интересная математика. Например, существуют два разных гендерных разрыва — нескорректированный и скорректированный. Первый — простая процентная разница в средней почасовой зарплате женщин и мужчин. Но нужно учитывать разное соотношение мужчин и женщин в конкретной профессии, меньшее среднее количество рабочего времени у женщин и т. д. По причине этого показатель корректируют, вводят «объяснимую» и «необъяснимую» части и т. д. Получить представление о ситуации в ЕС можно из документа комиссии ООН unesco.org/fileadmin/DAM/stats/documents/ece/ces/ge.30/2019/mtg1/WP12_Alcantara_RUS.pdf.

компетентными. Женщины, которым преподают только профессора-мужчины, с меньшей вероятностью продолжат изучать предмет, чем если бы у них были профессора-женщины⁴⁶. Один эксперимент со старшеклассниками показал, что девочки, занимающиеся наукой в классах, где имеются такие стандартные вещи, как предметы из вселенной «Звездных войн», технические журналы, видеогames, научно-фантастические книги и т. д., гораздо реже продолжают проявлять интерес к преподаваемому предмету, чем в случае занятий в классах с менее стандартной обстановкой (изображения природы, картины, ручки, кофеварка, журнал общего содержания и т. д.)⁴⁷. В канадских средних школах девочки, как правило, считают себя слабее в математике, чем мальчики, хотя на экзаменах демонстрируют такие же результаты⁴⁸. В ходе эксперимента с переговорами, проведенного в одном американском университете, оказалось, что женщины настолько же эффективны в этом деле, как и мужчины, когда действуют от имени другого лица, но менее эффективны, когда говорят от собственного имени. Эти различия объясняются страхом отрицательной реакции, если они победят в споре, — тем страхом, который мужчины так остро не ощущают⁴⁹.

Это всего лишь некоторые из многочисленных исследований, которые рассматривали Сапна Черян и ее коллеги — с целью установить препятствия, относящиеся к гендеру⁵⁰. Женщинам и девушкам труднее свободно выражать свое мнение; их недооценивают и мужчины, и другие женщины; им предлагают меньше ролевых моделей; они неявно подвергаются дискриминации, когда подают заявки на определенную работу: это статистически верный способ оценить наши школы и рабочие места. Поскольку большинство, включая Джордана Питерсона, согласны с тем, что нужно стремиться к равенству возможностей, ответ прост: нам нужно информировать людей о результатах тех работ, которые выявляют предрассудки в нашем обществе.

Как ни странно, Питерсон делает противоположный вывод. Он нападает на научные исследования по вопросам меньшинств и гендер, заявляя, что те левые и что их проводят марксисты. Здесь он неправ. Социологи — например, Моя Бёрседл, Катрин Ауспург

и Сапна Черян — осознанно изучают возможности, а не результаты. Возможно, их мотивация — желание добиться справедливости для всех, но именно поэтому очень важно, чтобы на их работу не влияли политические взгляды, которых они придерживаются. Цель всех вышеупомянутых и многих других работ — выяснить, где нет равенства возможностей, чтобы решить проблему. Нет никаких свидетельств идеологической предвзятости этих исследователей.

Питерсон никогда не упоминает об этих работах. Он сосредоточивается на психологических различиях между мужчинами и женщинами. В январе 2018 года в интервью Кэти Ньюман в новостях британского канала Channel 4, которое позже стало вирусным на YouTube, он утверждал: «Покладистые доброжелательные люди вежливы и проявляют сочувствие. Им меньше платят, чем неприятным людям, за ту же работу. Женщины более покладисты, чем мужчины»⁵¹.

Есть несколько причин, почему подобные психологические объяснения не так убедительны, как проверка конкретных моделей с доверительными интервалами. Рациональные основания, стоящие за прямыми и зависящими от контекста вопросами вроде «Как бы вы оценили это резюме?» или за наблюдением, как мужчины и женщины ведут переговоры, заключаются в том, что, понимая действия отдельных людей, мы можем объяснить, как возникает неравенство⁵². А вот покладистость и доброжелательность определяются через тесты самооценки, когда люди отвечают на общие вопросы вроде «Я проявляю сочувствие к эмоциям других людей». Сказать, что кто-то «доброжелателен» и «покладист», — всего лишь способ суммировать ответы из таких анкет. Вовсе не очевидно, почему покладистость — препятствие для повышенной зарплаты. Это работает в обе стороны. Может быть, покладистые люди получают вознаграждение за свое дружелюбие или они плохие переговорщики? Покладистость может оказывать разное влияние в зависимости от карьеры, навыков, необходимых для работы, и стажа.

Сами по себе личностные тесты не дают никакого объяснения, поэтому для проверки связи покладистости и зарплаты требуются

дополнительные изыскания. Одно исследование 59 недавних выпускников в США показало, что на ранних этапах карьеры покладистые люди действительно получают пониженную зарплату⁵³. Но оно также установило, что женщинам платили значительно меньше, чем мужчинам. Даже с учетом всех остальных черт характера, общих умственных способностей, эмоционального интеллекта и успехов покладистость была единственным фактором, который помогал объяснить разрыв в зарплате, и то лишь отчасти. Неприятным женщинам платили все равно меньше, чем неприятным мужчинам, а покладистым мужчинам — больше, чем покладистым женщинам. Это исследование показало, что — вопреки тому, что Питерсон говорил в интервью Channel 4, — не связанные с гендером факторы не могут объяснить разрыв в зарплате.

Отсутствие четкой модели того, как личность влияет на зарплату, очень затрудняет конкретный разговор о том, как личность влияет на возможности. Даже если мы установили, что существует дискриминация покладистых людей в отношении зарплаты, а не прямая дискриминация женщин, нам все равно нужно искать ответ на вопрос, справедливо это или нет. Какие-то соображения (например, что покладистые люди недостаточно эффективно действуют в интересах своей компании) можно счесть справедливыми, а другие (руководство извлекает выгоду из покладистости сотрудников, платя им меньше) — нет. Разговор о личных качествах в общих чертах не помогает нам понять реальные проблемы.

Также важно ближе присмотреться к тому, что на самом деле подразумевает Питерсон, когда говорит, что женщины покладистее мужчин. Здесь мы можем применить уравнение уверенности. Психологи проводили личностные исследования на сотнях тысяч людей. Как мы видели раньше в этой главе, чем больше наблюдений, тем лучше мы можем обнаружить сигнал, скрытый в шуме. Например, если мы провели $n = 400$ анализов людей, то обнаружим разницу даже при отношении сигнала к шуму 1/10. Имея много данных, мы способны использовать уравнение уверенности для определения даже очень малых различий в покладистости мужчин и женщин. И ровно их мы обнаруживаем, когда дело касается личных качеств. Отношение сигнала к шуму для покладистости — той

черты, которая сильнее всего отличает мужчин от женщин, — составляет примерно 1/3. На каждые три единицы шума приходится одна единица сигнала.

Чтобы понять, насколько слаб сигнал, представьте, что из популяции выбираются наугад одна женщина и один мужчина. Вероятность того, что женщина более покладиста, составляет всего 63%. Подумайте, что это означает на практике. Вы стоите перед закрытой дверью и собираетесь познакомиться с Джейн и Джеком. Разумно ли для вас войти в комнату и сказать: «Джейн, я думаю, что ты больше склонна соглашаться со мной, чем Джек, потому что ты женщина», а Джеку сказать: «Ну а с тобой мы, видимо, поспорим»?

Нет. Это статистически неверно. Вероятность ошибки составляет 37%⁵⁴.

Питерсон утверждает (например, в скандинавском ток-шоу *Skavlan*), что психологи «усовершенствовали, по крайней мере отчасти, измерение личностных качеств с помощью улучшенных статистических моделей»⁵⁵. Затем он упоминает огромное количество анкет по личностным характеристикам, которые применялись к сотням тысяч людей. Потом он справедливо признаёт, что женщины и мужчины больше схожи, чем различны. Потом просит нас подумать, «в чем главные различия», и говорит, что «мужчины менее покладисты... а женщины больше склонны к негативным эмоциям или невротизму»⁵⁶.

Хотя это не совсем неверно, рассуждение несколько обманчиво. Подразумевается, что ученые выявили масштабные гендерные различия специфичного рода, использовав большие объемы данных. Правильная интерпретация такова: исследователи придумали сотни различных способов опрашивать мужчин и женщин, как они себя видят, и опросили сотни тысяч людей, но почти все исследования не выявили никаких гендерных различий в личностных качествах. По сути, самый примечательный результат последних тридцати лет личностных исследований — гипотеза о гендерном сходстве. Эта гипотеза, впервые высказанная в 2005 году в журнале *American Psychologist* профессором психологии Висконсинского университета в Мадисоне Джанет Хайд, которая занимается

исследованиями гендеров и женщин, не утверждает, что мужчины и женщины одинаковы. Она говорит, что между личностными качествами есть очень мало статистических различий, зависящих от гендеров. Хайд рассмотрела 124 теста личностных различий и установила, что в 78% из них обнаружены несущественные или небольшие различия между гендерами (отношение сигнала к шуму меньше 0,35)⁵⁷. Эта гипотеза выдержала проверку временем: спустя десять лет новый независимый обзор установил, что только 15% тестов дали для связи между гендером и личностью отношение сигнала к шуму больше 0,35⁵⁸.

В плане невротизма, экстраверсии, открытости, позитивных эмоций, грусти, гнева и многих других черт мужчины и женщины различаются очень слабо или не различаются вообще. В более позднем обзоре Хайд обнаружила, что гендерные различия также малы в математических способностях, навыках устной речи, добросовестности, реакции на вознаграждение, непрямой агрессии, использовании положительных слов, отношении к мастурбации и внебрачным связям, эффективности руководства, самоуважении и самооценке⁵⁹. Сильнее всего полы различаются по интересу к вещам и людям, физической агрессии, использованию порнографии и отношению к случайному сексу. Оказывается, некоторые предрассудки всё же верны. Наши личностные характеристики сильно разнятся. Мужчины отличаются друг от друга. Женщины тоже. Но в целом статистически неверно утверждать, что личностные характеристики мужчин и женщин очень разные.

Опасность заявлений Джордана Питерсона в том, что исследования гендеров каким-то образом извращены «левыми» или «марксистскими» силами. Всё наоборот. Джанет Хайд, имеющая диплом бакалавра по математике, — часть статистической революции, измеряющей равенство возможностей и изгоняющей идеологическое мышление из психологии и социальных наук. За свои работы она получила несколько наград, включая три от Американской психологической ассоциации — крупнейшей научной и профессиональной организации психологов в США. Гендерные различия изучены так подробно, что документировано даже мельчайшее различие. Аналогичные результаты были получены в нейробиологии:

различия среди представителей каждого пола гораздо больше, чем различия между женским и мужским мозгом⁶⁰. Ирония в том, что Питерсон для поддержки своей точки зрения с идеологической мотивацией может выковыривать подходящие ему результаты только из этого обширного массива строгих исследований.

Уравнение уверенности учит нас заменять байки наблюдениями. Никогда не полагайтесь на историю одного человека — даже собственную. Когда вы побеждаете, тщательно поразмыслите, длилась ли полоса успехов достаточно долго, чтобы вы могли ее приписать своему умению. Всегда есть тот, кому везет, и, возможно, на этот раз этим человеком были вы. Ищите другие истории и собирайте статистику. По мере того как вы изучаете информацию, рассуждайте в рамках правила о квадратном корне из n : чтобы обнаружить сигнал вдвое слабее, нужно вчетверо больше наблюдений. Если вы «вверху» и это статистически лучше, чем у окружающих, проверьте свое привилегированное положение с помощью статистического интервала. Будьте корректными: осознавайте плюсы и минусы того, что у вас есть. Будьте уверенными, не обманывая себя, а понимая, как именно общество формирует вашу жизнь. Только тогда вы сможете найти преимущество и заявить о нем.

Глава 4

Уравнение умений

$$P(S_{t+1}|S_t) = P(S_{t+1}|S_t, S_{t-1}, S_{t-2}, \dots, S_1)$$

Я сижу вечером в кафе и вижу, как он заходит. Пожимает руку официанту, затем руку бариста, обмениваясь улыбками и парой слов. Сначала он меня не видит, а когда я встаю, чтобы подойти к нему, замечает еще одного знакомого. Теперь раунд объятий. Я снова сажусь, ожидая, пока он закончит.

Его известность здесь частично объясняется его прошлым профессионального футболиста и тем, что он часто появляется в телевизоре, но он популярен и благодаря тому, как себя держит: уверенно, дружелюбно, находит время поговорить с людьми, уделив несколько слов каждому.

Усевшись со мной, уже через несколько минут он вовсю разглагольствует: «Я думаю, что делаю хорошее дело, поскольку показываю людям свои способы работы. Мне кажется, иногда этого не замечают. Я просто делаю свое дело, говорю как есть, я честен, как и нужно в этой игре. У меня куча контактов. Полно встреч, как вот эта с вами, чтобы поддерживать связи. Люди желают поговорить со мной, поскольку у меня уникальный способ смотреть на всё. Из-за моего прошлого, ни у кого другого этого нет. Вот что я хочу донести, когда сижу здесь с вами...» Наблюдения перемежаются байками о временах, когда он играл, несколькими отрепетированными историями, где упоминаются громкие имена, и своевременными шутками.

Он улыбается, смотрит прямо в глаза и временами заставляет меня чувствовать, что я просил у него всю эту информацию. Но я просил не это.

Я хотел поговорить об использовании данных — как в СМИ, так и в футболе. К сожалению, пока я не слышу ничего толкового.

Я называю людей такого рода «Мой путь» — по песне, которую прославил Фрэнк Синатра*. Суть каждой из их историй — двигаться осторожно, встречать вызовы с поднятой головой и доводить дело до конца. Из этого может получиться красивая мелодия, и в те две-три минуты, пока очередной господин «Мой путь» обнимается и здороваются, проходя по кафе, окружающие получают удовольствие.

Но это работает, пока он переходит от одного человека к другому.

А я сейчас застрял в этом положении, и мне некуда деться.

Стыдно признаться, но первые несколько раз, когда я разговаривал с футболистами вроде господина «Мой путь», я верил их историям. С момента публикации моей книги «Футболоматика» в 2016 году** меня приглашали в некоторые ведущие футбольные клубы мира, а их представители посещали меня. Меня приглашали на радио и телевидение, чтобы рассказывать об игре вместе с бывшими профессионалами. Переход из научной среды к общению с бывшими футболистами, телеперсонами, скаутами и членами советов директоров клубов Премьер-лиги опьяняет. Мне нравилось слушать закулисные истории об играх и больших матчах и узнавать о жизни на тренировочных полях. Если использовать самое распространенное футбольное клише, то мечта стала реальностью: я превратился из болельщика в того, которому доверяют люди, близкие к самому действию.

Я до сих пор люблю слушать такие истории и воочию наблюдать реальный мир футбола. Но интересные истории слишком часто сопровождаются «геройскими» рассказами о взглядах очередных господ типа «Мой путь», о том, как их успехам помешали жульничающие соперники и как бы они могли сделать всё лучше, если бы им дали такую возможность.

* В рамках песни слова My Way точнее переводить «по-своему».

** Переведена на русский язык: Самптер Д. Футболоматика. Как благодаря математике «Барселона» выигрывает, Роналду забивает, а букмекеры зарабатывают состояния. М. : Бомбара, 2018.

Поскольку я математик, эти парни часто решают, что должны объяснить мне ход своей мысли. Они начинают рассказывать, что я смотрю на мир не так, как они, — при этом не выяснив, как же я воспринимаю жизнь.

«Думаю, статистика отлично подходит для размышлений о прошлом, — скажет мне такой человек, — но у меня есть предвидение и интуиция для будущего».

Потом он объяснит, что обладает уникальной способностью определять свое конкурентное преимущество. Или как его уверенность в себе и сильный характер помогают принимать правильные решения. Или как он обнаружил метод находить закономерности в данных, которые я (по его мнению) упустил. В своих рассказах он, как правило, обходит неудобные моменты. «Ошибки я начал делать, только потеряв концентрацию», — говорит он мне. И всегда подчеркивает свои сильные стороны: «Когда я сосредоточен, я все делаю правильно».

Однако, начав работать в индустрии футбола, я уж точно не осознавал, сколько времени мне придется выслушивать рассказы людей о том, почему они считают себя особенными.

Мне следовало бы это понимать лучше, поскольку такое происходит не только в футболе. Я встречал то же и в промышленности и бизнесе: специалисты по инвестициям говорили мне о своих уникальных качествах. Им не нужна математика, ведь у них есть чутье, которого нет у специалистов по количественному анализу («квантов»*). Руководители технических компаний объясняли мне, что их стартап преуспел благодаря уникальным взглядам и талантам. Даже ученые занимаются тем же. Исследователи-неудачники рассказывали, как другие украдли их идеи, а успешные ученые утверждали, что придерживались своих принципов. Каждый из них делал это по-своему.

Встает вопрос, на который сложно ответить: как узнать, говорят мне что-то полезное или нет?

Парень, с которым я сейчас сижу, очевидно, просто переполнен. Он без остановки говорит о себе последние полтора часа. Но

* Quant — разговорное сокращение от quantitative («количественный»).

многим другим людям есть что сказать полезного, причем включая иногда и господина «Мой путь». Вопрос в том, как отделить ценные сведения от самолюбования.

*

Подход специалистов по прикладной математике к этому вопросу таков: разделить все, что вам говорят, на три категории. Первые две обсуждались в предыдущих главах: это модель и данные. Модели — наши гипотезы о мире, а данные — опыт, который позволяет нам установить истинность или ложность наших гипотез. Господин «Мой путь», о котором я сейчас рассказываю, создает третью категорию: бес-смыслица. Он рассказывает истории о своих триумфах, неудачах и ощущениях, но ничего конкретного о том, как он думает или что он знает.

Я использую дефис в слове «бес-смыслица», чтобы заставить вас задуматься. Этот трюк позаимствован у философа Алфреда Джулса Айера из Оксфордского университета, который способствовал становлению моего взгляда на математику. Айер признавал, что «бес-смыслица» — очень провокационное слово, но использовал его для описания информации, которая не исходит от наших чувств. То, что чувствует «Мой путь», как он воспринимает свои успехи и неудачи, не основано на наблюдениях или на том, что можно измерить. Айер внес предложение: когда «Мой путь» или кто-то еще рассказывает вам что-то, вам нужно спросить, проверяется ли его утверждение. Можете ли вы в принципе определить, истинно ли какое-то утверждение или нет, используя данные, которые получаете от органов чувств?

Проверяемые утверждения, например, «наш самолет вот-вот потерпит крушение», «Рэйчел — стерва», «чудеса случаются», «Ян и Мариус обладают определенным преимуществом на рынке ставок», «шведские работодатели принимают расово предвзятые решения о том, кого приглашать на интервью», «Джесс нужно поменять работу, если она желает ощущать себя более счастливой» и т. д. Утверждения именно такого рода я формулировал в качестве моделей в этой книге. Когда мы сравниваем модели с данными, то проверяем, насколько они истинны.

Мы не требуем, чтобы у нас имелся доступ к данным для проверяемости нашей модели. Когда Айер в 1936 году опубликовал книгу «Язык, истина и логика»*, которая объясняла принцип проверяемости, еще не было фотографий обратной стороны Луны. А потому нельзя было утверждать, что гипотеза о существовании там гор истинна или ложна. Однако она относилась к проверяемым; и когда советский космический аппарат «Луна-3» облетел вокруг нашего спутника в 1959 году, гипотеза была протестирована.

Утверждения об ощущениях господина «Мой путь» и его вере в себя — дело другое. Его байки могут содержать интересные подробности, имена реальных людей и события, которые происходили на самом деле, но они не поддаются проверке. Мы не можем придумать тест, который бы подтвердил (или опроверг) то, что он обладает «уникальным способом смотреть на вещи»; или проверил бы, что именно есть у него такого, чего «ни у кого другого нет»; или откуда он знает, что относится к делу, а что нет. Такой тест невозможен, поскольку человек неспособен правильно объяснить основу своих утверждений. Он не может отделить свое ощущение от фактов, и мы не можем переформулировать его утверждения в виде модели, которую несложно проверить на соответствие данным. Песню господина «Мой путь» определяет мешанина личных идей. То, что он говорит, не данные и не модель. Это бессмыслица в буквальном смысле слова.

*

«La Masia», Барселона. Когда дело касается глубокого интеллектуального подхода к этой прекрасной игре, ничто не сравнится с тренировочным комплексом футбольного клуба «Барселона». Академию для молодых игроков основал в 1979 году легендарный Йохан Кройф; здесь разработана философия, которая пронизывает все, что делается в клубе.

Я прошел мимо небольшой группы болельщиков, каждый из которых надеялся взглянуть на игроков, пока они будут входить или выходить через главные ворота, и добрался до бокового входа новой

* Издана на русском языке: Айер А. Дж. Язык, истина и логика. М. : Канон+РООИ «Реабилитация», 2010.

«Ла Масии». Так же, как многие университеты переехали из старинных традиционных зданий в новые сверкающие сооружения, «Ла Масия» перекочевала из фермерского дома, где изначально располагалась, в современный корпус со стеклянными фасадами.

В «Ла Масию» меня пригласил Хавьер Фернандез де ла Роза, глава отдела спортивной аналитики клуба «Барселона» и докторант в области искусственного интеллекта. Он попросил меня провести презентацию моей недавней работы и рассказать о способах анализа игры.

Внутри новая «Ла Масия» также похожа на современный факультет университета, поскольку здесь ведется не только преподавание, но и исследовательская работа. Только что закончили тренировку игроки основной команды, молодежь занималась на другом поле. Хавьер сидел в ярко освещенном офисе с батареей мониторов перед собой и рядами книг позади. В других клубах, где я бывал, доминировали места для тренировок, а аналитиков можно было найти разве что в труднодоступном закутке. Здесь же и у игроков было все, что может им понадобиться, и у исследователей — собственное пространство для работы и размышлений, чтобы планировать и совершенствовать стиль игры команды. Организация пространства в «Ла Масии» отражала тот футбол, каким видел его я: разум играет вместе с телом.

Мы с Хавьером тут же взялись за работу. Отправились в его офис, уселись за свои ноутбуки и стали сравнивать записи. Как вы оцениваете передачи? Как отслеживаете движение игроков? Как делите матч на разные игровые состояния? Каким определением контратаки пользуетесь? Как моделируете контроль поля? Вопросы и ответы летали туда-сюда. Данные, модель, данные, модель, данные, модель, потом еще больше моделей. Это продолжалось и продолжалось.

В какой-то момент — как мне показалось, довольно неожиданно — Хавьер сказал, что мне пора провести семинар для остальной части его группы. Мы прошли в просторный зал для семинаров, подключили мой ноутбук к большому экрану, и я начал говорить снова — на этот раз перед аудиторией из тренеров, скаутов и аналитиков. Затем пять или шесть человек из первого ряда

стали прерывать меня и задавать вопросы об использованных данных, моих предположениях и результатах. Они рассказали о собственных находках и о том, что я мог бы улучшить.

Команда аналитиков «Барселоны» давала мне именно то, что я люблю в исследованиях: глубокое погружение непосредственно в модель и данные. Идеальный день исследований увенчался первым рядом мест стадиона, и я смог вечером увидеть в действии Лионеля Месси и компанию. Пока солнце садилось за стадионом «Камп Ноу», я, вероятно, был как никогда близок к тому же движущемуся телу, которое видел ранее тем же днем в координатах в виде кривой на экране своего компьютера.

*

Я сосредоточил свою презентацию в «Барселоне» на одном конкретном игроке. В то время, через несколько месяцев после чемпионата мира 2018 года, меня очень интересовал Пол Погба. И «Барселона», если верить газетным слухам, тоже им интересовалась.

Я был поклонником Погба очень долго, поскольку он больше других определяет игру своей команды. Хотя Месси — талисман «Барселоны», философия клуба состоит в том, чтобы быть чем-то большим, нежели просто суммой частей, и не сосредоточиваться на каком-то одном человеке. Безусловно, Криштиану Роналду заметен на поле, но по большому счету он традиционный, очень атлетичный бомбардир. Стиль мадридского «Реала» или «Ювентуса» не строится вокруг его способностей.

Когда Пол Погба играет за «Манчестер Юнайтед», он и есть эта команда, а во время чемпионата мира он определял игру сборной Франции, выигравшей трофей. Такова моя гипотеза. Но как ее проверить? В отличие от Месси или Криштиану Роналду, Погба не забивает прорву голов. На чемпионате мира он отправил мяч в ворота всего один раз. Да, в финале, что само по себе достижение, но множество игроков забили больше его. Так что одни только голы не объясняют его мастерство.

Математическая идея, которую я использовал для оценивания Поля Погба, — сосредоточиться на его вкладе в команду, забивающей гол, а не на голах, которые он забивает сам. Здесь футбольный

болельщик может спросить, не имею ли я в виду голевые передачи — те, которые привели к успеху. Первая передача — пас игроку, который забил мяч; вторая (предголевая) — пас игроку, который отдал затем голевую передачу, и т. д. Учет передач в голевых комбинациях — часть моего подхода, но небольшая. Вместо того чтобы придавать особый статус голам и голевым передачам, я оцениваю все действия на поле: отборы, пасы, перехваты и т. д. Моя цель — измерить, как каждое из них повышает вероятность того, что своя команда забьет гол, и снижает вероятность того, что гол забьет соперник.

Чтобы добиться этого, сначала нам нужно подумать, как описать футбольный матч числами. Допустим, пас делается из точки (x_1, y_1) в точку (x_2, y_2) . Чтобы понять такие координаты, представьте поле с высоты птичьего полета, и пусть ось x проходит по боковой линии, а ось y — по линии ворот. Началом координат $(0, 0)$ будем считать угловой флагок справа от ворот атакующей команды. Тогда у противоположного углового флагка будут координаты $(105, 68)$ (типичное футбольное поле имеет длину 105 метров и ширину 68 метров: такие размеры рекомендует ФИФА). Каждый пас во время матча можно описать примерно так: вратарь выбивает далеко мяч на фланг: $(10, 30) \rightarrow (60, 60)$, передача мяча к центру поля: $(60, 60) \rightarrow (60, 34)$, подача в чужую штрафную площадь: $(60, 34) \rightarrow (90, 40)$. Представьте себе игру в виде последовательности координат, определяющих передачи и дриблинг игроков. Каждый фрагмент игры (цепочка владения) разбивается на отдельные действия, которые можно описать в координатах x и y .

Теперь мы хотим определить, как любое действие отдельного игрока в таких цепочках владения увеличивает шансы команды на гол и/или снижает шансы на гол у соперника. Для этого я делаю математическое предположение. Как правило, когда математик говорит вам, что он собирается «сделать предположение», это означает, что он желает сказать что-то ложное и просит вас попридержать скепсис и использовать свое воображение. Это несколько отличается от повседневного использования данного слова. Например, я могу сказать жене о гостях, приглашенных на ужин: «Предполагаю, они появятся около семи». Или: «Предполагаю, что мы снова проиграем», если моя команда проигрывает два мяча за пять минут до

конца игры. И то и другое, скорее всего, верно, но это не математические предположения.

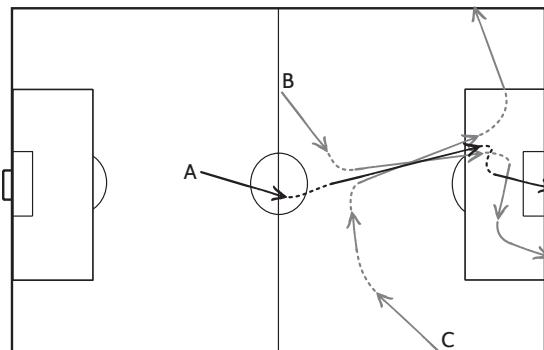
В математике слово «предполагать» используется, чтобы описать некоторое множество вещей, которые необязательно верны, но о которых мы не желаем сейчас беспокоиться. Я просто хочу, чтобы вы придержали свое неверие, а затем посмотрим, к чему нас приведет наше предположение, — при этом не станем обсуждать само предположение. Однако важно, чтобы оно было сделано с самого начала, поскольку это основание нашей модели, и когда мы сравниваем модели с реальностью, нам нужно быть честными в отношении их ограничений.

Мое предположение таково: качество паса в футболе зависит от координат его начальной и конечной точек, а не от того, что происходит до и после паса, или от того, какие игроки находились на поле во время этого паса, или от чего-то в таком духе. Иными словами, если Погба может отдать пас с середины поля, скажем из точки (60, 34), в штрафную площадь — в точку (90, 40), эта передача всегда будет иметь одинаковое влияние на шансы Франции забить гол — независимо от того, что еще происходит в матче.

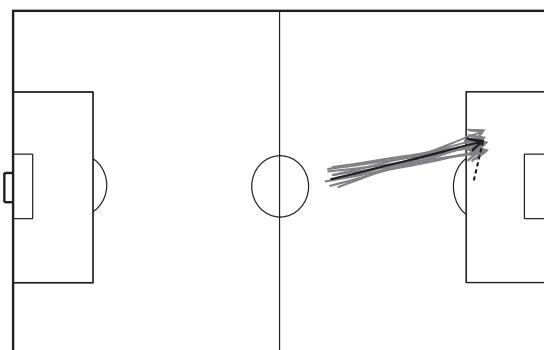
Впрочем, очевидно, что такое предположение некорректно. Например, в игре чемпионата мира против Перу за одну минуту Погба сделал две передачи в штрафную примерно из одной точки поля. В первом случае мяч перелетел через защитников и попал к Мбаппе, который пытался ударить пяткой в прыжке, но не смог направить мяч мимо вратаря. Во втором случае мяч, пущенный низом, дошел до Оливье Жиру, удар которого заблокировал защитник, но после рикошета Мбаппе открыл счет в матче. Согласно моему предположению, эти две передачи — одна из которых привела к голу, а другая нет — имеют одинаковую ценность для Франции как команды.

Отбросив недоверие, мы можем построить модель всего, что происходит в футбольном матче. Вместе с моим коллегой Эмри Долевом я использовал базу данных, куда входят начальные и конечные координаты всех передач, сделанных за много сезонов на высшем уровне — в английской Премьер-лиге, Лиге чемпионов, испанской Ла Лиге, на чемпионатах мира и т. д. Мы проверили, приводил ли каждый пас в итоге к удару по воротам. Это позволило нам создать

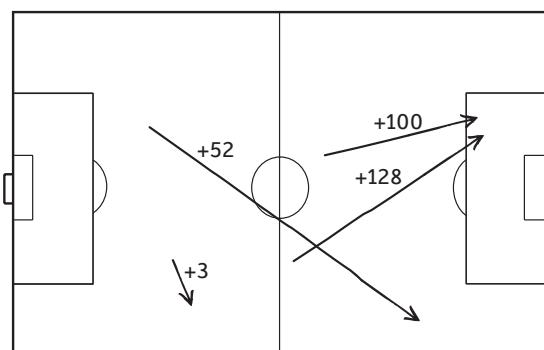
статистическую модель, которая связывает координаты начальной и конечной точек паса с вероятностью забить гол (см. рис. 4). Так мы смогли присвоить определенную оценку каждому пасу — независимо от того, что происходило перед ним или после него.



Каждая цепочка владения получает значение 1, если она завершилась голом, 0, если мяч вышел из игры. Таким образом, цепочке А присвоено значение 1, а у цепочек В и С значения равны 0.



Из предположения о марковости следует, что оценка какого-то паса — доля времени, когда он входит в цепочку, ведущую к голу. В данном случае из 10 схожих пасов 1 ведет к голу, то есть мы получаем значение 0,1.



Умножаем полученную долю на 1000, чтобы каждой передаче Поля Погба присвоить значение от 0 до 1000. Общая сумма таких чисел определяет умение Погба делать передачи.

Рис. 4. Как предположение о марковости может оценивать передачи в футболе

Когда мы с Эмри присвоили определенное значение любому действию в любом матче, мы смогли в итоге оценить Поля Погба. Он выделяется двумя умениями: вернуть мяч в центр поля и мгновенно превратить защиту в атаку с помощью длинных точных передач. Во время чемпионата мира он сделал несколько потрясающих пасов, получая мяч в центре поля, разворачиваясь и доставляя его в ноги партнеров по команде далеко на половине противника. Он увеличивал шансы Франции забить больше, чем любой другой игрок команды.

У «Барселоны» уже есть человек, который играет на поле схожую роль: это Серхио Бускетс. Месси — звездный атакующий игрок, имя которого известно всем, а Бускетс — мотор, который приводит команду в движение, начиная атаки с середины поля. Бускетс и Погба во многом разные, однако их роднит способность производить сильное впечатление игрой в полузащите. Бускетс на пять лет старше Погба; когда двигатели стареют, они становятся менее эффективными.

Разработанную нами с Эмри модель можно применить к любому профессиональному игроку в любом матче. Она может за секунды оценить их так же, как оценила Погба. Это позволяет командам подбирать себе игроков, которые точно соответствуют требованиям. Когда один игрок уходит, ему можно подобрать сделанную на заказ замену.

Традиционный способ оценивать игроков — заставлять скаутов смотреть матчи и писать отчеты. Технический директор одного известного клуба недавно показал мне базу данных для потенциальных кандидатов. Он мог искать 17-летних футболистов из пятой лиги Швеции или 15-летних игроков в бразильских юношеских лигах. Зеленая галочка рядом со спортсменом означала, что его игру просматривал скаут. Директор мог щелкнуть по любому игроку мира и прочитать отчеты разных скаутов о нем.

Наша модель дополняет этот подход. Она рассматривает конкретное умение перемещать мяч из одной точки поля в другую. Когда какой-нибудь скаут оценивает игрока, он благодаря своему опыту оценивает его расположение на поле, его понимание обстановки и взаимодействие с партнерами. Однако каким бы хорошим

ни был скаут, он не может заявить, что оценивает каждый пас игрока в Премьер-лиге. А вот модель вполне может.

Когда я говорю с тренерами и скаутами, то свои предположения описываю именно так. Вместо «статистика показывает, что Погба был лучшим полузащитником на чемпионате мира» я говорю: «Если нас интересует, насколько далеко игрок продвигает мяч от центра поля, то и при игре на чемпионате мира, и при игре за «Манчестер Юнайтед» Погба — один из лучших в мире по этому показателю».

Изложение наших предположений и выводов в разговоре с другими людьми очень важно; и не только когда мы говорим о футболе, но и когда обсуждаем всё, что нас волнует. Разделение мира на модели, данные и бессмыслицу требует от нас быть честными в отношении того, что мы предполагали, когда делали выводы. Оно просит нас хорошенько подумать о точке зрения — как нашей собственной, так и других людей.

*

Основа большинства математических моделей для измерения умений — уравнение, выражающее так называемое марковское свойство (марковость)*. Будем называть его марковским предположением. Вот оно:

$$P(S_{t+1}|S_t) = P(S_{t+1}|S_t, S_{t-1}, S_{t-2}, \dots, S_1) \quad (\text{Уравнение 4}).$$

Величина $P(S_{t+1}|S_t)$ определяется так же, как в уравнении 2 в главе 2. P означает вероятность того, что мир находится в состоянии S_{t+1} , а символ $|$ — «при условии». Дополнительный компонент в формуле — индексы $t+1, t, t-1$ и т. д. для каждого из таких событий. Иными словами, $P(S_{t+1}|S_t)$ означает «вероятность того, что мир будет находиться в состоянии S_{t+1} в момент времени $t+1$, при условии, что ранее он находился в состоянии S_t в момент времени t ».

* Название дано в честь русского математика Андрея Андреевича Маркова, который первым стал изучать случайные процессы и последовательности случайных величин, обладающие таким свойством. Они тоже носят его имя (марковские процессы и цепи Маркова).

Ключевая идея марковости в том, что будущее зависит только от настоящего, но не от прошлого, которое привело нас к этому настоящему. Уравнение 4 говорит, что будущее состояние в момент времени $t + 1$ зависит только от состояния в настоящее время t , так что мы предполагаем, что прошлые состояния $S_{t-1}, S_{t-2}, \dots, S_1$ роли не играют.

Для примера представим Эдварда, бармена в оживленном заведении. Он стремится обслуживать своих клиентов максимально быстро. Количество гостей может меняться, но Эдвард старается взять как можно больше заказов. Обозначим S_t количество людей, которые ожидают заказа в минуту t .

Эдвард приступает к работе. В момент начала его смены обслуживания ждут $S_1 = 2$ человека. Нет проблем. Он наливает пару пинт первому парню в очереди и приносит бокал вина женщине за ним. Пока он обслуживает этих двоих, в очереди появляются еще трое, поэтому в минуту $t = 2$ в очереди $S_2 = 3$ человека. Эд всех их обслуживает и видит, что в минуту $t = 3$ ждут уже $S_3 = 5$. На этот раз он успевает обслужить только трех человек, а двое остаются, плюс за следующую минуту к ним добавляются еще четверо, и $S_4 = 6$.

Марковское предположение говорит, что для измерения умений Эда достаточно знать скорость обслуживания клиентов: нам требуется знать, как S_{t+1} зависит от S_t . Количество людей, которые ждали ранее в тот вечер — $S_{t-1}, S_{t-2}, \dots, S_1$ — не играет роли в анализе его умений в данный момент. Для оценки мастерства бармена это вполне разумное предположение. Эдвард может обслуживать примерно двух-трех человек в минуту — это разность между S_{t+1} и S_t .

Руководитель Эда, незнакомый с марковским предположением, может заглянуть в бар, заметить много ожидающих в очереди клиентов и заключить, что Эд плохо выполняет свою работу. Эд мог бы разъяснить боссу марковское свойство и сказать, что нужно учитывать два показателя: скорость появления людей в баре и скорость их обслуживания. Эд несет ответственность только за второй. Или он мог бы просто сказать: «Сегодня здесь действительно напряженно. Просто посмотрите, как усердно я работаю». Так или иначе, Эд использует марковское предположение, чтобы объяснить, как корректно измерять его умения.

Уравнение 4 отличается от виденных нами ранее тем, что не дает ответа сразу. В уравнениях 1–3 мы закладывали данные в модель и улучшали свое понимание настоящего или ближайшего будущего. Уравнение 4 — предположение. Это шаг к ответу, однако еще не ответ. Когда мы имеем дело с работой бармена, марковское предположение предлагает нам смотреть на скорость, с которой Эдвард обслуживает клиентов. Аналогичное предположение мы делали в модели футбольных пасов: предполагали, что можно забыть все, что происходило до получения мяча Полем Погба, и все, что происходит после передачи. Это позволяло нам измерить, как конкретные передачи помогли его команде.

Важно быть честным и до создания модели (при предположениях), и после (сработали ли они так, как мы думали). В этом отличие от господина «Мой путь», который объясняет свои несчастья чужими ошибками или просто невезением. Мастерство создателя модели в том, чтобы определить, какие события включить в нее, а какие можно безопасно игнорировать. Какие события и измерения действительно характеризуют истинную ситуацию в баре, футбольной команде или в какой-то другой организации?

Мы могли ошибиться в своем предположении. Пока мы подбадривали Эда, на максимальной скорости смешивающего и подающего свои коктейли, его босс высовывает голову из офиса во второй раз. Теперь он видит огромную кучу грязных стаканов. Эд забыл включить посудомоечную машину! Наша ошибка и досадная ошибка Эда объясняются неверным предположением. Мы сочли, что единственное важное в баре — посетители, и забыли о мытье посуды.

Управляющий показывает Эду, как работать с посудомоечной машиной, и говорит ему, что теперь будет оценивать его и по скорости обслуживания посетителей, и по скорости мытья посуды. Вместе они переделывают свою модель; и теперь, например, состояние $S_t = \{S, 83\}$ говорит, что в баре 5 ожидающих посетителей и 83 грязных стакана. Теперь и Эд, и управляющий счастливы. Пока менеджер не заметил, что Эд забыл взять еду, которую заказали клиенты...

Когда вы рассматриваете свою жизнь, ключ к успеху — честность в тех аспектах, которые вы пытаетесь улучшить. Например, можно считать самым важным фактором успеха свою заработную плату. Марковское предположение говорит, что нужно меньше беспокоиться о тех повышениях, которые у вас были в прошлом и больше не относятся к делу, а больше заботиться о том, как ваши нынешние действия улучшают ваш доход. Честно скажите себе, что для вас важна именно зарплата, но если из-за того, что вы тратите слишком много времени на работу, начнут страдать отношения с близкими, то объясните им, что вы ошиблись в своих предположениях. Пересмотрите их и начните заново.

*

Принцип проверяемости, изложенный Айером в книге «Язык, истина и логика», возник из мышления группы философов, известных как «Венский кружок». У его истоков стояли физик Мориц Шлик, возглавлявший группу, и Рудольф Карнап, ученик выдающегося логика и математика Готлоба Фреге⁶¹. Сумрачным героем этого движения был Людвиг Витгенштейн. Он учился у Бертрана Рассела в Кембридже и не участвовал активно в работе кружка, но именно его «Логико-философский трактат»* наиболее четко продемонстрировал идею, что все значимые утверждения должны проверяться с помощью данных. Седьмое положение Витгенштейна — «О чем невозможно говорить, о том следует молчать» — было окончательным «заткнитесь» для всех, кто сомневался в силе верификации.

В 1933 году 22-летний Айер как-то сумел получить приглашение поучаствовать в дискуссиях «Венского кружка», а спустя три года появилась его книга. Благодаря ему подход кружка, известный как логический позитивизм, распространился из континентальной Европы в Англию. Вторая мировая война привела Карнапа и его идеи в США. К моменту победы в войне почти весь западный мир принял принципы экспериментальной проверки.

* «Логико-философский трактат» был опубликован на немецком языке в 1921 году. В 1922-м появился английский перевод.

В первой половине XX столетия мышление в духе логического позитивизма трансформировало «Десятку». В фокусе всех научных исследований оказались модели, а Альберт Эйнштейн переписал законы физики, используя новую математику. Теперь такой подход должен был получить уникальную авторитетность. Модели и данные становились не просто одним из способов познания мира, а единственно возможным.

Дело не в том, что участники «Десятки» создавали свои исследовательские кружки, чтобы лучше понять Витгенштейна, Рассела, Карнапа и Айера. Некоторые из них читали философские труды, но большинство следовали собственным суждениям о том, как применять модели, и приходили к тем же выводам, что и эти философы. Помните, что в сознании участников «Десятки» никакой «Десятки» нет, так что не может быть и сораний для определения ее принципов. Но логический позитивизм хорошо вписывался в собственное мышление общества. Он описывал именно то, что они делали с тех времен, когда де Муавр создал уравнение уверенности.

В Европе настал золотой век «Десятки». В России на рубеже веков это общество основал Андрей Марков (прославившийся цепью), а после революции в СССР его деятельность продолжил другой Андрей — Колмогоров. Он определил аксиомы для вероятности, объединив в общую структуру работы де Муавра, Байеса, Лапласа, Маркова и других математиков. Теперь код можно было передавать от учителей небольшим группам студентов. Летом Колмогоров приглашал самых способных учеников к себе на дачу, где предлагал им задачи, обсуждая их по очереди, оттачивая умения учеников и совершенствуя код. Несмотря на чистки, советское руководство раз за разом доверяло «Десятке» — чтобы развивать социальные идеи, вести космическую программу и разрабатывать новую экономику.

Схожий дух интеллектуальной свободы и доверия к «Десятке» распространился в Европе. В Великобритании центром математики был Кембридж. Именно здесь Рональд Фишер переписал в уравнениях теорию естественного отбора; Аллан Тьюринг описал свою универсальную вычислительную машину и изложил основы информатики; Джон Мейнард Кейнс использовал свои студенческие

математические исследования, чтобы изменить то, как правительства принимают решения в экономике; а Берtrand Рассел вел синтез западной философии. И именно в Кембридж приехал учиться в конце войны Дэвид Кокс.

В Австрии, Германии и Скандинавии «Десятка» лихо разбиралась с физическими вопросами. Эрвин Шрёдингер написал уравнения квантовой механики, Нильс Бор предложил математику атома, Альберт Эйнштейн — ну, сделал все, чем он знаменит. Французы, изгнавшие де Муавра двести лет назад, не были полностью убеждены в принципе проверяемости до окончания войны (а может, и позже). Но именно французский математик Анри Пуанкаре заложил основы области математики, позднее ставшей известной как теория хаоса.

Разделение «Десяткой» модели и данных было важнейшим событием еще и потому, что на него не влияли религиозные убеждения. От христианства, которому Ричард Прайс приписывал уравнение суждений, потихоньку отказались. Проверка здесь была невозможна. Предположение, что математические истины нам дал Бог, считалось бессмысленным. Идея, что мы можем жить в аллегорической пещере Платона, была нонсенсом. Тот факт, что уравнение уверенности берет начало в азартной игре, ничего не менял в его применимости и поэтому тоже не имел значения. Все представления о религии и этике следовало отбросить и заменить строгим проверяемым мышлением.

*

Нынешняя группа участников «Десятки» совместно обсуждает текущие проблемы: теорию относительности, изменения климата, бейсбол или опросы по поводу брексита. В последние сто лет темы менялись, но характер дискуссий оставался тем же. Он отличается точностью. Члены общества честны в своих предположениях. Они обсуждают, какие аспекты мира их модель объясняет, а какие нет. Когда они не согласны, то сравнивают предположения и внимательно изучают данные. Создатель модели может ощущать гордость, когда она наилучшим образом интерпретирует данные, или испытывать легкое разочарование, когда она не срабатывает, но

важнее всего сама цель — найти объяснение, которое будет в наименьшей степени неверным.

Тех, кто не говорит на языке моделей и данных, либо потихоньку предостерегают, либо вежливо игнорируют — от предубежденных политиков, орущих футбольных тренеров и гневных болельщиков до сверхусердных экоактивистов и невежественных отрицателей, толкователей культурных войн и марксистов-фундаменталистов, Дональда Трампа и женоненавистников. «Десятка» — небольшая группа, которая постепенно приближается к истине, пока остальное человечество движется все дальше по спирали.

*

Люк Борнн расслаблен, носит футболку и улыбается в камеру своего компьютера. Мы разговариваем по скайпу в феврале 2019 года: он сидит в своем ярко освещенном офисе в Сакramento; я на своем темном цокольном этаже в Швеции. За ним я вижу растянутую майку баскетбольного клуба «Сакramento Кингз» с его фамилией, а в другой части офиса — обязательный ряд научных книг. Разница во времени заметна по нашей энергичности. Я сижу и пытаюсь вспомнить свои вопросы, а Люк крутится в офисе на своем стуле, выхватывая книги с полки и показывая их мне на экране.

Люк — вице-президент по стратегии и аналитике в «Кингз». Он не следовал тому, что сам называет традиционным путем в баскетбол и к его нынешней должности. Он тыкает пальцем за спину и говорит:

— Когда мы последний раз дали объявление о вакансии на роль аналитика сюда, у нас было больше тысячи заявлений, и большинство мечтали работать в спорте с того времени, как были вот такими, — его рука показывает рост четырехлетнего ребенка. — Я только начал научную работу в Гарварде, моделируя передвижение животных и системы климата, и случайно встретился с Кирком Голдсберри [аналитик НБА, бывший специалист по стратегии в баскетбольном клубе «Сан-Антонио Спёрс»], где он показал мне все эти данные.

Люка очаровал масштаб наблюдений. Здесь была информация о здоровье и форме игроков, нагрузке на их суставы, схемы перемещений всех спортсменов во время тренировок и матчей, данные о передачах и бросках. Фиксировалось все относящееся к баскетболу, однако тренеры команды использовали это очень слабо.

— Для меня, — говорит Люк с нарастающим возбуждением в голосе, — это был не просто крутой спортивный проект, а буквально самая интересная научная задача, с которой я когда-либо сталкивался.

Он обладал идеальным набором умений для того, чтобы справиться с этой задачей, и быстро выдал результат. В докладе, представленном на конференции по спортивной аналитике в Школе менеджмента Слоуна при Массачусетском технологическом институте в 2015 году, они с Кирком ввели новый показатель для игры в защите, который называли *counterpoints**. Успех подхода Люка вызвал интерес владельца футбольного клуба «Рома», который назначил его директором по аналитике. В «Роме» он быстро научился представлять информацию визуально (с помощью графиков и диаграмм), и это стало эффективным способом донести математические идеи до скаутов и тренеров. Именно во время его пребывания в «Роме» клуб подписал двух игроков мирового класса — нападающего Мохаммеда Салаха и голкипера Алиссона Бекера, которые позже перешли в «Ливерпуль» и выиграли Лигу чемпионов в 2019 году.

— Я точно не ставлю себе в заслугу их подписание, — говорит он мне, — в механизмах футбольных трансферов слишком много подвижных деталей. Но скажу, что Салах с Алиссоном были проданы после того, как я ушел в «Кингз». Я не несу ответственности за их продажу!

Когда я говорю Люку, что мне интересно поговорить о марковском предположении, его лицо светится даже сильнее, чем при разговоре о футбольных трансферах.

— Мы использовали его с самого начала в нашей статье о набираемых очках, — говорит он.

* Этот показатель учитывает очки, набранные против конкретного защитника.

Система очков Люка автоматически показывает, кто кого как опекает, и позволяет измерить, какие игроки лучше всего действуют в ситуациях один на один. Например, во время рождественской игры 2013 года между «Сан-Антонио» и «Хьюстон Рокетс» позиция в защите Джеймса Хардена из «Рокетс» лучше всего прогнозировалась позицией форварда «Спёрс» Кавая Леонарда. В том противостоянии лучше был Леонард, набравший в матче 20 очков. Алгоритм Люка приписал 6,8 из этих очков оценке защиты Хардена.

— Всем нам хотелось бы знать Модель Бога, — говорит Люк с понимающей улыбкой. — Такая модель подсказывала бы Леброну Джеймсу, что ему делать дальше, чтобы иметь максимальные шансы на набор очков. Но все мы знаем, что такое невозможно.

Ключ к построению полезной модели — решить, что принять «за данность», какие предположения сделать. Модель Бога брала бы в таком качестве все, что когда-либо происходило в прошлом: каждую тренировку, на которую приходил Леброн Джеймс, каждый матч, где он играл, что он ел на завтрак в течение всей жизни, как он завязывал шнурки перед матчем. Это правая сторона уравнения 4. В качестве данных берется вся жизнь Джеймса до момента его броска. Люк как создатель модели должен проявить свое умение решить, что можно игнорировать. Он приходит к выводу, что из этого нужно оставить в левой части уравнения 4, когда он делает марковское предположение.

Люк продолжает: «Когда мы моделируем Леброна Джеймса, то учитываем, где он на площадке, жестко ли против него защищаются, где его партнеры по команде. Затем делаем предположение, что все происходившее более чем за несколько секунд до этого не важно. И обычно такое предположение работает».

Я спрашиваю Люка: а как же комментаторы, которые говорят о последних 5–10 минутах матча что-то вроде «Он сегодня выглядит не идеально» или «Игрок поймал кураж». Люк отвечает:

— Это субъективность. Лучший прогноз дает не средний результат игрока за последние пять бросков, а его позиция на площадке вместе с позициями оппонентов в момент броска, а также общий уровень баскетболиста.

Важный вопрос в баскетболе — когда атакующий игрок должен отдать пас на дугу партнеру, который бросит трехочковый (броски с более близкого расстояния приносят по два очка). Марковское предположение позволяет Люку воспроизвести весь сезон НБА с помощью компьютерного моделирования. В одной из таких моделей «альтернативной реальности» виртуальных баскетболистов внутри дуги «заставляют» пасовать или вести мяч в позицию для трехочкового броска. Результаты такой симуляции понятны: если игрок не рядом с кольцом, то лучше вывести мяч за линию трехочковой дуги и бросить с расчетом набрать максимум.

Именно здесь Джеймс Харден показывает свою истинную ценность. Он провел больше матчей с минимум 50 набранными очками, чем любой из ныне действующих игроков в НБА, включая Леброна Джеймса, и этот результат достигнут в основном благодаря дальним броскам. Он довел до совершенства любимое движение: когда изображает перемещение внутрь дуги, словно собирается вести мяч, а затем отшагивает назад и бросает трехочковый.

В модели Люка команда Джеймса Хардена «Рокетс» была ближе всего к трехочковому математическому совершенству. Возможно, здесь нет ничего удивительного, если учесть, что генеральный менеджер клуба Дэрил Мори окончил Северо-Западный университет по специальности «информатика и статистика». Люка с его выводами уже опередил другой математик. Харден уже использовал трехочковую стратегию, которую с тех пор окрестили *Moreyball**.

Баскетбол стал не только спортивным сражением на площадке, но и битвой математических умов за ее пределами. Теперь бой идет за то, у кого в модели будут лучшие предположения. Сейчас Люк встраивает в свое марковское предположение давление в защите и время на бросок, используя методику, именуемую тензором переходных вероятностей; он рассчитывает определить, когда есть смысл бросать отчаянный двухочковый на исходе отведенного

* Отсылка к названию книги Майкла Льюиса *Moneyball*. (Переведена на русский язык: Льюис М. *Moneyball. Как математика изменила самую популярную спортивную лигу в мире*. М. : Манн, Иванов и Фербер, 2013.) Но «Морибол» не просто стратегия опоры на трехочковые броски, а активное использование в игре сложной статистической аналитики.

времени. Возможно, тензор переходных вероятностей Люка и не так зрешичен, как дальний бросок Хардена после финта с отшагиванием назад, но он, безусловно, столь же элегантен.

*

Фильм *Moneyball* (в российском прокате «Человек, который изменил всё») — история бейсбольного тренера Билли Бина, сыгранного Брэдом Питтом, об одном из самых фанатичных применений статистики всех времен. Фильм рассказывает, как генеральный менеджер клуба-аутсайдера «Окленд Атлетикс» с небольшим бюджетом собрал команду из малоизвестных игроков на основе их статистических показателей. Команда в итоге выдала серию из двадцати побед подряд.

Фильм заканчивается тем, что Бин отказывается от предложенной ему высокооплачиваемой должности в знаменитом клубе «Бостон Ред Сокс» и остается со своим клубом. Эта романтическая концовка не особо отражает то, что произошло в бейсболе после успеха «Окленда».

Бин — бывший игрок, а не статистик и не экономист по образованию. Но когда владельцы бейсбольных клубов стремились повторить его успех, они обращались непосредственно к математикам*. Билл Джеймс, создатель статистического подхода, примененного Бином, действительно получил должность в «Бостон Ред Сокс» и работает там с 2003 года. «Ред Сокс» также назначили своим директором по информационным сервисам математика Тома Типпетта.

Еще один клуб, добившийся успеха благодаря статистическому подходу, — «Тампа-Бэй Рейс», который в 2010 году нанял Дуга Фиринга, занимавшегося исследованием операций в Гарвардской школе бизнеса. За пять лет, которые Дуг проработал в команде, она трижды доходила до четвертьфиналов плей-офф Главной лиги бейсбола, хотя имела одну из самых маленьких зарплатных ведомостей в лиге⁶². Затем Дуг перешел в «Лос-Анджелес Доджерс»,

* Справедливости ради стоит сказать, что Бин тоже не сам вел расчеты: в «Атлетикс» этим занимался экономист Пол ДеПодест, прототип Питера Брэнда в фильме.

где в его аналитической группе трудилось двадцать человек, из которых минимум семеро имели магистерские либо докторские степени по статистике или математике. Они анализировали всё — от расположения в защите и порядка отбивания до длительности контрактов игроков. Я встретился с Дугом вскоре после того, как он в феврале 2019 года ушел из «Доджерс» и создал свою компанию, занимающуюся спортивной аналитикой. Прежде всего я спросил его, был ли он большим фанатом бейсбола.

— Я бы сказал, что по сравнению с людьми, которые трудятся в сфере спорта, я, может, и не фанат, — пошутил Дуг, — однако относительно всего населения можно сказать, что я большой фанат.

Дуг всю жизнь болел за «Доджерс» и мечтал работать в клубе.

Современная бейсбольная аналитика уходит корнями в работу статистиков-любителей, интересующихся этой игрой. Когда я спросил Дуга о теории «Манибала», он сказал, что «во многом успех игроков “Окленда” принадлежал Полу ДеПодесте (в фильме его сыграл Джона Хилл), который взял методы, уже использовавшиеся открыто, и реализовал их во внутреннем процессе принятия решений в рамках клуба».

Дуг описал, как в бейсболе на смену менеджерам с успешными игровыми карьерами и интуицией постепенно пришли выпускники университетов из Лиги плюща* с опытом работы с данными.

— Бейсбол можно упрощенно представить как ряд матчей один на один между питчером и бэттером**, — говорил Дуг, — и марковское предположение во многих ситуациях работает очень хорошо.

Потому-то этот вид спорта анализировать проще, чем другие. И математики здесь победили быстрее.

* Лига плюща — группа из восьми частных престижных вузов на северо-востоке США (Йельский, Принстонский, Колумбийский, Пенсильванский, Корнеллский, Гарвардский, Брауновский университеты и Дартмутский колледж).

** В бейсбольном матче участвуют две команды, которые по очереди играют в нападении и в защите. Игрок защищающейся команды (питчер) бросает мяч в сторону игрока противника (бьющего, или бэттера), который старается отбить мяч битой в поле. Если это удалось, бьющий может бежать по точкам, которые называются базами. Другие игроки защищающейся команды в поле стараются поймать мяч и разными способами не дать закончить такие пробежки. Побеждает команда, которая совершила больше пробежек. Реальная суть игры — противостояние питчера и бэттера: первый старается сделать сложный бросок, чтобы второму было труднее нанести хороший удар; бэттер старается выбить мяч как можно дальше.

Дуг с энтузиазмом отзывался о ранних научных статьях, посвященных бейсбольной аналитике, выходивших в 1960-е и 1970-е. Джордж Линдси в своей работе 1963 года использовал одну статистическую модель, чтобы отвечать на разные вопросы — например, когда нужно постараться украсть базу*, а когда игроки защищающейся команды должны располагаться ближе к бьющему. Его марковское предположение состояло в том, что на игру влияют количество выбывших игроков и расположение бейсболистов по базам. Он нашел оптимальные стратегии для бьющего и игры в поле, опробовав свои модели на данных, которые вручную собрал его отец, полковник Чарльз Линдси; они включали 6399 полуиннингов** в сезонах 1959–1960 годов. Свои выводы Линдси предварял предупреждением: «Повторим, что эти расчеты относятся к мифической ситуации, когда все игроки “средние”»⁶³.

Такая слегка преувеличеннная честность — рассматривать свою модель одновременно как нечто мифическое и полезное — признак настоящего специалиста по математическому моделированию. Точное изложение предположений так же важно, как и точное изложение результатов.

Такие математические модели создавались в основном теми, кто находился вне игры, — людьми, очарованными статистикой и стремившимися ее применить. Как только в каком-то виде спорта признаётся мощь чисел, там начинают приветствовать тех, кому этот код знаком; людей же без таких умений просят проходить мимо. В бейсболе этот переход закончен. В баскетболе он продолжается, в футболе начинает распространяться. «Ливерпуль», который подписал лучших игроков «Ромы», принадлежит американскому бизнесмену Джону Генри — человеку, который привел Билла Джеймса в «Бостон Ред Сокс». Для поиска людей «Ливерпуль» нанял физика Иэна Грэма. Когда клуб выиграл Лигу чемпионов в 2019 году, газета The New York Times спрашивала в интервью у него и его коллеги — физика и аналитика Уильяма Спирмена об их ролях⁶⁴. Разрешая такое интервью, клуб радостно подтверждал

* Кража базы — одна из игровых ситуаций в бейсболе.

** Матч в бейсболе обычно состоит из 9 периодов (иннингов). Каждый из них делится на две половины: половину времени каждая команда играет в защите, а половину — в нападении.

их участие в успехах команды. Победитель английской Премьер-лиги сезона 2018/19 — «Манчестер Сити» также имеет серьезную бригаду аналитиков; как мы уже видели, у чемпиона испанской Ла Лиги — «Барселоны» она тоже есть. Другие команды, в первую очередь «Манчестер Юнайтед», кажется, еще этого не уловили. Похоже, они понятия не имеют, сколько реально стоит для них Поль Погба.

Красная сторона «Манчестера» получила предупреждение*. Правила, которые применимы везде, действуют и в спорте. Модели побеждают. Бессмыслица проигрывает.

*

Когда вы собираетесь измерить свои умения или умения других, вам нужно четко сформулировать собственные предположения. Каково состояние дел до того, как вы начнете действовать, и какими они станут потом? Определитесь, какие сферы вашей жизни вы хотели бы улучшить. Возможно, желаете лучше познакомиться с математикой или чаще выходить на пробежки? Будьте честны: отметьте, какие уравнения знаете или сколько километров в неделю пробегаете. Зафиксируйте текущее положение и начните работать. Через месяц снова изучите ситуацию. Уравнение умений предлагает вам быть честными в отношении начальных предположений. Не оправдывайте свою неудачу, утверждая, что пытались достичь чего-то другого, и не преуменьшайте успехи, отвлекаясь на неудачи в других сферах своей жизни. Однако прежде чем продолжить, заново проанализируйте свои предположения. Оцените, в чем вы хотите стать лучше. Не зацекливайтесь на истории. Используйте марковское предположение, чтобы забыть прошлое и сосредоточиться на будущем.

В разговоре с Люком Борнном я осознаю, что мне самому нужно кое-что усовершенствовать. Мне надо быть терпеливее, когда я разговариваю с господином типа «Мой путь». Если отказаться от карикатурного изображения таких людей, изложенного в начале главы, им действительно есть что предложить. Такой человек обладает опытом и энергией. Он хорошо ладит с людьми. И он знает

* Цвета «Манчестер Юнайтед» — красные, «Манчестер Сити» — голубые.

и любит свой вид спорта. Как мне удержать господина «Мой путь» от такого количества бессмыслицы и помочь ему сосредоточиться на данных и моделях?

Люк рассказывал мне, что ему нередко приходится сидеть на собраниях скаутов, где обсуждение увиденного недавно игрока начинается с фразы:

- Тебе нравится этот парень?
- Да, нравится, — отвечает другой скаут.
- Да, он крутой, — добавляет третий.
- Ну а мне он не нравится, — замечает первый.

В таких ситуациях Люк пытается для контекста добавить статистику. Скажем, третьему скауту он говорит: «Ну, вы его видели 22 ноября, а статистика показывает, что у него тогда выдался матч всей жизни. Так что...»

Теперь обсуждение может принять новое, более информированное направление. Скауты могут совместно посмотреть видео и побеседовать о том, насколько характерно оно для игры баскетболиста.

Люка поразило, насколько открыт персонал в спортивных клубах. Каждый встреченный им скаут желает получить всю доступную информацию. Как и математики, они жаждут данных. Люк пытается представить способ организовать всю такую информацию в форме базовой модели.

— Мы стараемся быть честными в отношении того, что говорят наши модели, — сказал мне Люк. — Мы выкладываем на стол всё и сообщаем скаутам свои предположения. Это становится фундаментом в их дискуссиях.

Он предоставляет своей организации статистические сводки, диаграммы, газетные отчеты — всё, что им нужно. Он старается не использовать в разговоре слово «данные», поскольку оно, как правило, фигурирует в спорах о надежности данных или о человеческом знании; он считает себя поставщиком информации. Как риторически спросил меня Люк, «кто не желает дополнительной информации?».

Меня заинтриговало то, что себя он рассматривает в основном как ресурс, и я не мог не высказаться по этому поводу:

— Вы так говорите, будто ставите себя на уровень ниже скатов...

— Может, в каком-то смысле, — ответил Люк, немного подумав. — Мне нет нужды быть единственным умным человеком в «Кингз». Я предпочел бы быть тем, кто делает всех остальных умнее.

Мой опыт показывает, что такая скромность характерна для многих статистиков и специалистов по прикладной математике.

Я вспомнил свою беседу с сэром Дэвидом Коксом. Когда мы обсуждали понятие гениальности, он стал очень задумчивым.

— Я не пользуюсь словом «гений», — сказал он мне. — Оно очень сильное... И я никогда не слышал, чтобы кто-то называл кого-то гением, за исключением, возможно, Фишера.

Кокс имел в виду кембриджского математика Роналда Фишера, признанного отцом современной статистики.

— И даже тогда, — добавил он, — вероятно, это слово использовалось слегка саркастически. Может быть, это очень по-английски, но я считаю, что оно уже перебор.

Сэр Дэвид назвал нескольких человек, которых он действительно мог быть счастье гениями: Пикассо, Моцарт, Бетховен.

Слово «гений» часто употребляют применительно к точным и прикладным наукам: гений Альберта Эйнштейна в физике, Джона Нэша в экономике или Алана Тьюринга в информатике. Их вклад, безусловно, впечатляет; но это слово не отражает того, как стоит смотреть на их работу. Кажется, она становится недоступной для всех остальных, а сам математик превращается в господина «Мой путь», который считает себя умнее вас или меня.

В «Барселоне» есть гении. Это Лионель Месси, Серхио Бускетс, Самюэль Умтити и другие. Они видят то, чего мы никогда не увидим. Их игра создает произведения искусства, которые мало кто может повторить.

Члены «Десятки» не гении. Мы создаем идеи, которые можно повторить и измерить. Мы сортируем и организуем данные. Избавляемся от бессмыслицы. Предоставляем модели. И когда мы добиваемся наилучших результатов, мы делаем это незаметно.

Глава 5

Уравнение влияния

$$A \cdot p_{\infty} = p_{\infty}$$

Вы когда-нибудь задумывались о вероятности того, что вы — именно вы, а не кто-то еще? Я не имею в виду кого-то слегка другого — скажем, человека, который был (или не был) в Диснейленде или видел все (или не все) фильмы «Звездных войн». Нет, кто-то совсем другой: родившийся в другой стране или даже в другое время.

Население нашей планеты составляет примерно 8 миллиардов человек. Это означает, что вероятность оказаться конкретным человеком — примерно 1 к 8 миллиардам. Шансы угадать все номера в лотерее 6 из 49, которая проводится в Великобритании, равны примерно 1 к 14 миллионам, поэтому вероятность выиграть в такой лотерее по единственному билету в 570 раз выше, чем вероятность того, что вы — это вы.

Иногда я представляю себе Вселенную, в которой каждый день просыпаюсь случайным человеком. Вышеприведенные вычисления говорят: мы можем почти забыть о том, что дважды подряд проснемся собой (шансы на это тоже равны 1 на 8 миллиардов), но чему равна вероятность, что мы проснемся в том же городе, где заснули? В шведском Уппсале, где я живу, примерно 200 тысяч жителей. Значит, вероятность того, что я проснусь завтра здесь же, составляет всего около 1 / 40 000. Если я продолжу свое путешествие в течение пятидесяти лет, каждое утро просыпаясь случайным человеком, то шансы, что однажды снова окажусь в Уппсале, составляют

примерно 50%*. Можно сказать, что подбрасывание монетки определит, увижу ли я снова восход в моем родном городе.

В своем случайному путешествии каждые два года я проводил бы день в Лондоне и день в Лос-Анджелесе. Нью-Йорк, Каир и Мумбай посещал бы почти раз в год⁶⁵. Большой Токио с его населением 38 миллионов был бы моим домом почти дважды в год. Хотя вероятность проснуться в любом конкретном населенном пункте мала, у меня будет гораздо больше шансов оказаться в густонаселенных городских районах, чем в сельской местности. Больше всего шансов проснуться в Китае, затем в Индии. Если и можно найти какую-то стабильность во всей этой суматохе обитания в случайных телах, то она в этих двух странах. В совокупности в них проживает 2,8 миллиарда человек, поэтому примерно 2,5 дня в неделю я жил бы в одном из этих государств. В Африке окажусь примерно раз в неделю, а в США — чуть чаще раза в месяц. Мое путешествие, которое, вероятно, никогда не вернет меня в исходную точку, напоминает мне, что я одновременно и исключительно маловероятен, и до невозможности незначителен.

А теперь представьте, что я просыпаюсь не как человек, выбиравшийся из всего населения планеты, а как один из тех, на кого подписан в Instagram. Я не активный пользователь этой социальной сети для фотографий и поэтому подписан только на нескольких френдов, которые нашли время, чтобы отыскать меня. Так что я проснусь одним из них: возможно, это будет школьный приятель или коллега-ученый из другого университета. Я получу на день их тело, узнаю, каково быть ими (возможно, даже отправлю какое-нибудь сообщение старому себе), а затем отправлюсь в постель и проснусь снова другим случайному человеком — одним из тех, на которых подписаны уже они.

Я могу даже снова проснуться Дэвидом Самптером. Типичные пользователи Instagram имеют 100–300 подписчиков, поэтому

* Если использовать данные автора о населении мира и Уппсалы более скрупулезно, то вероятность оказаться завтра в другом месте равна $\frac{39999}{40000}$, вероятность просыпаться в разных местах целый год (365 раз подряд) равна $\left(\frac{39999}{40000}\right)^{365}$. Пятьдесят лет с учетом високосных — это примерно $365 \times 50 + 13 = 18\,263$ дня. Так что вероятность просыпаться не в Уппсале пятьдесят лет подряд равна $\left(\frac{39999}{40000}\right)^{18\,263} \approx 0,63$, или 63%. Соответственно, вероятность оказаться в Уппсале хотя бы раз за 50 лет равна 37%.

с учетом симметричных отношений (я подписан / на меня подписаны) есть вполне разумный шанс (примерно 1 / 200), что я пронесусь собой. В любом случае крайне вероятно, что проведу несколько дней, путешествуя по своей социальной группе (мои френды, френды френдов и в целом люди, которые близки мне по культуре и происхождению).

Затем происходит то, что навсегда меняет мою жизнь. Я просыпаюсь Криштиану Роналду. Ну, не обязательно КриРо. Может, это будет Кайли Дженнер, Дуэйн «Скала» Джонсон или Ариана Гранде*. Знаменитости могут быть разными, но само превращение в звезду гарантировано. Примерно через неделю после начала путешествия я стану одним из самых известных людей в Instagram. У этих знаменитостей сотни миллионов подписчиков, среди них есть люди из моей социальной группы, и вскоре я прыгну в их тела.

Я вполне могу оставаться в мире знаменитостей неделю или даже больше. Криштиану Роналду подписан на Дрейка, Новака Джоковича, Снуп Догга и Стефена Карри**, так что мне предстоит перемещаться между спортивными звездами и рэперами. Из Дрейка я прыгаю в Фаррелла Уильямса, а затем в Майли Сайрус; она же ведет меня к Уиллоу Смит и Зендае***. Теперь я свободно перемещаюсь по миру музыкантов и кинозвезд.

Затем после двух недель славы происходит еще одна трансформация — даже более драматичная, чем пробуждение в теле Снуп Догга. Однажды утром, проведя весь прошлый день на съемках боевика, я просыпаюсь школьным другом Дуэйна «Скалы» Джонсона. В этот момент осознаю ужасную истину. Я потерялся. Сейчас почти нет шансов, что я когда-нибудь снова стану Дэвидом Самптером. Совсем скоро я опять вернусь в круг знаменитостей, буду тусоваться со звездами и вывешивать фотографии своего полуобнаженного тела. Иногда эти периоды будут прерываться путешествиями по списку звезд рангом ниже, а изредка — кратковремен-

* Кайли Дженнер — американская модель и бизнесвумен. Дуэйн «Скала» Джонсон — американский актер. Ариана Гранде — американская певица и актриса.

** Дрейк — канадский рэпер, Новак Джокович — сербский теннисист, Снуп Дог — американский рэпер, Стефан Карри — американский баскетболист.

*** Фаррелл Уильямс — американский рэпер, Майли Сайрус — американская певица, Уиллоу Смит — американская актриса и певица, Зендае — американская актриса, модель и певица.

ным пребыванием в телах обычных людей; но после этого я снова вернусь в сияющий мир звездной жизни.

Вероятность того, что я стану собой завтра, становится крайне мала — возможно, один на триллион, а то и меньше. Все случайные путешествия по Instagram сходятся к знаменитостям и остаются там.

*

Одно важное уравнение XXI века выглядит так:

$$A \cdot p_{\infty} = p_{\infty} \quad (\text{Уравнение 5}).$$

Забудьте о миллиардных заработках при использовании логистической регрессии в азартных играх. Это уравнение — основа индустрии с триллионами долларов. Это Google. Это Amazon. Это Facebook. Это Instagram. Это суть любого интернет-бизнеса. Оно создает суперзвезд и подавляет повседневное и обыденное. Оно создает авторитетов и возводит на престол королей и королев социальных сетей. Оно причина нашей непрерывной потребности во внимании, одержимости восприятием себя, разочарования и увлечения модой и побудительными мотивами знаменитостей. Из-за него мы ощущаем себя потерянными в море рекламы и продукт-плейсмента. Оно сформировало все аспекты нашей онлайн-жизни.

Это уравнение влияния.

Вы можете подумать, что такое важное уравнение очень трудно объяснить или понять. Это не так. Фактически я уже объяснил его, когда представил, как стал Криштиану Роналду, Дуэйном Джонсоном или Уиллоу Смит. Достаточно связать символы A (обозначает переходную матрицу) и p_t (вектор, определяющий вероятность оказаться тем или иным человеком в какой-то социальной группе в момент времени t) с тем путешествием, которое мы только что совершили по населению всего мира.

Чтобы наглядно представить переходную матрицу, вообразите себе электронную таблицу, в которой и строки, и столбцы подписаны именами людей. На пересечении стоит вероятность пронуться завтра другим человеком. Вообразите, что в мире живет всего пять человек: я, Дуэйн «Скала» Джонсон, певица Селена

Гомес и еще двое, о ком я никогда не слышал: назову их Ван Фан и Ли Вэй. Если я предположу (как в своем первом мысленном эксперименте), что каждое утро просыпаюсь случайным человеком, то матрица A будет выглядеть так:

$$A = \begin{pmatrix} \text{ДС} & \text{ДД} & \text{СГ} & \text{ВФ} & \text{ЛВ} \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \end{pmatrix} \begin{matrix} \text{ДС} \\ \text{ДД} \\ \text{СГ} \\ \text{ВФ} \\ \text{ЛВ} \end{matrix}$$

Столбцы и строки матрицы помечены инициалами пяти обитателей мира. Каждый день я смотрю на столбец того, кем сегодня оказался, а затем число в каждой строке сообщает мне вероятность того, что завтра я стану тем, с кем связана эта строка. Поскольку все числа равны $1/5$, то это означает, что завтра я с равной вероятностью стану одним из пяти людей (включая меня самого).

Если я предположу (как во втором мысленном эксперименте), что буду просыпаться в теле того человека, на кого подписан в Instagram, то матрица A будет выглядеть иначе. Чтобы было интереснее, пусть Скала застрял на какой-то математической задачке и решил подписаться на меня в Instagram. Предположим, Селена Гомес познакомилась с Ван Фан и Ли Вэем на одном из своих концертов, подумав, что они здорово смотрятся вместе (я забыл упомянуть, что они пара), и подписалась на них. Разумеется, все подписаны на Селену и на Дуэйна. Тогда мы имеем:

$$A = \begin{pmatrix} \text{ДС} & \text{ДД} & \text{СГ} & \text{ВФ} & \text{ЛВ} \\ 0 & 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 1/3 & 1/3 & 1/3 \\ 1/2 & 1/2 & 0 & 1/3 & 1/3 \\ 0 & 0 & 1/3 & 0 & 1/3 \\ 0 & 0 & 1/3 & 1/3 & 0 \end{pmatrix} \begin{matrix} \text{ДС} \\ \text{ДД} \\ \text{СГ} \\ \text{ВФ} \\ \text{ЛВ} \end{matrix}$$

Когда я Дэвид Самптер, для меня есть только два возможных варианта на завтрашний день: Селена или Скала. Поэтому в каждой из соответствующих клеток в моем столбце стоит $1/2$. То же верно и для Скалы Джонсона. Остальные жители планеты могут перейти в трех других людей. Нули по диагонали матрицы отражают тот факт, что мы не можем остаться собой второй день подряд, потому что не подписаны на себя.

Теперь обратите внимание, что для создания моей модели я использовал марковское свойство (уравнение 4 из главы 4): предположил, что то, кем я был два дня назад, никак не влияет на то, кем оказался сегодня. По сути, матрица A определяет цепь Маркова: она дает нам переходные вероятности для перемещения из одного состояния в другое, при этом следующее состояние зависит только от нынешнего, но не более ранних.

Будем постепенно двигаться по дням с помощью нашей матрицы A . Предположим, в первое утро я проснулся Дэвидом Самптером. Теперь вычислим вероятности того, кем я стану завтра.

$$\begin{array}{ccccccccc}
 & \text{ДС} & \text{ДД} & \text{СГ} & \text{ВФ} & \text{ЛВ} & & \text{День 1} & \text{День 2} \\
 \left(\begin{array}{ccccc}
 0 & 1/2 & 0 & 0 & 0 \\
 1/2 & 0 & 1/3 & 1/3 & 1/3 \\
 1/2 & 1/2 & 0 & 1/3 & 1/3 \\
 0 & 0 & 1/3 & 0 & 1/3 \\
 0 & 0 & 1/3 & 1/3 & 0
 \end{array} \right) & \cdot & \left(\begin{array}{c} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{array} \right) & = & \left(\begin{array}{c} 0 \\ 1/2 \\ 1/2 \\ 0 \\ 0 \end{array} \right) & \text{ДС} \\
 & & & & & & \text{ДД} \\
 & & & & & & \text{СГ} \\
 & & & & & & \text{ВФ} \\
 & & & & & & \text{ЛВ}
 \end{array}$$

Я объясню, как умножаются матрицы, в примечаниях⁶⁶, но важнее всего обратить внимание на два столбца чисел в скобках по обеим сторонам от знака равенства. Они называются векторами, и каждый элемент вектора — число от 0 до 1, которое определяет вероятность того, что я окажусь определенным человеком в определенный день. В день 1 я Дэвид Самптер, так что число в моей строке равно 1, а остальные элементы вектора — 0. В день 2 я могу оказаться либо Селеной Гомес, либо Дуэйном Джонсоном (поскольку Дэвид Самптер подписан только на них), и в этом векторе есть два числа $1/2$ для них, а остальные равны 0.

В день 3 все становится интереснее. Мы имеем:

$$\begin{array}{ccccccccc}
 & \text{ДС} & \text{ДД} & \text{СГ} & \text{ВФ} & \text{ЛВ} & & \text{День 2} & \text{День 3} \\
 \left(\begin{array}{cccccc}
 0 & 1/2 & 0 & 0 & 0 \\
 1/2 & 0 & 1/3 & 1/3 & 1/3 \\
 1/2 & 1/2 & 0 & 1/3 & 1/3 \\
 0 & 0 & 1/3 & 0 & 1/3 \\
 0 & 0 & 1/3 & 1/3 & 0
 \end{array} \right) \cdot \left(\begin{array}{c}
 0 \\
 1/2 \\
 1/2 \\
 0 \\
 0
 \end{array} \right) = \left(\begin{array}{c}
 1/4 \\
 1/6 \\
 1/4 \\
 1/6 \\
 1/6
 \end{array} \right) & \text{ДС} \\
 & \text{ДД} \\
 & \text{СГ} \\
 & \text{ВФ} \\
 & \text{ЛВ}
 \end{array}$$

Я могу оказаться кем угодно из пяти человек. Скорее всего, я буду Дэвидом Самптером или Селеной Гомес, но с вероятностью $1/6$ могу оказаться также и Джонсоном, и одним из китайских френдов Селены. Давайте произведем умножение еще раз, чтобы найти, кем я могу оказаться в день 4.

$$\begin{array}{ccccccccc}
 & \text{ДС} & \text{ДД} & \text{СГ} & \text{ВФ} & \text{ЛВ} & & \text{День 3} & \text{День 4} \\
 \left(\begin{array}{cccccc}
 0 & 1/2 & 0 & 0 & 0 \\
 1/2 & 0 & 1/3 & 1/3 & 1/3 \\
 1/2 & 1/2 & 0 & 1/3 & 1/3 \\
 0 & 0 & 1/3 & 0 & 1/3 \\
 0 & 0 & 1/3 & 1/3 & 0
 \end{array} \right) \cdot \left(\begin{array}{c}
 1/4 \\
 1/6 \\
 1/4 \\
 1/6 \\
 1/6
 \end{array} \right) = \left(\begin{array}{c}
 6/72 \\
 23/72 \\
 23/72 \\
 10/72 \\
 10/72
 \end{array} \right) & \text{ДС} \\
 & \text{ДД} \\
 & \text{СГ} \\
 & \text{ВФ} \\
 & \text{ЛВ}
 \end{array}$$

Теперь мы видим, как знаменитости выходят на центральные роли. Глядя на вектор вероятностей, вычисленный для дня 4, мы обнаруживаем, что вероятность оказаться Скалой или Селеной Гомес — $23/72$ — почти вчетверо выше, чем вероятность оказаться Дэвидом Самптером (всего $6/72$).

Каждый раз, умножая нашу матрицу с переходными вероятностями на вектор очередного дня, мы переходим на один день в будущее. А теперь вопрос, который стал движущей силой для всех моих путешествий по населению мира: насколько часто я буду одним из этих пяти людей через очень большой промежуток времени?

Именно на этот вопрос и отвечает уравнение 5. Чтобы увидеть как, заменим матрицу и векторы символами. Матрицу мы уже

назвали A , а вектор назовем p_t и p_{t+1} . Тогда наше умножение матриц примет сжатую форму:

$$A \cdot p_t = p_{t+1},$$

где p_t — вектор, определяющий вероятность быть тем или иным человеком в нашей социальной группе в момент t . Мы станем использовать индекс t , чтобы обозначать время, как в предыдущей главе. Мы уже видели, что

$$p_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, p_2 = \begin{pmatrix} 0 \\ 1/2 \\ 1/2 \\ 0 \\ 0 \end{pmatrix}, p_3 = \begin{pmatrix} 1/4 \\ 1/6 \\ 1/4 \\ 1/6 \\ 1/6 \end{pmatrix} \text{ и } p_4 = \begin{pmatrix} 6/72 \\ 23/72 \\ 23/72 \\ 10/72 \\ 10/72 \end{pmatrix} \begin{array}{l} \text{ДС} \\ \text{ДД} \\ \text{СГ} \\ \text{ВФ} \\ \text{ЛВ} \end{array}$$

Переходим к уравнению 5, которое я повторю здесь:

$$A \cdot p_{\infty} = p_{\infty}.$$

Мы считаем, что прошло бесконечно много времени, поэтому разницы между t и $t + 1$ нет. Можем заменить эти индексы значком бесконечности ∞ . Задумайтесь об этом. Отсюда следует, что если мы прыгали между телами достаточно много дней, то не имеет значения, прыгнем ли мы еще один раз: вероятность оказаться в некотором теле будет постоянной и определяться вектором p_{∞} . Назовем его стационарным распределением. Оно дает нам неизменное распределение вероятностей нахождения в том или ином состоянии (в теле того или иного человека) через очень большое время, когда то, кем мы были изначально, уже забылось.

Уравнение 5 определяет вероятность того, что я буду определенным человеком в какой-то день в отдаленном будущем. Осталось решить уравнение. Для вселенной из пяти человек, в которой я сейчас обитаю, мы находим, что:

$$\begin{pmatrix} 0 & 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 1/3 & 1/3 & 1/3 \\ 1/2 & 1/2 & 0 & 1/3 & 1/3 \\ 0 & 0 & 1/3 & 0 & 1/3 \\ 0 & 0 & 1/3 & 1/3 & 0 \end{pmatrix} \begin{pmatrix} 8/60 \\ 16/60 \\ 18/60 \\ 9/60 \\ 9/60 \end{pmatrix} = \begin{pmatrix} 8/60 \\ 16/60 \\ 18/60 \\ 9/60 \\ 9/60 \end{pmatrix} \begin{array}{l} \text{Дэвид Самптер} \\ \text{Дуэйн Джонсон} \\ \text{Селена Гомес} \\ \text{Ван Фан} \\ \text{Ли Вэй} \end{array}$$

Обратите внимание, что два вектора слева и справа от знака равенства одинаковы. Это значит, что, сколько бы раз я ни умножал на этот вектор переходную матрицу A , результат не изменится. Именно таковы в отдаленной перспективе мои вероятности оказаться тем или иным человеком.

Вывод? У меня вдвое больше шансов проснуться Дуэйном Джонсоном, чем Дэвидом Самптером, и еще больше — Селеной Гомес. Больше даже шансов стать Ван Фан и Ли Вэем, чем Дэвидом. Если перевернуть вероятности, можно узнать, сколько времени мы проведем в телаах всех жителей нашего мира. Шестьдесят дней — примерно два месяца, и стационарное распределение говорит нам, что в среднем 8 дней из них я буду Дэвидом, 16 — Джонсоном, 18 — Гомес, по 9 — Ван Фан и Ли Вэем. Когда время сдвигается к бесконечности, более половины жизни я проведу как знаменитость.

*

Ясно, что мы не просыпаемся каждое утро разными людьми, но Instagram дает нам возможность заглянуть в чужую жизнь. Каждая увиденная фотография — момент, когда подписчик несколько секунд ощущает, каково быть кем-то еще.

Twitter, Facebook и Snapchat тоже дают возможность распространять информацию и влиять на чувства и мысли подписчиков. Стационарное распределение p_∞ измеряет такое влияние; и не только с точки зрения того, кто на кого подписан, но и с точки зрения скорости, с которой тот или иной мем или идея распространяется среди пользователей. Люди с большими вероятностями в векторе p_∞ влиятельнее и распространяют мемы быстрее. Люди с маленькими вероятностями в векторе p_∞ менее влиятельны.

Вот почему уравнение 5 — уравнение влияния — так ценно для сетевых гигантов. Оно говорит им, кто в их соцсети самые важные люди, и при этом компании ничего не знают о том, кто они в реальности и чем занимаются. Измерение влиятельности — всего лишь вопрос матричной алгебры, и этим бездумно и некритично занимается компьютер.

Изначально уравнение влияния применила Google при разработке своего алгоритма ранжирования страниц PageRank — незадолго до рубежа веков. Компания вычисляла стационарные распределения для сайтов в предположении, что пользователи случайным образом щелкают по ссылкам на посещенных сайтах, чтобы выбрать следующий, на который перейдут. По этой причине в результатах поиска они выше ставили сайты с более высокими значениями p_∞ . Примерно в то же время Amazon стала создавать матрицу смежности A для своего бизнеса. В ней связывались те книги, а позже игрушки, фильмы, электроника и другие товары, которые люди покупали вместе. Определив тесные связи в матрице, компания смогла давать рекомендации для клиентов под заголовком вроде «Вам также может понравиться». Twitter использует стационарное распределение в своей соцсети, чтобы найти и предложить вам людей, на которых можно подписаться. Facebook применяет те же идеи при обмене новостями, а YouTube — чтобы рекомендовать видеоролики. Со временем подход развивался, появлялись дополнительные детали, но базовым инструментом для нахождения влиятельных лиц в соцсети остается матрица смежности A и ее стационарное распределение p_∞ .

За последние два десятилетия это привело к неожиданному результату. Система, которая первоначально создавалась для измерения влияния, превратилась в его *создателя*. Алгоритмы на базе уравнения влияния определяют, какие публикации должны занимать видное место в социальных сетях. Идея в том, что если некто популярен, то этого человека желают выслушать больше людей. Результатом становится цикл обратной связи: чем влиятельнее человек, тем большую заметность дает ему алгоритм, а от этого его влияние еще больше растет.

Один бывший сотрудник Instagram рассказал мне, что изначально основатели компании очень неохотно применяли в бизнесе

алгоритмы и математику. «Они видели в Instagram нечто очень нишевое, артистичное и считали алгоритмы негодными», — говорил он. Эта платформа предназначалась для обмена фотографиями между близкими друзьями. Все изменилось после успеха Facebook. «За последние пару лет платформа стала совершенно иной. Один процент ее пользователей имеет более 90% подписчиков», — заметил мой собеседник.

Вместо того чтобы побуждать пользователей подписываться только на друзей, компания применила к своей сети уравнение влияния. Оно еще сильнее раскручивало самые популярные аккаунты. Возникла обратная связь, и аккаунты знаменитостей росли всё сильнее. Едва Instagram стал использовать уравнение влияния, как и все платформы социальных сетей до него, его популярность резко возросла — в нем более миллиарда пользователей.

*

Математические методы, используемые при конструировании соцсетей, появились задолго до возможности создания таких приложений. Вовсе не Google изобрела уравнение влияния: его происхождение восходит к Маркову, который предложил свойство, получившее его имя, для рассмотрения цепей состояний, где каждое новое состояние зависит только от предыдущего. Именно это и происходит при моем случайному путешествии по Instagram.

Решая уравнение 5 для моего мира из пяти человек, я слегка поленился. Я нашел ответ — вероятность, с которой буду тем или иным человеком в отдаленном будущем, — многократно перемножая матрицу A и вектор p , пока тот не перестал меняться. Сделав это, я нашел p_∞ .

Такой метод в итоге приводит к правильному ответу, но он не особо изящен. И Google им не пользуется. Свыше ста лет назад математики Оскар Перрон и Георг Фробениус показали, что для любой цепи Маркова с матрицей A существует единственное стационарное распределение p_∞ . Итак, независимо от структуры социальной группы мы всегда можем вычислить, сколько времени будем проводить тем или иным человеком, если станем случайным образом перемещаться между людьми. Это стационарное распределение

можно найти с помощью метода Гаусса, который, как и нормальная функция, связан с именем Карла Фридриха Гаусса, однако восходит к более древним временам: китайские математики решали системы линейных уравнений с помощью этого метода больше двух тысяч лет назад. К матрице A применяются так называемые элементарные преобразования, чтобы привести ее к более удобному виду и найти в итоге p_∞ . Это дает возможность быстро и эффективно находить влиятельных людей даже для сетей с миллионами участников.

На протяжении XX века «Десятка» изучала свойства сетей; соответствующая область математики называется теорией графов. Еще в 1922 году Джордж Удни Юл описал математику, стоящую за ростом популярности в Instagram, в терминах процесса, который позже был назван «предпочтительным присоединением»: чем больше у человека подписчиков, тем больше он привлекает к себе внимания, и его известность растет. Затем, в начале XXI века, всего за несколько лет до появления Facebook, эта область исследований расширилась и стала именоваться наукой о сетях: сейчас она описывает распространение мемов и фейковых новостей, изучает, как соцсети создают маленький мир, в котором все соединены через шесть рукопожатий, и потенциальную возможность поляризации⁶⁷.

«Десятка» была готова. Ее участники оказались среди основателей и первых работников будущих гигантских соцсетей. В основе их бизнеса лежит уравнение влияния. Зарплата, предложенная людям с нужными умениями, была достаточно соблазнительной даже для самых идеалистичных участников общества. И что еще важнее, эта работа давала им свободу для творческого мышления, создания новых моделей и их практического применения.

Вскоре перед участниками «Десятки» поставили задачу установить, как мы реагируем на соцсети. Они управляли лентой новостей Facebook, чтобы посмотреть, как пользователи реагируют, получая только негативные новости; они создавали кампании в сетях, чтобы побудить людей голосовать на выборах; конструировали фильтры, чтобы пользователи читали больше тех новостей, которые им интересны. Они контролировали то, что мы видели, определяя, видим ли мы публикации друзей, новости (фейковые и реальные),

знаменитостей или рекламу. Именно участники «Десятки» стали влиятельными — не из-за сказанного ими, а благодаря их решениям о том, как нам связываться друг с другом. Они знали о нас даже то, чего мы не знали сами...

*

Вероятно, ваши друзья намного популярнее вас. Я ничего не хочу сказать о вас как о человеке, не желаю быть к вам несправедливым, но могу утверждать это с определенной уверенностью.

Математическая теорема, известная как парадокс дружбы, утверждает, что большинство людей в любой социальной сети, включая Facebook, Twitter и Instagram, менее популярны, чем их друзья⁶⁸. Начнем с примера. Представьте, что мы убрали Дуэйна Джонсона из той социальной сети, которую описывали ранее. Теперь нас четверо — я, Селена Гомес, Ван Фан и Ли Вэй, и они имеют, соответственно, 0, 3, 2 и 2 подписчика. Вероятно, китайская пара чувствует себя довольно популярной за счет компании Гомес, но у меня есть для нее сюрприз. Я прошу каждого участника сети посчитать среднее число подписчиков у их друзей. Я подписан только на Селену Гомес, у нее три подписчика, поэтому среднее число подписчиков у моих друзей равно 3. Гомес подписана на двух людей, у каждого по 2 подписчика, так что среднее число подписчиков у ее друзей равно 2. Ван Фан и Ли Вэй подписаны друг на друга и на Гомес, поэтому их друзья в среднем имеют 2,5 подписчика. Таким образом, среднее число подписчиков у друзей в этой сети равно $(3 + 2 + 2,5 + 2,5) / 4 = 2,5$. Только Селена Гомес имеет больше друзей, чем ее подписчики. У Ван Фан, Ли Вэя и у меня число подписчиков ниже среднего.

Причина парадокса дружбы заключается в разнице между случайным выбором человека и случайным выбором отношения дружбы (см. рис. 5). Для начала выберем наугад одного человека. Ожидаемое (среднее) количество подписчиков у него — сумма числа подписчиков у всех пользователей, поделенная на общее число пользователей платформы. Для Facebook это число составляет около 200. Для нашей сети из четырех человек это $(0 + 3 + 2 + 2) / 4 = 1,75$. В теории графов это называется средним числом ребер, входящих

в узел*. Выше мы показали, что среднее число подписчиков у друзей в этой сети равно 2,5, что больше, чем 1,75 — среднее число подписчиков.

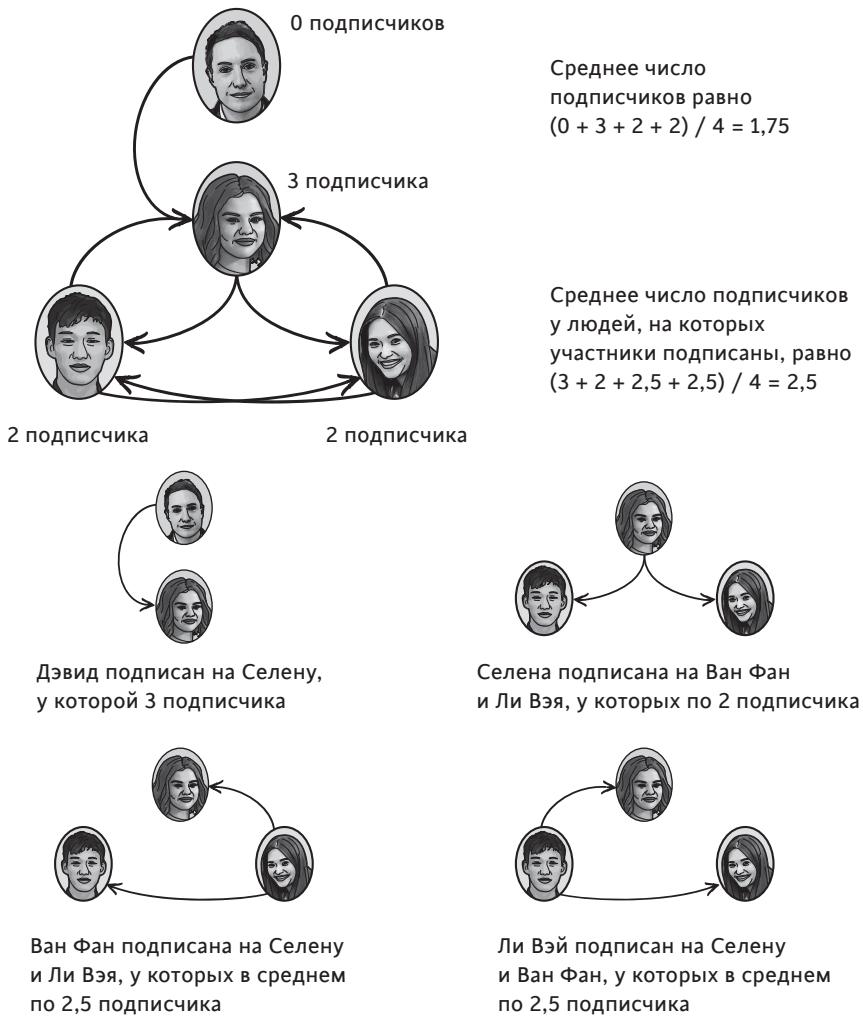


Рис. 5. Парадокс дружбы для четырех участников

* Узлы графа — люди в соцсети. Соединяющие их ребра изображают «подписанность» одного человека на другого. В узел графа, соответствующий мне, входят несколько ребер из других узлов (тех, кто подписан на меня), так что число моих подписчиков — это число входящих ребер. В свою очередь, из узла, изображающего меня, выходит несколько ребер, ведущих к узлам, соответствующим тем людям, на которых подписан я сам.

То же произойдет, если мы снова вернем в нашу сеть Дуэйна Джонсона. Результат сохранится, даже если Селена Гомес подпишется на меня. Парадокс дружбы можно доказать для любой соцсети, в которой каждый подписан на одно и то же количество людей. Доказательство таково. Сначала выберите наугад одного человека во всей сети; затем — кого-нибудь, на кого этот человек подписан. Если представить ситуацию иначе, то мы, выбрав двух связанных людей, взяли какую-то случайную связь среди всех изображающих отношения «подписанности» людей в соцсетях. В теории графов такие связи называются ребрами графа. Теперь, поскольку популярные люди имеют (по определению) больше входящих ребер, на конце любого данного ребра мы с большей вероятностью найдем популярного человека, чем если бы выбирали человека наугад. Таким образом, случайно выбранный друг случайно выбранного человека (человек на конце ребра), вероятно, имеет больше друзей, чем случайно выбранный человек. Это и показывает, что парадокс дружбы справедлив⁶⁹.

Такова математическая теория. Как все это работает на практике? Кристина Лерман, исследователь из Университета Южной Калифорнии, решила выяснить это. Она и ее коллеги взяли социальную сеть пользователей Twitter в 2009 году (на такой ранней стадии развития соцсети в ней было всего 5,8 миллиона пользователей) и рассмотрели отношения «подписанности»⁷⁰. Ученые обнаружили, что люди, на которых был подписан типичный пользователь Twitter, имели примерно вдвадцать раз больше подписчиков, чем он. Только 2% пользователей были популярнее своих подписчиков.

Лерман и ее коллеги пришли еще к одному выводу, который полностью противоречит интуиции. Оказалось, что подписчики случайно выбранного пользователя Twitter были в среднем вдвадцать раз лучше связаны! Хотя кажется разумным, что люди, на которых мы подписаны, популярны (в конце концов, многие из них — знаменитости), гораздо труднее понять, почему люди, подписанные на нас, оказываются намного популярнее нас. Если они подписаны на *вас*, как они могут быть популярнее? Это не выглядит справедливо.

Ответ — в нашей склонности создавать взаимные отношения. Когда кто-то подписывается на *вас*, появляется определенное

социальное давление, заставляющее сделать то же в ответ. Отказ выглядит грубостью. В среднем люди, подписанные на вас в Instagram или отправившие вам запрос в друзья в Facebook, также с большой вероятностью отправят аналогичные запросы другим людям. В результате они составят большую часть нашей социальной группы. И это еще не всё. Исследователи также обнаружили, что ваши друзья выкладывают посты чаще, получают больше лайков, репостов и охватывают больше людей, чем вы.

Как только вы примете математическую неизбежность непопулярности, ваши отношения с соцсетями начнут улучшаться. Вы не одиноки. По оценке Кристины Лерман и ее коллег, 99% пользователей Twitter в том же положении. В самом деле, у популярных людей положение может оказаться еще хуже. Подумайте об этом. В нескончаемом поиске лучшего социального положения «крутые ребята» пытаются добиться взаимных отношений с теми, кто успешнее их. Чем чаще они это делают, тем скорее оказываются в окружении людей, которые популярнее, чем они сами. Это слабое утешение, но приятно осознавать, что те, кто кажется успешным, вероятно, ощущают то же, что и вы. Возможно, за исключением Пирса Моргана и Джоан Роулинг, оставшийся 1% пользователей Twitter — либо аккаунты знаменитостей, которыми управляют пиар-службы, либо, весьма вероятно, люди, которые почти обезумели от стремления постоянно находиться в соцсетях.

*

Я не собираюсь присоединяться к тем, кто предлагает вам уйти из соцсетей. Мантра математика — не сдаваться. Нужно разделить все на три составные части: данные, модель и бессмыслицу.

Рекомендую начать сегодня. Прежде всего посмотрите на данные. Проверьте сами, сколько подписчиков или взаимных друзей имеют ваши друзья в Facebook или Instagram. Я только что сделал это для Facebook и обнаружил, что 64% моих друзей популярнее меня. Затем вспомните модель. Популярность в соцсетях создается с помощью обратной связи, благодаря которой люди привлекают больше подписчиков. Это статистическая иллюзия, порожденная парадоксом дружбы. Затем убирайте бессмыслицу.

Не жалейте себя и не ревнуйте к другим: просто осознайте, что мы часть Сети, которая искажает нашу самооценку самыми разными способами.

Психологи пишут и говорят о наших когнитивных искажениях, при которых отдельные люди или общество в целом воспринимают субъективную реальность, не совпадающую с реальным миром. Список таких искажений растет: «горячая рука», эффект присоединения к большинству, ошибка выжившего, предвзятость подтверждения, эффект фрейминга*. Участники «Десятки», разумеется, не отрицают существования таких искажений, но для них важнее не ограничения человеческой психологии. Вопрос в том, как убрать фильтр и увидеть мир четче. Для этого они придумывают сценарии «Что, если?». Что, если я буду просыпаться каждый день другим человеком? Что, если я путешествую по Snapchat как интернет-мем? Что, если буду читать только те новости, которые мне предлагает Facebook, и смотреть только фильмы, рекомендованные Netflix? Как тогда будет выглядеть для меня мир? И как это отличается от «более справедливого» мира, в котором я уделяю равное внимание всей доступной мне информации?

«Десятка» предлагает вам представить себе невероятные, фантастические сценарии. Затем они становятся математическими моделями. После этого можно начинать цикл. Модель сравнивается с данными, а те используются для совершенствования модели. Медленно, но верно участники «Десятки» могут убрать фильтр и раскрыть нашу социальную реальность.

*

Лина и Микаэла открывают свои аккаунты в Instagram и показывают мне свои телефоны. «Это реклама или селфи?» — спрашиваю я Лину.

* Заблуждение «горячей руки» — представление, что успех в данный момент повышает шансы на успех в дальнейших попытках. Эффект присоединения к большинству — ориентация на мнение и поведение других. Ошибка выжившего — вид систематической ошибки отбора, когда есть данные только по одной группе объектов. Предвзятость подтверждения — склонность человека к подтверждению своей точки зрения, восприятию только той информации, которая согласуется с его точкой зрения. Эффект фрейминга — когнитивное искажение, когда форма подачи информации влияет на ее восприятие.

Лина показывает фотографию местного пекаря, который демонстрирует на камеру поднос, полный пирожных в виде сердечек. Фотография выглядит настоящей, но идея здесь в приглашении подписчиков в магазин. Лина отвечает, что это селфи, но она классифицирует аккаунт как компанию.

Лина и Микаэла работают над своей дипломной работой по математике⁷¹. Они изучают, как Instagram представляет им мир. Незадолго до того, как они начали свой проект, Instagram снова обновил алгоритм, который определяет порядок показа фотографий. Компания заявила, что акценты сместились в сторону приоритетов снимков от друзей и семьи.

В результате многие популярные люди ощутили угрозу. Шведская блогерша и гуру соцсетей Анита Клеменс (65 тысяч подписчиков) сказала: «Психологически трудно видеть, как мои подписчики исчезают. Мне почти сорок, интересно, каково это для всех молодых авторитетов?»⁷²

Клеменс ощущала, что она активно работала «для своих подписчиков», а новый алгоритм не доносил до них ее сообщения. Чтобы проверить ограничения, она опубликовала свою фотографию с новым партнером: легко было ошибочно решить, что на снимке она беременна.

Эта фотография широко разошлась по платформе, а затем Клеменс раскрыла, что ее основной целью была проверка того, какие снимки срабатывают, а какие нет. Заявление о своей беременности в Instagram, похоже, работает, если вы желаете привлечь больше внимания.

Хотя реакция на один снимок с ложной беременностью дает очень мало информации, можно сказать, что Клеменс проводит своего рода эксперимент. Келли Коттер из Университета штата Мичиган обнаружила, что многие популярные люди в Instagram пытаются понять алгоритм и манипулировать им⁷³. Они открыто обсуждают расходы и выгоды от лайков и комментирования максимально возможного количества постов, проводя A/B тестирование для разных стратегий (вспомните уравнение ставок). Часто такие люди пытаются определить, не «затеняет» ли их сайт, опуская их посты ниже в ленте подписчиков. Когда Instagram изменил свои

алгоритмы, многие из этих людей поспешили написать в социальных сетях заявления с хештегом #RIPInstagram.

Теперь Лина и Микаэла планировали тщательнее изучить алгоритм Instagram исходя из собственной позиции типичных пользователей. В течение месяца они будут открывать свои аккаунты только один раз в день в 10 утра, смотреть на порядок показываемых фотографий и отмечать тип каждого поста и сообщения. Так они могут проверить гипотезу популярных лиц, что алгоритм затеняет их и снижает их приоритет.

«Я думаю, что более редкое появление в аккаунте пойдет нам на пользу», — говорит Микаэла, имея в виду то, что они собирают данные всего один раз в день. Как и многие из нас, эти молодые женщины проверяют свои соцсети чаще, чем на самом деле хотят.

Задача в том, чтобы провести обратную разработку (реверс-инжиниринг) алгоритма, который использует Instagram: выяснить, что сеть от них скрывает (если она это делает). В математике это называется обратной задачей, и один из примеров использования — современная компьютерная томография. Когда пациент лежит в аппарате, во всех направлениях производятся снимки с помощью рентгеновских лучей. Материалы разной плотности в нашем теле поглощают рентгеновские лучи по-разному, что позволяет видеть скелет, легкие, мозг и другие системы нашего тела. Обратная задача в этом случае — собрать все снимки воедино, чтобы получить полное изображение наших внутренних органов. Математический метод, лежащий в основе этого процесса, — преобразование Радона (названное в честь предложившего его австрийского математика Иоганна Радона), которое дает возможность восстановить точную трехмерную картину по отдельным двумерным изображениям.

У нас нет преобразования Радона для социальных сетей, но имеется — в форме уравнения 5 — ясное понимание того, как они поглощают и изменяют социальную информацию. Чтобы провести обратную разработку процесса деформации данных, который применяет Instagram, Лина и Микаэла использовали статистический метод под названием бутстрэпинг. Каждый

день они брали первые 100 сообщений в своей ленте и случайным образом перемешивали их, создавая новый порядок. Они повторяли этот процесс 10 000 раз, получая распределение перестановок, которое имело бы место, если бы Instagram просто показывал посты случайно, не устанавливая для них какие-то приоритеты. Сравнивая положение популярных людей в реальной ленте Instagram с таким рандомизированным ранжированием, они смогли определить, сдвигаются ли такие люди в их ленте вверх или вниз.

Результаты резко противоречили мнениям, вызвавшим появление хештега #RIPInstagram. Не нашлось подтверждений тому, что каких-то популярных людей угнетали: их положение в лентах Лины и Микаэлы статистически не отличалось от того, каким оно было бы при случайном порядке. Оказалось, что Instagram нейтрально относится к таким лицам. Однако некоторые аккаунты получили существенный приоритет: друзья и семья поднимались на верх ленты. Это происходило в основном за счет новостных сайтов, политиков, журналистов и организаций. Instagram не столько снижал влияние популярных людей в ленте, сколько увеличивал влияние друзей и семьи и затенял аккаунты, которые не платили за рекламу.

Самым показательным в кампании #RIPInstagram было ощущение незащищенности популярных людей. Они внезапно осознали, что у них не так много контроля над социальным положением, как они думали. Их позиции определялись алгоритмом, который способствовал популярности, а теперь они беспокоились, что ее отнимет другой алгоритм, ориентированный на друзей.

Это исследование показало, что настоящие влиятельные лица в сети не те, кто публикует снимки своей еды и моментов жизни. Скорее, это программисты Google, Facebook и Instagram, конструирующие фильтры, через которые мы смотрим на мир. Они решают, кто популярен и что популярно.

Для Лины и Микаэлы эксперимент оказался терапевтическим. Лина сказала мне, что изменился ее взгляд на Instagram. Она ощущала, что теперь лучше использует свое время в этом приложении. «Вместо того чтобы прокручивать ленту в поисках интересного,

я останавливаюсь, увидев посты друзей. Я знаю, что дальше будет просто скучота», — говорит она.

Уравнение влияния справедливо не для одной социальной сети; оно относится ко всем. Его сила — способность показать, как структура онлайн-сетей формирует ваш взгляд на мир. Когда вы ищете товары на Amazon, то попадаете в систему «вам также может понравиться», поскольку самые популярные продукты показываются первыми. Twitter сочетает полярные мнения с возможностью оспаривания ваших взглядов людьми со всего мира. В Instagram вас окружают друзья и семья, вы ограждены от новостей и мнений. Используйте уравнение 5, чтобы честно рассмотреть, кто и что на вас влияет. Напишите матрицу смежности для вашей социальной сети и посмотрите, кто находится в вашем онлайн-мире, а кто вне его. Подумайте о том, как эта соцсеть влияет на вашу самооценку и как контролирует информацию, к которой вы имеете доступ. Подвигайтесь в ней и посмотрите, как это затрагивает других людей, с которыми вы связаны.

*

Через несколько лет Лина и Микаэла, которые планируют стать преподавателями математики, будут объяснять подросткам, как алгоритмы телефонов фильтруют взгляды на мир. Большинству детей этот урок поможет разобраться в сложной социальной группе, в которую они встроены. А некоторые увидят другую возможность — карьерную. Они станут серьезно учиться, глубже понимать математику и узнавать, как применять алгоритмы, используемые Google, Instagram и прочими компаниями. Некоторые из них могут пойти дальше и стать частью богатой и влиятельной элиты, которая контролирует, как нам подается информация.

В 2001 году сооснователь Google Ларри Пейдж получил патент на использование уравнения 5 при поиске в интернете⁷⁴. Изначально тот принадлежал Стэнфордскому университету, в котором тогда работал Пейдж, и Google приобрел его за 1,8 миллиона акций компании. В 2005 году Стэнфорд продал эти акции за 336 миллионов долларов. Сегодня они стоили бы в десять раз больше. Применение уравнения 5 — это лишь один из многих патентов Google

или Facebook, где к интернету применяют математику XXI века. Теория графов стоит миллиарды для технических гигантов, ее применяющих.

Тот факт, что математические формулы, созданные почти за сто лет до подачи патента, могут принадлежать какому-нибудь университету или компании, похоже, противоречит духу «Десятки». У участников всегда были секреты, однако ими обычно делились, их использовали все желающие. Должны ли у общества быть принципы, которые не позволяют его участникам ограждать свои открытия или получать сверхприбыль от своих тщательно собранных знаний?

Оказывается, ответ на этот вопрос далеко не очевиден...

Глава 6

Уравнение рынка

$$dX = hdt + f(X)dt + \sigma \cdot \varepsilon_t$$

Разделение мира на модели, данные и бессмыслицу дало участникам «Десятки» чувство уверенности. Им больше незачем было беспокоиться о последствиях; достаточно применять свои умения на практике. Они превращали любую задачу в числа и данные, четко формулировали предположения, рационально рассуждали и отвечали на поставленные вопросы.

Первоначально участники трудились на государственной службе и в государственных исследовательских организациях. В 1940-е и 1950-е они продолжали работу Ричарда Прайса, создавая системы национального страхования и обеспечивая услуги здравоохранения. Именно в это время Дэвид Кокс стал работать в текстильной промышленности, используя математику для промышленного роста. В 1960-е и 1970-е участники «Десятки» заняли должности в таких исследовательских учреждениях, как Bell Labs в Нью-Джерси, NASA, в министерствах обороны в странах — участницах холодной войны, а также в элитных университетах и стратегических аналитических центрах вроде корпорации RAND. Внутри групп происходила консолидация знаний. В 1980-е и 1990-е результатами заинтересовалась финансовая индустрия, которая нанимала участников «Десятки» для управления своими фондами.

Освободившись от бессмыслицы, общество полагало, что его участники сами способны решить мировые проблемы. Болтые и влиятельные соглашались, платя им огромные деньги за

управление инвестиционными фондами. Правительства полагались на них, планируя экономическое и социальное будущее своих стран. Межправительственные организации отдавали им центральную роль в прогнозировании изменений климата и определении целей развития.

Но кое о чем математики «Десятки» забыли. Алфред Айер писал об этом в книге «Язык, истина и логика», однако это не было понято так же хорошо, как другие идеи логического позитивизма, которые двигали общество вперед. Когда Айер использовал принцип проверяемости, чтобы отделить математику и науку от бессмыслицы, он обнаружил, что категория бессмыслицы гораздо шире, чем признают большинство ученых. Он продемонстрировал, что мораль и этика также относятся к категории бессмыслицы.

Айер доказывал это поэтапно. Он начал с классификации религиозных истин. Он показал, что вера в Бога непроверяема: нет эксперимента, который мог бы подтвердить или опровергнуть существование Господа. Он писал, что верующий может заявлять, будто Бог — загадка, выходящая за рамки человеческого понимания, или мы имеем дело с актом веры, или Бог — объект мистической интуиции. Айер принимал все эти объяснения, пока верующему было понятно, что они — «бес-смыслица». Верующие не должны и не могут подразумевать, что Бог или любое иное сверхъестественное существо играет какую-то роль в наблюдаемом мире. Религиозные убеждения любого человека или учение пророка нельзя подтвердить данными, поэтому они непроверяемые. Если же какой-то верующий утверждает, что его убеждения поддаются верификации, то их можно проверить данными и (с большой вероятностью) доказать их ложность. Религиозные убеждения были бессмыслицей.

До этого момента большинство участников «Десятки» принимали и понимали рассуждения Айера. Они соответствовали их взглядам и убеждениям. Они уже отвергли чудеса и больше не нуждались в Боге. Но Айер пошел дальше. С атеистами, выступавшими против религий, он обращался с таким же холодным пренебрежением, как и с верующими. Атеисты обсуждали бессмыслицу,

поэтому и сами участвовали в ее создании. Эмпирически обоснованными утверждениями о религии могли быть только те, что касались анализа психологических аспектов отдельного адепта и/или роли веры в обществе. Оспаривание же убеждений было таким же бессмысленным, как и их отстаивание.

Этим дело тоже не закончилось. Далее Айер отверг практический аргумент, выдвинутый кем-то из других участников «Венского кружка», что нам следует стремиться к величайшему счастью для всех. Он утверждал, что невозможно использовать одну только науку, чтобы решить, что такое «благо» или «добродетель», либо чтобы оправдать наши действия в балансе между счастьем сейчас и состоянием в будущем. Мы можем моделировать скорость, с которой онлайн-казино забирают деньги у клиентов, слишком бедных для азартных игр, но неспособны использовать нашу модель, чтобы сказать: нехорошо, что игроки тратят деньги так, как хотят. Специалист по моделированию климата может сказать: «Если мы не сократим выбросы СО₂, будущим поколениям грозят неустойчивый климат и нехватка продовольствия»; но это не скажет нам ничего о том, стоит ли оптимизировать свою жизнь сейчас или лучше подумать о благополучии наших внуков. Для Айера все стимулирование к моральному поведению, все заявления вроде «Мы должны помогать другим», «Нам нужно действовать ради общего блага», «Мы несем моральную ответственность за сохранение мира для будущих поколений» и «Вам не следует патентовать математические результаты» были эмоциональными всплесками, относящимися к сфере психологии, которые не содержали никакого значимого смысла.

Точно так же рассуждение Айера не допускает эмпирической проверяемости таких индивидуалистических эмоциональных заявлений, как «жадность — это хорошо» или «позаботьтесь в первую очередь о себе». И снова это бессмыслица, хотя и глубоко укоренившаяся в нашей психике. Нет способа проверить такие утверждения на основе нашего опыта, кроме как обсудить относительный финансовый и социальный успех людей, следующих таким максимам. Мы можем смоделировать факторы, которые привели какого-то человека к богатству и славе. Или измерить личностные

характеристики тех, кто добился успеха. Либо говорить о том, как эти черты развились в результате естественного отбора. Но мы не можем использовать математику, чтобы доказать: те или иные ценности по сути — благо или добродетель. Принцип верификации, так здорово помогающий участникам «Десятки» моделировать мир, оказался бесполезным при определении морального пути.

Если «Десятка» не может найти мораль изнутри, откуда у нее чувство уверенности? И чьим интересам она на деле служит, если нет моральной ориентации? Может, все не так добродетельно и правильно, как воображал Ричард Прайс?

*

Я сидел за столиком в одном из первоклассных ресторанов Гонконга, глядя на порт. Один из крупнейших инвестиционных банков мира пригласил меня пообедать с его лучшими аналитиками рынка. Первоклассным было все — от перелета с моей женой до пятизвездочной гостиницы и еды, которую мы сейчас поглощали.

Обсуждение дошло до одного из самых серьезных противоречий в их мире — разницы между долгосрочными и краткосрочными инвестициями. Эти мужчины (и одна женщина) работали в основном над долгосрочными вопросами, управляя ростом пенсионных фондов. Их решение инвестировать в какую-то компанию определялось ее базовыми принципами, структурой управления, планами и положением на рынке. Это был мир, который они понимали и в котором были уверены. Если бы они не знали, что делают, мы бы не сидели в ресторане с такой великолепной панорамой.

А вот в краткосрочных перспективах аналитики были не так уверены. Торговля стала алгоритмической, и они не понимали, что делают алгоритмы. Участники обеда спрашивали меня: какие языки программирования должны изучать их новые сотрудники? Какими математическими навыками обладать? Какие университеты дают лучшие магистерские степени в области науки о данных?

Я пытался ответить на эти вопросы как можно лучше, но понял, что упускаю очевидное. Я считал само собой разумеющимся то,

чего у этих людей не было. Я предполагал — из-за красивой панорамы и ресторана с мишленовскими звездами, — что эти парни походят на меня, смотрят на мир через призму математики и именно поэтому так богаты. Когда вначале рассказал им о том, как использую марковское предположение для анализа последовательностей владения в футбольных матчах, аналитики кивали и выглядели весьма осведомленными. Они бросали модные словечки: «машинное обучение» и «большие данные». Я, конечно, понимал, что они не знали всех деталей того, над чем я работал, но верил, что они улавливали ключевые идеи.

Аналитики не хотели подпортить впечатление. И я внезапно понял это в тот момент, когда они спросили о навыках, которые необходимы их новым работникам. Они понятия не имели, о чем я говорю, и очень мало знали об уравнениях. Аналитики не умели программировать и считали статистику не наукой, а списком чисел в приложении к годовому бухгалтерскому отчету. Один из них спросил меня, можно ли считать математический анализ важным навыком для выпускника математического факультета.

Как я мог быть таким наивным? Почему я не заметил этого раньше? В тот день мы слушали одного парня, который написал книгу о том, почему нам следует «думать медленно»*. Это было очень «вдохновляюще». Он повторял слово «ме-е-е-е-е-едленно» очень неторопливо, чтобы мы понимали, что перед принятием решения нам следует притормозить. И рассказывал истории о том, как до-о-о-олго держал акции и они выросли в цене или как устанавливал большо-о-о-о-ой интервал времени для оценки своих активов. И он вешал нам о спорах с каким-то другим парнем, который считал, что нужно действовать быстро. Один из примеров, который он приводил в доказательство своей точки зрения, касался калифорнийской компании, занимающейся автоматической торговлей. В этой сфере даже времена пересылки данных о цене с западного побережья в торговый зал в Чикаго слишком велико. По этой причине компания переехала, передвинув свои основные структуры

* Вероятно, речь о Даниэле Канемане, авторе книги «Думай медленно... решай быстро» (М. : АСТ, 2013). Прим. ред.

к фондовой бирже. Однако производительность используемого алгоритма упала. Он работал лучше, когда расстояние было больше.

Вывод выступавшего был таков: данный случай подтверждает, что медленнее — значит лучше. Это было явно неверно. История, по сути, повествовала о том, как алгоритм, настроенный на одно время, может не работать при других параметрах. Банальное наблюдение. В лучшем случае это можно считать историей о том, что если ваш алгоритм настроен не на такую шкалу времени, как другие, то у вас может быть какое-то преимущество. Все алгоритмы торговали поблизости от фондовой биржи были настроены на использование неэффективностей для коротких промежутков времени, а трейдеры западного побережья могли эксплуатировать неэффективности на несколько более продолжительной шкале. Так было, пока они не перенесли сервер. Но в более медленных временных масштабах нет ничего уникального.

Безусловно, в сфере принятия решений — как в экономике, так и в психологии — существуют высококачественные исследования, но докладчик не соблюдал базовые научные стандарты. Он давал полу продуманные рекомендации по инвестициям, используя ложную дилемму времени, — и казалось, что у него есть какая-то теория. Но моя цель не разбить его аргументы, меня беспокоило то, как его рассказы и истории других докладчиков на конференции воспринимались присутствующими. Аналитики рынка, которые почти ничего не знали об алгоритмах в основе бизнеса, рассказывали друг другу случаи, чтобы другие чувствовали себя умнее.

Я позволил себе стать частью этого. Казалось, моя роль здесь — предложить такие же истории: о ставках в Премьер-лиге, скаутах из футбольных клубов и алгоритмах Google. Они подтвердили бы убеждения хозяев, что они понимают, как работает высокочастотная торговля или спортивная аналитика. Но еще больше, чем незнание ими технических деталей высокочастотной торговли, меня тревожило то, что из алгоритмов, которые использовали эти трейдеры, можно было извлечь полезные уроки. Они могли бы помочь найти более сбалансированный подход к их работе. Однако эти люди рассматривали алгоритмы как черный ящик, методы небольшой

горстки специалистов по количественному анализу, которым они платят зарплату, чтобы эти «кванты» приносили им прибыль; и они, похоже, даже не пытались понять то, что знают эти специалисты и чего до сих пор не могли ухватить они сами.

Более того, они опасались задавать вопросы, когда могли не понять ответов. Я ощущал за столом этот страх и, к своему стыду, потворствовал ему. Вместо того чтобы рассказать им то, что им нужно было знать, я продолжал пичкать их историями, которые они ожидали услышать: рассказывал о своем визите в Барселону, о Яне и Мариусе и о том, как футбольные скауты находят новых игроков. Они выглядели заинтересованными, и вечер прошел приятно. У них тоже было что рассказать и хватало действительно интересных историй. Один из них недавно встречался с Нассимом Талебом, которого я очень уважаю. У другого дочь занималась математикой в Гарварде. Я пил вино и пропитывался этой атмосферой. Я радовался, когда подходила моя очередь рассказывать разные случаи, и старался излагать как можно лучше.

Не судите меня. Нет причин, по которым я не могу наслаждаться компанией людей, которые не разбираются в математических деталях торговли. Иногда они могут быть даже забавнее, чем те, кто в этом разбирается.

*

А вот что я мог бы рассказать, если бы не был таким лицемером.

Нельзя раскрыть секреты финансовых рынков, не начав с этого фундаментального уравнения:

$$dX = hdt + f(X)dt + \sigma \cdot \varepsilon_t \quad (\text{Уравнение 6}).$$

Уравнения упрощают мир, концентрируя много знаний в небольшом количестве символов, и уравнение рынка — блестящий пример. Если мы хотим распаковать знание, заключенное в этом уравнении, нам нужно шаг за шагом разобраться с ним.

Это уравнение описывает, как меняется величина X , которая представляет «ощущение» инвесторов в отношении текущей стоимости какой-то акции. Это ощущение может быть положительным

или отрицательным, так что $X = -100$ означает реально плохое ощущение о будущем, а $X = 25$ — довольно хорошее. Экономисты говорят о «быках» и «медведях» на рынке. В нашей модели бычий рынок в будущем положителен ($X > 0$), а медвежий отрицателен ($X < 0$). Если мы хотим большей конкретики, то можем представлять себе нашу величину X как количество «медведей» (людей, играющих на повышение) минус количество «быков» (играющих на понижение). Но на этом этапе мы не хотим связываться с конкретной единицей для измерения X . Вместо этого думайте об X примерно как об улавливании эмоций. Это могут быть не инвесторы, а чувства на собрании, когда объявляют о сокращении рабочих мест, или ощущение после того, как ваша компания получила большой заказ.

В математике принято ставить слева в уравнениях то, что мы желаем объяснить, а справа — то, что, по нашему мнению, дает объяснение. Именно это мы делаем в нашем случае. Слева стоит величина dX . Буква d обозначает изменение. Соответственно, dX — «изменение в ощущении». Обратите внимание, как атмосфера в помещении мрачнеет, когда вы обнаруживаете, что ваша работа под угрозой. Такая угроза сокращения может быть $dX = -12$. Если новый заказ поможет вашей компании работать еще несколько лет, то $dX = 6$. А если он еще и большой, то может оказаться, что $dX = 15$.

Не надо фокусироваться на единицах, которые я использую для этих величин. Когда мы решаем школьные задачки, то обычно складываем и вычитаем реальные предметы — яблоки, апельсины или деньги; сейчас же мы можем позволить себе большую свободу. Я понимаю, что не существует такой вещи, как изменение эмоций ваших коллег $dX = -12$, но это не означает, что мы не можем написать уравнение, которое пытается уловить изменения в ощущениях группы людей. Именно это и есть цена какой-нибудь акции: это то, как инвесторы ощущают будущую стоимость компании. Мы хотим объяснить изменения в наших коллективных ощущениях по поводу инвестиций в конкретные активы или в отношении к какому-то политику либо потребительскому бренду.

В правой части уравнения три слагаемых — hdt , $f(X)dt$ и $\sigma \cdot \varepsilon_t$. Самая важная часть здесь — сигнал h , далее обратная связь $f(X)$

и стандартное отклонение σ (или шум). Коэффициенты, на которые умножены эти величины, указывают, что мы интересуемся изменениями (d) во времени (t). Шум умножается на ε_t (или небольшие случайные отклонения во времени). Эти слагаемые моделируют наши ощущения в виде комбинации сигнала, социальной обратной связи и шума. Сейчас мы уже способны понять нечто фундаментальное, но сначала рассмотрим конкретный пример.

*

Возможно, вас интересует, можно ли применить уравнение рынка для выбора пенсионной программы. Боюсь, здесь ответа придется подождать. Есть более насущные вопросы — например, стоит ли вам отправиться смотреть новый фильм из вселенной Marvel. Или какой тип наушников лучше купить. Или куда поехать в отпуск в следующем году.

Рассмотрим решение о покупке новых наушников. У вас припасено 200 фунтов, и вы просматриваете сайты в поисках оптимального варианта. Вы идете на страницу компании Sony и читаете технические характеристики; вы смотрите отзывы о японском бренде Audio-Technica; вы видите, что все знаменитости и спортивные звезды пользуются Beats. Что выбрать?

Я не могу посоветовать вам, какие наушники купить, но могу рассказать, как нужно подходить к этой задаче. Проблемы такого рода относятся к разделению сигнала h , обратной связи $f(X)$ и шума σ . Начнем с Sony и будем использовать переменную X_{Sony} , чтобы измерить, насколько потребители любят этот бренд. Мой первый качественный кассетный плеер и пара наушников, купленные в 1989 году у Ричарда Блейка уже подержанными, были выпущены как раз Sony. Они классические и надежные. Для уравнения 6 продукты Sony имеют фиксированное значение $h = 2$, а в качестве промежутка времени возьмем $dt = 1$ год. Поскольку единицы «ощущения» произвольны, сама по себе величина 2 не важна. Значима величина сигнала относительно социальной обратной связи и шума. Для Sony мы выберем $f(X) = 0$ и $\sigma = 0$. Иными словами, есть только сигнал.

Если мы начинаем с $X_{\text{Sony}} = 0$ в 2015 году, то для 2016-го в силу равенства $dX_{\text{Sony}} = h \cdot dt = 2$ получаем $X_{\text{Sony}} = 2$. В 2017-м у нас $X_{\text{Sony}} = 4$ и так далее — до 2020-го, в котором $X_{\text{Sony}} = 10$. Позитивное ощущение от Sony растет, потому что сигнал положителен.

О другом бренде, Audio-Technica, вы знаете намного меньше. На паре каналов в YouTube у него хорошие отзывы. Один фанат в вашем местном магазине аппаратуры утверждает, что именно этот бренд самый популярный у японских диджеев, но информации у вас немного. Если пользоваться советами только из одного-двух источников, возникает риск, и именно он создает шум. Поскольку рекомендации об этих наушниках японских диджеев поступили от небольшого числа лиц, мы дадим им $\sigma = 4$, итого шум будет вдвое больше сигнала.

Уравнение рынка для бренда Audio-Technica получилось таким: $dX_{\text{AT}} = 2dt + 4\epsilon_t$. Мы можем считать, что компонент ϵ_t дает нам каждый год какое-то случайное число*. Иногда оно положительное, иногда отрицательное, но в среднем $\epsilon_t = 0$, а его дисперсия равна 1.

Выбирая случайные значения для ϵ_t , мы можем моделировать случайную природу информации о бренде Audio-Technica. Именно это, как правило, делают специалисты по количественному анализу, когда моделируют изменения в цене акций. В любой конкретной задаче они запускают миллионы имитаций и смотрят на распределение результатов.

Сейчас, чтобы посмотреть, как работает такое моделирование, я «запущу» имитацию с выбранными ранее случайными значениями. Предположим, в 2015 году случайная величина $\epsilon_t = -0,25$. Тогда $dX_{\text{AT}} = 2 - 4 \cdot 0,25 = 1$. Если в следующем году $\epsilon_t = 0,75$, то $dX_{\text{AT}} = 2 + 4 \cdot 0,75 = 5$, а если в 2017 году $\epsilon_t = -1,25$, то $dX_{\text{AT}} = 2 - 4 \cdot 1,25 = -3$. Ваше доверие к бренду Audio-Technica со временем растет (в 2016, 2017 и 2018 годах величина X_{AT} составляет 1, 6 и 3 соответственно), однако более хаотично, чем в случае Sony.

Наконец, у нас есть продукт с социальной обратной связью: Dr Dre's Beats. Эти наушники предназначены для того, чтобы вы

* Если время (как в нашем случае) дискретно, то ϵ_t — последовательность случайных величин; если непрерывно, то ϵ_t — случайный процесс.

выкладывали свои фото в социальных сетях и создавали у других ощущение, что нужно последовать вашему примеру. Beats заставляют вас верить в рекламную шумиху. По мере того как такое ощущение возникает у знаменитостей и сетевых авторитетов, оно привлекает все больше людей и становится сильнее. В нашей модели мы могли бы установить, например, $f(X) = X$: ощущение для наушников Beats растет пропорционально своей величине. Чем больше любят Beats, тем больше любви к ним производится. Это дает нам уравнение рынка $dX_B = 2dt + X_B dt + 4\epsilon_t$: изменение отношения для Beats составляют 2 единицы роста, X_B для социальной обратной связи и 4 единицы для шума.

Вообразим, что бренд Beats начался с плохого года в 2015-м, случайный фактор $\epsilon_t = -1$. Предположим, что изначально у нас $X_B = 0$. После применения уравнения рынка получаем $dX_B = 2 + 0 - 4 \cdot 1 = -2$. К началу 2016-го наше доверие к X_B отрицательно, $X_B = -2$. В следующем году дела с шумом обстоят лучше, $\epsilon_t = 0,25$, но обратная связь ограничивает улучшение ситуации: $dX_B = 2 - 2 + 4 \cdot 0,25 = 1$. Поэтому в 2017-м $X_B = -1$. В течение 2018 года $\epsilon_t = 1$, и отношение к Beats начинает улучшаться: $dX_B = 2 - 1 + 4 = 5$. Теперь $X_B = 4$, и социальная связь определяет взлет: несмотря на плохонький 2019-й с $\epsilon_t = 0$, мы имеем $dX_B = 2 + 4 + 0 = 6$, и ощущение от Beats сохраняется: $X_B = 10$. Слагаемое социальной обратной связи усиливает и хорошее, и плохое: сначала это может мешать популярности товара, но когда какое-то отношение установилось, то оно усиливается все быстрее.

Я, конечно, создаю карикатурное изображение Sony, Beats и бренда японских диджеев. И прежде чем кто-то из них подаст на меня в суд, покажу вам реальную трудность для вас как потребителя. Просматривая сайты и расспрашивая друзей, вы измеряете ощущения людей от тех или иных наушников. Для нашей модели они сведены воедино на рис. 6.

В различные периоды верхнюю строчку в соответствии с ощущениями потребителей занимают разные товары. В 2016 и 2018 годах лучшей считалась Sony. В 2017-м — наушники Audio-Technica. А в 2019 и 2020-м лучшим товаром выглядит Beats.

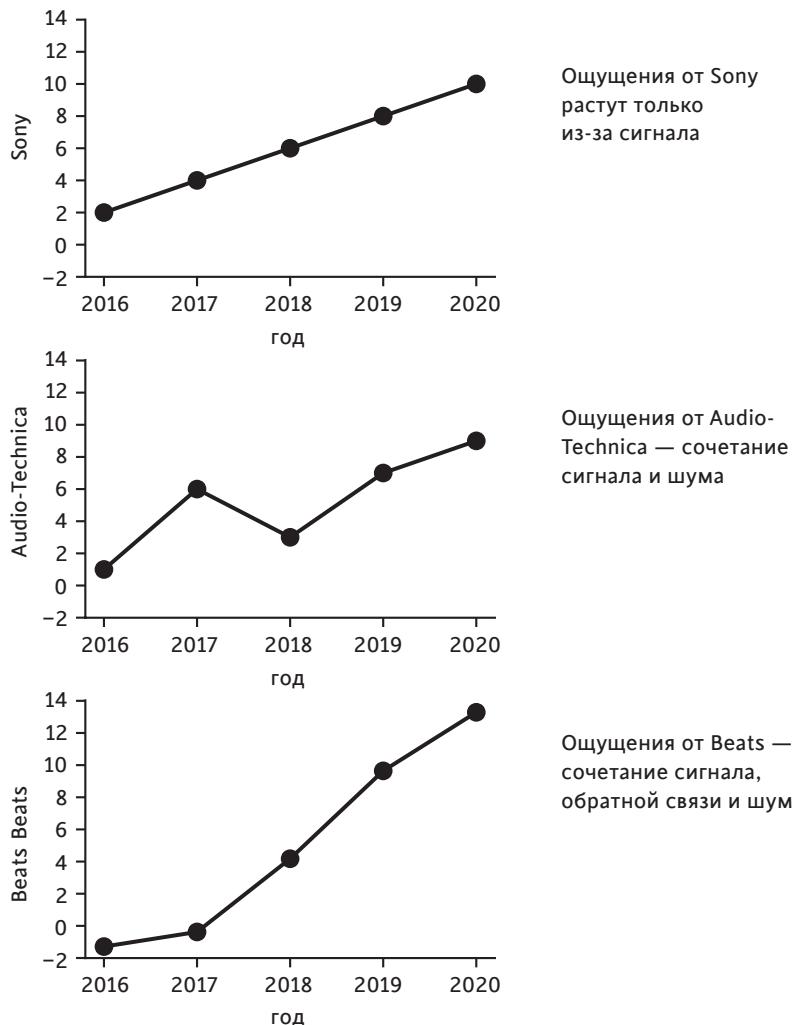


Рис. 6. Как меняется со временем отношение к трем продуктам

У вас может возникнуть соблазн сделать из вышенаписанного вывод, что следует отдать предпочтение Sony, у которой самый надежный сигнал. Но помните, что у всех остальных товаров тоже есть реальный сигнал, и в нашем примере он одинаков для всех: $h = 2$. Правая сторона уравнения рынка говорит вам, что нужно копать глубже. Для любого бренда в уравнение будут входить все три компонента. Ваша задача как слушателя — обнаружить сигнал

в обратной связи и шуме. То же рассуждение можно применять ко всем потребительским продуктам — от последнего блокбастера или онлайн-игры до кроссовок и дамских сумочек. Чаще всего вы руководствуетесь ощущением о продуктах, но в реальности хотели бы знать об их качествах.

Задача для фондовых рынков ровно та же. Мы часто знаем только рост цены акции dX , но нам хочется узнать и силу сигнала. Может быть, здесь важная составляющая — социальная обратная связь, создающая ажиотаж? Каковы источники смущающего шума?

*

Веками участники «Десятки» видели только сигнал. Вдохновленный неизбежной силой ньютоновского тяготения, шотландский экономист XVIII века Адам Смит описал невидимую руку, которая приводит рынок к равновесию. Обмен продуктами уравновешивает спрос и предложение. Итальянский инженер Вильфредо Парето формализовал взгляды Смита с помощью математики, описав наше постоянное экономическое развитие в сторону оптимальности. Сигнал прибыли неизбежно ведет нас к стабильному процветанию. Во всяком случае так они считали.

Первые редкие признаки нестабильности — тюльпанная лихорадка в Голландии или крах Компании Южных морей — давали мало реальных оснований для беспокойства. Только после распространения капитализма по всему миру для взлетов и падений потребовались объяснения. От Великой депрессии 1929 года до биржевого краха 1987-го повторяющиеся кризисы демонстрировали обществу, что рынки несовершены — они могут быть беспорядочными, со значительными колебаниями. Шум стал таким же сильным, как сигнал.

Как физика в начале XX века из ньютоновской превратилась в эйнштейновскую, так эволюционировала и математика рынков.

В 1900 году французский математик Луи Башелье опубликовал свою диссертацию «Теория спекуляций», в которой описывались два компонента уравнения 6. Эйнштейн в 1905 году, еще до публикации своей теории относительности, объяснил, как движение пыльцы в воде вызывается случайными ударами молекул воды.

Казалось, эта новая математика случайностей способна идеально описать удары, которые внешние события наносят экономическому процветанию, и участники «Десятки» начали создавать новую теорию. На протяжении большей части XX века новым источником прибыли стал шум. Для разработки и определения цены деривативов, фьючерсов и опционов используются дальнейшие развития базовой теории — например, модель Блэка — Шоулза. Участников общества привлекали как для создания, так и для управления этими новыми моделями. По сути, им поручили управлять денежной массой мира.

Точно так, как ньютоновский детерминистический анализ был неправильной моделью финансовых рынков, во взгляде на бомбардировку рынка шумами отсутствовал один крайне важный элемент: мы участники рынка. Мы не частицы, на которые воздействуют какие-то события; мы активные агенты, одновременно и рациональные, и эмоциональные. Мы ищем сигнал в шуме и в процессе влияем на других людей, учимся у них, манипулируем ими. Математическая теория не имеет права пренебрегать человеческой сложностью.

Вдохновленные таким открытием, некоторые участники «Десятки» выбрали новое направление исследований. Институт Санта-Фе в Нью-Мексико собрал математиков, физиков и других ученых со всего мира. Они начали набрасывать новую теорию сложности, которая пыталась учесть наши социальные взаимодействия. Она прогнозировала большие непредсказуемые колебания цен на акции, вызванные стадными действиями трейдеров. Модели утверждали, что по мере увеличения волатильности следует ожидать еще более серьезных всплесков и спадов, чем те, что мы видели в прошлом. Исследователи публиковали одно предупреждение за другим в профильных научных журналах⁷⁵. Секреты «Десятки», как всегда, были открыты, их могли прочитать все. К сожалению, очень немногие удосужились это сделать.

Один из исследователей в Санта-Фе, Дж. Дойн Фармер, покинул институт, чтобы претворять эти идеи на практике. Позже он рассказывал мне, что работа была очень тяжелой, гораздо сложнее, чем он мог себе представить, но она окупилась. Во время азиатского

кризиса 1997–1998 годов, пузыря доткомов в 2000-м и финансового кризиса 2007 года инвестиции Фармера оказались в безопасности: они были защищены от потрясений, обрушивших финансовые учреждения и правительства и посевавших семена политического недовольства по всей Европе и в США.

Математики могли бы с определенным основанием сказать, что они с самого начала знали о будущих катастрофах. В отличие от других, они были к ним готовы. Пока многие теряли, участники «Десятки» по-прежнему получали прибыль.

*

Я забежал вперед. Рассказ о том, как математика перешла от понимания рынков в терминах простого сигнала к оценке шума и в итоге к учету социальной обратной связи, — это хорошо, но здесь упускается важный момент. Звучит слишком похоже на сказку о том, как на каком-то этапе человеческое безумие было исправлено с помощью новых методов мышления.

Да, математики извлекли уроки из своих ошибок за последнее столетие и по-прежнему опережают всех, но есть еще одно необходимое важное замечание: они на самом деле понятия не имеют, как найти истинный сигнал в шуме финансовых рынков.

С моей стороны это довольно смелое заявление, придется его объяснить шаг за шагом. Секрет первоначального успеха уравнения рынка можно найти в главе 3. Математики, работавшие в сфере финансов в 1980-е, смогли разделить сигнал и шум, собрав достаточно наблюдений. В первой версии уравнения рынка, созданной Луи Башелье, не было $f(X)$: оно просто описывало рост уверенности в компании (и, соответственно, в цене ее акций) как комбинацию сигнала и шума. Такое математическое знание уже позволяло трейдерам снижать воздействие случайностей, которым подвергались их клиенты. Это давало им преимущество над теми, кто не понимал случайности, смешивал сигнал и шум.

Более чем за десять лет до кризиса 2007 года одна группа физиков-теоретиков заявила, что уравнение рынка, основанное только на сигнале и шуме, опасно. Они показали, что такая модель не дает достаточных колебаний цены акций, которые объяснили бы

значительные всплески и спады, наблюдавшиеся в предыдущее столетие. И пузырь интернет-компаний, и азиатский кризис 1998 года привели к обвалу цен до значений, которые не могла предсказать простая модель сигнала и шума.

Чтобы понять масштаб таких больших отклонений, вспомните де Муавра и его опыт с подбрасыванием монеты. Он обнаружил, что количество орлов после n бросков обычно лежит в интервале размера, пропорционального \sqrt{n} . Центральная предельная теорема расширила результат де Муавра и говорит, что то же правило с \sqrt{n} приложимо ко всем играм и даже ко многим жизненным ситуациям, например опросам общественного мнения. Ключевое предположение, необходимое для применения центральной предельной теоремы, — независимость событий*. Мы складываем результаты независимых вращений рулетки или спрашиваем независимые мнения множества разных людей.

Простая рыночная модель сигнала и шума также предполагает независимость при ценообразовании. Согласно ей, будущая стоимость акций должна следовать правилу $\sigma\sqrt{n}$ и нормальному распределению. В реальности дела обстоят не так. Физики-теоретики из Санта-Фе и других научных центров мира показали, что колебания будущих цен на акции могут оказаться пропорциональными более высоким степеням n , например $n^{2/3}$ или даже самой величине n^{76} . Это делает рынок крайне волатильным, а прогнозы почти невозможными: акция за день способна потерять всю стоимость. Это эквивалентно тому, что де Муавр подкинул монету 1800 раз и получил 1800 решек.

Причина этих колоссальных колебаний в том, что трейдеры не действуют независимо друг от друга. В рулетке одно вращение колеса не зависит от другого, и можно применять центральную предельную теорему. Однако на фондовом рынке один продающий трейдер заставляет другого потерять уверенность и тоже продавать. Это делает недействительными предположения теоремы де Муавра, поэтому она неверна; и колебания в цене акций уже

* На самом деле центральная предельная теорема верна не только для независимых, но и для слабо зависимых случайных величин.

нельзя считать небольшими и предсказуемыми. Участники фондового рынка — стадные животные, которые следуют друг за другом и в подъемах, и в спадах.

Не все специалисты по финансовой математике понимали, что центральную предельную теорему нельзя применять к рынкам. Когда я встретился с Фармером в 2009 году, он рассказал мне об одном коллеге из брокерской компании, которая (в отличие от собственной компании Фармера) потеряла много денег по время кризиса 2007–2008 годов. Тот специалист назвал банкротство инвестиционного банка Lehman Brothers «событием двенадцати сигм». Как мы видели в главе 3, отклонение более чем на 1σ бывает примерно один раз из трех; на 2σ — один раз из двадцати, а на 5σ — один раз на 3,5 миллиона случаев. Отклонение на 12σ случается один раз из... не могу даже назвать уверенно это число, потому что мой калькулятор дает сбой, если я пытаюсь найти что-то больше 9σ . В любом случае это крайне маловероятное событие, и оно никак не могло произойти — если только модель была правильной.

Эти физики-теоретики, возможно, обнаружили математику, стоящую за большими уклонениями, но явно были не единственными, кто описывал стадный менталитет трейдеров. Две книги Нассима Талеба — «Одураченные случайностью» и «Черный лебедь» — содержат занятно высокомерный, но исключительно прозорливый анализ финансового мира до 2007 года. Книга американского экономиста Роберта Шиллера «Иrrациональный оптимизм», написанная примерно в то же время, дает более академичное и тщательное изложение близких идей⁷⁷. Когда физики-теоретики, pragматичные инвесторы с количественным подходом и экономисты из Йельского университета говорят о недостатках модели одно и то же — вероятно, прислушаться к их словам вполне разумная идея.

На рубеже тысячелетий многие физики-теоретики, пришедшие на работу в финансовую сферу, нашли на рынке определенное преимущество. Они сохраняли это преимущество в течение всего финансового кризиса и выиграли, когда рынки рухнули. Добавив слагаемое $c f(X)$ в свои уравнения рынка, они оказались готовыми к таким событиям, как крах Lehman Brothers, когда трейдеры занимали друг за другом экстремальные и рискованные позиции.

Сейчас все специалисты по финансовой математике знают, что рынки — комбинация сигнала, шума и стадности: их модели показывают, что крахи будут происходить, и позволяют им хорошо представлять, насколько масштабными они будут в долгосрочной перспективе. Однако математики не знают, почему или когда происходят такие крахи; по крайней мере не более того, что это имеет какое-то отношение к стадной ментальности. Они не понимают фундаментальных причин подъемов и падений. Когда я сидел в онлайн-казино в главе 3, я знал, что игра несправедлива и у казино есть преимущество. Я знал, что сигнал составляет 1/37 — это средние потери игрока при одном запуске рулетки. Я мог узнать это, просто посмотрев, как устроено колесо. В главе 4 Люк Борн рассматривал вклады всех баскетболистов в общую игру команды, чтобы измерять умения игроков. Он нашел сигнал умений, сочетая свои знания игры и хорошо выбранные предположения. В главе 5, когда Лина и Микаэла занимались обратной разработкой алгоритма Instagram, они могли начать с понимания, как социальные сети искажают их взгляд на мир. Во всех этих примерах модель давала представление, как действуют рулетка, баскетбол и социальные сети соответственно. Уравнение рынка само по себе понимания не дает.

В разное время исследователи пытались понять, смогут ли они сделать еще один шаг и найти на рынках истинный сигнал. В 1988 году, после Черного понедельника 1987-го, Дэвид Катлер, Джеймс Потерба и Ларри Саммерс из Национального бюро экономических исследований написали статью под названием «Что движет биржевыми ценами?»⁷⁸ Они установили, что такие факторы, как промышленное производство, процентные ставки и дивиденды, которые влияют на доходность фондового рынка, могут объяснить только примерно треть колебаний на рынке акций. Затем проверили, не играли ли какую-то роль важные новости — например, войны или смена президента. В дни таких серьезных новостей значительные колебания стоимости акций действительно происходили, но имелось изрядное количество дней, когда никаких новостей не было, а рынок активно двигался. Подавляющее большинство движений на фондовой бирже нельзя объяснить внешними факторами.

В 2007 году Пол Тетлок, профессор экономики Колумбийского университета, предложил «коэффициент пессимизма СМИ» для колонки «В ногу с рынком» газеты *Wall Street Journal*, которая составляется ежедневно непосредственно после закрытия торгов⁷⁹. Этот коэффициент учитывает, сколько раз в колонке использовались разные слова, и тем самым измерялось общее настроение в отчете автора о дневных торгах. Тетлок обнаружил, что существовала связь между пессимистическими словами и падением акций на следующий день, но позже на неделе эти падения обращались вспять. Он заключил, что колонка «В ногу с рынком» вряд ли содержит полезную информацию о долговременных трендах. Другие исследования показали, что сплетни в чатах интернета и даже то, что люди говорят друг другу в биржевых залах, могут предсказать объемы торговли, но не направление движения рынка⁸⁰. Надежных правил для прогнозирования будущей цены акций просто не существует.

Хочу сразу прояснить два момента. Во-первых, эти результаты не означают, что новости о компании не влияют на стоимость ее акций. Акции Facebook упали после скандала с Cambridge Analytica. Акции BP — после катастрофы с разливом нефти на буровой платформе Deepwater Horizon. Однако в этих случаях события, вызвавшие изменения в цене акций, были еще менее предсказуемы, чем сама их стоимость, что делало их более-менее бесполезными для инвестора, стремящегося к прибыли. Когда вы слышите новости, их слышат все. Возможность получения преимущества пропадает.

Во-вторых, еще раз подчеркну, что модели, основанные на уравнении рынка, обеспечивают полезное долгосрочное планирование рисков. Мой друг, математик Майя, работает в крупном банке. Она использует уравнение 6, чтобы оценивать различные риски, которым подвержен банк, затем покупает страховку, чтобы защитить его от неизбежных подъемов и спадов. Майя видит, что нематематики редко понимают ограничения той модели, которую она использует. Когда мы в последний раз обедали с ней и ее коллегой Пейманом, она сказала:

— Самая большая проблема, с которой сталкиваюсь у нематематиков, — они воспринимают результаты моделей буквально.

Пейман согласился:

— Вы показываете какой-нибудь доверительный интервал для промежутка времени в будущем, а они считают его истиной. Очень немногие понимают, что наша модель основана на весьма слабых предположениях.

Майя и Пейман борются с восприятием, что раз это математика, то это должно быть истиной. Уравнение рынка не таково. Его основной посыл в том, что нам нужно быть осторожным, поскольку в будущем может произойти почти что угодно.

Многие трейдеры разделяют такой взгляд на финансовые рынки: мы можем застраховаться от колебаний, но неспособны понять, почему они случились. Когда в начале 2018 года рынки временно упали, а потом подскочили опять, Манодж Наранг, генеральный директор компании MANA Partners, занимающейся алгоритмической торговлей, рассказал деловому изданию Quartz: «Понять, почему на рынке что-то произошло, лишь немногим проще, чем понять смысл жизни. Обоснованные предположения есть у многих, но знаний нет»⁸¹.

Если трейдеры, банкиры, математики и экономисты не понимают причин движения рынка, почему вы думаете, что вы понимаете? Что заставляет вас считать, будто акции Amazon достигли своего пика, а бумаги Facebook продолжат падать? Что придает вам уверенность, когда вы говорите о выходе на рынок в нужное время?

Летом 2018 года меня пригласили на Power Lunch на канале CNBC — одну из крупнейших программ деловых новостей в США. Я и раньше бывал в новостных студиях, но сейчас был совершенно другой размах: огромный зал размером с хоккейный стадион, заполненный журналистами, которые бегали во все стороны между столами. Повсюду располагались экраны с изображениями сияющих офисов в Сиэтле, подземных залов с мощными компьютерами в Скандинавии, крупных заводских комплексов в Китае и деловой встречи в одной из африканских столиц. Ведущие передачи взяли меня в монтажную комнату, чтобы посмотреть, как вся эта информация объединяется для эфира. Видеоролики со всего мира перекрывались бегущими цифрами цен на акции и заголовками последних новостей.

Уравнение рынка научило меня тому, что почти всё на этих экранах — бессмысленный шум или социальная обратная связь. Это бессмыслица. Нельзя получить ничего полезного, наблюдая за ежедневными изменениями цены акций или за экспертами, объясняющими, почему вам нужно или не нужно покупать золото. Есть много инвесторов (в том числе кое-кто из тех, с кем я встретился тогда в Гонконге), умеющих определить удачные инвестиции путем тщательного изучения принципов бизнеса, в который они инвестируют. Однако если исключить систематическое исследование того, как работает компания, то все советы по инвестициям — случайный шум. Они включают мотивационные раздумья гуру, которые иногда зарабатывали деньги в прошлом.

Эта неспособность предсказать будущее на основании прошлого относится и к нашим личным финансам. Покупая дом, не беспокойтесь о том, как менялись цены в этом районе в последние несколько лет. Вы не можете использовать этот тренд для прогнозирования будущего. Вы должны четко осознавать, что цены на жилье подвержены колоссальным колебаниям — в соответствии с настроениями рынка. Убедитесь, что вы морально и финансово готовы к обоим случаям. И тогда купите тот дом, который вам нравится больше всего и который вы можете себе позволить. Найдите район, который вам по душе. Решите, сколько времени можете вложить в ремонт. Обратите внимание на время на дорогу до работы и до школы. Важны общие принципы рынка, а не то, находится ли ваш дом в «перспективном районе».

Когда дело доходит до покупки акций, не мудрствуйте слишком много. Найдите компании, которым доверяете, вложите деньги и смотрите, что происходит. Кроме того, вложите некоторую сумму в какой-нибудь индексный инвестиционный фонд, распределяющий ваши средства по акциям большого числа разных компаний. Убедитесь, что у вас хорошая пенсия. Вы не можете сделать больше. Не зацикливайтесь на этом.

Проверить настоящее качество трех видов наушников очень просто. Составьте плейлист из любимых десяти песен и прослушайте

их в каждой паре наушников по очереди. Порядок песен для каждой пары сделайте случайным. Оцените звук. Не спрашивайте друзей и не смотрите на мнения в интернете. Слушайте сигнал.

*

Математики — люди хитрые и коварные. Как только мы говорим вам, что все случайно, то обнаруживаем какое-нибудь новое преимущество. Когда мы устанавливаем, что с помощью математики невозможно предсказать долговременные тенденции для стоимости акций, тодвигаемся в противоположном направлении. Мы рассматриваем всё более короткие промежутки времени. Мы находим преимущество там, где люди неспособны производить расчеты.

15 апреля 2015 года на биржу вышла компания Virtu Financial. Она была основана семью годами ранее трейдерами Винсентом Виолой и Дугласом Сифу и разработала новаторские методы для высокочастотной торговли — покупки и продажи акций в течение миллисекунд после совершения сделки на фондовой бирже на другом конце страны. Вплоть до момента предложения своих акций Virtu была крайне скрытна в отношении своих методов и заработанных сумм. Но для выхода на биржу через IPO (первичное публичное размещение акций) ей требовалось раскрыть для проверки свои финансовые дела и детали бизнеса.

Секрет вышел наружу. В течение пяти лет торговли Virtu понесла убытки всего в один день. Этот результат был поразителен по любым меркам. Финансовые трейдеры обычно имели дело со случайностью; они привыкли, что неотъемлемой частью получения прибыли оказываются недели или месяцы убытков. Virtu убрала из трейдинга спады, оставив только подъемы.

Первоначальная оценка Virtu на фондовой бирже составляла 3 миллиарда долларов.

Заинтригованный гарантированными ежедневными прибылями, профессор астрономии Йельского университета Грэг Лафлин захотел узнать, почему деятельность Virtu была такой надежной⁸². Дуглас Сифу говорил в интервью Bloomberg, что только 51–52% сделок Virtu прибыльные⁸³. Сначала это заявление озадачило Лафлина: если 48–49% сделок убыточны, то для того, чтобы

гарантированно ежедневно оставаться в плюсе, требуется очень большое их количество.

Лафлин подробно изучил тип сделок, которыми занималась Virtu. Компания получала прибыль, зная об изменении цен раньше конкурентов. Ее документы для IPO раскрыли, что компания контролировала корпорацию Blueline Communications LLC, разработавшую технологию высокочастотной связи, благодаря которой информация о ценах проходила между фондовыми биржами Иллинойса и Нью-Джерси примерно за 4,7 миллисекунды. В своей книге 2014 года о высокочастотном трейдинге *Flash Boys** Майкл Льюис указывал, что при использовании волоконной оптики для связи между биржами задержка составляет примерно 6,65 миллисекунды. В результате Virtu имела преимущество примерно в 2 миллисекунды над волоконной оптикой.

В масштабе 1–2 миллисекунд маржа прибыли составляет около 0,01 доллара на акцию. Это означает, что часто сделки совершаются без прибыли или без убытка, то есть оказываются «нулевыми». Предположив, что 24% сделок убыточные, а 25% — «нулевые», Грег Лафлин вычислил среднюю прибыль от одной сделки: $0,51 \cdot 0,01 - 0,24 = 0,0027$ доллара на одну акцию. Предоставленная компанией Virtu отчетность говорила о доходе в 440 000 долларов в день, откуда следовало, что Virtu ежедневно совершала 160 миллионов сделок с акциями⁸⁴. Это составляло 3–5% всех сделок на фондовом рынке США. Они получали крохотный процент от вполне заметной части всех сделок. Мельчайшее преимущество на максимально возможной скорости приносило большую гарантированную прибыль.

Я связался с Винсентом Виолой и Дугласом Сифу и попросил о беседе. Ни тот ни другой не ответили. Тогда я позвонил своему другу Марку⁸⁵, который работает на другого крупного алгоритмического трейдера, и спросил, не может ли он рассказать мне секрет компаний, подобных Virtu. Он выделил пять способов получения преимуществ трейдерами, занимающимися высокочастотной торговлей.

* Имеется перевод на русский язык: Льюис М. *Flash Boys: Высокочастотная революция на Уолл-стрит*. М. : Альпина Паблишер, 2015.

Первый — скорость. Благодаря обладанию высокоскоростными каналами коммуникации, такими как технология микроволновой связи, разработанная Blueline, трейдеры узнавали направление сделок раньше конкурентов.

Второй — вычислительная мощность. Загрузка расчетов по трейдингу в центральный процесс компьютера требует времени, поэтому группы из сотен разработчиков используют видеокарты в своих машинах для обработки сделок по мере поступления.

Третий — тот, что наиболее часто используют Марк и его команда, — основан на уравнении 6. В последние годы популярной формой инвестиций были биржевые инвестиционные фонды (exchange-traded funds, ETF), которые представляют собой «корзины» инвестиций в различные компании на более крупном рынке, например S&P 500 (индекс, который отражает капитализацию 500 крупнейших компаний США). Марк объяснял мне: «Мы ищем арбитражные сделки для разницы между стоимостью отдельных акций в ETF и самим ETF». Арбитражная сделка — это возможность заработать деньги без риска, используя разницу в цене на один товар. Если в течение достаточного количества миллисекунд индивидуальная стоимость всех акций в ETF не отражает стоимости самого ETF, то алгоритмы Марка могут определить последовательность покупок и продаж, которые обеспечат прибыль из такой разницы в ценах. Команда Марка определяет арбитражные сделки не только для текущей стоимости акций, но и для будущих цен. Чтобы оценить опционы на покупку или продажу какой-то акции для срока в неделю, месяц или год в будущем, используется некий вариант уравнения рынка. Если Марк и его команда способны вычислить будущую стоимость ETF и всех отдельных акций раньше других, то они могут получить прибыль без риска.

Четвертый способ получения преимущества состоит в том, что такая компания — крупный игрок. «Чем больше вы торгуете, тем дешевле обходятся транзакции, — объясняет Марк. — Еще одна польза состоит в наличных кредитах или кредитах ценными бумагами, которые можно использовать для покрытия инвестиций, для выплат по которым уходит три или четыре месяца». По сути, богатые становятся богаче, потому что их капитал больше, а расходы ниже.

Пятый способ — тот, который сам Марк за пятнадцать лет трейдинга на самом высоком уровне с капиталами в миллионы долларов никогда не использовал: он заключается в попытке спрогнозировать истинную стоимость продаваемых акций и товаров. Существуют трейдеры, которые для принятия решений по инвестициям смотрят на принципы различного бизнеса, используя опыт и здравый смысл. Марк не из них. «Я придерживаюсь мнения, что рынок умнее меня в отношении цен, когда я смотрю, верно ли установлены цены на фьючерсы или опционы, если предполагать, что рынок правильный».

Марк вернулся к тому, что, на мой взгляд, является самым важным уроком из уравнения рынка — тем, который применим не только к нашим экономическим инвестициям, но и к вкладам в дружбу, взаимоотношения, работу и свободное время. Не верьте, будто вы сможете надежно предсказать, что произойдет в жизни. Принимайте решения, которые кажутся вам разумными, в которые вы действительно верите (здесь вам, конечно, следует применить уравнение суждений). Затем используйте три слагаемых уравнения рынка, чтобы морально подготовиться к неопределенному будущему. Помните о слагаемом для шума: будет множество подъемов и спадов вне вашего контроля. Помните о социальном слагаемом: не дайте ажиотажу поймать себя и не огорчайтесь, если толпа не разделяет ваших убеждений. И помните слагаемое для сигнала: истинная стоимость ваших инвестиций существует, даже если вы не всегда можете увидеть ее.

*

«Десятка» контролирует случайности с растущей в течение последних 300 лет уверенностью, забирая деньги у инвесторов, которые не владеют таким кодом. Те, кто не знает математических секретов, видя рост курса, считают, что это сигнал, и вкладывают деньги. Видя падение курса, продают. Либо поступают наоборот, пытаясь оспорить рынок. В обоих случаях они не учитывают возможность, а в основном руководствуются шумом и обратной связью.

Посторонние начинают всё лучше понимать особенности финансовой игры. Члены «Десятки» терпеливо слушают, когда полупрофессиональные игроки или инвесторы-любители рассуждают

о сигнале и шуме. Широко используются фразы вроде «одураченные случайностью», «обнаружение сигнала», «отношение сигнала к шуму» и «две сигмы» — они вылетают легко и с явной уверенностью. Пока идет такой разговор, «Десятка» продолжает находить новые преимущества на все более коротких временных промежутках — причем даже без поиска сигнала. Их алгоритмы используют возможность арбитража почти при любой сделке. Грег Лафлин внимательнее изучил сделки Virtu после прочтения книги *Flash Boys* Майкла Льюиса и статьи по этой теме в *The New York Times*, написанной американским экономистом Полом Кругманом⁸⁶. В электронном письме Грег сообщил мне: «Идея [статьи Кругмана] заключается в том, что при высокочастотной торговле трейдеры используют более изощренные и морально сомнительные методы для нечестного получения денег на рынке». Однако данные от Virtu не соответствуют такой точке зрения: компания забирала менее 1% от сделки для улучшения общей эффективности рынка. «Если у кого-то есть законное основание для покупки каких-то активов, а именно для долгосрочной выгоды и на основании здравых экономических принципов, расходы на транзакцию крайне низки, — говорил мне Грег. — Если кто-то пытается обыграть рынок в ежедневной торговле или паникует и желает разгрузить портфель в момент высокой волатильности, высокочастотная торговля получит выгоду от такого поведения».

Пока трейдеры играли на фондовом рынке подобно любителям, делающим спортивные ставки, математики старались использовать то, что эти трейдеры не понимали случайности. Как всегда, «Десятка» зарабатывала свои деньги благодаря крохотным преимуществам, анонимно, незаметно и без суэты.

*

Последний вопрос, который я задал Марку, касался моральной стороны: что он ощущал, получая быстрые прибыли от операций других людей? Я пояснил ему, что, когда его команда находит возможность арбитражной сделки, ее прибыль исходит от пенсионных фондов и инвесторов, которые торгуют не так быстро и точно, как он. Я спросил, как он относится к тому, что получает деньги

от инвестирования моих пенсионных средств и средств других людей.

Мы разговаривали по телефону, Марк стоял в своем саду в зеленом пригороде крупного европейского города. Я слышал пение птиц, пока он тщательно обдумывал свой ответ. И мне было очень неудобно задавать ему вопрос, который, как я знал, лежал за пределами технических аспектов работы Марка: о его вкладе в общество. Такие люди, как Марк, — которые зарабатывают свои деньги анонимно, без суеты, раз за разом применяя уравнения, — глубоко честны по своей природе. Он вынужден анализировать свой вклад с той же строгостью, с которой анализирует фондовый рынок, так же тщательно, как вообще все. Я знал, что сказанное им было верным с точки зрения фактов.

— Я бы спрашивал себя не о моральности отдельной сделки, а о том, становятся ли рынки более или менее эффективными из-за моей торговли. Несет ли ваш пенсионный фонд больше или меньше издержек, — сказал он. — До начала высокоскоростной торговли, если бы вы позвонили своему брокеру и спросили его о цене продажи и покупки, разница между ними была бы больше, чем сейчас.

Марк описал сомнительную практику, когда брокеры получали довольно высокий процент при совершении сделок. «Сейчас гораздо меньше более изощренных компаний, которые получают крохотную долю от каждой сделки». Брокеры старой школы, не умевшие правильно рассчитывать и бравшие большую долю, исчезли из бизнеса. И в целом складывается впечатление, что рынки эффективнее, чем были когда-то, но уверенности нет, поскольку объемы торговли тоже увеличились. Марк признал, что у него нет всех цифр и он не может сказать больше, однако его ответ соответствовал тому, что мне говорил Грег Лафлин.

Ответ Марка о роли высокочастотной торговли был неоднозначным, но честным. Он не прибегал к самооправданию, извинениям, идеологии или содержательным аргументам. Он превратил моральный вопрос в финансовый. Это был ответ, который бы одобрил Алфред Айер. Это был ответ участника «Десятки»: нейтральный и без бессмыслицы.

Глава 7

Уравнение рекламы

$$r_{x,y} = \frac{\sum_i (M_{i,x} - \bar{M}_x)(M_{i,y} - \bar{M}_y)}{\sqrt{\sum_i (M_{i,x} - \bar{M}_x)^2 \sum_i (M_{i,y} - \bar{M}_y)^2}}$$

Сначала я подумал, что это электронное письмо — спам. Оно начиналось с приветствия: «Мистер Самптер:», а в мире мало реальных людей, которые используют двоеточие в начале письма. Даже когда я прочитал текст — просьбу комитета по предпринимательству, науке и транспорту Сената США в Вашингтоне, округ Колумбия, о беседе со мной, — я оставался скептиком. Странным показался уже сам факт, что просьба пришла в форме электронного письма. Не знаю, чего стоило ожидать, но я с подозрением отнесся к соседству длинного и подробного названия комитета и неформального обращения за помощью. Не сходилось.

Однако все было правильно. Комитет Сената действительно хотел побеседовать со мной. Я отправил короткий положительный ответ, и через несколько дней мы общались по скайпу с людьми из республиканской части комитета. Они желали узнать о компании Cambridge Analytica, которую Дональд Трамп нанял для обращения к избирателям в соцсетях и которая предположительно собирала данные о десятках миллионов пользователей Facebook. В СМИ уже имелись две стороны истории Cambridge Analytica. Одна сторона — блестящее представление Александра Никса, тогдашнего CEO, который заявлял, что использует алгоритмы в политических

кампаниях для микротаргетинга. Другая — разоблачитель Крис Уайли с крашеными волосами, который утверждал, что помогал Никсу и его компании создать инструмент для «психологической войны». Впоследствии Уайли сожалел о своих действиях, которые позволили избрать Трампа, а Никс создавал свой бизнес в Африке, опираясь на свой «успех».

В 2017 году, за год до скандала, я детально исследовал алгоритм, который использовала Cambridge Analytica, и пришел к заключению, противоречащему обеим версиям событий — и Никса, и Уайли. Я сомневался, что компания могла повлиять на президентские выборы в США. Она, конечно, пыталась, но я обнаружил, что методы, которые, по их словам, использовались для таргетинга избирателей, были с изъяном⁸⁷. Мои заключения привели к странной ситуации, когда я оспаривал оба имеющихся варианта изложения. Вот почему комитет Сената желал поговорить со мной. Больше всего республиканцы из администрации Трампа весной 2018 года хотели узнать, что делать с грандиозным скандалом вокруг рекламы в социальных сетях.

*

Прежде чем мы сможем помочь сенаторам, нам нужно понять, как нас видят создатели соцсетей. Для этого мы будем рассматривать людей как наблюдения (так делают и компании) и начнем с самых активных и важных: подростков. Эта группа желает увидеть как можно больше и как можно быстрее. Каждый вечер можно наблюдать, как они — либо вместе на диване, либо (всё чаще) в одиночестве в спальне — быстро щелкают и листают странички на своих любимых платформах в соцсетях: Snapchat и Instagram. Через окошко своих телефонов они могут видеть невероятные картины мира: гномов, падающих со скейтбордов; пары, идущие на свидания «правда или действие»; собак, играющих в Fortnite; маленьких детей, сующих руки в пластилин Play-Doh; девочек-подростков, стирающих макияж; или «сцепленные» истории из текстовых диалогов между воображаемыми студентами колледжа. Они перемежаются сплетнями о знаменитостях, крайне редкими реальными новостями и, конечно, регулярной нескончаемой рекламой.

Внутри Instagram, Snapchat и Facebook создается матрица наших интересов. Это набор чисел в виде таблицы, где в строках — люди, а в столбцах — типы «постов» или «снимков», на которые они нажимают. В математике мы представляем таблицу подростковых кликов в виде матрицы, которую обозначим M . Вот пример для иллюстрации в гораздо меньшем масштабе: так выглядит матрица некой соцсети для двенадцати пользователей.

	Еда	Косметика	Кайли Дженнер	Пьюдипай	Fortnite	Дрейк	
$M =$	8	6	6	0	9	2	Мэдисон
	1	6	1	4	9	9	Тайлер
	2	0	0	9	5	3	Джейкоб
	5	0	9	8	7	2	Райан
	5	9	7	1	0	1	Алисса
	3	6	9	1	2	3	Эшли
	5	7	7	1	2	4	Кайла
	6	3	3	5	6	9	Морган
	6	0	0	0	2	8	Мэтт
	1	4	9	8	2	1	Хосе
	8	7	8	2	3	1	Сэй
	2	0	1	8	7	4	Лорен

Каждое число в матрице показывает, сколько раз подросток кликнул по конкретному типу постов. Например, Мэдисон посмотрела 8 постов о еде, по 6 о косметике и Кайли Дженнер, ни одного о ютюбере Пьюдипае и видеоигре Fortnite и 2 публикации о рэпере Дрейке.

Просто глядя на эту матрицу, мы можем получить хорошее представление о том, что за человек Мэдисон. Попробуйте представить ее себе, а потом потратите несколько секунд, чтобы вообразить некоторых других персонажей, которых я ввел здесь, используя в качестве ориентиров просмотренные ими снимки. Не беспокойтесь. Это не настоящие люди. Вы можете быть сколь угодно категоричными.

В матрице есть еще несколько человек, похожих на Мэдисон. Например, Сэм любит косметику, Кайли Дженнер и еду, но проявляет незначительный интерес к другим категориям. Есть и люди,

которые резко отличаются от Мэдисон. Джейкоб, как и Лорен, предпочитает Пьюдипая и Fortnite. Некоторые не совсем вписываятся в два этих типа. Скажем, Тайлер любит Дрейка и косметику, но интересуется Пьюдипаем.

Уравнение рекламы — математический способ автоматически определять тип людей. Оно имеет следующую форму:

$$r_{x,y} = \frac{\sum_i (M_{i,x} - \bar{M}_x)(M_{i,y} - \bar{M}_y)}{\sqrt{\sum_i (M_{i,x} - \bar{M}_x)^2 \sum_i (M_{i,y} - \bar{M}_y)^2}} \quad (\text{уравнение 7}).$$

Оно измеряет корреляцию между различными категориями снимков. Например, если люди, которые обычно ставят лайк Кайли Дженнер, также ставят лайк и косметике, то $r_{\text{косметика,Кайли}}$ будет положительным числом. В этом случае мы говорим, что существует положительная корреляция между Кайли и косметикой. Но если люди, которые ставят лайки Кайли, обычно не ставят их Пьюдипаю, $r_{\text{Пьюдипай,Кайли}}$ будет отрицательным числом, и мы назовем это отрицательной корреляцией.

Чтобы понять, как работает уравнение 7, разберем его шаг за шагом начиная с $M_{i,x}$. Это число в строке i и столбце x нашей матрицы M . Мэдисон 6 раз просматривала посты о косметике, поэтому $M_{\text{Мэдисон,косметика}} = 6$: у нас строка $i = \text{Мэдисон}$, а столбец $x = \text{косметика}$. В общем случае каждый раз, когда мы смотрим на число в строке i и столбце x матрицы, то видим $M_{i,x}$. Взглянем на \bar{M}_x . Эта величина — среднее число постов в категории x , приходящееся на одного пользователя. Например, среднее число просмотренных публикаций о косметике для наших подростков таково: $\bar{M}_{\text{косметика}} = (6 + 6 + 0 + 9 + 6 + 7 + 3 + 0 + 4 + 7 + 0) / 12 = 4$.

Если мы вычтем среднюю заинтересованность в косметике из общего числа публикаций, просмотренных Мэдисон, то получим $M_{i,x} - \bar{M}_x = 6 - 4 = 2$. Это говорит нам, что Мэдисон интересуется косметикой выше среднего. Аналогично, вычислив $\bar{M}_{\text{Кайли}} = 5$, мы видим, что она также (слегка) выше среднего интересуется Кайли Дженнер, поскольку $M_{i,y} - \bar{M}_y = 6 - 5 = 1$, если $i = \text{Мэдисон}$, а $y = \text{Кайли}$.

А теперь переходим к мощной интересной идее, лежащей в основе уравнения 7: если мы перемножим $(M_{i,x} - \bar{M}_x) \cdot (M_{i,y} - \bar{M}_y)$, то определим те интересы, которые, как правило, у людей общие. Для Мэдисон мы получаем:

$$(M_{\text{Мэдисон,косметика}} - \bar{M}_{\text{косметика}}) \cdot (M_{\text{Мэдисон,Кайли}} - \bar{M}_{\text{Кайли}}) = \\ = (6 - 4) \cdot (6 - 5) = 2 \cdot 1 = 2.$$

Это говорит нам о том, что между ее интересом к Кайли и косметике существует положительная корреляция.

Для Тайлера взаимоотношения между косметикой и Кайли отрицательные — $(6 - 4) \cdot (1 - 5) = 2 \cdot (-4) = -8$, — поскольку он проявляет интерес только к первой. Для Джейкоба величина снова положительна: $(0 - 4) \cdot (0 - 5) = (-4) \cdot (-5) = 20$, так как ему не нравятся ни первая, ни вторая (см. рис. 7). Обратите внимание на один нюанс. И у Джейкоба, и у Мэдисон положительное значение, хотя у них противоположные взгляды на Кайли и косметику. Однако их взгляды предполагают, что Кайли и косметика коррелируют между собой, хотя Джейкоб вообще никогда не смотрел ни на ту ни на другую. Поведение Тайлера в социальной сети не соответствует такой закономерности.

Мы можем произвести расчеты для каждого из подростков и сложить все такие величины. Получится сумма

$$\sum_i (M_{i,x} - \bar{M}_x)(M_{i,y} - \bar{M}_y).$$

Знак Σ_i указывает, что мы берем сумму по всем двенадцати тинейджерам. Сложив все произведения, где перемножены отношения подростков к косметике и к Кайли, получим:

$$2 - 8 + 20 - 16 + 10 + 8 + 6 + 2 + 20 + 0 + 9 + 16 = 69.$$

Большая часть слагаемых положительна: это показывает, что дети имеют схожее отношение к Кайли и косметике. Среди тех, кто вносит свой положительный вклад в сумму, — Мэдисон

и Джейкоб: 2 и 20 соответственно. Исключения — Тайлер, которому не нравится Кайли, и Райан, которому не нравится косметика; зато Кайли Дженнер по душе. Именно эта пара дала слагаемые -8 и -16 .

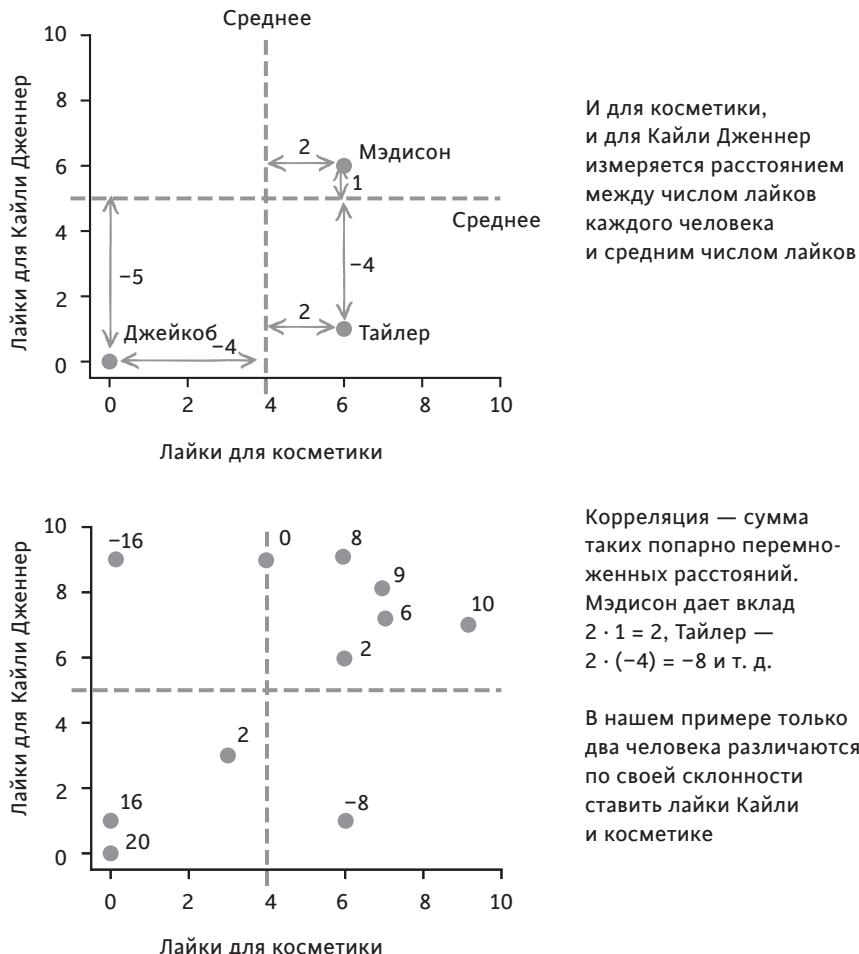


Рис. 7. Иллюстрация к вычислению корреляции между Кайли и косметикой

Математики не любят больших чисел вроде 69. Мы предпочитаем, чтобы они были меньше, лучше между 0 и 1, так их удобно сравнивать. Для этого мы добавим в уравнение 7 знаменатель (нижнюю

часть дроби)*. Я не стану подробно разбирать это вычисление, но если мы подставим все наши числа, то получим

$$r_{\text{косметика,Кайли}} = \frac{69}{\sqrt{120 \cdot 152}} = 0,51.$$

Мы получили одно-единственное число 0,51, которое измеряет корреляционную зависимость между косметикой и Кайли. Значение 1 показывало бы идеальную корреляцию между этими двумя типами постов, значение 0 говорило бы об отсутствии связи**. Так что реальное значение 0,51 дает нам среднюю корреляцию между любовью к косметике и к Кайли Дженнер.

Я понимаю, что провел уже довольно много вычислений, но мы нашли только одно из пятнадцати важных чисел, отражающих предпочтения подростков! Нам бы хотелось узнать корреляцию не только между косметикой и Кайли, но и между всеми категориями: еда, косметика, Кайли, Пьюдипай, Fortnite и Дрейк. К счастью, мы уже в курсе, как вычислить один коэффициент корреляции с помощью уравнения 7, — остается только подставлять в это уравнение каждую пару категорий. Именно это я сейчас и сделаю. Получится то, что известно под названием корреляционной матрицы, которую мы обозначим как R . Если вы посмотрите на пересечение строки «Кайли» и столбца «Косметика», то увидите найденное нами ранее число 0,51. Точно так же заполняются и остальные строки матрицы — для всех пар категорий. Например, Fortnite и Пьюдипай дают корреляцию 0,71. Но есть и такие пары,

* Число $r_{x,y}$, определяемое уравнением 7, называется выборочным коэффициентом корреляции. Он соответствует коэффициенту корреляции для случайных величин X и Y , определяемому как $r_{x,y} = \frac{E(X - EX)(Y - EY)}{\sqrt{E(X - EX)^2} \sqrt{E(Y - EY)^2}}$, где E означает математическое ожидание. Стоит уточнить, что и коэффициент корреляции, и выборочный коэффициент корреляции заключены в промежутке от -1 до 1 , а не от 0 до 1 , как можно понять из текста автора.

** Здесь нужно заметить, что корреляционная зависимость и математическая зависимость случайных величин — разные вещи. Действительно, если коэффициент корреляции для случайных величин X и Y равен 1 (или -1), это означает их сильную зависимость; отсюда следует линейная зависимость $X = aY + b$. Однако если коэффициент корреляции равен 0, это ничего не говорит о независимости случайных величин X и Y . Даже зависимые случайные величины вполне могут иметь $r_{x,y} = 0$, и поэтому малое значение коэффициента корреляции еще ничего не значит. А вот обратное верно всегда: если случайные величины X и Y независимы, то $r_{x,y} = 0$.

как Fortnite и косметика, которые дают коэффициент $-0,74$, то есть коррелируют отрицательно. Это означает, что геймеры, как правило, не особо интересуются косметикой.

	Еда	Косметика	Кайли	Пьюдипай	Fortnite	Дрейк	
$R =$	1	0,24	0,23	-0,61	-0,1	-0,11	Еда
	0,24	1	0,51	-0,63	-0,74	-0,26	Косметика
	0,23	0,51	1	-0,17	-0,17	-0,69	Кайли
	-0,61	-0,63	-0,17	1	0,71	-0,08	Пьюдипай
	-0,1	-0,74	-0,17	0,71	1	0,06	Fortnite
	-0,11	-0,26	-0,69	-0,08	0,06	1	Дрейк

Корреляционная матрица группирует людей по типам. Когда я просил вас представить себе этих подростков и не стесняться быть категоричными, я предлагал вам самим построить такую матрицу. Корреляция Кайли/косметика относит к одному типу таких подростков, как Мэдисон, Алисса, Эшли и Кайли, а корреляция Пьюдипай/Fortnite относит к другой группе Джейкоба, Райана, Моргана и Лорен. А вот Тайлер и Мэтт не вполне подходят под такую простую категоризацию.

В мае 2019 года я спрашивал Дуга Коэна, специалиста по данным из Snapchat, о той информации о пользователях, которую они хранят в корреляционных матрицах. «Ну, это почти всё, что вы делаете в Snapchat, — отвечал он. — Мы смотрим, как часто наши пользователи разговаривают в чатах с друзьями, сколько у них полос общения, какими фильтрами пользуются, как долго разглядывают карты, в скольких групповых чатах сидят, сколько времени тратят на просмотр контента или когда читают истории своих друзей. И мы смотрим, как эти действия коррелируют друг с другом».

Данные анонимны, поэтому Дуг не знает, чем занимаетесь конкретно вы. Но такие корреляции позволяют Snapchat категоризировать пользователей — от «одержимых селфи» и «документалистов» до «див макияжа» и «королев фильтров», если пользоваться внутренней терминологией компании⁸⁸.

Как только компания узнаёт, что привлекает определенного пользователя, она дает ему это в большом количестве. Слушая, как

Дуг описывает свою работу по привлечению людей, я не мог не прокомментировать: «Погодите! Я, как родитель, стараюсь, чтобы мои дети пользовались телефоном меньше, а вы трудитесь, чтобы повысить их вовлеченность!»

Дуг парировал, слегка уколов конкурентов: «Мы не просто стараемся максимизировать время, проведенное в приложении, как традиционно делал Facebook. Мы следим за уровнем участия, смотрим, как часто пользователи возвращаются. Мы помогаем им общаться с друзьями».

Snapchat не претендует на то, чтобы мои дети проводили у них всё свое время, но компания желает, чтобы они снова и снова возвращались. И по личному опыту могу сказать, что это работает.

*

Большинство хотели бы, чтобы их уважали как личностей, а не изображали в виде каких-то стандартных типов. Уравнение 7 полностью игнорирует наши желания. Оно сводит нас к корреляциям между вещами, которые нам нравятся.

Математики, работающие в Facebook, осознали силу корреляций еще на ранних стадиях разработки платформы. Каждый раз, когда вы ставите лайк на странице или комментируете какую-то тему, ваши действия предоставляют соцсети сведения о вас как о личности. Со временем Facebook стал использовать эти сведения иначе. В 2017 году, когда я впервые начал наблюдать, как аналитики отслеживают наши действия, категории казались забавными: среди коробок, по которым нас раскладывали, были «Брит-поп»*, «королевские свадьбы», «буксиры», «шея» или «верхушка среднего класса».

Такие категории оставляли у многих пользователей Facebook ощущение дискомфорта, и, что важнее для выгоды компании, они не были полезны рекламодателям. К 2019 году Facebook пересмотрел свою категоризацию, сделав ее более ориентированной на конкретный продукт. Знакомства, воспитание, архитектура, ветераны

* «Брит-поп» — движение в британской музыкальной культуре 1990-х, возрождавшее гитарный стиль рок-музыки 1960-х.

войн, защита окружающей среды — вот некоторые из нескольких сотен категорий, которые использует компания для описания пользователей.

Одна из возможных реакций на такую типизацию — говорить, что все это неправильно, и кричать: «Я не наблюдение, не точка данных, я человек, личность!» Не хочу вас огорчать, но вы не так уникальны, как вам может показаться. Вас разоблачает то, как вы ведете себя в интернете. Есть и другие люди с тем же сочетанием интересов, что и у вас, с тем же любимым фильтром для фотографий, которые делают столько же селфи, что и вы, подписаны на тех же знаменитостей и щелкают по той же рекламе. По сути, это не один человек, а много людей, объединенных Facebook, Snapchat или другими используемыми приложениями.

Бессмысленно расстраиваться или сердиться на то, что вы наблюдение в матрице. Вы должны принять это. Чтобы понять почему, необходимо подумать о группировании людей с какой-нибудь другой точки зрения, об ином способе разделения на категории.

Представьте, что матрица M содержит не социальные интересы Мэдисон, Тайлера и других детей, а их гены. Современные генетики действительно рассматривают нас как отдельные наблюдения: матрицы из 1 и 0, которые указывают, есть ли у нас определенные гены. Взгляд на людей с помощью такой корреляционной матрицы спасает жизни. Этот подход позволяет ученым определять причину заболеваний, подбирать лекарства, подходящие для вашей ДНК, и лучше понимать развитие различных форм рака.

Это также помогает нам ответить на вопросы о наших предках. Ной Розенберг со своими коллегами из Стэнфордского университета построили матрицу из 4199 различных генов и 1056 людей со всего мира. Для любого из этих генов в исследовании найдутся минимум два человека, у которых они отличны. Это важно, поскольку у всех людей есть множество общих генов (они делают нас людьми). Розенберг специально искал различия между людьми и то, как на них влияет место рождения и проживания. Чем африканцы отличаются от европейцев? А выходцы из разных частей Европы? Можно ли объяснить различия в генах тем, что мы обычно называем расой?

В поисках ответа Ной сначала использовал уравнение 7, чтобы найти корреляцию между людьми в терминах общих генов⁸⁹. Затем он применил метод, который называется дисперсионным анализом (ANOVA, от Analysis of Variance), чтобы посмотреть, объясняются ли эти различия географическим местом происхождения. На такой вопрос не бывает ответов «да» или «нет». ANOVA дает ответ в виде определенного процентного соотношения — от 0 до 100 процентов. Хотите попробовать угадать, какая часть нашей генетики объясняется нашими предками? 98 процентов? 50? 30? 80?

Ответ — примерно 5–7 процентов. Не больше. Результаты Ноя подтверждены и другими исследованиями. Да, некоторые гены создают весьма заметные изменения между расами, самые яркие примеры — гены, регулирующие производство мелатонина и цвет кожи. Но когда дело доходит до категоризации людей, концепция расы приводит к путанице. Географическое происхождение наших предков не объясняет различия между нами.

Возможно, с моей стороны слегка снисходительно объяснять наивность расовой биологии в XXI веке, но, к сожалению, некоторые верят, что определенные расы по своей природе обладают, например, сниженным интеллектом. Эти люди расисты, и они ошибаются. Есть и другие типы — «Я не расист, но...» — которые считают, что признание равенства между расами нам навязано учителями или обществом. Одним из таких людей был ушедший на пенсию профессор, который писал для Quillette (см. главу 3). Они считают, что мы подавляем дискуссии о различиях между расами ради политкорректности.

По сути, место, откуда пришли наши предки, отвечает только за крохотную часть разнообразия наших генов. Более того, гены не полностью определяют нашу личность. Наши ценности и поведение формируются опытом и встреченными людьми. То, кем мы стали, имеет мало общего с биологической расой или нашими предками.

Люди, не достигшие двадцати лет, — как мои воображаемые тинейджеры Джейкоб, Алисса, Мэдисон и Райан, — представляют новое поколение Z, следующее за миллениалами. Для них крайне

важно быть личностями. Они явно не желают, чтобы на них смотрели с точки зрения гендера или сексуальной ориентации. Один из опросов 300 представителей поколения Z в США показал, что только 48% из них идентифицировали себя как абсолютных гетеросексуалов, а треть опрошенных предпочли отвести себе место на шкале бисексуальности⁹⁰. Более трех четвертей согласились с утверждением, что «гендер не определяет человека в той степени, как раньше». Можно слышать, как моя возрастная группа — поколение X — выражает скептицизм по поводу «отказа» поколения Z видеть гендерные различия. Опять же есть мнение, что представители Z пытаются быть политкорректными и при этом отрицают биологические факты.

Существует другой способ увидеть эту смену поколений. У поколения Z гораздо больше информации, чем было в молодости у моего поколения. «Иксы» росли в рамках ограниченного количества типов, предоставленных телевидением и умеренным личным опытом, а «зеты» засыпаны примерами разнообразия и индивидуальности.

Поколение Z считает эту индивидуальность более важной, чем поддержание гендерных стереотипов.

Успех рекламных категорий Facebook, основанный на корреляции наших интересов, предполагает, что взгляд поколения Z на мир статистически правильный. Дуглас Коэн, который работал в рекламной службе Facebook до перехода в Snapchat, рассказал мне, что рекламодатели его бывшего работодателя конкурируют друг с другом за гипертаргетинг своей рекламы по небольшим группам по интересам, определенным в матрице корреляций компании. Цена за охват целевой аудитории может удваиваться или утраиваться, поскольку рекламодатели конкурируют за право общаться непосредственно с любителями рукоделия, боевиков, серферами, игроками в онлайн-покер и многими другими группами. Для рекламодателей идентификация людей стоит больших денег.

Верная категоризация пользователей в соответствии с тем, что они реально любят, и деятельностью, которая им нравится, может быть крайне эффективной и справедливой. Корреляции помогут

нам найти группы с общими интересами и целями, так же как учёные используют корреляции между генами, чтобы установить причины заболеваний.

*

Парламент может быть страшным местом для молодого специалиста по данным.

— Прошло не так много времени с тех пор, как таблички в Вестминстере называли обычных граждан «посторонними», — сказала мне Николь Нисбетт, когда я встретился с ней в Лидсском университете. — Все меняется, и персонал Вестминстера активно обращается к общественности и исследователям, но эти таблички показывают историческую настороженность по отношению к посторонним.

Николь уже два года работает над диссертацией, проводя половину времени в Лидсе, а половину — в Палате общин. Теперь у нее есть пропуск с «доступом к большинству мест» в Палате. Ее задача — улучшить то, как члены парламента и остальной постоянный персонал прислушиваются к внешнему миру и взаимодействуют с ним. До того как Николь начала свой проект, многие сотрудники, занятые в повседневной работе правительства, ощущали, что все комментарии, оставленные публикой в Facebook или даже на их собственных дискуссионных форумах, были слишком обременительными, чтобы с ними работать.

— Оставалось также ощущение, будто они уже знали, что люди собираются им сказать, — сказала Николь, — и отсеивание всех негативных комментариев и оскорблений было трудной задачей.

Работа с данными дала Николь опыт и, соответственно, другую точку зрения. Она понимала, что количество комментариев в Twitter и Facebook может быть ошеломительным для любого, но также умела находить корреляции. Николь показала мне созданную ею карту, в которой резюмировались споры о запрете изделий из натурального меха. Она заложила в матрицу все слова, употребленные в дискуссии, и посмотрела, как коррелировалось их использование. Связями соединялись слова, которые

использовались вместе. «Мех» обнаруживался в одном кластере с «продажей», «торговлей», «промышленностью», которые далее соединялись с терминами «варварский» и «жестокий». Другой кластер объединял слова «страдание», «убийство» и «красивый». Третий связывал «благосостояние», «законы» и «стандарты». В каждом кластере объединялась определенная ветвь аргументации.

В одной области карты Николь рядом стояли слова «смерть от тока» и «аналльный». Их соединяла толстая линия. Я уставился на них, пытаясь понять, какой вывод от меня желает услышать Николь.

— Сначала мы думали, что эти слова используют тролли, — сказала она.

В любой дискуссии всегда найдутся люди, которые пытаются накручивать оппонентов, часто с помощью оскорбительной лексики. Однако в спорах с руганью, как правило, наблюдалась меньшая корреляция между словами — оскорбления были буквально случайны, — а эти два слова повторялись широким кругом разных людей. Когда Николь взглянула на фразы, их содержащие, она обнаружила информированную группу людей, обсуждавших реальную проблему: выращенных на ферме лисиц и енотов убивали, вставляя в тела электроды и подавая высокое напряжение. Это добавило новое измерение в обсуждение парламентского персонала, которое они никогда бы не заметили без работы Николь.

— Я избегаю предположений о том, что напишут люди; моя работа — объединить тысячи мнений, чтобы парламент мог быстрее реагировать на такую дискуссию, — говорит Николь.

В ее анализе собраны разные взгляды — не потому, что «политически корректно» выслушивать все стороны в споре, а потому, что статистически верно выделять важные мнения. Взгляды меньшинства получают свое выражение, потому что они реально способствуют разговору. Корреляции дают беспристрастное представление обо всех сторонах дискуссии, нам не нужно занимать какую-то политическую позицию, решать, какие взгляды стоит или не стоит рассматривать.

— Это детские шаги. Нельзя с помощью статистики решить всё, — заметила Николь. И рассмеялась: — Никакой анализ данных не сможет помочь с брекситом!

*

Социологи делают всё возможное, чтобы найти статистически верное объяснение данным. Я познакомился с Би Пуранен из Института исследований будущего в Стокгольме, когда мы вместе ехали в Санкт-Петербург на конференцию по политическим переменам. Исследователи в институте, который мы посетили, финансировались непосредственно из средств, выделенных Дмитрием Медведевым, который был президентом России, когда Владимир Путин занимал пост премьер-министра. Однако настроения молодых докторантов были направлены резко против истеблишмента. Они нуждались в демократических переменах и страстно рассказывали, как подавляют их взгляды. Я сам был свидетелем, как тщательно Би разбиралась в конфликтах, симпатизируя собеседникам, но принимая при этом реалии проведения исследовательского проекта в современной России.

Для Би было жизненно важно, чтобы российские исследователи, с которыми она работала, безотносительно к их политическим взглядам вели проект «Всемирный обзор ценностей» точно так же, как и во всех странах-участницах (их почти сто). Задавая людям со всего мира один и тот же набор вопросов, многие из которых относились к таким деликатным темам, как демократия, гомосексуальность, иммиграция и религия, Би и ее коллеги хотели понять, как меняются ценности в разных странах планеты⁹¹. Даже самый политически мотивированный исследователь быстро понимал: данные нужно собирать как можно более нейтрально.

В обзоре было 282 вопроса, поэтому корреляции давали удобный способ обобщать сходства и различия между ответами. Двое из коллег Би — Рональд Инглхарт и Кристиан Вельцель — обнаружили, что люди, делающие упор на семейные ценности, национальную гордость и религию, чаще имеют моральные возражения против разводов, абортов, эвтаназии и самоубийства. Корреляции в ответах на такие вопросы позволили Инглхарту и Вельцелю

классифицировать граждан разных стран по шкале традиционализм — светскость⁹². Такие страны, как Марокко, Пакистан или Нигерия, склонны к традиционализму, а Япония, Швеция и Болгария более светские. Этот результат ни в коем случае не означает, что все люди в какой-то стране придерживаются одинаковых взглядов; однако он дает статистически верную суммарную сводку мнений, доминирующих в этом государстве.

Крис Вельцель нашел еще один вариант корреляции вопросов и ответов. Люди, которых беспокоит свобода слова, тоже склонны ценить воображение, независимость и гендерное равенство в образовании, а также терпимы к гомосексуализму. Ответы на эти вопросы, которые Вельцель отнес к эмансипационным ценностям, коррелировали положительно. К странам с высокими эмансипационными ценностями относятся, например, Великобритания, США и Швеция.

Действительно важным моментом здесь оказалось то, что первая ось (традиционизм — светскость) не коррелировала со второй — эмансипацией. Например, в начале нового тысячелетия русские и болгары отличались высокими светскими ценностями, но не ценили эмансипацию. В США свобода и эмансипация важны почти для всех, но при этом страна остается традиционалистской в том смысле, что многие граждане отдают приоритет религиозным и семейным ценностям. Скандинавские страны — яркие примеры государств и со светскими, и с эмансипационными ценностями, а Зимбабве, Пакистан и Марокко находятся на противоположном конце спектра, где ценят и традиции, и послушание властям.

Разделение ценностей по двум независимым осям натолкнуло Би Пуранен на одну идею. Она захотела узнать, как изменились ценности мигрантов после их появления в Швеции. В 2015 году в стране попросили убежища 150 тысяч человек, в основном из Сирии, Ирака и Афганистана (это примерно 1,5% населения Швеции); и все они прибыли из трех стран с совершенно иным набором культурных ценностей по сравнению с ценностями их нового дома.

Когда уроженцы западных стран смотрят на этих иммигрантов, то часто замечают проявления, связанные с их традиционными ценностями (хиджабы, новые мечети). Такие наблюдения приводят

некоторых к выводам, что, например, мусульмане не могут приспособить свои ценности к новой родине. Внешний вид может демонстрировать, как иммигранты пытаются сохранить свои традиции, однако единственный статистически верный способ понять их внутренние ценности — поговорить с ними и спросить, что они думают. Именно этим занимались Би и ее коллеги. Они опросили 6501 из тех людей, которые приехали в Швецию за последние десять лет, задавая им вопросы об их ценностях.

Результаты были поразительными. Многие иммигранты разделяли типичное европейское стремление к гендерному равенству и терпимость к гомосексуальности, не принимая при этом крайнего шведского секуляризма. Они сохраняли свои традиционные ценности — те, что видны посторонним, — которые касаются важности семьи и религии. По сути, типичная иракская или сомалийская семья, живущая в Стокгольме, имеет очень схожие ценности с обычной американской семьей, проживающей в Техасе.

Мусульмане не единственная группа меньшинств, которую не воспринимают статистически верно. Я часто слышу, как в разговоре об американцах-христианах смешивают противников абортов и гомофобов. Мишель Дилон, профессор социологии из Университета Нью-Гэмпшира, показала, что одни религиозные группы, выступающие против абортов, одобряют однополые браки, а другие против них⁹³. В целом абORTы и права сексуальных меньшинств внутри религиозных групп рассматриваются как разные вопросы.

*

По мере того как наша жизнь все больше перемещается в интернет, доступные данные о нас также расширяются: с кем взаимодействуем в Facebook, что нам нравится, куда ходим, что покупаем — и список можно продолжать. Facebook, Google и Amazon сохраняют все социальные взаимодействия, все запросы и потребительские решения. Это мир больших данных. Нас теперь определяют не возраст, гендер или место рождения, а миллионы наблюдений, фиксирующих все наши действия и мысли.

«Десятка» быстро приступила к проблеме больших данных. Ее участники составили матрицу населения мира. Они объединяли

людей по интересам. Им казалось, будто расизм и сексизм ушли в прошлое. Они измеряли, как общество развивается, чтобы стать более толерантным: справедливым и уважающим личность людей. «Десятка» действовала статистически правильно.

Новый порядок во многом финансировался рекламой, адаптированной под конкретного человека. Рекламодатели начали тендерные войны за право показывать свои продукты небольшим целевым группам пользователей Facebook. Чтобы доставлять информацию с хорошей точностью, стали привлекать других статистиков и специалистов по данным. Родилась новая область — реклама с микротаргетированием. Сначала выявляются потенциальные клиенты и им подается только нужная информация в нужное время — чтобы максимизировать их заинтересованность.

Участники «Десятки» снова победили, добавив к списку решенных ими задач рекламу и маркетинг. На этот раз даже кажется, что отчасти мораль на их стороне. Но возникла одна проблема. На числа в этой матрице смотрели не только участники «Десятки». И не все люди, видевшие ее корреляции, правильно понимали наблюдаемые закономерности...

*

Исследование Ани Ламбрехт посвящено правильной интерпретации больших данных. В качестве профессора маркетинга в Лондонской школе бизнеса она изучала, как данные используются повсюду — от брендовой одежды до спортивных сайтов. Аня объяснила мне в электронном письме, что, несмотря на очевидные выгоды от использования больших массивов данных в рекламе, важно учитывать и ограничения. «Данные без умения извлекать надлежащие выводы не особо полезны», — сказала она.

В одной из своих научных статей, из которой взят следующий пример, Ламбрехт с коллегой Кэтрин Такер объясняет возникающие проблемы, используя сценарий из торговли в интернете⁹⁴. Представьте себе розничного продавца игрушек, обнаруживающего, что пользователи, которые видят больше его рекламы в Сети, покупают у него активнее. При этом он установил корреляцию между рекламой и приобретением игрушек, в итоге его

маркетинговый отдел «больших данных» пришел к выводу, что кампания работает.

А теперь посмотрим на эту рекламу под другим углом. Возьмем Эмму и Джулию: они незнакомы друг с другом, но у них есть семилетние племянницы. Бродя в Сети в последнее воскресенье перед Рождеством, они, независимо друг от друга, видят рекламу компании игрушек. Эмме предстоит напряженная рабочая неделя, у нее нет времени на магазин. Джулия в отпуске и большую часть свободного времени тратит на поиски рождественских подарков. Увидев три-четыре раза рекламу игры «Четыре в ряд», Джулия переходит по ссылке и решает приобрести ее. Эмма же отправляется в какой-то магазин днем 23 декабря, видит модель дома на колесах фирмы Lego и покупает ее.

Джулия видела рекламу гораздо чаще Эммы, но можно ли сказать, что та эффективна? Нет. Мы понятия не имеем, что делала бы Эмма, будь у нее время смотреть рекламу. Те специалисты по «большим данным», которые заключили, что их кампания действует, перепутали корреляцию с причинностью. Мы не знаем, реклама ли заставила Джулию купить игру «Четыре в ряд», и поэтому не можем сделать вывод об ее эффективности.

Разделить корреляцию и причинно-следственную связь сложно. Та корреляционная матрица, которую я заполнил выше для Мэдисон, Райана и их друзей, основана на очень малом числе наблюдений, и мы не можем делать глобальных выводов только из нее (помните уравнение уверенности?). Однако представьте, что такую матрицу построили по очень большому количеству пользователей Snapchat, и мы обнаружили, что Пьюдипай коррелирует с Fortnite. Можно ли сделать вывод, что рост количества подписчиков Пьюдипая увеличит число игроков в Fortnite? Нет. Это снова смешивание корреляции и причинности. Дети не играют в Fortnite из-за того, что смотрят youtube-канал Пьюдипая. Если какая-то кампания по увеличению подписчиков Пьюдипая сработает, то она увеличит количество времени, которое дети будут тратить на его ролики. Она не заставит их играть в Fortnite после просмотра его видео.

А если Fortnite покупает рекламное место на канале Пьюдипая? Это может сработать: например, если некоторые игроки в Fortnite

вернулись к Minecraft, Пьюдипай может привлечь их обратно. Но может ничего и не получиться. Не исключено, что интерес к Fortnite среди аудитории Пьюдипая уже находится в точке насыщения и нужно, чтобы этой игрой увлеклась Кайли Дженнер!

Если мы немного подумаем, то увидим разного рода проблемы с потенциальными выводами из наличия в наших данных корреляции Пьюдипай/Fortnite. Но когда началась революция больших данных, многие из них игнорировались. Компаниям рассказывали, что их данные весьма ценные, поскольку теперь они знают о своих пользователях всё. Но это не так.

*

Cambridge Analytica — яркий пример компании, которая не сумела разобраться с причинно-следственной связью.

Комитет Сената внимательно слушал то, что я им говорил по скайпу. «Cambridge Analytica собрала массу сведений о пользователях Facebook, в частности о продуктах и сайтах, где те щелкали кнопку “Нравится”. Они намеревались использовать данные для целевого информирования в соответствии с личностью пользователей. Они хотели, чтобы нервным людям показывали сообщения о защите семьи с помощью оружия, а традиционалистам рассказывали о передаче оружия от отца к сыну. Каждое рекламное сообщение должно было быть адаптировано под конкретного избирателя».

Я понимал, что мои собеседники — республиканцы из комитета, — вполне могли вообразить себе выгоды от такого инструмента для следующих выборов. А потому быстро перешел к сути. «Но по нескольким причинам это не могло сработать, — сказал я. — Во-первых, невозможно надежно определить свойства личностей по их лайкам. Таргетирование приводило к ошибкам в определении личностных качеств ненамного реже, чем давало правильные результаты. Во-вторых, тот тип невротизма, который можно найти у пользователей Facebook — любителей группы Nirvana и стиля эмо, — отличается от невротизма, связанного с защитой семьи с помощью оружия».

Я прошелся по проблемам, появляющимся из-за смешивания корреляционной зависимости и причинно-следственной связи.

Когда Cambridge Analytica создавала свой алгоритм, выборы еще не состоялись. Как же они могли проверить, работает ли их реклама?

Далее я рассказал о неэффективности фейковых новостей для влияния на избирателей — еще одной теме, которую исследовал в своей предыдущей книге «В меньшинстве»⁹⁵. Я также рассказал им, что, вопреки теории эхо-камеры*, демократы и республиканцы на выборах 2016 года должны были слышать все стороны. Моя точка зрения противоречила взглядам либеральных СМИ, которые считали победу Трампа манипуляцией с онлайн-избирателями. Его избиратели обвиняли в наивности и в том, что эти люди оказались жертвами промывки мозгов. Cambridge Analytica стала олицетворением легкости, с которой соцсети могут влиять на общественное мнение. Я не разделял такую точку зрения.

Координатор звонка сказал: «Сейчас я отключу у вас звук, пока мы обсудим то, что услышали».

Им потребовалось секунд тридцать, чтобы принять решение. «Мы хотели бы, чтобы вы прилетели в Вашингтон для дачи показаний в комитете Сената. Сможете?»

Я не ответил сразу. Я пробормотал что-то о запланированном отпуске и сказал, что мне надо подумать.

В тот момент я действительно не был уверен, стоит ли мне ехать. Но, хорошенько все обдумав, пришел к выводу, что не стоит: мой приезд в США был им нужен не для того, чтобы я объяснял сенаторам причинность и корреляцию. Они не хотели понять используемые мною модели. Они лишь хотели услышать те мои выводы, которые соответствовали их представлениям — что Cambridge Analytica и фейковые новости не сделали Трампа президентом. И я не поехал.

Однако тем летом я оказался в США. Я был в Нью-Йорке и встретился с Алексом Коганом сразу после того, как он давал показания на слушаниях в Сенате. Алекс, исследователь из Кембриджа

* Эхо-камера — ситуация, когда человек окружен людьми, взгляды которых совпадают с его собственными (например, внутри субкультуры или среди друзей в соцсети). В этом закрытом пузыре общие идеи ходят по кругу, все друг с другом соглашаются, а альтернативная информация внутрь не попадает.

ского университета, считался одним из плохих парней в истории Cambridge Analytica. Он собрал данные по 50 миллионам пользователей Facebook и продал их Cambridge Analytica. Не особо мудрый поступок, о котором он потом сожалел.

Мы познакомились с Алексом, когда я начал проверять точность методов Cambridge Analytica. Он интересный собеседник. Возможно, с ним не стоит вести бизнес, но он отлично понимает, как можно и как нельзя использовать данные. Алекс действительно пытался создать то, что Крис Уайли назвал инструментом «психологической войны», для точной фиксации избирателей, но пришел к выводу, что такое оружие невозможно разработать. Данных недостаточно.

Работая в компании, он пришел к тому же выводу о деятельности Cambridge Analytica, что и я: «Эта дрянь не работает». На слушаниях в Сенате он сказал сенаторам то же, но в более вежливой форме.

*

Основная «проблема» с алгоритмом компании Cambridge Analytica состояла в том, что он не работал.

В начале эры «больших данных» многие так называемые эксперты предполагали, что корреляционные матрицы могут вести непосредственно к лучшему пониманию пользователей и клиентов. Но все не так просто. Алгоритмы, основанные на корреляциях в данных, использовались не только для политической рекламы, но и при рекомендациях для тюремного заключения, оценке работ школьных учителей и для обнаружения террористов. Название книги Кэти О'Нил «Оружие математического поражения»* хорошо отражает возникающие проблемы⁹⁶. Как и ядерные бомбы, алгоритмы не избирательны. Термин «таргетированная реклама» подразумевает жесткий контроль над тем, кому она показывается, но на самом деле эти методы имеют очень ограниченные возможности по надлежащей классификации людей.

* В оригинале использован каламбур — mass и math. Русский перевод вышел под названием: «Убийственные большие данные. Как математика превратилась в оружие массового поражения». М. : ACT, 2017.

Для интернет-рекламы это невелика беда. Жизнь игрока в Fortnite не рухнет, если ему покажут рекламу косметики. Но совсем другое дело — если алгоритм назовет вас преступником, плохим учителем или террористом. Это может изменить карьеру и жизнь. Алгоритмы на базе корреляций изображались объективными, потому что основаны на данных. На деле же, как я обнаружил при написании своей последней книги — «В меньшинстве», — многие алгоритмы делают почти столько же ошибок, сколько и точных прогнозов.

Нашлось и множество других проблем, которые могут возникнуть при конструировании алгоритмов, основанных на корреляционных матрицах. Например, метод представления слов Google в его поисковой системе и сервисе перевода основан на корреляции между использованием слов⁹⁷. Для определения того, когда те или иные группы слов употребляются вместе, применяются также Википедия и базы данных для новостных статей⁹⁸. Когда я посмотрел, как эти алгоритмы рассматривают мое имя Дэвид по сравнению с именем Сьюзен (самым популярным именем для женщин моего возраста в Великобритании), я пришел к нелестным выводам. Там, где я в качестве «Дэвида» был «смышенным», «сообразительным» и «умным», алгоритм давал для Сьюзен определения «изворотливая», «жеманная» и «сексуальная». Основная причина проблемы в том, что эти алгоритмы построены на корреляциях в наших исторических текстах, которые набиты стереотипами.

Алгоритмы, используемые для больших данных, нашли корреляции, но не поняли их причин. В результате они совершали колossalные ошибки.

*

Последствия перехваливания «больших данных» были сложны, но причины просты. Помните, как мы разделили мир на данные, модели и бессмыслицу? Произошло вот что: компаниям и общественности рассказали о данных, но не обсудили модели. Когда моделей нет, верх берет бессмыслица. Александр Никс и Крис Уайли говорили бессмыслицу о таргетировании по личностным качествам и инструментам психологической войны. Компании,

прогнозирующие качество работы учителей и создающие программы для вынесения приговоров, говорили бессмыслицу об эффективности их продуктов. Facebook подкреплял ложные стереотипы своей рекламой, ориентированной на этническую принадлежность⁹⁹.

У Ани Ламбрехт есть ответ. Она решает проблему причинности, вводя модель — создавая историю, например, Эммы и Джулии с их способом совершения покупок. Мы можем оценить успех рекламной кампании, если рассмотрим точку зрения потребителей, а не просто взглянем на собранные данные. Ламбрехт разделяет проблему на модель и данные (хотя и не использует таких терминов) — именно эту стратегию мы применяли в данной книге. Сами по себе данные говорят очень мало, но при их объединении с моделью можно получить многое.

Такой базовый подход к определению причинности известен как А/В-тестирование. Я уже описывал этот метод в главе 1, а сейчас мы можем применить его на практике. Компания должна попробовать на своих клиентах две разные рекламы: А) оригинальную, эффективность которой нужно проверить, и В) контрольную, например благотворительной организации, где не будет никаких отсылок к компании игрушек. Если компания продает столько же товаров пользователем, которые смотрят благотворительную рекламу, сколько тем, кто видел оригинал, то она будет знать, что ее реклама не оказывает никакого эффекта.

Исследование Ани Ламбрехт описывает множество примеров, как нам стоит подходить к причинности. В одной работе она изучала распространенную в рекламном деле идею, что если на ранней стадии привлечь внимание известных людей в соцсетях, то это может помочь продукту стать вирусным¹⁰⁰. Если рекламодатель нацелен на людей, которые быстро принимают новые тренды, реклама должна иметь больший эффект. Звучит разумно — не так ли?

Чтобы проверить надлежащим образом эту идею, Ламбрехт и ее коллеги сравнили: а) группу пользователей, которые на ранней стадии делились хештегами свежих трендов, например #RIPNelsonMandela и #Rewind2013, и б) группу пользователей, которые присоединялись к тем же трендам позже. Обеим группам

показывали ссылку на рекламу, и исследователи измеряли, как часто люди нажимали на рекламу или делали ретвит (делились ею).

Оказалось, что теория «авторитетов на ранней стадии» неверна. Группа А с меньшей вероятностью щелкала по рекламе или делилась ею по сравнению с группой В. Результаты были одинаковы и для рекламы благотворительной организации для бездомных, и для модного бренда. На самих авторитетных людей трудно повлиять. Они авторитетны, потому что разбираются в том, чем делятся. На них подписываются из-за их независимости и здравомыслия. Тот, кто быстро принимает что-то без применения здравого смысла, — просто спамер, и никто не желает подписываться на таких.

В Snapchat Дуг Коэн со своей командой применяют А/В-тестирование ко всему. Когда я разговаривал с ним, они работали над уведомлениями, проверяя сложный ряд схем и пытаясь выяснить, какая из них побуждает пользователей открывать приложение. Но когда Дуг говорил о том, хорошо ли они понимают своих пользователей, он был осторожен: «Утром вы не тот человек, что в конце дня, поэтому мы можем отнести вас к более широкой категории, но вы меняетесь в течение месяца и года, становясь старше». Он также подчеркнул, что люди не хотят видеть одно и то же все время: «Мы можем классифицировать человека как интересующегося спортом, но это не значит, что он хочет видеть кучу мужских вещей». Пользователи раздражаются, если считают, что алгоритм приписал их к какой-то определенной категории.

Уравнение рекламы говорит нам, что неизбежным следствием организации больших объемов данных становится некоторая степень стереотипизации. Так что не огорчайтесь из-за того, что стали частью корреляционной матрицы. Это верное представление о том, кто вы есть. Ищите корреляцию с интересами ваших друзей и стройте связи. Когда корреляции истинны (а не построены на расовых или гендерных стереотипах), они облегчают поиск общих точек. Если получаются исключения из правила, примите их и скорректируйте свою модель. Ищите закономерности в разговорах, как Николь Нисбетт в политических дискуссиях, и используйте их для упрощения споров. Внимательно ищите небольшие

кластеры новых точек зрения и обращайте на них особое внимание. Но не смешивайте корреляцию и причинно-следственную связь. Приглашая на обед друзей, проведите А/В-тестирование меню. Не предлагайте снова пиццу только потому, что в прошлый раз, по словам приятелей, она оказалась вкусной. Стройте статистически правильную модель своего мира.

*

В манхэттенском метро после разговора с Алексом Коганом в Нью-Йорке я понял кое-что о себе. Я больше не знал своей политической позиции. Я всегда был левым; прочтя в девятнадцать лет Жермен Грир, стал феминистом. Я всегда был тем, кого сейчас называют термином «воук» («активный, пробудившийся»). Как минимум настолько, насколько это может сделать белый мальчик, выросший в маленьком рабочем поселке в 1980-х. Многие считают, что быть левым, не одобрять Дональда Трампа и обвинять его в манипулировании людьми с помощью соцсетей — части единого пакета. СМИ изображали их коррелированными признаками левого взгляда на мир.

Однако математика заставила меня выбрать другую модель. Я был вынужден принять победу Трампа, поскольку модель говорила: победа заслуженная. Я не одобрял его политических взглядов, но не поддерживал и того, что его избирателей рисовали тупыми и легко управляемыми. Они такими не были.

Вместо того чтобы искать реальные причины роста националистических настроений, которые привели Трампа к власти, спровоцировали брексит, несли ответственность за «Движение пяти звезд» в Италии, Виктора Орбана в Венгрии и Жайра Болсонару в Бразилии, все бросились искать какого-то злодея из бондианы — некую личность, отравившую воды политики. Их Доктор Но явился в образе Александра Никса и компании Cambridge Analytica. Каким-то образом этот человек, обладающий всего лишь базовым пониманием моделей и данных, на их взгляд, должен был манипулировать всей современной демократией.

Самая большая угроза для любого тайного общества — угроза его раскрытия. Компанию Cambridge Analytica разоблачили.

Оказалось, что заурядной рекламной фирме Трамп заплатил максимум миллион долларов — и это во время выборов, на которые было потрачено 2,4 миллиарда. Эффект оказался примерно пропорционален размеру вклада: мизерный.

Конспиративной теории с бондовским злодеем не хватило здравого смысла. Данные были слишком скучными, чтобы дать нам хоть какую-то уверенность. А «Десятка» продолжала работать незамеченной. Ее участники управляют банками, финансовыми учреждениями и букмекерскими конторами. Они создают наши технологии и контролируют соцсети. В каждой такой деятельности они получают небольшую долю: 2–3 цента на доллар при азартных играх, цент на доллар в сетевой торговле и даже меньше, когда выдают рекламу в сочетании с результатами поиска в интернете. Со временем эти маленькие доли складываются, и прибыль участников «Десятки» растет. Во всех сферах жизни математики переигрывают тех, кто не знает уравнений.

Сидя в метро по дороге домой, я думал о людях, которые умоляли меня о футбольных советах и о том, как компании, занимающиеся онлайн-ставками, предлагают одиноким мужчинам возможность пообщаться с женщинами во время проигрывания своих денег. Я думал о призме Instagram, через которую мы смотрим на образ жизни, основанный на потребительстве и известности. Я думал о рекламе кредитов на мобильные телефоны с высокими процентами, ориентированной на самые бедные слои общества.

Люди, которые объясняли результаты президентских выборов в США и брексит фейковыми новостями и применением данных, украшенных из Facebook компанией Cambridge Analytica, игнорируют базовые противоречия в нашем обществе между богатыми и бедными. Подобные мне люди — математики и другие ученые — тоже сыграли значительную роль в том, что общество стало менее справедливым. Ирония здесь в том, что одна теория заговора верна. Иллюминаты действительно существуют — это «Десятка». И это общество настолько скрыто, что его не видят даже другие участники-конспираторы.

Глава 8

Уравнение вознаграждения

$$Q_{t+1} = (1 - \alpha)Q_t + \alpha R_t$$

В течение пятнадцати лет в начале своей карьеры я исследовал, как животные ищут и собирают вознаграждение. Я не принимал какого-то решения, что должен этим заниматься, — просто так уж меня угораздило.

Два друга-биолога — Эймон и Стивен — как-то спросили меня, не хотел бы я съездить на день на Портленд-Билл (узкий полуостров на южном побережье Англии). Им нужны были муравьи. Эймон показывал мне, как аккуратно раскрыть небольшую трещину в скале, где насекомые могут прятаться. Казалось, всякий раз он оказывался прав, находя муравьев под выбранными камнями. Он быстро всасывал их импровизированным пылесосом в пробирку, которую нужно было взять в лабораторию.

В конце концов этого добился и я — встав чуть дальше от моих коллег под вздымающимся маяком. Это было приятное удовлетворение — втягивать колонию муравьев через оранжевую пластиковую трубку. Мы потратили пять лет на изучение того, как эти насекомые выбирают новое место для своего дома. Я создавал модели; они собирали данные.

Аналогичную цель имели прогулки по йоркширским болотам с Мадлен, которая тогда была в постдокторантуре и занималась биологией в Шеффилде: сюда летали медоносные пчелы от своего улья, чтобы собрать пыльцу с вереска. Здесь наши головы очистились от густого офисного воздуха, и я пытался связать мои

уравнения с ее описаниями того, как пчелы и муравьи сообщают друг другу о пище. Мы работали вместе более десяти лет, наблюдая за тем, как разные виды общественных насекомых определяют, какие источники пищи использовать.

Многие обсуждения проходили в менее плenительных местах. Дора — тогда аспирантка в Оксфорде и первый мой друг после переезда сюда — рассказывала мне о своих голубях, когда мы сидели на холодной ступеньке у фургона с кебабами. Через несколько дней мы в кафе Jericho изучали результаты отслеживания ее птиц с помощью системы GPS. Через год вносили финальные штрихи в работу о том, как пары птиц меняют маршруты на пути домой.

Эшли тщательно конструировал Y-образные лабиринты для колюшек (мелких рыб). Я встретил его в пабе с Иэном, и мы обсудили, как можем моделировать их решения в группах. Вместе мы смотрели, как они избегают хищников и следуют друг за другом к еде.

Далее мой путь лежал за пределы Англии. Большеголовые муравьи в Австралии с Одри. Аргентинские муравьи с Крисом и Таней. Муравьи-листорезы на Кубе с Эрнесто. Домовые воробыи на юге Франции с Мишелем. Саранча из Сахары с Джеромом и Иэном. Слизевики в Японии с Тоси (а также с Одри и Таней). Цикады в Сиднее с Тедди.

Все мои коллеги тех лет теперь профессора в университетах по всему миру. Но это никогда не было нашей единственной целью. Мы люди, которые беседуют друг с другом, учатся друг у друга и решают проблемы совместно. Мы получали удовлетворение, отвечая на вопросы, и постепенно всё лучше понимали мир природы. Через пятнадцать лет я знал почти все о том, как животные принимают решения в группах. Тогда в моей голове еще не было ясности, однако, оглядываясь сейчас назад, понимаю, что практически за всем, что я тогда делал, стоит одно уравнение.

*

Чтобы выживать, животным нужны еда и убежище. И партнер для воспроизведения.

В основе этих трех требований для жизни лежит нечто еще более фундаментальное, что должны получать животные: информация.

Они собирают информацию о еде, убежище и половой жизни исходя из собственного опыта и опыта других особей. Затем применяют ее для выживания и воспроизведения.

Муравьи — один из моих любимых примеров. Многие из них используют химический маркер-феромон, чтобы показать другим насекомым, где они были. Когда они находят лежащую на земле сладкую пищу, то оставляют метку. Другие муравьи ищут ее и следуют за ней до еды. В итоге срабатывает механизм обратной связи: все больше муравьев оставляют свои феромоны, и находить пищу можно все быстрее.

Людям тоже нужны пища, кров и партнер для воспроизведения. В прошлом мы тратили уйму времени на поиск информации, которая позволила бы нам получить и сохранить три этих важнейших элемента. В современном обществе такой поиск изменил форму. Для значительной части населения мира поиск предметов первой необходимости уже завершен, однако поиск информации о еде, жилье и сексе продолжается и расширяется: теперь он принимает форму просмотра кулинарных передач и реалити-шоу *Love Island* («Остров любви»); чтения сплетен о знаменитостях; изучения выставленного на продажу жилья и цен на недвижимость. Мы публикуем фотографии наших партнеров, обеда, детей и домов. Показываем всем, куда едем и что делаем. Как муравьи, мы стремимся поделиться тем, что нашли, и следовать полученным советам.

Мне неудобно признать масштаб своих ежедневных поисков информации. Я захожу в Twitter проверять уведомления; открываю почту в поисках новых писем; читаю политические новости, а затем пробегаю по спортивным. Я иду на онлайн-платформу Medium, где публикую свои тексты — чтобы увидеть, нравятся ли мои истории и нет ли там интересных комментариев.

Математический способ интерпретации моего поведения возвращает нас к игровым автоматам, упомянутым в главе 3. Каждое приложение в телефоне — все равно что дергание ручки автомата в надежде получить вознаграждение. Я тяну ручку Twitter: семь ретвитов! Теперь электронной почты: письмо с приглашением выступить. Класс, я популярен! Я тяну ручку новостей и спорта:

очередная заморочка брексита или слух о трансфере. Заглядываю на платформу Medium — но никто не поставил лайк моим постам. О боже, эта ручка плохо работает.

Сейчас я представляю такую жизнь с приложениями — игровыми автоматами в виде уравнения. Вообразите, что я открываю Twitter раз в час. Вероятно, это заниженная величина, но начнем с простого предположения.

Обозначим как R_t вознаграждение, которое я получаю за час t . Для простоты скажем, что $R_t = 1$, если кто-то сделал ретвит моего поста, и $R_t = 0$, если не было ни одного ретвита.

Представим вознаграждения за рабочий день с 9 до 17 часов в виде последовательности единиц и нулей. Например, она может выглядеть так:

$$R_9 = 0, R_{10} = 1, R_{11} = 1, R_{12} = 0, R_{13} = 1, R_{14} = 1, R_{15} = 0, R_{16} = 1, R_{17} = 1.$$

Эти вознаграждения моделируют ретвิตы внешнего мира.

Теперь нужно учесть мое внутреннее состояние. Обращаясь к этому приложению, я улучшаю свою оценку качества Twitter, его способности дать мне мгновенное ощущение самоутверждения, которое предоставляет только ретвит или лайк. Здесь мы можем использовать уравнение вознаграждения:

$$Q_{t+1} = (1 - \alpha)Q_t + \alpha R_t \quad (\text{Уравнение 8}).$$

Кроме времени t и вознаграждения R_t сюда входят еще два символа: Q_t отражает мою оценку качества вознаграждения, а α определяет, насколько быстро я теряю уверенность при его отсутствии. Эти буквы требуют дополнительного пояснения.

Если я пишу: $Q_{t+1} = Q_t + 1$, это означает, что я увеличил Q_t на единицу. Эта идея используется в программировании внутри «цикла со счетчиком»: мы увеличиваем Q_t на 1 каждый раз, когда проходим цикл. Та же идея применяется и в уравнении вознаграждения. Но вместо прибавления 1 мы изменяем Q_t , добавляя два разных слагаемых. Первый компонент — $(1 - \alpha)Q_t$ — понижает оценку качества вознаграждения. Например, если мы выберем $\alpha = 0,1$,

на каждом шаге наша оценка будет снижаться на $1 - 0,1 = 90\%$ по сравнению с предыдущим уровнем. Это то же уравнение, которое мы сейчас используем, например, для описания того, как автомобиль каждый год падает в цене; далее мы увидим, как оно описывает испарение феромонов и других химических веществ. Второй компонент — αR_t — повышает нашу оценку стоимости вознаграждения. Если вознаграждение равно 1, добавляем α к нашей оценке.

Сложив оба компонента, можем увидеть, как работает уравнение в целом. Представьте, что я начинаю работу утром в 9 часов с оценкой $Q_9 = 1$. Иными словами, я на 100% уверен, что Twitter даст мне вознаграждающий ретвит. Открываю его, но с разочарованием обнаруживаю, что $R_9 = 0$. Нет ретвитов. Нет вознаграждения. И я использую уравнение 8, чтобы изменить мою оценку качества на $Q_{10} = 0,9 \cdot 1 + 0,1 \cdot 0 = 0,9$. Теперь я немного меньше уверен, когда открываю Twitter в 10 часов утра, однако на этот раз получаю то, что искал: $R_{10} = 1$. Ретвит! Моя оценка качества не возвращается к исходному состоянию, но чутьдвигается вверх: $Q_{11} = 0,9 \cdot 0,1 + 0,1 \cdot 1 = 0,91$.

В 1951 году математики Герберт Роббинс и Саттон Монро доказали, что уравнение 8 всегда дает верную оценку среднего значения вознаграждения¹⁰¹. Чтобы понять этот результат, предположим, что вероятность получения вознаграждения (ретвита) за любой конкретный час равна \bar{R} , и пусть $\bar{R} = 0,6$, или 60%. Перед тем как начать ежечасную проверку Twitter, я понятия не имел о значении \bar{R} . Моя цель — оценить значение этой величины по последовательности вознаграждений, которые я получаю после открытия Twitter. Они у нас представлены в виде последовательности из 0 и 1 — 011001011... Если та продолжается бесконечно, средняя частота единиц будет $\bar{R} = 60\%$.

Уравнение 8 быстро начинает отражать вознаграждения: $R_{11} = 0$, и поэтому $Q_{12} = 0,919$; $R_{12} = 0$, и поэтому $Q_{13} = 0,827$ и т. д., так что к концу дня мы получаем $Q_{47} = 0,724$. Каждое наблюдение приближает меня к истинному значению \bar{R} . По этой причине Q_t часто называют отслеживающей переменной: она отслеживает значение \bar{R} . Рисунок 8 иллюстрирует этот процесс.

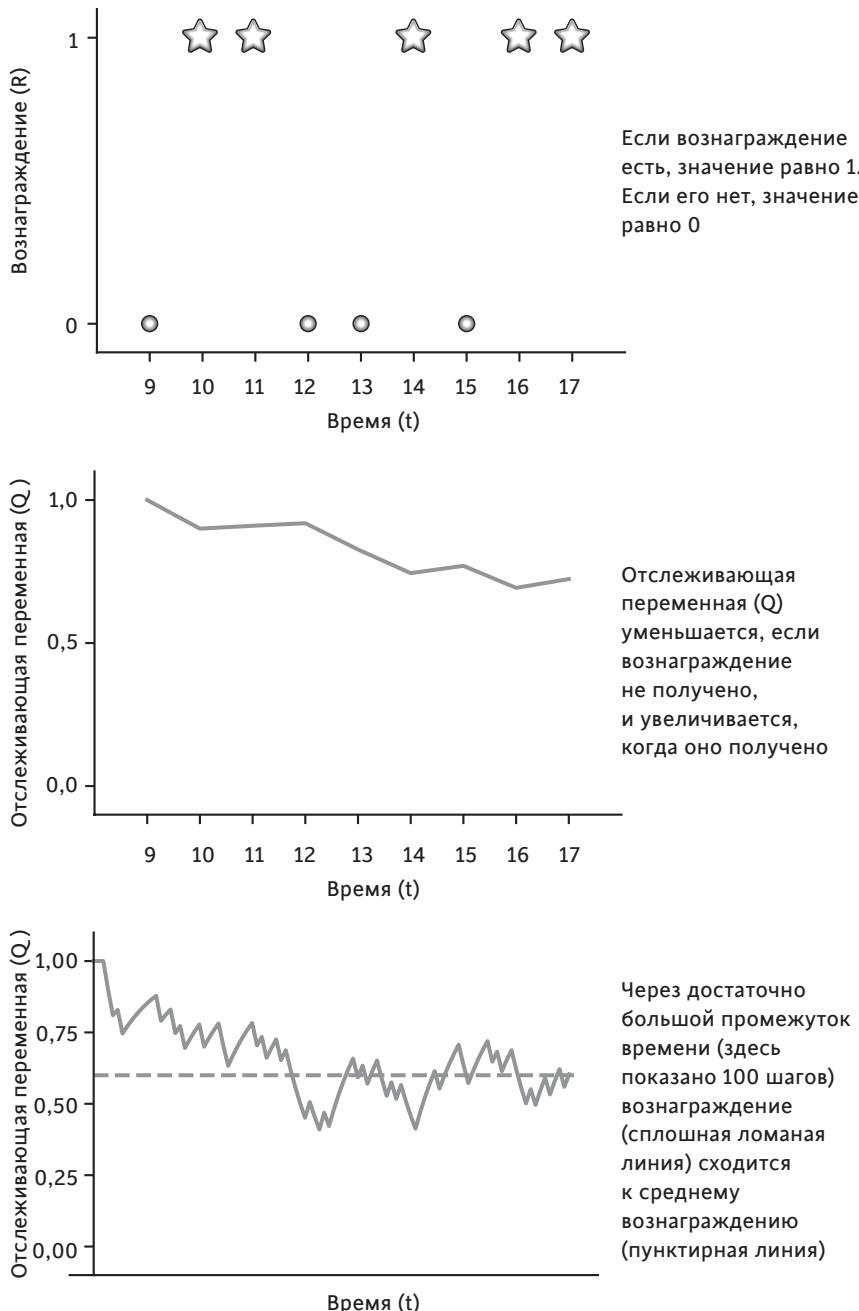


Рис. 8. Как отслеживающая переменная отслеживает вознаграждение

Роббинс и Монро показали, что для надежной оценки \bar{R} не нужно хранить всю последовательность нулей и единиц. Чтобы получить очередную оценку Q_{t+1} , надо знать текущую оценку Q_t и следующее вознаграждение в последовательности R_t . Если я все вычислил правильно вплоть до этого момента, то могу забыть о прошлом и сохранять только отслеживающую переменную.

Есть оговорки. Роббинс и Монро показали, что нам нужно очень медленно уменьшать со временем значение α . Помните, что α (греческая буква) — параметр, которые управляет скоростью забывания. Изначально у нас доверия нет, поэтому нужно уделять много внимания последним величинам, и поэтому α получает значение, близкое к 1. Со временем нам нужно понижать α , так что эта величина стремится к 0. Именно медленное уменьшение гарантирует, что наша оценка сходится к вознаграждению.

*

Представьте, что вы лежите на диване и вознаграждаете себя просмотром телевизора. На экране какой-то сериал Netflix. Первая серия — отличная (как всегда), вторая — средняя, третья — чуть лучше. Вопрос таков: сколько времени вам следует смотреть, прежде чем бросить сериал? Вашему мозгу это не особо важно, но вас это заботит. Вы хотите смотреть в выходной что-то хорошее.

Решение — использовать уравнение вознаграждения. Для телесериала хорошим значением для нашего показателя снижения доверия будет $\alpha = 0,5$, или половина. Это очень высокая скорость забывания прошлого, но хорошее шоу должно постоянно дарить новые идеи.

Вот ваши действия. Вы ставите первому эпизоду оценку по 10-балльной шкале — скажем, 9. Итак, $Q_1 = 9$. Если смотрите серии подряд, то держите в голове число 9 и начните следующую серию. Поставьте ей оценку. Предположим, это 6. Теперь имеем $Q_2 = 9/2 + 6/2 = 7,5$. Удобно каждый раз округлять, так что новая оценка будет 8. Смотрим следующий эпизод. Пусть на этот раз мы ставим 7. Берем $Q_3 = 8/2 + 7/2 = 7,5$, снова округляем до 8.

Продолжаем в том же духе и дальше. Сила этого метода в том, что не нужно помнить, насколько сильно вам понравились предыдущие

эпизоды. Вы отмечаете Q_t для последней серии в голове. Сохранять отслеживающую переменную Q_t можно не только при просмотре телесериала, но и при оценке того, нравится ли вам ходить на разные мероприятия, читать различных авторов или заниматься в классе йоги. Это единственное число для каждого занятия позволяет понимать общее вознаграждение за различную деятельность, не возвращаясь к конкретным моментам, когда вас втянули в разговор с одним скучным математиком во время выпивки после работы или когда вы повредили седалищный нерв во время йоги.

Когда бросать просмотр? Чтобы ответить на этот вопрос, нужно установить личный порог. Я использую число 7. Если качество серий падает до 7, я останавливаюсь. Это довольно жесткое правило, поскольку оно означает, что если текущий уровень 8, а следующий эпизод получает оценку 6, то у меня получается $8/2 + 6/2 = 7$ и я вынужден бросить просмотр. Но мне кажется, что это справедливо. Хороший сериал должен регулярно выдавать эпизоды уровня 8, 9 и 10. Если он достигает таких высот, то переживет оценку 6 и даже 5. Например, если текущее значение $Q_t = 10$ и я вижу серию, которая тянет только на 5, то $Q_{t+1} = 10/2 + 5/2 = 7,5$, округляю до 8 и могу смотреть дальше. Однако для продолжения просмотра следующий эпизод должен оказаться хорошим. На основании этого правила я посмотрел три с половиной сезона «Форс-мажоров», два сезона «Большой маленькой лжи», полтора сезона «Рассказа служанки» и два эпизода сериала «Ты».

*

В большинстве компьютерных игр для отслеживания ваших успехов используется только одно число — набранные очки или уровень. Очки подобны Q_t в уравнении вознаграждения: они отслеживают ваши вознаграждения. Вы выбираете, что делать дальше: каким маршрутом поедете в гоночной игре Mario Kart, какого противника хотите выследить и убить в Fortnite, какой ряд переместите в игре «2048», на какую арену отправитесь за покемоном в игре Pokémon Go — и ваш счет меняется в зависимости от ваших успехов.

Ваш мозг действует аналогично. Химическое вещество дофамин часто называют элементом системы вознаграждения мозга,

и иногда люди говорят, что были «вознаграждены» приливом дофамина. Однако такая картина недостаточно детализирована. Более двадцати лет назад немецкий нейробиолог Вольфрам Шульц рассмотрел экспериментальные данные по дофамину и пришел к выводу, что «дофаминовые нейроны активируются событиями, которые лучше прогноза, не затрагиваются теми, которые настолько же хороши, как он, и подавляются теми, которые хуже его»¹⁰². Итак, дофамин не вознаграждение R , а отслеживающий сигнал Q . Мозг использует его для оценки вознаграждения: он дает нам наш игровой счет.

Игры удовлетворяют многие наши базовые психологические потребности — например, демонстрацию умений решать задачи или работать совместно в группах¹⁰³. Одна из причин того, что мы играем в игры, возможно, в том, что те измеряют степень выполнения задач. В реальной жизни все запутано. Когда мы принимаем какие-то решения на работе и дома, результаты порой непросты, и о вознаграждении судить сложно. В играх же все легко: если мы действуем правильно, то получаем вознаграждение; если плохо, то терпим убытки. Игры устраниют неопределенность и позволяют нашей дофаминовой системе заниматься тем, что ей больше всего нравится: отслеживать вознаграждения. Простота ведения счета в какой-нибудь игре, которая обеспечивается одной-единственной отслеживающей переменной, отражает работу наших биологических систем вознаграждения.

Индустря компьютерных игр разбралась с уравнением вознаграждения. Одно исследование, в котором изучались британцы, ежедневно после работы игравшие в *Block!* (головоломка типа «Тетриса») или пользовавшиеся приложением для медитации *Headspace*, установило, что игроки в *Block!* лучше восстанавливались после рабочих стрессов. Эмили Коллинз, постдокторант из Университета Бата, проводившая эксперимент, позже заметила: «Медитация порой хороша для расслабления, но видеоигры обеспечивают психологическую отстраненность. Вы получаете внутренние вознаграждения и реальное ощущения контроля»¹⁰⁴.

Один разработчик компьютерных игр — Niantic — использовал наше желание получать вознаграждение для разработки игр,

которые заставляют нас двигаться. В самой знаменитой их игре — *Pokémon Go* — нужно выходить в реальный мир и «собирать» маленьких созданий, покемонов, с помощью мобильного телефона. Эта игра поощряет к прогулкам, поиску монстриков, высиживанию их яиц и работе в команде. Если вы видели группу людей, стоящих перед местной церковью или библиотекой и бездумно стучавших по своим телефонам, то наверняка это были охотники за покемонами, собравшимися для захвата арены.

А сейчас я вам расскажу кое-что личное. О моей жене. Ловиса Самптер — очень успешная женщина. Она доцент в департаменте математического образования Стокгольмского университета. Учит студентов, которые сами однажды станут преподавателями в школах; организует крупные международные конференции и выступает на них; курирует магистров и аспирантов; пишет документы, определяющие политику образования, и ведет вдохновляющие беседы с учителями. Ловиса также квалифицированный преподаватель йоги. Я мог бы написать целую книгу о том, какая замечательная у меня жена, и большая часть была бы посвящена тому, как она терпела меня все это время и организовала нашу семейную жизнь.

Но личное — другое. Все, кто встречался с Ловисой, знают, какая она замечательная. Ее таланты — не секрет, а установленный эмпирически факт. Однако Ловиса добилась всего этого, живя с 2004 года с постоянными болями. В 2018 году ей поставили диагноз «фибромиалгия» — состояние, которое характеризуется хронической болью во всем теле; в первую очередь это проблема нервной системы. Организм Ловисы постоянно посыпает в мозг болевые сигналы. Затем система отслеживания болей отправляет предупреждения вместо вознаграждения. Любая незначительная боль усиливается, из-за чего трудно спать, концентрироваться и оставаться терпимыми к близким. Лекарства неизвестны. Именно поэтому игра *Pokémon Go* стала важной частью жизни Ловисы: там она может получить какие-то вознаграждения, в которых ей отказывает ее организм.

Игра позволяет Ловисе сосредоточиться во время болей на чем-то другом, а также гарантирует долгие прогулки каждый день.

Благодаря ей Ловиса нашла новых друзей, вместе с которыми «сражается на аренах» и участвует в «рейдах». Многие из таких игроков имеют работу, связанную со стрессами, — например, медсестры или врачи. Среди них есть также учителя и ИТ-специалисты, студенты и другая молодежь. Благодаря группе Ловисы появилась минимум одна семейная пара. Есть также немало людей, которых другая социальная среда, возможно, выталкивала, — например, безработная молодежь, безвылазно игравшая дома в PlayStation, которую Pokémon Go снова вывела во внешний мир.

У всех поклонников Pokémon Go имеется своя история о том, как игра им помогла. Одна пенсионерка начала играть, чтобы вместе со внуками делать то, что им может понравиться. Как и многие ее ровесники, она сравнивает это с хоровым обществом. «Идете в рейд и делаете свое дело. Преимущество в том, что вы можете хоть беседовать с другими, хоть просто спокойно стоять».

Еще у одного из товарищей Ловисы по игре есть партнер, большой раком. Игра для него — шанс выбраться наружу и подумать о чем-то другом. Некоторые страдали от долгой депрессии и получают удовольствие от помощи новичкам. У новой подруги Ловисы по игре — Сесилии — синдром Аспергера и синдром дефицита внимания и гиперактивности, один из симптомов — желание копить вещи, например рецепты и журналы. «Теперь я могу собирать и организовывать, не становясь скопидомом. И при этом получаю физическую нагрузку!» — сказала она Ловисе. Сесилия грубовато-пряма, с юмором, и это помогает Ловисе справляться со своими ощущениями.

Pokémon Go приносит стабильность в жизнь Ловисы и других людей. Вознаграждения идут устойчивым потоком — пусть и в не-предсказуемое время, зато регулярно. «Это не лекарство, а скорее своего рода управление симптомами, — говорит мне Ловиса. — Это механизм выживания».

Ловиса и ее друзья всего лишь одна группа из многих разбросанных по планете любителей покемонов, жизнь которых улучшилась благодаря прогулкам и поиску вознаграждения. Йенни Солхейм Фуллер, старший менеджер, отвечающий в Niantic за гражданское и социальное влияние, рассказала мне об одном игроке, который

страдал от посттравматического стрессового расстройства после возвращения из зарубежной поездки: «Стремление к успехам в игре вынуждало его выходить из дома и сосредоточиваться не на ПТСР, а на чем-то еще». Йенни добавляет: «Еще одна большая группа — аутисты. Мы встречали множество родителей, дети которых отличались невероятной чувствительностью к шуму и хаосу вокруг и не могли выходить на улицу. Сейчас же они стоят перед художественной школой, проводя рейды и разговаривая с другими людьми».

Йенни получала сообщения от онкобольных, благодаривших ее за игру, которая помогла им пережить тяжелое время. Она прочитала письмо от сына одного человека, пятнадцать лет страдавшего диабетом, а потом начавшего играть в *Pokémon Go*. «Сейчас он добрался до сорокового уровня, самого высокого, — писал сын. — Он стал одним из самых чутких пожилых игроков. Диабет больше не угрожает его здоровью, и ему уже не нужны инъекции».

Это всего лишь одна из историй, которые заставили Йенни и ее коллег плакать; когда она читала их мне, я тоже плакал. Ловиса достигла сорокового уровня летом 2018 года. И пусть внешнему миру это может показаться не самым впечатляющим ее достижением, я вижу, как она использовала вознаграждения, чтобы справиться с болью.

*

Результат Герберта Роббинса и Саттона Монро стал начальной точкой для области математики, используемой при обнаружении сигналов; она стала развиваться в 1950-х и 1960-х. Математики показали, что отслеживающую переменную Q , можно использовать для оценки изменений в нашей среде. Вознаграждения — хорошие и плохие — можно наблюдать. В 1960 году Рудольф Эмиль Калман опубликовал основополагающую работу, показывающую, как надежно отфильтровать шум в вознаграждениях, чтобы раскрыть истинный сигнал¹⁰⁵. Его метод использовали, чтобы оценивать скорость и положение объектов в роторах¹⁰⁶, что стало существенным шагом в разработке автоматических датчиков.

Затем теория обнаружения сигналов была соединена с новой областью — математической теорией управления. Ирмгард Флюгге-

Лотц уже разработала систему прерывистого автоматического управления, которая давала возможность автоматически реагировать включением и выключением на изменения в температуре воздуха или турбулентности воздуха¹⁰⁷. Ее работа, наряду с трудами других специалистов по теории управления, позволила инженерам проектировать автоматизированные системы, которые отслеживают изменения в окружающей среде и реагируют на них. Сначала их применили в терmostатах, реагирующих на температуру в наших холодильниках и домах. Эти же уравнения легли в основу крейсерского режима для самолетов. Их также использовали для выравнивания зеркал в мощных телескопах, заглядывающих далеко во Вселенную. Именно эта математика управляла двигателем, который отвечал за начальные стадии торможения при приближении лунного модуля «Аполлона-11» к поверхности нашего спутника. Сегодня эти методы используются в роботах, которые трудятся на производственных линиях Tesla и BMW.

Теория управления создала мир устойчивых решений. Инженеры написали уравнения и потребовали, чтобы мир следовал этим правилам. Во многих случаях это прекрасно работало. Но мир неустойчив: бывают флюктуации и случайные события.

1960-е заканчивались новой контркультурой, бросавшей вызов установленвшемуся порядку, и «Десятка» тоже претерпела революцию. Акценты сместились с устойчивых линейных систем к неустойчивым, хаотичным и нелинейным. Именно такая математика повлияла на меня, когда я был молодым аспирантом в конце 1990-х, и я взялся за изучение всех этих математических теорий с экзотическими названиями: бабочка хаоса, модели песчаных лавин, критические лесные пожары, бифуркации в седловых точках, самоорганизация, степенные законы, критические точки... Каждая новая модель помогала объяснить сложность, которую мы видели вокруг.

Ключевой оказалась такая идея: устойчивость желательна не всегда. Новые математические модели описывали, как меняются экологические и социальные системы: не всегда возвращаясь обратно к тому же устойчивому состоянию, но иногда колеблясь между состояниями. Они описывали, как муравьи создают

тропинки к пище, как нейроны синхронно возбуждаются, как рыба плавает в косяках и как взаимодействуют биологические виды. Они рассказывали, как люди принимали решения, — и с точки зрения процессов внутри мозга, и в плане обсуждения в группах. В результате участники «Десятки» смогли занять должности на кафедрах биологии, химии и физиологии.

Именно эти математические методы я применял к данным, собранным биологами, с которыми работал.

*

На своем телефоне я открываю не только Twitter, но и множество приложений. Так и у муравьев и пчел есть не один источник пищи, а сразу несколько, и животные могут выбирать. На игровых автоматах есть много рычагов, и у нас нет времени дергать все. Задача в том, чтобы определить, какой рычаг тянуть. Мы знаем, что, потянув за один из них, можем получить достаточно хорошее представление о доступных вознаграждениях от этого игрового автомата. Но если мы проведем все время, дергая за этот рычаг, то не узнаем, что могут предложить другие машины. Это явление известно под названием «дилемма разведки/эксплуатации». Сколько времени тратить на использование уже известного, а сколько на изучение менее знакомых альтернативных вариантов?

Для решения этой задачи муравьи применяют химические вещества — феромоны. Их количество отражает оценку Q_t того, что имеется какой-то источник пищи. Теперь представьте, что у насекомых есть два источника с различными феромоновыми следами к каждому. Чтобы выбрать, каким путем идти, каждый муравей сравнивает количество феромона на обоих следах. Чем его больше, тем выше вероятность, что муравей последует этим маршрутом.

Выбор каждого следующего насекомого приводит к процессу подкрепления: чем больше муравьев пойдет по определенному пути и получит вознаграждение, тем вероятнее, что их сотоварищи отправятся тем же маршрутом. Пути с активным трафиком получают подкрепление; другие забрасываются. Это наблюдение можно сформулировать в рамках уравнения 9, где дополнительный коэффициент учитывает выбор муравьев¹⁰⁸. Вот один из примеров:

$$Q_{t+1} = (1 - \alpha)Q_t + \alpha \left(\frac{(Q_t + \beta)^2}{(Q_t + \beta)^2 + (Q'_t + \beta)^2} \right) R_t.$$

Новый коэффициент учитывает, как муравей выбирает между двумя альтернативными путями. Величину Q_t можно представлять как количество феромона на пути к одному потенциальному источнику пищи, а Q'_t — на пути к другому. Теперь у нас две отслеживающие переменные (Q_t и Q'_t) — по одной для каждого источника или (если мы моделируем использование соцсетей) по одной для каждого приложения в телефоне¹⁰⁹.

Когда сталкиваешься с новым запутанным уравнением с кучей параметров, всегда полезно рассмотреть сначала более простой вариант. Взглянем на новый коэффициент без квадратов:

$$\frac{Q_t + \beta}{Q_t + \beta + Q'_t + \beta}.$$

Если $\beta = 0$, это просто доля, которую одна отслеживающая переменная составляет от их суммы. Соответственно, вероятность того, что муравей использует конкретное вознаграждение, пропорциональна доле отслеживающей переменной для него. Теперь посмотрим, что произойдет при $\beta = 100$. Поскольку Q_t заключена между 0 и 1, она невелика по сравнению с числом 100, так что вышеуказанная дробь примерно равна $100 / (100 + 100) = 1/2$. Вероятность того, что муравей будет использовать определенное вознаграждение, равна 0,5 (или пятьдесят на пятьдесят).

Проблема баланса между разведкой и эксплуатацией превращается в проблему нахождения оптимального уровня подкрепления маршрута. Это то же, что задача нахождения правильного значения β . Если подкрепление сильное (значение β очень мало), муравьи всегда следуют по пути с самым сильным запахом. Очень быстро второй источник забрасывается (насекомые перестают его посещать), и даже если он станет лучше, никто о нем не узнает. В результате муравьи оказываются прикованными к тому источнику,

который казался лучше первоначально, даже если потом качество изменилось.

Слишком слабое подкрепление (значение β очень велико) приводит к противоположной беде. В этом случае насекомые выбирают маршруты наугад и не пользуются своими знаниями о том, какой из них лучше.

Ответ на задачу разведки и эксплуатации включает неожиданный поворот. Оказывается, решение дилеммы оптимального подкрепления связано с другим понятием, которое обычно возникает в совершенно другом контексте: критическими точками.

Поясню: критические (переломные) точки — моменты, когда накапливается какая-то критическая масса и система резко переходит из одного состояния в другое: например, мода внезапно распространяется после того, как авторитетные люди стали рекламировать этот бренд, или вспыхивает бунт, когда маленькая группа агитаторов заводит протестующих¹¹⁰. В каждом из этих и во многих других примерах подкрепление представлений приводит к внезапным переменам состояния. То же у муравьев — формирование феромонового маршрута происходит при достижении критической точки: путь начинается, когда небольшая группа муравьев решает двигаться к пище одной дорогой.

И вот удивительный вывод: наилучший способ сбалансировать разведку и эксплуатацию — чтобы муравьи оставались в состоянии, близком к критической точке. Если насекомые отойдут от нее, то слишком многие из них будут замкнуты на один источник пищи; они не смогут переключиться, когда появится что-то лучшее. Но если этому источнику будет привержено недостаточно насекомых и ситуация не дойдет до критической точки, то муравьи не смогут сосредоточиться на оптимальной пище. Они должны найти между разведкой и эксплуатацией золотую середину.

Муравьи эволюционировали так, чтобы оставаться в критической точке. Один из моих любимых примеров того, как они добиваются этого равновесия, обнаружила биолог Одри Дюссютур, работавшая с большеголовыми муравьями (этот вид получил свое название за необычно крупную голову). У них много поводов гордиться своей головой: они колонизировали большую часть тропического

и субтропического мира, выиграв конкуренцию у других местных видов. Одри выяснила, что они используют два вида феромонов: один испаряется медленно и дает слабое подкрепление, другой же испаряется быстро и дает очень сильное подкрепление¹¹¹.

Мы с математиком Стэром Николисом разработали модель с двумя уравнениями вознаграждения: одно для слабого, но длительно действующего феромона, а другое — для сильного, но коротковивущего. Мы показали, что комбинация этих двух феромонов позволила муравьям оставаться в районе критической точки. В нашей модели муравьи могли отслеживать два разных источника, переключаясь между ними всякий раз, когда качество пищи менялось. Одри подтвердила наши прогнозы экспериментально: когда она меняла качество еды, большеголовые муравьи переключали свои усилия на лучший источник.

Жизнь около критической точки характерна не только для муравьев. Для многих животных это бесконечное казино с игровыми автоматами. Не сидит ли в тех кустах какой-нибудь хищник? Есть ли еда там, где я нашел ее вчера? Где можно найти убежище на ночь? Чтобы выжить в этих условиях, эволюция довела животных до критических точек. Этот феномен я обнаруживал раз за разом в течение пятнадцати лет, когда изучал их поведение: саранча перемещается с такой плотностью, которая позволяет быстро менять направление движения; косяки рыб внезапно рассыпаются при нападении акул; скворцы в стаях дружно уворачиваются от ястреба. Двигаясь вместе, жертвы путают хищника.

Животные эволюционировали до критической точки. Они находятся в состоянии постоянной коллективной сознательности — переключаются от одного решения на другое, быстро реагируют на изменения. Для них это вопрос выживания.

А как насчет людей? Застряли ли мы на какой-нибудь критической точке? И если да, должны ли мы там оставаться?

*

В 2016 году Тристан Харрис обрушился на социальные сети. Преждущие три года он работал в Google, но похоже было, что он этим сыт. Харрис ушел из Google и опубликовал на платформе Medium

свой манифест. Заглавие утверждало: «Технологии захватывают ваш разум», а двенадцатиминутное содержание объясняло, как они это делают¹¹².

Аналогия, которую он выбрал для соцсетей, вам уже знакома: игровой автомат. По его словам, технологические гиганты фактически положили такие автоматы в карманы нескольких миллиардов человек. Уведомления, твиты, электронные письма, ленты Instagram и пролистывания-свайпы Tinder просят нас «потянуть за рычаг» и узнать, выиграли ли мы. Они прерывали наш день постоянными напоминаниями, а затем пропитывали нас страхом пропустить что-нибудь, если мы не дернем рычаг вовремя. Они соблазняли нас одобрением от наших друзей и побуждали отвечать взаимностью, ставя лайки и делясь публикациями, когда мы дергали рычаги одновременно с другими. Все это происходило в соответствии с собственными планами технологических компаний — заставить вас просмотреть рекламу и переходить по оплаченным ссылкам. Google, Apple и Facebook создали гигантское онлайн-казино и загребали прибыли. Наши карманные игровые автоматы вызывают такое привыкание, поскольку постоянно ставят нас перед необходимостью решать задачу разведки и эксплуатации. При этом социальные сети не простой игровой автомат: здесь тысячи ручек, и за все нужно потянуть, чтобы узнать, что происходит.

Ученым давно известно, какие проблемы доставляет мозгу животных наличие нескольких ручек. В 1978 году Джон Кребс и Алекс Касельник провели в Оксфордшире эксперимент с большими синицами¹¹³. Они предложили птицам две разные жердочки. Насесты были сконструированы так, что, когда синица садилась туда, иногда выпадала еда. Вероятность этого события была разной: на одной жердочке больше, чем на другой. Кребс и Касельник обнаружили, что в случаях, когда один насест был намного выгоднее другого, птицы быстро сосредоточивались только на нем. Но когда насесты были схожими, задача для птиц оказывалась сложной. Они перелетали между жердочками туда и обратно, пытаясь проверить, какая лучше. В моей терминологии синицы подошли близко к критической точке.

Математик Питер Тейлор показал, что уравнение вознаграждения полностью согласуется с этим результатом. Чем труднее выбрать вознаграждение, тем больше исследований-разведки требуется. Мы делаем то же, что и эти синицы, только у нас больше вариантов выбора. Мы открываем приложение за приложением. Но проблема не в доступности всех этих вознаграждений — она в желании мозга разведывать и эксплуатировать. Мы желаем быть уверенными в том, что знаем, где можно найти каждое из потенциальных вознаграждений. Нас подталкивают к критической точке.

Междуд использованием одного источника и нескольких огромная разница. Когда вы читаете какую-то книгу, играете в Mario Kart или Pokémon Go, смотрите «Игру престолов», играете в теннис с другом или идете в спортзал, вы сосредоточены только на одном источнике вознаграждения. Вам нравятся повторяющиеся сигналы о найденном предмете или звонки, отмечающие пройденный круг в гонке.

Ваше уравнение вознаграждения сходится к устойчивому состоянию. Это уравнение 1950-х, для которого Роббинс и Монро доказали устойчивую сходимость. Вы узнаёте, что судит такая деятельность, и ваша уверенность медленно, но верно начинает соответствовать ожидаемому вознаграждению. Именно эта привычная стабильность и дарит удовольствие.

А когда вы пользуетесь соцсетями, то исследуете-разведываете и эксплуатируете много разных источников вознаграждения. На деле вы вообще не собираете вознаграждения, а отслеживаете неопределенную среду. Вспомните, что дофамин не вознаграждение, поэтому вы не получаете удовольствия. Вы в режиме выживания собираете как можно больше информации. Проблема не обязательно состоит в неограниченной доступности вознаграждений, она в том, что вам надо следить за всеми потенциальными источниками вознаграждения; это и делает вашу жизнь трудной. Вы подводите свой мозг к критической точке — ставите его на порог хаоса, перемещаете в переходную стадию. Неудивительно, что у вас стресс.

И в критической точке не только ваш мозг: там все общество в целом. Мы подобны муравьям, которые неистово носятся вокруг,

пытаюсь отслеживать источники информации. А те постоянно двигаются, меняют качество, иногда исчезают. Что же вы можете сделать с этой проблемой?

Центр гуманных технологий — организация, одним из сооснователей которой был Тристан Харрис, — предлагает рекомендации, как взять под контроль и увести разум от критической точки. Отключите все уведомления на телефоне, чтобы вас не прерывали постоянно. Измените настройки экрана, чтобы значки были неяркими и меньше бросались в глаза.

По большей части я согласен с советами Харриса. Это здравый смысл. Однако из того, как муравьи пользуются уравнением вознаграждения, можно извлечь не такие очевидные, но, возможно, даже более полезные идеи.

Во-первых, вы должны осознать невероятную мощь того, что ваш разум и общество в целом находятся в районе критической точки. Не случайно самыми экологически успешными видами муравьев стали те, которые используют феромоны наиболее эффективно. То же касается и перехода людей к переломным точкам. Они могут вызывать у вас стресс как у индивидуума, но общество на пороге критической точки способно быстрее генерировать и распространять новые идеи. Подумайте о множестве идей, которые появились вследствие движения #MeToo или #BlackLivesMatter. Эти кампании действительно заставили людей узнать о проблемах и смогли привести к переменам. Или, если вы придерживаетесь других политических убеждений, посмотрите на выборы Трампа или #MakeAmericaGreatAgain*. Подумайте о том, как появились эти идеи, и о реакции на них с обеих сторон.

Сегодня мы чаще вовлечены в политические дискуссии. Когда дело доходит до политических причин, то молодежь сейчас активнее, чем когда-либо, — как в интернете, так и в реальной жизни¹¹⁴. Мы словно стая птиц, кружящаяся в вечернем небе. Мы будто косяк рыб, носящийся при приближении хищника. Мы как пчелиный рой, летящий к новому дому; как колония муравьев, прочесывающая

* Make America Great Again («Вернем Америке былое величие») — слоган Трампа во время рекламной кампании 2016 года.

лес в поисках пищи. Мы человеческая толпа, просматривающая новости.

Оказывайтесь в этой критической точке и наслаждайтесь свободой в ней. Читайте одну новостную статью за другой. Получайте информацию, принимайте новые идеи и следуйте своим интересам. При написании этой книги я провел бесчисленные часы в поисковой системе Google Scholar, просматривая научные статьи, уточняя, кто кого цитировал, и решая, какие научные вопросы важны. Общайтесь с людьми в Сети. Спорьте при необходимости. Отправьте письмо какому-нибудь полоумному старому профессору, который пишет для Quillette. Станьте частью информационного потока. Затем, когда вы проведете около часа в этой критической точке, я расскажу вам о второй идее, на которую наткнулся, глядя на муравьев.

Я понимаю, что мог создать у вас впечатление, будто муравьи — гиперактивные фанаты игровых автоматов. Это правда, когда они трудятся, и некоторые муравьи действительно работают очень усердно. Однако многие из них весьма ленивы. В любой конкретный момент большинство муравьев ничего не делают¹¹⁵. Пока меньшая часть сообщества бегают как сумасшедшие, оценивая и собирая пищу, большинство их сотоварищей просто бездействуют. Отчасти это связано с посменной работой; не все муравьи активны одновременно. Но в колониях также много насекомых, которые трудятся очень мало, редко выходят наружу и не участвуют в уборке. Почему муравьи в своей эволюции дошли до такой апатии, никто пока точно не знает, но если мы собираемся восхищаться меньшинством за его высочайшую активность, то нужно отдать должное и большинству за их расслабленное отношение к жизни.

Итак, когда вы уже некоторое время побыли в критической точке, поживите как ленивый муравей. Отрешитесь от всего. Поставьте «Игру престолов» на автоматическое воспроизведение. Посмотрите снова сериал «Друзья» от первого эпизода до последнего. Проведите неделю или месяц, собирая всех покемонов. Конечно, следует добавить все высоконравственные занятия вроде рыбалки, прогулок или сидения на крыльце. Но главное при этом — расслабляться без телефона. Не обращайте внимания на новости

и бесконечные электронные письма, где вас указывают в качестве адресата копии. Не волнуйтесь, о них побеспокоится кто-то другой. Это вовсе не обязательно должно быть вашей заботой.

Уравнение вознаграждения говорит вам, что нужно сосредоточиться на настоящем, а не зацикливаться на прошлом. Помните то единственное число, что держите в голове. Если что-то происходит, обновите его; если нет, пусть ваша оценка чуть-чуть уменьшится. Убедитесь, что вы ощущаете разницу между устойчивыми вознаграждениями, которые продолжают поступать (пусть и нерегулярно) независимо от ваших действий, и неустойчивыми, которые со временем меняют свою природу. Устойчивые вознаграждения можно найти в дружбе и в личных отношениях; в книгах, фильмах, телевизоре; в долгих прогулках и рыбной ловле; в играх «2048» и *Pokémon Go*. Неустойчивые есть в социальных сетях, поиске партнера в *Tinder*, на большинстве рабочих мест и часто (хотим мы это признавать или нет) — в нашей семейной жизни. Не бойтесь в таких ситуациях вести разведку и эксплуатацию, однако помните, что максимальную отдачу от вознаграждений можно найти, находясь в критической точке. Так что прежде чем неустойчивые вознаграждения опрокинут вас в нежелательное место, найдите дорогу обратно к устойчивости.

Глава 9

Уравнение обучения

$$-\frac{d(y - y_{\theta})^2}{d\theta}$$

Вероятно, вы слышали, что в технологиях будущего станет доминировать искусственный интеллект (ИИ). Ученые уже натренировали компьютеры побеждать в го, а сейчас испытывают беспилотные автомобили. Да, я объясняю некоторое количество уравнений в этой книге, но не забыл ли я что-нибудь? Не стоит ли мне также рассказать вам секреты, стоящие за ИИ, который используют Google и Facebook? Не следует ли мне объяснить, каким образом мы можем заставить компьютеры учиться так же, как мы сами?

Я открою вам секрет, который не совсем соответствует содержанию фильмов «Она» или «Из машины»*. Он также не увязывается с опасениями Стивена Хокинга или шумихой Илона Маска. Тони Старк, вымышленный супергерой — Железный человек из комиксов Marvel, — не обрадовался бы тому, что я скажу: искусственный интеллект в его современной форме не больше (и не меньше) чем десять уравнений, которые инженеры используют совместно и творчески. Но прежде чем я объясню, как работает ИИ, сделаем рекламную паузу.

*

Примерно во времена песни Gangnam Style у YouTube возникла одна проблема. Шел 2012 год; хотя сотни миллионов людей щелкали по видеороликам и посещали этот сайт, они не оставались

* В фильмах Спайка Джонса Her («Она») и Алекса Гарленда Ex machina («Из машины») речь идет об искусственном интеллекте.

там надолго. Новые ролики вроде «Чарли укусил меня за палец», «Двойная радуга», «Что говорит эта лиса?» или Ice Bucket Challenge удерживали их внимание всего на тридцать секунд, а дальше они снова возвращались к телевизору или к другим занятиям. Чтобы получать доход от рекламы, YouTube должен был стать местом, где пользователи будут зависать.

Значительную часть проблемы представлял алгоритм сайта. Он использовал систему рекомендации видео, основанную на уравнении рекламы из главы 7. Для роликов, которые смотрели и отмечали лайками пользователи, строилась корреляционная матрица. Однако этот метод не учитывал, что молодежь хотела смотреть самые свежие видео, и не уточнял, насколько интересен пользователям ролик. Он просто показывал видео, которые смотрели другие. В результате в списках рекомендуемых продолжала появлятьсяся норвежская армия, исполняющая Harlem Shake, а пользователи с сайта уходили.

Владельцы YouTube обратились к специалистам Google: «Эй, Google, как помочь детям найти те видеоролики, которые им нравятся?» — спросили (наверное) они. Три разработчика, получившие эту задачу, — Пол Ковингтон, Джей Адамс и Эмре Саргин — вскоре поняли, что самый важный критерий для оптимизации YouTube — время просмотра. Если бы сайт мог заставить пользователей смотреть как можно больше роликов как можно дольше, то легче было бы вставлять рекламу через регулярные промежутки времени и зарабатывать больше денег. При этом короткие свежие ролики были не так важны, как целые каналы, обеспечивающие постоянное появление свежего и длительного контента. Задача состояла в том, чтобы найти способ выявить этот контент на платформе, где каждую секунду загружаются часы видеороликов¹¹⁶.

Ответ разработчиков имел форму воронки. Это приложение брало сотни миллионов видеороликов и сводило их примерно к десятку рекомендаций, представленных сбоку на странице сайта. Каждый пользователь получал собственную персонифицированную воронку с роликами, которые, возможно, он захочет посмотреть.

«Воронка» — нейронная сеть, которая изучает наши предпочтения при просмотре. Такие сети лучше всего представлять в виде

столбца входных нейронов слева и выходных справа. Между ними находятся слои соединительных нейронов, известных как скрытые (см. рис. 9). В сети могут быть десятки или даже сотни тысяч нейронов. Она не реальна в физическом смысле: это компьютерные коды, которые моделируют взаимодействие нейронов. Однако аналогия с мозгом полезна, потому что именно прочность связей между нейронами позволяет их сетям изучать наши предпочтения.

Каждый нейрон кодирует определенные аспекты того, как сеть реагирует на входные данные. В «Воронке» нейроны фиксируют взаимосвязи между разными элементами контента и каналов YouTube. Например, люди, которые смотрят правого комментатора Бена Шапиро, также склонны смотреть и видео Джордана Питерсона. Я знаю это, потому что после завершения своего исследования для главы 3 об уравнении уверенности YouTube с маниакальным упорством подсовывает мне ролики Шапиро. Где-то внутри «Воронки» есть нейрон, который представляет связь между этими двумя иконками «Темной сети интеллектуалов». Когда он получает входной сигнал, что я заинтересовался роликом Питерсона, то дает на выходе вывод, что меня могут интересовать и ролики Шапиро.

Мы можем понять, как «обучаются» искусственные нейроны, узнав, как формируются связи внутри сети. Нейроны кодируют отношения в виде параметров — регулируемых величин, которые измеряют прочность отношений. Рассмотрим нейрон, отвечающий за определение того, сколько пользователей будут тратить время на просмотр выступления Бена Шапиро. Внутри этого нейрона имеется параметр θ , который соотносит время, потраченное на видео Шапиро, с количеством просмотренных роликов с Джорданом Питерсоном. Например, мы можем спрогнозировать, что количество минут, которое пользователь тратит на видео Шапиро (обозначим его y_θ), равно θ , умноженному на количество просмотренных роликов Питерсона. Скажем, если $\theta = 0,2$, то прогнозируется, что человек, просмотревший десять роликов Питерсона, потратит $y_\theta = 0,2 \cdot 10 = 2$ минуты на видео Шапиро. Если $\theta = 2$, то прогнозируется, что тот же человек потратит $y_\theta = 2 \cdot 10 = 20$ минут на Шапиро, и т. д. Процесс обучения включает корректировку параметра θ для улучшения прогнозов для времени просмотра.

Предположим, первоначальное значение нейрона $\theta = 0,2$. Здесь появляюсь я, который видел 10 выступлений Питерсона и трачу на просмотр Шапиро $y = 5$ минут. Квадрат разности между прогнозом (y_θ) и реальностью (y) составляет:

$$(y - y_\theta)^2 = (5 - 2)^2 = 3^2 = 9.$$

Мы уже видели идею квадрата разности — в главе 3, когда измеряли стандартное отклонение. Вычислив $(y - y_\theta)^2$, получаем меру того, насколько хороши (или плохи) прогнозы нейронной сети. Расхождение между прогнозом и реальностью равно 9, так что, похоже, предсказание не особо хорошее.

Чтобы чему-то научиться, искусственный нейрон должен знать, что он делал неправильно, когда прогнозировал, что я буду смотреть только две минуты. Поскольку прочностью связи между количеством видеороликов Питерсона и типичным временем, которое пользователь тратит на просмотр Шапиро, управляет параметр θ , его увеличение также увеличит и предсказанное время y_θ . Поэтому, например, если мы возьмем для θ небольшое увеличение $d\theta = 0,1$, то получим $y_{\theta+d\theta} = (\theta + d\theta) \cdot 10 = (0,2 + 0,1) \cdot 10 = 3$ минуты. Такой прогноз будет ближе к реальности:

$$(y - y_{\theta+d\theta})^2 = (5 - 3)^2 = 2^2 = 4.$$

Именно это улучшение и использует уравнение 9 — уравнение обучения*.

$$-\frac{d(y - y_\theta)^2}{d\theta} \quad (\text{Уравнение 9}).$$

Это выражение говорит, что мы рассматриваем, как маленькое изменение $d\theta$ увеличивает или уменьшает квадрат расстояния $(y - y_{\theta+d\theta})^2$. Конкретно в нашем примере получаем:

$$-\frac{d(y - y_\theta)^2}{d\theta} = -\frac{(y - y_{\theta+d\theta})^2 - (y - y_\theta)^2}{d\theta} = -\frac{4 - 9}{0,1} = 50.$$

* Снова отметим, что это выражение формально уравнением не является.

Поскольку эта величина положительна, увеличение θ улучшает качество прогноза — и расстояние между ним и реальностью уменьшается.

Математическая величина, задаваемая уравнением 9, известна как производная по θ или градиент*. Она измеряет, приближает или отдаляет ли нас изменение θ от хорошего прогноза. Процесс медленной корректировки θ на основании производной часто называют градиентным подъемом**, что вызывает в мозге образ человека, движущегося по крутому уклону холма. Следуя по градиенту, мы можем медленно улучшать точность искусственного нейрона (см. рис. 9).

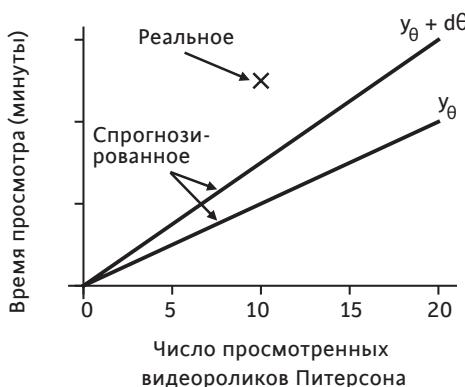
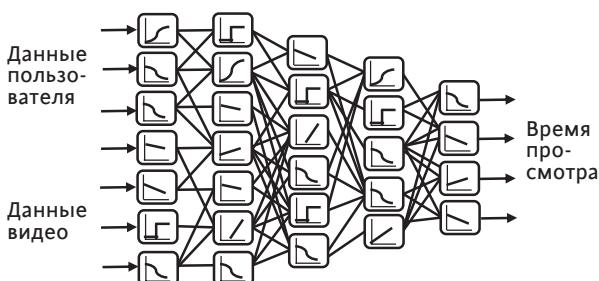
«Воронка» работает не только с одним нейроном, а сразу со всеми. Первоначально все параметры принимают случайные значения, и нейронная сеть делает очень плохие прогнозы о времени, которое люди потратят на просмотр видео. Затем инженеры начинают подавать на входные нейроны (широкий конец «Воронки») данные о просмотре роликов пользователями YouTube. Небольшое число выходных нейронов (узкий конец «Воронки») измеряет, насколько хорошо нейронная сеть предсказывает продолжительность просмотра роликов. Сначала ошибки в прогнозах очень велики. При применении метода обратного распространения ошибки отклонения, измеренные на узком конце, передаются обратно по слоям «Воронки». Каждый нейрон измеряет градиент и улучшает параметры. Медленно, но верно нейроны поднимаются по градиенту, и прогнозы постепенно улучшаются. Чем больше данных от пользователя YouTube подается в сеть, тем лучше будет прогноз.

Нейрон из моего примера, занимающийся отношением Шапиро/Питерсон, не нужно изначально кодировать в сети. Сила нейронных сетей в том, что нам не нужно сообщать им, какие отношения искать в данных: сеть сама находит эти взаимоотношения в процессе градиентного подъема. Поскольку связь Шапиро/Питерсон прогнозирует время просмотра, в итоге один или несколько нейронов начнут использовать эту связь. Они будут тесно

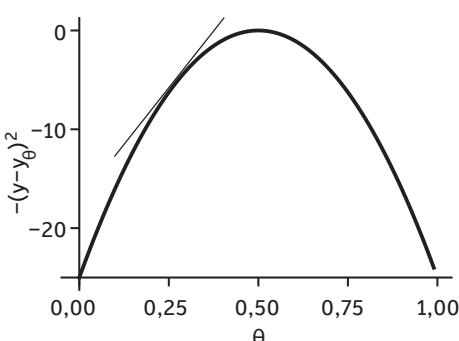
* Градиент — векторная величина, показывающая направление наибольшего возрастания исходной числовой величины.

** Либо градиентным спуском, если нам нужен минимум (в этом случае мы двигаемся по антиградиенту).

взаимодействовать с другими нейронами, связанными с другими авторитетами «Темной сети интеллектуалов» и даже с более правыми идеологиями. Это создает статистически корректное представление типа человека, который, вероятно, смотрит видеоролики Джордана Питерсона.



Внутри каждого нейрона функцию настраивают, чтобы прогноз был лучше. Здесь увеличение θ на $d\theta$ приближает спрогнозированное значение времени к реальному времени просмотра



Поднимаясь по градиенту к точке, в которой расстояние между прогнозом и наблюдением больше не улучшается, нейроны обучаются связям в данных

Рис. 9. Как обучается нейронная сеть

Уравнение 9 — основа методов, известных как машинное обучение. Постепенное улучшение параметров с помощью градиентного подъема можно рассматривать как процесс «обучения»: нейронная сеть («машина») постепенно «учится» делать всё лучшие прогнозы. Если предоставить ей достаточно данных (а у YouTube их навалом), нейронная сеть изучает закономерности внутри этих данных. Как только она «обучилась», «Воронка» может спрогнозировать, сколько времени пользователь YouTube проведет за просмотром видеороликов. Платформа воплотила эту методику в жизнь. Она берет видеоролики с самым большим прогнозируемым временем просмотра и ставит их в рекомендательные списки для пользователей. Если человек не выбрал новое видео, YouTube автоматически воспроизводит тот ролик, который, по его мнению, понравится пользователю больше всего.

Успех «Воронки» был ошеломительным. В 2015 году время, потраченное на просмотр в YouTube пользователями в возрасте от 18 до 49 лет, возросло на 74%¹¹⁷. К 2019 году число просмотров выросло в 20 раз по сравнению с моментом, когда специалисты Google начали свой проект, причем 70% из них берутся из рекомендованных видео¹¹⁸. Дуг Коэн, специалист по данным из Snapchat, был восхищен этим решением. «Google решил за нас проблему разведки и эксплуатации», — сказал он мне. Вместо того чтобы бродить по разным сайтам и пытаться найти лучшие видео или ждать, пока вам кто-нибудь пришлет интересную ссылку, теперь вы можете часами сидеть в YouTube, выбирая либо «Следующее» видео, либо один из десятка предлагаемых альтернативных вариантов.

Если вы считаете, что ведете на YouTube разведку по своим интересам, а обнаруживаете, что смотрите предлагаемые видеоролики, то, к сожалению, заблуждаетесь. «Воронка» превратила YouTube в подобие телевидения, только программу составляет искусственный интеллект. И многие приклеиваются к этому экрану.

*

Ной хотел бы стать более популярным в Instagram. У многих его друзей подписчиков (фолловеров) больше, чем у него, и он с зависимостью смотрит, как им летят лайки и комментарии. Ной смотрит на

аккаунт своего друга Логана: около 1000 подписчиков, на каждый пост сотни лайков. Он хочет быть похожим на Логана и ставит себе цель: $y = 1000$ подписчиков. При текущей стратегии у него $y_\theta = 137$. До нужного числа далеко.

В течение следующей недели Ной постепенно начинает размещать все больше постов. Он полагает, что чем больше публикаций, тем больше людей на него подпишутся, и вывешивает фотографии своего ужина, новых туфель или прогулки до школы, но не старается улучшить качество снимков. Ной просто фотографирует все, что видит, и выкладывает это в блоге. В терминах уравнения 9 параметр θ , который он регулирует, — отношение количества его сообщений к качеству. Ной увеличивает количество постов, поэтому $d\theta > 0$.

Отзывы неблагоприятные. «Зачем ты спамишь?» — пишет под одной из фотографий его подруга Эмма, добавляя к тексту озадаченный смайлик. Некоторые знакомые Ноя отписались от него. Его популярность упала: $y_{\theta+d\theta} = 123$, то есть произошло снижение на 14 человек. Расстояние до цели увеличилось. Он опустился по градиенту, а не поднялся*. В последующие месяцы Ной снижает количество публикаций и сосредоточивается на качестве. Несколько раз в неделю он делает фотографию какого-нибудь приятеля с мороженым или забавный снимок своей собаки, тщательно редактирует эти фотографии и с помощью фильтров добивается, чтобы его архив выглядел хорошо. И во время такого перехода от количества к качеству Ной еще и измеряет y_θ . Количество его подписчиков медленно, но уверенно увеличивается. Через шесть месяцев оно выросло до 371, но потом застыло, седьмой месяц не дал прибавки.

Так мы приходим к важному уроку из уравнения 9: Ною нужно расслабиться и перестать стремиться к цели в 1000 фолловеров. Хотя $(y - y_\theta)^2 = (1000 - 371)^2 = 395\,641$ — по-прежнему большое число, выражение в уравнении 9 больше не меняется:

$$-\frac{d(y - y_\theta)^2}{d\theta} = 0.$$

* Здесь термин употреблен не в математическом смысле. Градиент — направление максимального изменения величины, но очевидно, что терять подписчиков можно и быстрее.

Уравнение говорит, что Ною надо остановиться в своей стратегии работы с социальной сетью и довольствоваться тем, что он получил. Больше нет надобности сравнивать себя с Логаном: Ной достиг своего пика популярности.

Применяя уравнение 9, мы должны помнить об итоговой цели, но в первую очередь нужно руководствоваться тем, движемся мы вверх или нет. Как гласит мудрость, если вы забрались на вершину горы, наслаждайтесь видом. Математика подтверждает это.

Разница между оптимизацией, которую создают алгоритмы машинного обучения вроде «Воронки», и оптимизацией Ноя в том, что он пытался увеличить число подписчиков, а машинное обучение стремится оптимизировать точность прогнозов. В случае «Воронки» y_θ — это прогнозы, как долго пользователи будут смотреть видео, а y — реальная продолжительность просмотра. YouTube хотел бы уметь предсказывать предпочтения пользователей с максимально возможной точностью, но понимает, что его прогнозы никогда не будут идеальными. «Воронка» удовлетворена, когда осознаёт, что лучшего добиться не может.

Фокус в использовании уравнения обучения в том, чтобы честно говорить, как ваши действия увеличивают или уменьшают разницу между целью и реальностью. Некоторые могут обвинить Ноя в том, что в своих попытках оптимизировать влияние в соцсети он «поверхностен» или что-то «фальсифицирует». Я не согласен. Работа с такими влиятельными специалистами, как Кристиан Ихо, который ведет в соцсети блог об уличной моде, научила меня обратному. Кристиан использует аналитические инструменты Google, чтобы изучать, как отношение количества к качеству в публикациях создает поток пользователей; однако все равно понимает, что его данные относятся к людям. Когда семнадцатилетний подросток сделал селфи в дизайнерской футболке, лицо Кристиана засветилось. Он поставил лайк и написал комментарий: «Отлично смотрится!» Он не шутил. Нет противоречия между тем, чтобы изучать данные и быть на 100% искренним в том, кто вы и что делаете.

При аккуратном применении уравнение обучения помогает вам оптимизировать собственную жизнь. Пытаетесь ли вы

добраться успеха в соцсетях или готовиться к экзамену, вы всегда стремитесь постепенно подниматься по градиенту. Ставьте себе цель, но не фокусируйтесь на расстоянии до нее. Пусть вас не беспокоят люди, которые популярнее вас, или ровесники с более высокими оценками. Сосредоточивайтесь на ежедневных шагах. Сфокусируйтесь на градиенте: на дружбе, которую приобретаете, или на новых знаниях, полученных в процессе учебы. Если вы видите отсутствие прогресса, признайтесь себе в этом. Вы добрались до вершины холма, и теперь пора понаслаждаться видом. Однако имейте в виду, что следовать градиенту не всегда идеальное решение; иногда вы попадете в ловушку неоптимального решения*. Тогда нужно сбросить настройки и начать заново. Найдите новую гору для подъема или новые параметры, которые можно регулировать.

*

В 2019 году Джарвис Джонсон бросил работу специалиста по программному обеспечению. Его ютуб-канал, где он выкладывал ролики о своей жизни программиста, привлекал все больше подписчиков. Джонсон решил проверить, сможет ли стать полностью «интернетным человеком», как он теперь себя называет.

Чтобы стать ютюбом, нужны две вещи: интересный контент и глубокое понимание процессов «Воронки». У Джарвиса есть и то и другое, а его ролики сочетают эти элементы с юмором. Он исследует, как некоторые каналы используют этот алгоритм ради своей выгоды, манипулируя «Воронкой» так, чтобы она выдавала рекомендации в их пользу. Затем излагает свои результаты в завлекательных интересных видео на этой же платформе.

Расследования Джарвиса сфокусировались на цифровой студии, которая называется TheSoul Publishing. Они именуют себя «одной из крупнейших медиакомпаний в мире» и заявляют, что их задача — «вовлекать, вдохновлять, развлекать и просвещать». Джарвис начал с просмотра одного из самых успешных

* Проблема здесь в том, что, поднимаясь по градиенту, вы можете найти точку локального максимума, а не глобального.

каналов — 5-Minute Crafts*. Там предлагаются лайфхаки — советы, которые облегчают повседневную жизнь. В одном из видео со 179 миллионами просмотров сайт утверждал, что несмываемый маркер можно удалить с футболки, если использовать смесь санитайзера для рук, разрыхлителя для теста, лимонного сока и зубную щетку. Джарвис решил проверить это сам, написав текст на белой футболке и последовав инструкциям. Результат? После выполнения всех инструкций и машинной стирки текст остался виден. Джарвис проверял совет за советом — и оказывалось, что они либо тривиальны, либо не работают. Лайфхаки канала 5-Minute Crafts были бесполезны.

Другой из каналов TheSoul — Actually Happened — утверждал, что показывает реальные истории из жизни подписчиков**. Джарвис обнаружил, что контент создан «сценаристами», которые писали правдоподобные истории, рассчитанные на подростков из США, используя в качестве исходных материалов Reddit и другие социальные сети. Джарвис объяснил мне, что Actually Happened изначально копировал другой канал — Storybooth***, который действительно воспроизводил подлинные истории детей и подростков. На платформе Storybooth дети часто делятся собственными историями, что помогает обеспечить искренность и честность.

«Алгоритм YouTube не может определить разницу между Storybooth и Actually Happened», — сказал мне Джарвис в беседе в мае 2019 года. Канал Actually Happened применяет те же названия, описания и теги, что и Storybooth, и ютубовская «Воронка» считает их примерно эквивалентными и начинает

* Название канала означает примерно «Рукоделки за 5 минут». Существуют также ответвления канала (5-minute Crafts Family, 5-minute Crafts Like, 5-minute Crafts Recycle и т. п.) и версии на других языках; русскоязычная — «Бери и делай». Помимо огромного числа подписчиков (более 70 млн на конец марта 2021 года только на основном канале) более всего известен нерабочими лайфхаками, а также многочисленными проверками-«разоблачениями» от разных блогеров (по сути, «разоблачения» можно уже считать отдельным жанром роликов на платформе). *Прим. ред.*

** Название канала означает «Произошло на самом деле».

*** Название канала означает примерно «Кабина историй». Этим словом в США назывались маленькие общедоступные студии звукозаписи, где можно было записать свою историю. Первая кабина появилась в 2003 году.

создавать ссылки между этими сайтами. «Actually Happened на-воднили рынок историями. Они нанимали подрядчиков по цене ниже рыночной и выдавали по видео в день, — продолжал Джар-вис. — Затем они достигли уровня, когда им больше не требовалось копировать Storybooth». Как только они добились миллиона подписчиков, «Воронка» решила, что дети хотят смотреть Actually Happened. Алгоритм признал это вершиной холма для своих пользователей.

Джарвис считал, что моральные проблемы вокруг каналов компании TheSoul запутанны. «Я, конечно, создавал контент, как у других, надеясь привлечь аудиторию, но я бы не мог делать это так явно и массово. Что не дает этой компании проделывать такое со всеми жанрами на платформе?»

Ограниченност алгоритма YouTube в том, что его не заботит контент, который он продвигает, или работа, затраченная на его создание. Я сам ощущил это, когда YouTube решил, будто меня интересует Бен Шапиро. Любой, кто сажал своих детей перед YouTube на час, наверняка замечал, как их затягивает в странный мир видеороликов с показом игрушек, вазочек с мороженым из цветного пластилина и диснеевских головоломок. Посмотрите на ролик PJ Masks Wrong Heads for Learning Colors: выглядит так, как будто на его создание ушло полчаса, но он имеет 200 миллионов просмотров. «Воронка» может рекомендовать не просто видеоролики плохого качества; они бывают негодными. В 2018 году журнал Wired зафиксировал видео, где собаки из мультсериала «Щенячий патруль» пытались совершить самоубийство, а Свинку Пеппу обманом заставили есть бекон¹¹⁹. Одно из расследований The New York Times обнаружило, что YouTube рекомендовал семейные видео с голыми детьми, играющими в бассейнах-лягу-шатниках, своим пользователям, которые интересовались педо-филией¹²⁰.

YouTube может дать нам «введение воронки». Возможно, его цель — обеспечение лучших рекомендаций для вас, однако уравнение 9 выполнено, когда обнаруживается наилучшее решение для доступных данных. Происходит подъем по градиенту обучения, пока не будет достигнут пик, а затем остановка — чтобы вы

наслаждались видом, причем каким угодно. «Воронка» совершает ошибки, а ответственность за их исправление лежит на нас. YouTube не всегда преуспевал в решении этой задачи.

*

Некоторые люди со стороны могут решить, что участники «Десятки» похожи на Железного человека — Тони Старка (промышленники и талантливые инженеры, использующие технологии для преобразования мира). Но если бы любому из участников «Десятки» предложили выбрать какого-нибудь супергероя, чтобы описать себя, это, вероятно, был бы Человек-паук (Питер Паркер). У него нет плана, нет нравственной программы — он как подросток, пытающийся сохранить контроль над организмом, когда тот меняется неожиданно.

Напряженность среди участников «Десятки» можно рассматривать по-разному. Похожи ли они на наивного Марка Цукерберга из фильма «Социальная сеть» или роботоподобного Марка Цукерберга, дающего свидетельские показания в юридическом и торговом комитетах Сената США? Похожи ли они на того Илона Маска, который курит травку на камеру, или на того, который связывает наше будущее с перелетом на Марс?

С одной стороны, уравнения дают «Десятке» возможность выносить безупречные суждения, и им доверяют планировать глобальные перемены в нашем обществе. Они создали научный подход, который укрепляет уверенность в моделях, использующих данные. Они связали нас всех так, как мы не ожидали. Они оптимизируют и улучшают работу. Они приносят эффективность и стабильность. С другой стороны, эти участники будут придерживаться уравнения вознаграждения, которое предлагает брать то, что есть сейчас, и забыть о прошлом. Они создают преимущество перед теми, кто не может позволить себе платить.

Это как раз то, о чем Алфред Джулс Айер говорил в 1936 году: в математике нет морали, а если и была когда-то, то сейчас уже утрачена. Невидимость «Десятки» означает, что мы не можем найти даже подходящую аналогию с супергероем. Кто такие участники «Десятки»? Наивные подростки, осознающие, как Питер Паркер,

что с большой властью приходит и большая ответственность; или жадные до власти маньяки, которые хотят управлять миром «ради его же блага»? Может, они даже похожи на суперзлодея Таноса из вселенной Marvel, готового убить половину людей, — поскольку думают, что это станет оптимальным решением?

Что бы они ни думали о себе, нам нужно знать, что они замышляют, потому что, куда бы они ни отправились, они всё меняют.

*

Когда мы изучаем примеры современного ИИ — например, нейронную сеть компании DeepMind, которая стала лучшим в мире игроком в го, или искусственный интеллект, который научился играть в «Космических захватчиков» либо в другие игры для Atari, — мы должны считать их выдающимися достижениями инженерии. Некая группа математиков и программистов собрала воедино все фрагменты. За этим ИИ стоит не какое-то одно уравнение.

Но — и это важно для всего моего проекта по описанию десяти уравнений — компоненты искусственного интеллекта включают девять из них. Так что в финале я попробую объяснить, как DeepMind стала мастером игры, используя ту математику, которую мы уже изучили в этой книге.

Представьте сцену, где в окружении кольца столиков стоит шахматный гроссмейстер. Он подходит к одному столу, изучает позицию и делает ход. Затем переходит к следующему и делает ход там. В конце сеанса оказывается, что он выиграл все партии. Сначала может показаться невероятным, что гроссмейстер отслеживает столько шахматных партий одновременно. Неужели он может помнить, как развивалась игра до данного момента, и решать, что делать дальше? Но потом вы вспоминаете уравнение умений.

Ситуацию в партии можно увидеть непосредственно на доске: защитная структура пешек, качество убежища для короля, насколько хорош для атаки ферзь и т. д. Гроссмейстеру не нужно знать, как шла игра до настоящего момента, достаточно изучить позицию и выбрать следующий ход. Умения шахматиста можно измерить тем, как он берет текущее состояние доски и переводит его в новое, делая какой-то закономерный ход. Это новое состояние

уменьшает или увеличивает его шансы на победу в партии? При оценке гроссмейстеров применяется уравнение 4 (марковское предположение).

«Многие игры с полной информацией — например, шахматы, шашки, реверси или го — можно считать марковскими». Такой была первая фраза в разделе «Методы» статьи Дэвида Сильвера и других специалистов Google DeepMind об их программе, ставшей лучшим игроком в го в мире¹²¹. Это наблюдение упрощает задачу нахождения решения для этих игр, поскольку позволяет сосредоточиться на поиске оптимальной стратегии для текущего состояния на доске, не заботясь о том, что происходило до этого момента.

Мы уже анализировали математику отдельного нейрона в главе 1. Уравнение 1 брало текущие коэффициенты для какого-нибудь футбольного матча и преобразовывало в решение, стоит нам делать ставку или нет. По сути, это упрощенная модель того, что делает отдельный нейрон в вашем мозге. Он получает внешние сигналы — от других нейронов или из внешнего мира — и преобразует их в решение, что ему сделать. Такое упрощающее предположение легло в основу первых моделей нейронных сетей, а уравнение 1 использовалось для моделирования реакции нейронов. Сегодня это одно из двух очень похожих уравнений, которые используются для моделирования нейронов почти во всех сетях¹²².

Далее мы обратимся к одному из вариантов уравнения вознаграждения. В уравнении 8 величина Q_t была оценкой качества сериала Netflix или вознаграждения, получаемого от проверки аккаунта в Twitter. Вместо того чтобы оценивать один фильм или один аккаунт, сейчас мы просим нашу нейронную сеть оценить $1,7 \times 10^{172}$ разных состояний в игре го или 10^{172} сочетаний клипов и пользователей на YouTube. Обозначим $Q_t(s_t, a_t)$ качество состояния мира s_t при условии, что мы намереваемся произвести некоторое воздействие a_t . В игре го состояние s_t — решетка 19×19 , где у каждого узла (в го они называются пунктами) есть три возможных состояния: пуст, занят белым камнем или занят черным камнем. Возможные действия a_t — пункты, куда можно поставить очередной камень. Тогда величина, характеризующая качество — $Q_t(s_t, a_t)$ — говорит

нам, насколько хорош будет ход a_t в состоянии s_t . Для YouTube одно состояние — все пользователи в сети и все имеющиеся ролики. Действие — просмотр конкретным пользователем конкретного видеоролика, а качество — насколько долго он его просматривает.

Вознаграждение $R_t(s_t, a_t)$ — награда, которую мы получаем за выполнение действия a_t в состоянии s_t . В го награда появляется только с концом партии. Мы можем дать 1 за выигрывающий ход, -1 за проигрывающий и 0 за любой другой. Какое-нибудь состояние может иметь высокое качество, но нулевое вознаграждение: например, если некоторое расположение камней близко к победному.

Когда DeepMind использовала уравнение вознаграждения для игр Atari, она добавляла еще один компонент: будущее. Когда мы производим какое-нибудь действие a_t (ставим камень в го), то переходим в новое состояние s_{t+1} (на доске занят тот пункт, куда мы сделали ход). Уравнение вознаграждения DeepMind добавляет вознаграждение размером $Q_t(s_{t+1}, a)$ за наилучшее действие в этом новом состоянии. Это позволяет искусству интеллекту планировать будущие шаги в игре.

Уравнение 8 дает нам гарантию. Оно говорит, что если мы будем следовать такой схеме и обновлять качество нашей игры, то постепенно ее освоим. Более того, с помощью этого уравнения мы в итоге придем к оптимальной стратегии для любой игры — от крестиков-ноликов до шахмат и го.

Но есть одна проблема. Это уравнение не говорит нам, сколько времени придется играть, чтобы узнать качество всех различных состояний. В игре го $3^{19 \times 19}$ состояний, что дает примерно $1,7 \times 10^{172}$ различных возможных позиций на доске*. Даже очень быстрому компьютеру потребуется огромное время, чтобы сыграть их все, а ведь чтобы наша функция качества сходилась, нам нужно пройти через каждое состояние много раз. Нахождение наилучшей стратегии возможно теоретически, но не практически.

Ключевой идеей для специалистов Google DeepMind стало понимание, что качество $Q_t(s_t, a_t)$ можно изобразить в виде нейронной

* Для сравнения: число атомов во Вселенной, по разным оценкам, составляет от 10^{78} до 10^{82} . Иными словами, даже если бы существовало столько вселенных, сколько атомов в нашей, во всех таких вселенных вместе все равно было бы гораздо меньше атомов, чем это число.

сети. Вместо того чтобы узнавать, как искусственный интеллект должен играть во всех $1,7 \times 10^{172}$ возможных позициях в го, разбирающийся в этой игре ИИ был представлен в таком виде: на входе позиции на доске 19×19 , далее несколько уровней из скрытых нейронов и выходные нейроны, которые определяли следующий ход. Как только задача была переформулирована в форме нейронной сети, исследователи смогли для получения ответа использовать метод градиентного спуска (уравнение 9).

Возможно, наиболее мощная демонстрация силы такого подхода была такой: ученые взяли нейронную сеть AlphaZero, которая не умела играть в шахматы, и всего за четыре часа она освоила их на уровне лучших компьютерных программ мира, которые сами уже далеко превосходили возможности лучших шахматистов-людей. После этого AlphaZero продолжала учиться, бросая вызов себе, чтобы найти методы игры, которые непредставимы ни для человека, ни для любого компьютера.

При работе с нейронными сетями применяются все уравнения, с которыми мы встречались ранее. Мы уже использовали уравнения 1, 4, 8 и 9. Уравнение 5 применяют при изучении связей между сетями. Способ соединения нейронов — ключ к определению типа задач, которые может решить сеть. Название «Воронка» дано по структуре нейронной сети, используемой YouTube, когда есть много входных нейронов и небольшое количество выходных. Ученые обнаружили, что в других случаях более эффективными будут другие структуры. Например, для распознавания лиц и игр оптимальна ветвящаяся структура, известная под названием «сверточная нейронная сеть»¹²³. Для обработки языковых текстов лучшим вариантом будет так называемая рекурсивная нейронная сеть¹²⁴.

Уравнения 3 и 6 применяют, когда смотрят на время, необходимое для обучения сети, — чтобы убедиться, что она обучилась надлежащим образом. Уравнение 7 лежит в основе метода, который называется «обучением без учителя»: его можно использовать, когда у нас для анализа есть миллионы разных видео, изображений или текстов и мы желаем установить самые важные закономерности, на которые стоит обратить внимание. Уравнение 2 — база

байесовских нейронных сетей, важных при изучении игр, связанных с неопределенностью, — например, покера.

Итак, всего в девяти уравнениях мы находим основу для современного искусственного интеллекта. Изучите их, и вы сможете помочь создать искусственный интеллект для будущего.

*

Большинство из нас не знают, что все лучшие работы в области ИИ доступны всем, кто хотел бы узнать больше и уже понимает эти девять уравнений на том же уровне, что и вы сейчас. Статьи публикуются в журналах с открытым доступом, а библиотеки компьютерного кода доступны всем, кто желает создавать собственные модели. Открытость растет — от доктрины де Муавра и записных книжек Гаусса через взрывное развитие науки в конце прошлого века и до современных архивов веб-сервиса Github, где технологические гиганты складывают ИТ-проекты и делятся программными кодами.

Людям нужно не верить страшилкам или раздуваемой шумихе, что искусственный интеллект становится человеком, а посмотреть на историю Google. Эта компания, основанная двумя аспирантами из Калифорнии, проводит и финансирует высококачественные исследования; и почти все, что она делает, доступно широкой публике. Конечно, есть опасность, что лучшие умы уходят из университетов и идут трудиться для Google, Facebook и других компаний. Но многие из нас по-прежнему сидят в башнях из слоновой кости; и сегодня мы узнаём от Google почти столько же, сколько он извлек из нашей предыдущей работы.

Секреты «Десятки» не в самих уравнениях, а в умении их использовать и комбинировать. Если их применять бездумно, они ничего не решат.

Риск для человечества исходит не от враждебного ИИ, который захватит наш мир, — дворецкого Джарвиса из вселенной Marvel или «Саманты», соблазняющей каждого мужчину на планете в фильме «Она». Искусственный интеллект для этого недостаточно умен. Он застревает в своих ограниченных решениях. Риск скорее заключается в растущем разрыве между теми, кто обладает властью

над данными, и теми, кто ею не обладает. Небольшая группа людей, знающих уравнения, имеет уровень умственного развития, какого на планете еще не было.

У тех, кто разбирается в математике, есть преимущество. Два аспиранта создали поисковую систему на основании уравнения влияния (уравнение 5). Три разработчика Google создали нейронную сеть, которая заставляет миллионы людей смотреть отупляющие видеоролики и рекламу. Раз за разом повторяется одна и та же схема: небольшое количество программистов, финансистов и игроков используют математику, чтобы доминировать над другими; иными словами, небольшая элитная группа математиков контролирует жизни тех, кто не может или не желает изучать этот код.

Не неся ответственности за свои действия, «Десятка» преобразует все аспекты мира. Равнодушная к ограничениям, она ищет оптимальные ответы на все задачи. «Десятка» может не осознавать себя, но свидетельства ее существования неоспоримы.

Теперь, когда мы знаем, как работают девять из десяти уравнений, а также силу и ограничения каждого из них, возможно, сумеем ответить на самый важный вопрос. Это тайное математическое общество, управляющее миром, — добро или зло?

Лично я сделался богаче, умнее и успешнее, следя «Десятке», но стал ли я лучше как человек?

Глава 10

Универсальное уравнение

Если... то...

Я набрал в телефоне запрос: «Кто в этом сезоне лучше — Месси или Криштиану Роналду?»

Я посмотрел на Людвига, Олофа и Антона, которые стояли перед экраном ноутбука. Людвиг нервно переступил с ноги на ногу: сначала тестируется его часть кода. Сможет ли футбольный бот преобразовать мою английскую фразу в язык, который ему понятен?

На экране побежал текст — результат работы мозга этого бота. Мой вопрос стал таким:

{цель: сравнить; контакт: {Криштиану Роналду, Месси};
настроение: нейтральное; промежуток времени: сезон}.

Бот справился! Он понял, что я имею в виду. Теперь очередь Олофа нервничать. Он должен был смоделировать игровые качества. Я не определил интересующий меня промежуток времени, но бот по умолчанию взял недавние матчи. Алгоритм Олофа может классифицировать характеристики как «плохие», «средние», «хорошие» и «отличные». Сейчас бота попросили оценить и сравнить двух отличных игроков.

Бот решил взять {мера: удары; турнир: Лига чемпионов}, чтобы сообщить о произведенных ударах и забитых голах в единственном турнире, где они играли оба. Мы могли видеть ответы на экранах:

бот определил, кого из игроков он считает лучшим. Теперь осталось отправить эту информацию обратно на мой телефон — не в виде фигурных скобок, двоеточий и краткого текста, а в виде фраз, которые я мог бы читать.

Антон, который должен был конструировать ответы, сказал: «Существует больше 100 тысяч фраз, которые он может сказать. Различные способы составления предложений и выбора слов. Мне интересно, что он выберет».

Я смотрел на свой телефон. Потребовалось время. Нам определенно нужно поработать над пользовательским интерфейсом...

Наконец он ответил: «Из этих двух игроков я считаю лучшим Лионеля Месси. Он забил шесть раз и имеет отличные точки ударов в этом сезоне». Бот прислал мне ссылку на карту ударов со всеми ударами и голами футболиста в Лиге чемпионов. Конечно, по упоминанию «точек ударов» видно, что писал бот, но в такой формулировке было что-то очаровательное. И по сути моего вопроса он дал правильный ответ.

*

Студенты сконструировали этого футбольного бота частично с помощью математики, с которой мы познакомились в этой книге. Людвиг использовал уравнение обучения, чтобы научить программу понимать вопросы о футболе. Олоф применил уравнение умений, чтобы оценивать игроков, а уравнение суждений — чтобы их сравнивать. Затем Антон связал всё воедино с помощью одного финального уравнения — «Если... то...».

Прежде чем мы сосредоточимся на нем*, я хочу посмотреть, на какой стадии нашего путешествия в математику мы находимся. Немного подумаем о том, что мы узнали.

Понимание уравнений может происходить на разных уровнях. Вы способны погружаться в их математические глубины, чтобы разобраться, как именно они работают и как их использовать. Если ваша цель — стать специалистом по обработке данных или статистиком в Snapchat, баскетбольном клубе или инвестиционном

* Это также не уравнение в обычном смысле слова.

банке, вам необходимо совершить такое погружение. Эта книга — только начало.

Вы можете использовать десять уравнений иначе — менее техническим способом, мягче. Вы можете применить их для управления процессом принятия решений и для собственного представления о мире. Я полагаю, что вы можете использовать десять уравнений, чтобы стать лучше.

В западном мышлении исходные «если... то...» содержались в десяти заповедях. *Если сегодня день субботний, то свято чти его. Если у тебя будут другие боги, то ты не предстанешь предо мной. Если у твоего соседа привлекательная жена, то ты не должен желать ее.* И так далее. Проблема заповедей в том, что они негибкие и через несколько тысячелетий кажутся несколько устаревшими.

Девять уравнений, изученных нами, совсем другие. Они не устанавливают правил, что вам нужно делать в тех или иных ситуациях. Они предлагают определенный подход к жизни. Помните, как Эми в туалете услышала злые слова Рэйчел? Или как парадокс дружбы помог нам увидеть, что мы напрасно жаждем социального успеха других людей? Или когда вы раскладываете друзей по типам с помощью уравнения рекламы? В каждой из этих ситуаций я не говорил действующим лицам, что им следует делать на основании какого-то заранее определенного морального компаса. Я рассмотрел данные, определил корректную модель и пришел к разумному выводу.

Десять уравнений дают больше гибкости, чем десять заповедей. Они могут справляться с гораздо более широким кругом проблем и дать более детальные советы. Ставлю ли я десять уравнений выше десяти заповедей? Да, конечно. У нас были тысячулетия, чтобы развивать мышление: со времен исходной «Десятки» мы нашли способы лучше думать о проблемах. Я ставлю десять уравнений выше не только христианства, но и многих других подходов к жизни. То, как данные и модели сочетаются между собой, противостоя бесмыслице, придает математике чистую честность, которая позволяет ей подняться над любыми другими способами мышления.

Математическое знание — как дополнительный уровень интеллекта. Я также считаю (хотя это более спорно), что изучить десять

уравнений — наша моральная обязанность. Я даже полагаю, что в целом уже проделанная работа участников «Десятки» благо для человечества. Не всегда, но чаще да. Изучая эти уравнения, вы помогаете не только себе, но и другим.

Такой вывод может показаться неожиданным, если учесть, что участники «Десятки» часто имеют преимущество перед теми, кто не обладает такими навыками. Может также показаться, что этот вывод противоречит философской позиции проверяемости, описанной Алфредом Айером, который говорил, что нельзя ожидать, чтобы в математике нашлись разумные ответы на моральные вопросы. Но вот во что я верю и что буду сейчас отстаивать. «Десятка» — это сила добра.

*

Чтобы выяснить, где в математике можно найти мораль, сначала нужно понять, где ее найти нельзя. С помощью метода исключения мы должны определить, что именно в математическом мышлении может подсказать нам, какие поступки «правильные».

Наше последнее уравнение «Если... то...» не одно уравнение, а сокращенное наименование множества алгоритмов, которые можно записывать в виде последовательности операторов «если... то...» и циклов «повторять... пока». Такие операторы — основа программирования. Например, в футбольном боте, созданном Антоном, мы найдем команды вроде:

```
если ключевые передачи > 5,  
то печатать («Он сделал много важных передач»)
```

Такого рода команды в сочетании со входными данными определяют, что будет на выходе.

В 1950-х, 1960-х и 1970-х зарождавшаяся информатика обнаружила целый ряд алгоритмов для обработки и организации данных. Одним из самых первых примеров был алгоритм сортировки слиянием, впервые предложенный Джоном фон Нейманом в 1945 году для сортировки чисел или по алфавиту. Чтобы понять, как он работает, сначала подумайте, как объединить два уже отсортированных

списка. Например, у меня есть список букв в алфавитном порядке $\{A, G, M, X\}$ и другой $\{C, E, H, V\}$. Чтобы создать сортированный список, объединяющий оба компонента, мне нужно идти слева направо по обоим спискам и ставить букву, которая идет раньше по алфавиту, в новый, а затем выбрасывать ее из исходного.

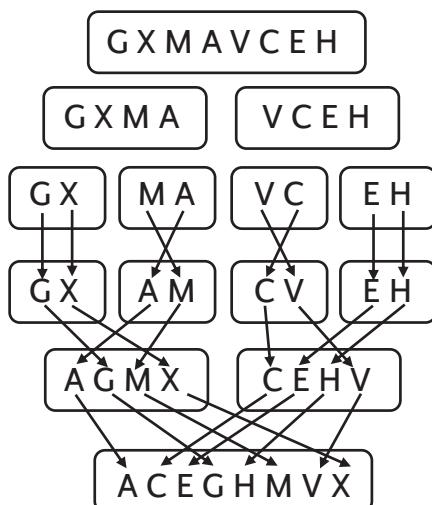
Попробуем. Сначала я сравниваю два первых элемента списков — A и C . Поскольку A в алфавите идет раньше, я помещаю ее в новый список и убираю из текущего. Теперь у меня три списка: новый $\{A\}$, первоначальные $\{G, M, X\}$ и $\{C, E, H, V\}$. Я снова сравниваю первые элементы исходных списков, G и C , и добавляю C в новый список. Он принимает вид $\{A, C\}$. Теперь я сравниваю G и E , отправляя в новый список E : $\{A, C, E\}$. Это повторяется, пока я не получу сортированный список $\{A, C, E, G, H, M, V, X\}$ и два пустых исходных.

Чтобы перейти от слияния уже отсортированных списков к сортировке произвольных, фон Нейман предложил стратегию, основанную на принципе «разделяй и властвуй». Полный список делится на более мелкие, и задача для каждого из них сводится к тому же методу — слиянию уже сортированных списков.

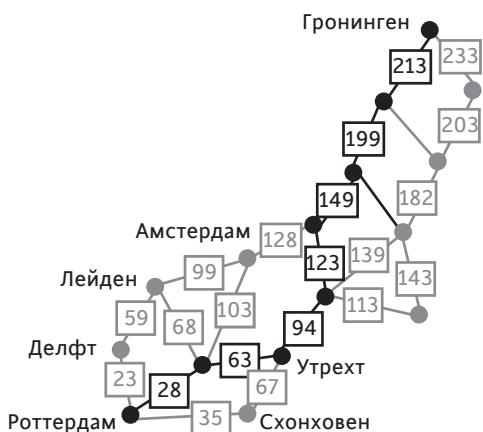
Предположим, мой первоначальный список таков: $\{X, G, A, M\}$. Сначала мы сливаем буквы $\{X\}$ и $\{G\}$, получая $\{G, X\}$, и буквы $\{A\}$ и $\{M\}$, получая $\{A, M\}$. Элегантность такого подхода в том, что на всех уровнях используется одна и та же техника. Разделив исходный список на достаточно мелкие части, мы в итоге приходим к списку, который гарантированно будет отсортирован, то есть к отдельным буквам. Затем, используя наше умение сливать два сортированных списка, мы гарантируем, что все создаваемые нами списки также будут отсортированы (см. рис. 10). Сортировка слиянием никогда не ошибается.

Еще один пример — алгоритм Дейкстры, определяющий кратчайший путь между двумя точками. Голландский физик и программист Эдсгер Дейкстра разработал свой алгоритм в 1953 году, чтобы продемонстрировать «некомпьютерным людям» (как он их называл), что компьютеры могут быть полезными при расчете самого быстрого маршрута между двумя нидерландскими городами¹²⁵. На придумывание алгоритма у него ушло всего двадцать

минут — он сидел в одном из амстердамских кафе. Позже он рассказывал журналу *Communications of the ACM**: «Одна из причин его красоты — отсутствие у меня карандаша и бумаги. Без них вы практически вынуждены избегать всех сложностей, которых можно избежать».



Сортировка слиянием работает по принципу «разделяй и властвуй». Сначала список букв делится на пары, которые сортируются. Затем они сливаются в сортированные списки из четырех букв и т. д. — пока не появится полностью отсортированный список



В алгоритме Дейкстры для нахождения кратчайшего пути мы двигаемся от одного города к другому, находя по дороге кратчайший путь между ними.

Черная линия —
кратчайший маршрут
между Роттердамом
и Гронингеном. Числа
указывают время
поездки в минутах

Рис. 10. Иллюстрация сортировки слиянием и алгоритма Дейкстры

* Ассоциация вычислительной техники (Association for Computing Machinery) — международная организация в сфере компьютерной техники.

Представьте, что вы хотите проехать из Роттердама в Гронинген. Алгоритм Дейкстры предписывает вам сначала определить время поездки от Роттердама до всех соседних с ним городов. Процесс показан на рисунке 10. Например, Делфт получит 23 минуты, Гауда — 28, а Схонховен — 35. Следующий шаг — рассмотреть все соседние для этих трех городов и найти кратчайшее время пути до них. Например, от Гауды до Уtrechta 35 минут, а от Схонховена до Уtrechta — 32 минуты; поэтому кратчайшее общее время пути до Уtrechta составляет $28 + 35 = 63$ минуты через Гауду (это меньше, чем $35 + 32 = 67$ минут через Схонховен). Алгоритм постепенно продвигается по территории Нидерландов, фиксируя кратчайшее расстояние до каждого города. Поскольку он каждый раз вычисляет кратчайший путь до любого города, то при добавлении нового города гарантировано, что до него тоже будет найден кратчайший путь. Алгоритм рассчитан не на конкретный Гронинген, он маркирует расстояния до всех городов; но когда в расчетах дело дойдет до него, алгоритм гарантирует, что нашел кратчайший путь.

Есть множество алгоритмов, похожих на сортировку слиянием Джона фон Неймана или нахождение кратчайшего пути¹²⁶. Вот несколько примеров: алгоритм Краскала для нахождения минимального остовного дерева (например, схема кратчайшей железнодорожной сети, соединяющей все города страны); расстояние Хэмминга для обнаружения разницы между двумя частями текста или данных; алгоритмы построения выпуклой оболочки для изображения фигуры, окружающей некоторое множество точек; алгоритмы обнаружения столкновения в 3D-графике; быстрое преобразование Фурье для обнаружения сигнала. Эти алгоритмы и их вариации — строительные блоки для аппаратного и программного обеспечения компьютеров. Они сортируют и обрабатывают наши данные, перемещают электронные письма, проверяют грамматику и позволяют Сири или Алексе за секунды идентифицировать передаваемую по радио песню.

Математика «Если... то...» всегда дает правильный ответ, и мы всегда знаем, что делать дальше. Возьмите, например, футбольного бота, созданного тремя студентами магистратуры. Я мог задавать ему простые вопросы о футболе, он отвечал; но Антона его ответы

не удивляли. Антон закодировал правила, которые определяют, что говорит бот, и программа четко следует им.

Почему я свалил в одно общее уравнение все алгоритмы «если... то...»? Потому что у них есть одно очень важное общее свойство: они универсальные истины. Алгоритм Дейкстры всегда найдет кратчайший путь; сортировка слиянием всегда отсортирует список имен от A до Z; выпуклая оболочка для множества точек всегда имеет одну и ту же конструкцию. Это утверждения, которые истинны независимо от того, что мы говорим или делаем.

В первых девяти главах этой книги мы использовали уравнения для проверки моделей, составления прогнозов и оттачивания нашего понимания реальности. Эти уравнения взаимодействуют с миром: мы позволяем прошлым данным информировать наши модели, а те прогнозируют будущие данные. Но алгоритмы «если... то...» негибкие. Они берут данные — скажем, список имен для сортировки или список точек, между которыми можно вычислить кратчайший путь, — и выдают ответ. Мы не пересматриваем свое знание мира на основе полученных ответов. Точно так же и наши наблюдения не влияют на истинность алгоритмов. Вот почему я называю их универсальными: их истинность доказана, они работают всегда.

Вышеприведенные примеры — алгоритмы из программирования, однако другие математические утверждения-теоремы из геометрии, анализа и алгебры также универсальные истины. В главе 5 мы видели пример парадокса дружбы. Сначала кажется невероятным, что наши друзья в среднем популярнее нас, но с помощью логических рассуждений мы показали, что это верно и иначе быть не может.

Математика полна удивительных результатов, которые сначала могут противоречить интуиции. Например, тождество Эйлера (названное в честь великого математика Леонарда Эйлера^{*}) $e^{\pi i} + 1 = 0$ говорит нам о связи между тремя хорошо известными вели-

^{*} Что тоже не совсем справедливо, поскольку в работах Эйлера этого равенства нет, хотя оно мгновенно следует из имеющейся там формулы $e^{ix} = \cos x + i \sin x$. Однако задолго до Эйлера Роджер Котс предложил геометрическое рассуждение, из которого следует эквивалентная формула $ix = \ln(\cos x + i \sin x)$.

чинами, основанием натуральных логарифмов $e = 2,71828\dots$, $\pi = 3,14159\dots$ и $i = \sqrt{-1}$. Этот факт так изящно сочетает фундаментальные константы, что его называют самой красивой формулой в математике¹²⁷.

Еще один пример — золотое сечение:

$$\varphi = \left(1 + \sqrt{5}\right) / 2 = 1,618\dots$$

Это число появляется, когда мы рисуем прямоугольник, который можно разрезать на квадрат и новый прямоугольник, подобный исходному. Иными словами, если сторона этого квадрата равна a , а стороны прямоугольника — a и b , то прямоугольник называется золотым, когда

$$\frac{a+b}{a} = \frac{a}{b} = \varphi.$$

Интересно, что число φ (фи) также появляется в связи с последовательностью чисел Фибоначчи: 1, 1, 2, 3, 5, 8, 13, 21, 34..., которая получается, когда мы складываем два предыдущих числа для получения следующего ($1 + 1 = 2$, $1 + 2 = 3$ и т. д.). Если мы рассмотрим отношение двух последовательных чисел Фибоначчи, то оно будет все сильнее приближаться к числу φ (например, $13/8 = 1,625$, $21/13 = 1,615\dots$, $34/21 = 1,619\dots$ и т. д.). Эти два примера — всего лишь начальная точка для путешествия в чистую математику, где повседневная интуиция начинает подводить и только строгие логические рассуждения могут вести вперед.

Огромное количество математических теорем, оказавшихся истинными, побудило французского математика Анри Пуанкаре написать в своей книге 1902 года «Наука и гипотеза»: «Если все предложения, которые математика выдвигает, можно вывести друг из друга по правилам формальной логики, то каким образом математика не сводится к бесконечной тавтологии? Логический вывод не может нас научить ничему существенно новому... Неужели мы можем допустить, что все эти теоремы, которые заполняют столько

томов, нужны только для того, чтобы окольным способом сказать, что A есть A ?» Вопрос Пуанкаре был риторическим, поскольку он считал, что проблемы, с которыми он и другие люди сталкиваются при разгадывании математических истин, должны содержать нечто более глубокое, нежели просто логические утверждения.

Аналогичная точка зрения в книге Дэна Брауна «Код да Винчи» — оригинальной, хотя и вымышенной теории математического заговора. В книге профессор Лэнгдон говорит: «Когда древние открыли число ϕ [φ], они были уверены, что наткнулись на строительный блок, который Бог использовал при создании мира... Загадочное волшебство, присущее божественной пропорции, было написано в начале времен». Лэнгдон переходит к примерам (некоторые верны, некоторые нет), как золотое сечение, которое он называет божественной пропорцией, можно увидеть в биологии, искусстве и культуре. На протяжении всей истории участники «Десятки» использовали ϕ в качестве кода, причем имя одного из главных персонажей романа — СоФИ Невё — содержало ключ.

Должен признать, что этот аспект математики меня очень привлекает. Я наслаждался «Кодом да Винчи». Есть что-то невероятное в обнаруживаемых нами неожиданных связях — не только в числах вроде ϕ , но и в алгоритме Дейкстры для кратчайшего пути, и в сортировке слиянием фон Неймана. Такая простая элегантность кажется выходящей за рамки обыденной реальности. Может, за всеми этими уравнениями скрыт какой-то глубокий код?

Правильный ответ на вопрос Пуанкаре гораздо более незатейлив и прямолинеен, чем он ожидал: «да». Все великие теоремы математики, как и все алгоритмы сортировки и организации из информатики, говорят всего лишь, что A равно A . Все они представляют собой одну колоссальную тавтологию; это очень полезные и неожиданные тавтологии, но тем не менее тавтологии. Пуанкаре был прав буквально и неправ риторически.

Рассуждение, которое утверждает, что Пуанкаре был прав, можно найти в книге «Язык, истина и логика». В ней Айер использовал пример с треугольником. Представьте, что друг рассказывает

вам о треугольнике, у которого сумма углов меньше 180 градусов¹²⁸. У вас есть два способа отреагировать: либо сказать ему, что он измерил неправильно, либо сообщить, что объект, который он имеет в виду, не треугольник. Ни при каких обстоятельствах вы не измените своего мнения о математических свойствах треугольника на основании данных вашего друга. Он не найдет в реальном мире треугольника, подрывающего устои геометрии.

И не существует списка слов любого языка, который невозможно отсортировать по алфавиту. Если я покажу вам список, где Айер стоит после Самптера, и скажу, что это — результат сортировки слиянием, то вы скажете, что либо мой алгоритм не работает, либо я не знаю алфавита. Он точно не станет доказательством, что сортировка слиянием не работает. Аналогично не существует компьютерной сети, в которой кратчайший путь длиннее, чем второй, по краткости.

К сожалению для профессора Лэнгдона, причина появления числа $\phi = 1,618\dots$ во всех этих геометрических и математических соотношениях состоит в том, что оно является собой положительный корень квадратного уравнения $x^2 - x - 1 = 0$. И изучение последовательности Фибоначчи, и нахождение золотого сечения приводят к решению одного и того же квадратного уравнения, поэтому дают один и тот же ответ. Нет никакого загадочного магического кода, скрытого в числе фи или в любом другом.

Точка зрения Айера: математические теоремы не зависят от данных. Математика непроверяема. Она состоит из тавтологических утверждений, которые можно доказать с помощью логики, но сами по себе они ничего не говорят о реальности. В ответ на риторический тон Пуанкаре Айер писал: «Способность логики и математики удивлять нас, как и ее полезность, зависит от ограниченности нашего разума».

Пуанкаре вводил в заблуждение тот факт, что заниматься математикой порой трудно даже ему. По сути, математические результаты не зависят от наших наблюдений. Вот почему я говорю, что они универсальны. Они верны для всей Вселенной — независимо от того, что мы говорим и делаем, от научных открытий, от того, сделал их Пуанкаре или другой математик.

Как мы видели в этой книге, сила десяти уравнений в том, что они взаимодействуют с реальным миром, комбинируя модели и данные. В отрыве от данных уравнения не имеют глубокого смысла. Они определенно не несут нам какой-то морали и не имеют ничего общего с Богом. Это просто набор крайне полезных результатов, которые оказались истинными.

Чтобы найти загадку и мораль в математике, нам придется поискать не в самих теориях, а в других местах.

*

Я все откладывал звонок Мариусу. Ян попросил, чтобы я позвонил и все уточнил у него, прежде чем публиковать финансовые подробности их операции со ставками, и я несколько опасался, что Мариус может ответить отказом. Возможно, ему хочется хранить все это в тайне от любопытных глаз.

Я напрасно беспокоился. Мариус был рад поговорить и выложил мне все подробности. Их прибыли продолжали расти, хотя ему приходилось сложно.

— Если изо дня в день смотреть на цифры, можно сойти с ума, — сказал он.

В какой-то момент у них был неудачный период, когда друзья потеряли 40 тысяч долларов.

— Это был худший момент за все время. Ты реально начинаешь сомневаться. Но мы продолжали работать и не торопились. И числа снова поднялись. Потом опять упали... и поднялись.

Мариус сказал, что азартные игры научили его быть более терпеливым и концентрироваться на вещах, которые он может изменить.

— Мы не можем управлять колебаниями. Я научился не смотреть матчи, как во время чемпионата мира. Сначала я проверял наши ставки в режиме реального времени. Сейчас просто прихожу в офис и работаю, а обзор мы смотрим раз в квартал.

— Вы когда-нибудь задумывались о моральной стороне того, что делаете? — спросил я. — О людях, которые теряют, когда вы выигрываете?

— Я думаю, что азартные игры — щекотливый вопрос из-за всей этой рекламы, которая обманывает, — ответил Мариус. — Но

в то же время всего лишь быстрый поиск в Сети по разным ставкам даст вам все, что нужно, чтобы оставаться в плюсе как любителю. Просто люди не хотят себя утруждать.

Он был прав. Это моральный урок из уравнения ставок. Если люди не хотят потратить несколько часов на поиск информации в интернете — почему Мариус должен за это отвечать? Мариус и Ян создали сайт, содержащий именно ту информацию, которая нужна этим людям для ставок у мягких букмекеров. Но очень мало кто делает ставки так, как они.

Я спросил Мариуса, что он будет делать, если рынки изменятся внезапно и он все потеряет.

— Никогда не знаешь, как все обернется. Такое всегда возможно, — ответил Мариус. — Но мне очень нравится то, что я делаю, и именно поэтому я счастлив. Мне не хотелось бы сидеть на пляже, пока какой-нибудь бот работает. Мне интересно копаться в данных и извлекать что-то оттуда.

Он обнаружил настоящий секрет «Десятки», и тот имел мало общего с количеством денег на его банковском счете. Его вознаграждением было количество новых знаний.

Можно ли это считать признаком морали? Думаю, да. Я считаю, что в подходе Яна и Мариуса есть интеллектуальная честность. Они не лгут о своих действиях. Они играют в игру по имеющимся правилам и выигрывают, потому что у них это получается лучше. То же можно сказать об операциях со ставками Уильяма Бентера и Мэтью Бенхэма в более крупном масштабе. Честность Бентера была поразительной. Все считали, что он осторегается раскрыть, сколько и где выиграл, однако он опубликовал статью о своих методах в научном журнале. Любой человек с математическими умениями сейчас может взять методы Бентера и применить их на практике.

Те, кто много работает, учится и проявляет настойчивость, — победители. Те, кто срезает углы, проигрывают. То же правило применимо ко всей «Десятке». Если выносим суждения, то вынуждены говорить, как данные сформировали наши убеждения. Когда строим модель умений, должны формулировать свои предположения. Когда инвестируем или делаем ставку — признаём свои прибыли и убытки, чтобы улучшить свою модель. Мы должны сообщать

друг другу, насколько уверены в своих выводах. Мы вынуждены признаться, что не мы центр нашей социальной группы, и нам не стоит жалеть, что мы не самые популярные. Когда мы видим какую-то корреляцию, вынуждены искать причинно-следственную связь. Когда создаем технологию, нам демонстрируют, как она вознаграждает и наказывает людей, которые ею пользуются. Такова суровая мораль математики. В итоге истина всегда побеждает.

Участники «Десятки» — настоящие защитники интеллектуальной честности. Они излагают свои предположения, собирают данные и сообщают ответы. Когда они не знают полного ответа, то говорят нам, чего не хватает. Они перечисляют правдоподобные альтернативы и указывают вероятность каждой из них; они начинают думать о следующих шагах, которые можно предпринять, чтобы узнать больше.

Верните честность в свою жизнь. Десять уравнений вам в помощь. Начните с размышлений о вероятностях — и с точки зрения азартных игр, когда вы стремитесь получить желаемое, и с точки зрения понимания риска неудачи. Прежде чем делать выводы, улучшайте свои суждения, собирая данные. Повышайте свою уверенность, не убеждая себя в своей правоте, а крутя колесо много раз. Каждый урок, которому учат нас уравнения — от обнаружения фильтра, созданного нашей социальной группой, до понимания того, как социальные сети ведут нас к критической точке, — снова и снова напоминает о важности честного подхода к модели и использования данных для улучшения нас самих.

Если вы будете следовать этим уравнениям, то заметите, как окружающие начнут уважать ваши суждения и ваше терпение. Это первый смысл, в котором математика может быть источником морали. Она сообщает нам суровые истины о нас и о людях вокруг.

*

В биографии Алфреда Айера, написанной Беном Роджерсом, излагается история о встрече философа с боксером Майком Тайсоном в 1987 году¹²⁹. Айер, которому тогда было семьдесят семь, находился на вечеринке на 57-й улице на Манхэттене, когда в комнату вбежала какая-то женщина со словами, что в спальне набросились

на ее подругу. Айер обнаружил, что Тайсон пытался приставать к молодой Наоми Кэмпбелл (будущей супермодели).

Роджерс пишет, что Айер сказал Тайсону прекратить. Боксер ответил:

— Черт возьми, ты знаешь, кто я? Чемпион мира в тяжелом весе.

Айер стоял на своем.

— А я бывший профессор логики в Оксфорде. В своих областях мы оба знамениты. Предлагаю поговорить об этом как разумные люди.

Тайсон, который, видимо, оказался любителем философии, был достаточно впечатлен и отступил.

Впрочем, если бы Тайсон захотел ударить Айера интеллектуальным апперкотом, он мог бы спросить его, на каком основании философ считает оправданным вмешиваться в его заигрывания с Наоми Кэмпбелл. В конце концов, как утверждал сам Айер в книге «Язык, истина и логика», мораль лежит вне рамок эмпирической дискуссии. Хотя вполне могло быть так, что Кэмпбелл боялась Тайсона. Железный Майк мог спросить: «Существуют ли логические причины, по которым мужчинам ненормально навязывать себя женщинам при их собственном сексуальном влечении?»

Айеру пришлось бы признать, что он навязывает общепринятые нормы того типа собраний, на котором они присутствовали в данный момент. К этому Тайсон мог бы добавить, что его нормы, приобретенные в криминальной юности на улицах Бруклина, отличались от норм Айера, получившего образование в Итоне; а это оставляло им мало общих позиций для продолжения дискуссии. «И если вы не возражаете, — мог бы продолжить Железный Майк, — то я хотел бы вернуться к ухаживанию за этой красивой дамой тем способом, который считаю наиболее подходящим».

Не могу знать, в этом ли направлении шел их разговор, но точно известно, что во время их беседы Наоми Кэмпбелл сбежала с вече-ринки. Майка осудили за изнасилование другой женщины спустя четыре года, и теперь он сексуальный преступник.

Эта история Тайсона и Айера иллюстрирует (среди прочего) фундаментальную проблему для любого последователя строго

позитивистского подхода. Невозможно решить даже самые очевидные моральные дилеммы. Несмотря на честность математики и логического мышления, каждый человек должен сам разбираться со своей нравственностью.

Очевидно, здесь в логическом позитивизме чего-то не хватает. Вопрос — чего? Чтобы отточить наши представления о роли математики в морали, английский философ Филиппа Фут разработала в 1967 году мысленный эксперимент, который стал известным под названием «проблема вагонетки»¹³⁰. Его можно описать так:

Эдвард управляет вагонеткой, у которой отказали тормоза. На рельсах перед ним пять человек, и насыпи настолько крутые, что люди не успеют сойти с пути. У рельсов есть ответвление вправо, и Эдвард может повернуть вагонетку туда. К сожалению, на правом пути тоже есть человек. Эдвард может либо повернуть вагонетку и убить одного человека, либо не поворачивать и убить пятерых.

Встает вопрос — что делать Эдварду: повернуть и убить одного либо ехать прямо и убить пятерых. Немного поразмыслив, большинство из нас обычно решают, что первый вариант лучше. Одно убийство лучше, чем пять. Пока все хорошо.

Теперь рассмотрим другой вариант проблемы вагонетки, предложенный в 1976 году Джудит Томсон — профессором философии из Массачусетского технологического института:

Джордж стоит на пешеходном мостике над рельсами для вагонетки. Они видят, что одна вагонетка приближается, а за мостом на рельсах стоят пять человек, и насыпи настолько крутые, что люди не успеют сойти с пути. Джордж понимает, что есть один способ остановить неуправляемую вагонетку: сбросить перед ней что-нибудь тяжелое на рельсы. Однако единственный имеющийся достаточно тяжелый объект — какой-то толстяк, который тоже смотрит на вагонетку с моста. Джордж может либо столкнуть его на рельсы перед вагонеткой, убив его при этом; либо не делать этого и позволить умереть пятерым.

Что делать Джорджу? С одной стороны, сбрасывать человека на рельсы явно неправильно. С другой — при бездействии он позволяет погибнуть пяти человекам (как если бы Эдвард не повернул вагонетку).

В опросе примерно 81% американцев сказали, что на месте Эдварда отклонили бы вагонетку и убили одного человека; однако всего 39% считали, что Джорджу нужно столкнуть толстяка и спасти пятерых¹³¹. Китайские и российские респонденты также были более склонны полагать, что Эдварду нужно действовать, а Джорджу — нет. Это поддерживает гипотезу, что у людей есть общее моральное представление для таких дилемм¹³². Однако культурные различия всё же имеются: китайцы скорее позволяют вагонетке двигаться своим курсом в обеих ситуациях.

Джудит Томсон разработала этот второй вариант проблемы вагонетки, чтобы дилемма стала прозрачной¹³³. Обе проблемы описывают одну математическую задачу — спасти пятерых или спасти одного; однако наша интуиция говорит, что они сильно отличаются друг от друга. Математическое решение простое — моральное гораздо сложнее. Проблемы вагонетки заставляют нас задуматься о действиях, которые мы готовы — или не готовы — предпринять для спасения жизней.

Проблема вагонетки лежит в центре многих произведений современной научной фантастики. Эта дилемма занимает от тридцати минут до часа в фильме (внимание: спойлер) «Мстители: война бесконечности» из вселенной Marvel. Философски настроенный суперзлодей Танос, став свидетелем того, как перенаселение полностью разрушило его родную планету, решает, что неплохо бы уничтожить половину населения вселенной. Он считает, что убийство половины людей сейчас спасет больше жизней в будущем, и — если рассуждать в терминах задачи Томсон — решает щелчком пальца сбросить миллиарды толстяков под вагонетку. В продолжении — «Мстители: финал» — Тони Старк сталкивается с более личной дилеммой того же типа: существование его дочери или возвращение друзей. Его просят сделать практически невозможный выбор.

Обычно в научной фантастике злодеи предпочитают столкнуть толстяка. Во многих случаях такое решение изображается в форме жесткой беспощадной логики. Роботы или ИИ принимают утилитарное решение спасти пятерых, а не одного — независимо от того, насколько ужасно действие, необходимое для достижения их цели. Для прагматичного робота, запрограммированного людьми на спасение максимального количества людей, в приоритете числа, а не чувства. При прочих равных условиях пользы от пятерых больше, чем от одного. Это точно тот же вопрос, который Филиппа Фут и Джудит Томсон иллюстрировали проблемой вагонетки: неверно считать, что мы можем решить такие проблемы с помощью утилитарного подхода. В фильмах роботы ошибаются, а если бы существовали в реальной жизни — ошибались бы тоже. Научная фантастика напоминает нам, что мы не можем создать универсальное правило, например

если $5 > 1$, то *печатать* («спасти 5»),

чтобы решить все проблемы. Если мы так сделаем, то совершим самые ужасные моральные преступления, которые никогда не сможем оправдать перед будущими поколениями.

Когда я был моложе, мог смотреть на отказ от действия — даже сталкивания толстяка на рельсы — как на указание логической слабости части человечества.

С моей стороны это было ошибкой, и не только потому, что я оказался слишком строг к человечеству. По сути, делать такой вывод — логическая слабость. Проблема вагонетки иллюстрирует два момента. Во-первых, она подтверждает, что на вопросы реальной жизни нет чисто математических ответов. Это то же, что Айер отвечал на риторический вопрос Пуанкаре об «универсальной» природе математики. Это причина того, что кода да Винчи не существует. Наше ощущение универсальности математики — результат ее тавтологической природы, а не более глубокой истинности. Мы не можем использовать ее как божественные заповеди. Только — как в этой книге — в качестве инструмента для организации своей работы вокруг моделей и данных.

Во-вторых, дилемма вагонетки сообщает нам, что чистый утилитаризм — идея, что мораль можно построить вокруг попытки максимизировать счастье, жизнь или любую другую переменную — одно из величайших зол среди всех (одинаково неверных) моральных кодексов, доступных нам¹³⁴. Правило «спасти как можно больше жизней» противоречит нашей нравственной интуиции и заставляет нас делать ужасные вещи. Если мы начинаем разрабатывать оптимальные моральные кодексы, то в итоге создаем моральный лабиринт.

Я пришел к выводу, что на эти дилеммы есть очень простой ответ: мы должны научиться доверять своей нравственной интуиции и пользоваться ею. Это и сделал Алфред Айер при стычке с Майком Тайсоном. То же сделала Моя Бёрсдел, когда решила изучать расизм, увидев, как нацисты преследуют ее друзей. Это то, чем руководствовался я, когда изучал дела компании Cambridge Analytica, фейковые новости и предвзятость алгоритмов или руководил написанием диссертации Бьорна по иммиграции и подъеме правых сил в Швеции. Это то, чем руководствовалась Николь Нисбетт, когда рассматривала коммуникацию в политике. И именно это делает Человек-паук: с помощью своей интуиции и умений убивает плохого парня.

Проблема вагонетки говорит нам, что нужен более мягкий способ размышлений для таких моральных и философских дилемм — дополняющий жесткую брутальную честность использования моделей и данных.

Участники «Десятки», которые вносят наибольший вклад в общество, мыслят и мягко (используя свою нравственную интуицию, чтобы определить, какие проблемы им решать), и жестко (сочетая модели и данные ради честности своих ответов). Они прислушались и поняли ценность окружающих. Они осознают, что не более любого другого человека компетентны, чтобы выбирать проблемы, но более компетентны в умении их решать. Они служители общества, сохранившие тот дух, что ввел в «Десятку» Ричард Прайс почти 260 лет назад. Прайс был неправ насчет чудес¹³⁵, но верно указал на необходимость нравственности, когда мы применяем математику.

У меня нет убедительных доказательств, но я полагаю, что после того, как логический позитивизм отбрасывает идею универсального утилитаризма, у нас остается направляющая нас моральная интуиция. Именно этот мягкий способ мышления говорит нам, какие проблемы нужно решать.

*

Участникам «Десятки» нужно разговаривать, учиться справляться с силой, нам данной, — точно так же, как Человек-паук осознаёт свои слабости в каждом новом воплощении.

Быть мягким означает, что бездумно приумножать деньги невежественных банкиров, занимающихся инвестициями, ненормально. Это значит, что нам не следует патентовать основные уравнения, чтобы зарабатывать деньги. Мы должны продолжать открыто рассказывать об используемых алгоритмах, делиться всеми своими секретами с теми, кто готов потрудиться, изучая их.

Мы должны использовать свою интуицию, чтобы двигаться к важным вопросам. Нам стоит прислушиваться к чувствам других людей и выяснить, что для них важно. Многие из нас уже делают это, но нам нужно быть открытыми в отношении того, кто мы есть и почему делаем то, что делаем.

Нам надо быть мягкими при определении своих проблем и жесткими в их решении.

*

Я сижу в комнате для семинаров на цокольном этаже математического факультета одного из университетов на севере Англии. Перед нами Виктория Спейсер, политолог из Лидского университета; она представляет сегодняшних докладчиков. Вместе со своим коллегой по исследованиям и спутником жизни Ричардом Манном она организовала двухдневный семинар об использовании математики в сфере гражданской активности. Цель — свести вместе математиков, специалистов по данным, представителей власти и бизнеса и найти способы использования математических моделей, чтобы сделать мир лучше.

Впервые я встретил Викторию почти восемь лет назад, Ричарда — чуть раньше. Вместе еще с одним моим аспирантом — Шьямом Ранганатаном — мы работали над моделированием расовой сегрегации в школах Швеции¹³⁶, демократических преобразований в странах мира и разработкой методов, с помощью которых ООН может достичь своих зачастую противоречивых целей устойчивого развития. Мы не всегда говорили об этом вслух, но всегда втайне считали, что математика должна не просто изучать мир, а менять его к лучшему. Выбор слова «активность» в названии собрания — наша первая попытка открыто заявить о наших целях.

Мы не одиноки. После того как Виктория открывает семинар, участники встают один за другим и говорят, чем они занимаются. Адам Хилл из британской компании DataKind создал сеть, показывающую связи между людьми, сидящими в правлениях анонимных компаний, которые созданы в Великобритании для сокрытия владения. Сопоставляя права собственности, он и его команда могут обнаружить коррупцию и потенциальное отмывание денег. Бетти Наньонга из Уганды рассказала нам, как ее коллеги используют математические модели, чтобы понять причины студенческих забастовок в Университете Макерере, Кампала, где она работает. Анна Оуэн, научный сотрудник Лидского университета, показала, что Грета Тунберг была права, когда заявила, что Великобритания лгала, говоря о сокращениях выбросов CO₂¹³⁷. Анна демонстрирует нам правильный расчет, учитывающий производство и транспортировку всех пластиковых товаров, которые мы импортируем из Китая. Среднестатистический человек в возрасте 60–69 лет — летая в отпуск или управляя большой машиной — создает в год на 64% больше выбросов CO₂, чем молодой человек до 30 лет. Именно старшее поколение, некоторые представители которого критикуют Грету Тунберг, должно тщательнее всего думать об углеродных выбросах.

Возможно, вы не знали о нас до сих пор, но теперь знаете.

Секрет раскрыт. «Десятка» — это мы.

Примечания

Глава 1

1. Эта статья была опубликована на платформе Medium: medium.com/@Soccermatics/if-you-had-followed-the-betting-advice-in-soccermatics-you-would-now-be-very-rich-1f643a4f5a23. Полное описание модели можно найти в моей книге «Футболоматика. Как благодаря математике “Барселона” выигрывает, Роналду забивает, а букмекеры зарабатывают состояния». Издана на русском языке (М. : Бомбара, 2018).
2. Каждая ставка умножает ваш капитал на 1,0003 (увеличение в 0,03% на ставку). Если вы делаете по 100 ставок в день, ваш ожидаемый капитал к концу года составляет $1000 \times 1,0003^{100 \times 365} = 56\,860\,593,80$.
3. Коэффициенты букмекера справедливы, если коэффициент на событие, умноженный на коэффициент против события, даст 1. Например, когда коэффициент на победу равен $3/2$, а на ничью или поражение $2/3$ — поскольку $3/2 \times 2/3 = 1$. Однако на практике букмекеры никогда не предложат справедливые коэффициенты. Например, в вышеприведенном примере они скорее дадут $7/5$ на победу фаворита и $4/7$ против, так что $7/5 \times 4/7 < 1$. Маржа букмекера в этом случае составит $1/(1 + 7/5) + 1/(1 + 4/7) - 1 \approx 0,05$ или 5%.
4. Ваша ожидаемая прибыль на одну ставку составляет:

$$\frac{2}{5} \times \frac{7}{5} + \frac{3}{5} \times (-1) = \frac{14}{25} - \frac{15}{25} = -\frac{1}{25},$$

то есть 4 пенса потеря на ставку в 1 фунт.

5. Даже после пяти неудачных попыток вам не следует слишком отчаиваться. Если вероятность успеха на одном

собеседований составляет $1/5$, вероятность того, что вам придется пройти минимум 5 собеседований, чтобы получить какую-то работу, составляет $1 - (1 - 1/5)^5 = 67\%$.

6. Benter W. Computer based horse race handicapping and wagering systems: a report // D. B. Hausch, V. S. Y. Lo, W. T. Ziemba, eds. Efficiency of Racetrack Betting Markets. Revised edn. Singapore: World Scientific Publishing Co. Pte Ltd, 2008. Pp. 183–198.
7. Chellel K. The gambler who cracked the horse-racing code // Bloomberg Businessweek, 3 May 2018 // bloomberg.com/news/features/2018-05-03/the-gambler-who-cracked-the-horse-racing-code.
8. Bolton R. N., Chapman R. G. Searching for positive returns at the track: a multinomial logit model for handicapping horse races // Management Science. 1986. August. Vol. 32. No. 8. Pp. 1040–1060.
9. Cox D. R. The regression analysis of binary sequences // Journal of the Royal Statistical Society: Series B (Methodological). 1958. Vol. 20. No. 2. Pp. 215–232.

Глава 2

10. Не следует считать, что это число — точное значение. Отчет Управления гражданской авиации «Мировая статистика катастроф со смертельным исходом, 2002–2011», CAP 1036 (июнь 2013), дает оценку 0,6 катастрофы со смертельным исходом на миллион совершенных полетов, не считая нападений террористов, для рейсов с 2002 по 2011 год. При этом в катастрофе со смертельным исходом погибают не все люди, и статистика отличается для разных стран, так что точную величину определить невозможно. В любом случае это число мало даже в масштабе миллионных долей.
11. Этот результат противоречит интуитивным представлениям, но математически он верен. Чтобы дополнительно убедить себя, предположим, что $\theta = 0,98$, а истинная вероятность восхода составляет 98%. В этом случае кажется неудивительным, если бы солнце вставало во все 100 дней его наблюдения. Вероятность 100 восходов подряд в этом случае равна $0,98^{100} \approx 13,3\%$. Число небольшое, но и не пренебрежимо малое. Ту же

логику можно приложить к числу $\theta = 0,985 \cdot (0,985^{100} \approx 22,1\%)$ и к другим значениям $\theta < 0,99$. Хотя более вероятно, что значение θ превосходит 99%, вполне правдоподобно, что оно меньше 99% (точнее, вероятность — 36,2%).

12. Юм Д. Исследование о человеческом разумении / Пер. С. И. Церетели. М. : РИПОЛ классик, 2017.
13. Рассуждения из этого абзаца взяты из работы Джеймса Оуэна «Юм против Прайса о чудесах и априорных вероятностях: свидетельство и байесовские вычисления» (Hume *versus* Price on miracles and prior probabilities: testimony and the Bayesian calculation // *Philosophical Quarterly*. 1987. April. Vol. 37. No. 147. Pp. 187–202).
14. Это вычисление предоставляется заинтересовавшемуся читателю. Используйте в помощь примечание 2 выше.
15. Аналогично. Используйте примечание 2 выше.
16. Zebrowski M. K. Richard Price: British Platonist of the eighteenth century // *Journal of the History of Ideas*. 1994. January. Vol. 55. No. 1. Pp. 17–35.
17. Price R. *Observations on Reversionary Payments... To Which Are Added, Four Essays on Different Subjects in the Doctrine of Life-Annuities... A New Edition, With a Supplement, etc.*, Vol. 2. London: T. Cadell, 1792.
18. Poitras G. Richard Price, miracles and the origins of Bayesian decision theory // *European Journal of the History of Economic Thought*. 2013. February. Vol. 20. No. 1. Pp. 29–57.
19. Price R., Turgot A.-R.-J. *Observations on the Importance of the American Revolution, and the Means of Making it a Benefit to the World*. London: T. Cadell, 1785.
20. Vernon I., Goldstein M., Bower R. G. Galaxy formation: a Bayesian uncertainty analysis // *Bayesian Analysis*. 2010. Vol. 5. No. 4. Pp. 619–669.
21. Carter C. Is screen time toxic for teenagers? // *Greater Good Magazine*, 27 August 2018 // greatergood.berkeley.edu/article/item/is_screen_time_toxic_for_adolescents.
22. Odgers C. L. Smartphones are bad for some adolescents, not all // *Nature*. 2018. February. Vol. 554. No. 7693. Pp. 432–434.

23. Этот результат исходно был получен при исследовании британских подростков, см.: Przybylski A. K., Weinstein N. A large-scale test of the Goldilocks hypothesis: quantifying the relations between digital-screen use and the mental well-being of adolescents // Psychological Science. 2017. January. Vol. 28. No. 2. Pp. 204–215.

Глава 3

24. Уравнение для кривой нормального распределения, которое де Муавр написал (в логарифмической форме) во втором издании своей книги по вероятности в 1738 году, таково:

$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

где μ — среднее, а σ — стандартное отклонение.

25. Третью, последнюю версию работы можно найти на Google Books. De Moivre A. The Doctrine of Chances: Or, A Method of Calculating the Probabilities of Events in Play. The Third Edition. London: A. Millar, 1756.
26. Найдем вероятность двух тузов в пятикарточной руке. Вероятность того, что первой картой вам сдали туз, равна $4/52$. Вероятность того, что после этого второй картой тоже сдали туз, равна $3/51$ (в колоде осталась 51 карта, из которых три — тузы). Следующие три карты должны оказаться не тузами, вероятности чего равны $48/50$, $47/49$ и $46/48$. Эти события должны произойти одновременно, поэтому их вероятности нужно перемножить. Однако мы учли пока всего лишь одно размещение двух тузов среди пяти карт (когда они находятся на первом и втором месте). Всего таких комбинаций 10, поэтому общая вероятность равна:

$$10 \cdot \frac{4 \cdot 3 \cdot 48 \cdot 47 \cdot 46}{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48} = \frac{259440}{6497400} = \frac{2162}{54145} \approx 4\%.$$

27. Walker H. M. De Moivre on the law of normal probability, 2006 // semanticscholar.org/paper/DE-MOIVRE-ON-THE-LAW-OF-NORMAL-PROBABILITY-Walker/d40c10d50e86f0ceed1a059d81080a3bd9b56ffd#citing-papers>.

28. Обзор истории центральной предельной теоремы можно найти в работе: Le Cam L. The central limit theorem around 1935 // Statistical Science. 1986, Vol. 1. No. 1. Pp. 78–91.
29. Статистика взята с сайта stats.nba.com/search/team-game/.
30. Just R. E., Weninger Q. Are crop yields normally distributed? // American Journal of Agricultural Economics. 1999. May. Vol. 81. No. 2. Pp. 287–304.
31. Сильвер Н. Сигнал и шум. Почему одни прогнозы сбываются, а другие — нет. М. : КоДибри, 2015.
32. $\sigma^2 = \frac{1}{3} \cdot (0 - (-1))^2 + \frac{2}{3} \cdot \left(0 - \frac{1}{2}\right)^2 = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}$. Поэтому стандартное отклонение $\sigma = 0,71$.
33. Эти величины дают нам число наблюдений, которое позволяет на 97,5% (а не на 95%) быть уверенными, что мы не сделаем ошибку и на самом деле $h = 0$ или меньше. Надежность в 97,5% возникает, поскольку наш 95%-й доверительный интервал покрывает и нижний, и верхний пределы для h . Существует также 2,5%-й шанс, что мы недооценили наше преимущество и оно больше, чем дает наш доверительный интервал. Но при оценке выгодности игры недооценка преимущества роли не играет, так что значение имеют только те 2,5%, что мы переоценили свое преимущество. Для этих величин также предполагается, что преимущество положительно, $h > 0$. Однако тот же результат получится при отрицательном преимуществе, если вместо h поставить $-h$.
34. Я проверил среднеквадратичное отклонение у нескольких гостиниц и обнаружил, что оно часто чуть меньше 1 — например, 0,8. Но для примерной оценки нам достаточно взять 1.
35. Arai M., Bursell M., Nekby L. The reverse gender gap in ethnic discrimination: employer stereotypes of men and women with Arabic names // International Migration Review. 2016. Vol. 50. No. 2. Pp. 385–412.
36. Bertrand M., Mullainathan S. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination // American Economic Review. 2004. September. Vol. 94. No. 4. Pp. 991–1013.

37. Bailey Z. D., Krieger N., Agénor M. et al. Structural racism and health inequities in the USA: evidence and interventions // *The Lancet*. 2017. April. Vol. 389. No. 10077. Pp. 1453–1463.
38. Вот эмпирическое правило, которое я считаю полезным, но оно требует математического обоснования. Для примера пусть среди населения в целом представлена доля p людей определенного типа (например, белых). Дисперсия максимальна при $p = 1/2$, так что для всех значений p она не превосходит $1/2(1 - 1/2) = 1/4$, и поэтому стандартное отклонение меньше $1/2$. Поскольку $1,96 \approx 2$, то это означает, что доверительный интервал для выборочной доли $p^* = 1,96 \frac{1/2}{\sqrt{n}} \approx 1/\sqrt{n}$. Отсюда и эмпирическое правило.
39. Этот вопрос обсуждается, например, в работе: Pearson K. Historical note on the origin of the Normal Curve of Errors // *Biometrika*. 1924, December. Vol. 16. No. 3–4. Pp. 402–404.
40. T. Zuberi, E. Bonilla-Silva, eds. *White Logic, White Methods: Racism and Methodology*. Lanham, MD: Rowman & Littlefield Publishers, 2008.
41. Staddon J. The devolution of social science // *Quillette*, 7 October 2018 // quillette.com/2018/10/07/the-devolution-of-social-science/.
42. Питерсон Дж. 12 правил жизни: Противоядие от хаоса. СПб. : Питер, 2019.
43. Например, в ноябре 2018 года в интервью для ток-шоу Skavlan, идущего на телевидении Норвегии и Швеции.
44. Auspurg K., Hinz T., Sauer C. Why should women get less? Evidence on the gender pay gap from multifactorial survey experiments // *American Sociological Review*. 2017. Vol. 82. No. 1. Pp. 179–210.
45. Moss-Racusin C. A., Dovidio J. F., Brescoll V. L. et al. Science faculty's subtle gender biases favor male students // *Proceedings of the National Academy of Sciences*. 2012. October. Vol. 109. No. 41. Pp. 16474–16479.
46. Bettinger E. P., Long B. T. Do faculty serve as role models? The impact of instructor gender on female students // *American Economic Review*. 2005. May. Vol. 95. No. 2. Pp. 152–157.
47. Master A., Cheryan S., Meltzoff A. N. Computing whether she belongs: stereotypes undermine girls' interest and sense

- of belonging in computer science // Journal of Educational Psychology. 2016. April. Vol. 108. No. 3. Pp. 424–437.
48. Ross J. A., Scott G., Bruce C. D. The gender confidence gap in fractions knowledge: gender differences in student belief-achievement relationships // School Science and Mathematics. 2012. May. Vol. 112. No. 5. Pp. 278–288.
49. Amanatullah E. T., Morris M. W. Negotiating gender roles: gender differences in assertive negotiating are mediated by women's fear of backlash and attenuated when negotiating on behalf of others // Journal of Personality and Social Psychology. 2010. February. Vol. 98. No. 2. Pp. 256–267.
50. Всеобъемлющий обзор этих вопросов в математике и инжиниринге см. в двух работах: Cheryan S., Ziegler S. A., Montoya A. K., Jiang L. Why are some STEM fields more gender balanced than others? // Psychological Bulletin. 2017. January. Vol. 143. No. 1. Pp. 1–35; Ceci S. J., Ginther D. K., Kahn S., Williams W. M. Women in academic science: a changing landscape // Psychological Science in the Public Interest. 2014. November. Vol. 15. No. 3. Pp. 75–141.
51. Цит. по Friedersdorf C. Why can't people hear what Jordan Peterson is saying? // The Atlantic, 22 January 2018 // theatlantic.com/politics/archive/2018/01/putting-monsterpaint-onjordan-peterson/550859/.
52. Хорошее научное введение в эту методологию — следующая работа: P. Hedström, P. Bearman, eds. The Oxford Handbook of Analytical Sociology. Oxford: Oxford University Press, 2011.
53. Rode J. C., Arthaud-Day M. L., Mooney C. H. et al. Ability and personality predictors of salary, perceived job success, and perceived career success in the initial career stage // International Journal of Selection and Assessment. 2008, September. Vol. 16. No. 3. Pp. 292–299.
54. Если вы педантично настаиваете на использовании информации от 63 до 37%, это нормально, но тогда вам нужно быть последовательными в своем подходе. Вам нужно вернуться к предыдущей главе этой книги и применить уравнение суждений, когда вы беседуете с Джейн и Джеком. Вам нужно начать с модели М — «Джейн покладистее Джека» при

$P\{M\} = 63\%$. Теперь войдите в комнату, улыбнитесь и поговорите с ними одинаково. Даже небольшой зрительный контакт и несколько фраз дадут вам изрядное количество данных Δ об их покладистости. Теперь вы можете обновить $P\{M|\Delta\}$ и сделать улучшенное суждение. Исходная вероятность $P\{M\}$ быстро потеряет актуальность.

55. Интервью в ноябре 2018 года для ток-шоу *Skavlan*.
56. Цитаты взяты из поста Питерсона в блоге, опубликованного в феврале 2019 года: *The gender scandal: part one (Scandinavia) and part two (Canada)* // jordanbpetersson.com/political-correctness/thegender-scandal-part-one-scandinavia-and-part-two-canada/.
57. Hyde J. S. The gender similarities hypothesis // *American Psychologist*. 2005. September. Vol. 60. No. 6. Pp. 581–592.
58. Zell E., Krizan Z., Teeter S. R. Evaluating gender similarities and differences using metasynthesis // *American Psychologist*. 2015. January. Vol. 70. No. 1. Pp. 10–20.
59. Hyde J. S. Gender similarities and differences // *Annual Review of Psychology*. 2014. January. Vol. 65. Pp. 373–398.
60. Rippon G. *The Gendered Brain: The New Neuroscience that Shatters the Myth of the Female Brain*. London: Bodley Head, 2019.

Глава 4

61. Вы можете увидеть, как эту историю рассказывает сам Айер: *A. J. Ayer on Logical Positivism and its legacy (1976)* // [youtube.com/watch?v=nG0EWNezFl4](https://www.youtube.com/watch?v=nG0EWNezFl4).
62. Reichard K. Measuring MLB's winners and losers in costs per win // *Ballpark Digest*, 8 October 2013 // [ballparkdigest.com/201310086690/major-league-baseball/news/measuring-mlbs-winner-and-losers-in-costs-per-win](https://www.ballparkdigest.com/201310086690/major-league-baseball/news/measuring-mlbs-winner-and-losers-in-costs-per-win).
63. Lindsey G. R. An investigation of strategies in baseball // *Operations Research*. 1963. July–August. Vol. 11. No. 4. Pp. 477–501.
64. Schoenfeld B. How data (and some breathtaking soccer) brought Liverpool to the cusp of glory // *New York Times Magazine*, 22 May 2019 // [nytimes.com/2019/05/22/magazine/soccer-data-liverpool.html](https://www.nytimes.com/2019/05/22/magazine/soccer-data-liverpool.html).

Глава 5

65. Я использую данные о населении, взятые из следующего документа ООН: The World's Cities in 2018 — Data Booklet (ST/ESA/SER.A/417), United Nations, Department of Economic and Social Affairs, Population Division (2018).
66. Матрицы умножаются так:

$$\begin{aligned}
 & \begin{pmatrix} 0 & 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 1/3 & 1/3 & 1/3 \\ 1/2 & 1/2 & 0 & 1/3 & 1/3 \\ 0 & 0 & 1/3 & 0 & 1/3 \\ 0 & 0 & 1/3 & 1/3 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \\
 & = \begin{pmatrix} 0 \cdot 1 + 1/2 \cdot 0 + 0 \cdot 0 + 0 \cdot 0 + 0 \cdot 0 \\ 1/2 \cdot 1 + 0 \cdot 0 + 1/3 \cdot 0 + 1/3 \cdot 0 + 1/3 \cdot 0 \\ 1/2 \cdot 1 + 1/2 \cdot 0 + 0 \cdot 0 + 1/3 \cdot 0 + 1/3 \cdot 0 \\ 0 \cdot 1 + 0 \cdot 0 + 1/3 \cdot 0 + 0 \cdot 0 + 1/3 \cdot 0 \\ 0 \cdot 1 + 0 \cdot 0 + 1/3 \cdot 0 + 1/3 \cdot 0 + 0 \cdot 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1/2 \\ 1/2 \\ 0 \\ 0 \end{pmatrix}.
 \end{aligned}$$

Попарно перемножаем элементы первой строки и вектора-столбца, а затем суммируем все такие произведения. Это дает нам первый элемент результирующего вектора-столбца (в нашем случае 0). Затем попарно перемножаем элементы второй строки и вектора-столбца, а потом суммируем все такие произведения и получаем второй элемент результирующего вектора (1/2), и т. д. В общем случае умножение производится по такому же правилу.

67. Чтобы узнать больше об этой теме, рекомендую книгу: Newman M. Networks, 2nd ed. Oxford: Oxford University Press, 2018.
68. Feld S. L. Why your friends have more friends than you do // American Journal of Sociology. 1991. Vol. 96. No. 6. Pp. 1464–1477.
69. Сейчас я докажу этот результат более строго. Пусть $P\{X_i = k\}$ — вероятность того, что человек i имеет k подписчиков. Пусть сначала мы выбираем человека j , а затем человека i

из числа тех, на кого j подписан. С помощью теоремы Байеса (уравнение 2) мы можем вычислить вероятность:

$$P\{X_i = k \mid j \text{ подписан на } i\} = \frac{P\{j \text{ подписан на } i \mid X_i = k\} \cdot P\{X_i = k\}}{\sum_k P\{j \text{ подписан на } i \mid X_i = k'\} \cdot P\{X_i = k'\}}.$$

Мы знаем, что $P\{j \text{ подписан на } i \mid X_i = k\} = k/N$, где N — общее число ребер в графе. Поэтому

$$\begin{aligned} P\{X_i = k \mid j \text{ подписан на } i\} &= \frac{(k/N) \cdot P\{X_i = k\}}{\sum_{k'} (k'/N) \cdot P\{X_i = k'\}} = \\ &= \frac{k \cdot P\{X_i = k\}}{\sum_{k'} k' \cdot P\{X_i = k'\}} = \frac{k \cdot P\{X_i = k\}}{E[X_i]}. \end{aligned}$$

Таким образом, когда $k > EX_i$, то $P\{X_i = k \mid j \text{ подписан на } i\} > P\{X_i = k\}$, и наоборот, когда $k < EX_i$, то $P\{X_i = k \mid j \text{ подписан на } i\} < P\{X_i = k\}$. Это говорит нам, что случайно выбранный человек, на которого подписан другой случайно выбранный человек, вероятно, имеет больше подписчиков, чем просто случайно выбранный человек.

Чтобы показать, что случайно выбранный человек имеет меньше подписчиков, чем средний человек, на которого он подписан, мы вычисляем ожидаемое (среднее) количество подписчиков для всех людей, на которых подписан j . Оно определяется так:

$$\begin{aligned} E[X_i = k \mid j \text{ подписан на } i] &= \sum_k k \cdot P\{X_i = k \mid j \text{ подписан на } i\} = \\ &= \sum_k \frac{k^2 \cdot P\{X_i = k\}}{E[X_i]} = \sum_k \frac{E[X_i]^2 + D[X_i]}{E[X_i]}. \end{aligned}$$

Таким образом:

$$E[X_i = k \mid j \text{ подписан на } i] = E[X_i] + \frac{D[X_i]}{E[X_i]} > E[X_i].$$

Поскольку $E[X_i] = E[X_j]$ — одно и то же значение для всех людей в социальной группе, ожидаемое количество подписчиков j будет меньше, чем у i (при условии, что j подписан на i).

70. Hodas N. O., Kooti F., Lerman K. Friendship paradox redux: your friends are more interesting than you // Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media, 2013.
71. Она опубликована: Norman M., Hahlin L. Hur tänker Instagram? En statistisk analys av två Instagramflöden [How does Instagram think? A statistical analysis of two Instagram accounts] (undergraduate dissertation). Mathematics department, University of Uppsala, 2019 // urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-388141.
72. Törner A., Anitha Schulman: “Instagram går mot en beklaglig framtid” [Инстаграм ведет к несчастливому будущему]. Dagens Media, 5 March 2018 // dagensmedia.se/medier/anithaschulman-instagram-gar-mot-en-beklaglig-framtid-6902124 (Клеменс — фамилия Аниты Шульман после замужества).
73. Cotter K. Playing the visibility game: how digital influencers and algorithms negotiate influence on Instagram // New Media & Society. 2019. April. Vol. 21. No. 4. Pp. 895–913.
74. Page L. Method for node ranking in a linked database. US Patent 6,285,999 B1, issued 4 September 2001 // patentimages.storage.googleapis.com/37/a9/18/d7c46ea42c4b05/US6285999.pdf.

Глава 6

75. См., например, работу: Bouchaud J.-P. Power laws in economics and finance: some ideas from physics // Quantitative Finance. 2000. September. Vol. 1. No. 1. Pp. 105–112; Mantegna R. N., Stanley H. E. Turbulence and financial markets // Nature. 1996. October. Vol. 383. No. 6601. P. 587.
76. Обратите внимание, что $\sqrt{n} = n^{1/2}$. Поэтому $n^{2/3}$ больше, чем $n^{1/2}$, когда $n > 1$.
77. Талеб Н. Н. Одураченные случайностью. М. : Манн, Иванов и Фербер, 2010; Талеб Н. Н. Черный лебедь. М. : Колибри, 2015; Шиллер Р. Иррациональный оптимизм. М. : Альпина Паблишер, 2017.
78. Cutler D. M., Poterba J. M., Summers L. H. What moves stock prices? // NBER Working Paper No. 2538, National Bureau of Economic Research, March 1988.

79. Tetlock P. C. Giving content to investor sentiment: the role of media in the stock market // *Journal of Finance*. 2007. Vol. 62. No. 3. Pp. 1139–1168.
80. Antweiler W., Frank M. Z. Is all that talk just noise? The information content of Internet stock message boards // *Journal of Finance*. 2004. Vol. 59. No. 3. Pp. 1259–1294.
81. Detrixhe J. Don't kid yourself — nobody knows what really triggered the market meltdown // *Quartz*, 13 February 2018 // qz.com/1205782/nobody-really-knows-why-stock-markets-went-haywire-last-week/.
82. Он опубликовал свои результаты в следующей статье: Laughlin G. Insights into high frequency trading from the Virtu initial public offering // online.wsj.com/public/resources/documents/VirtuOverview.pdf; см. также Hope B. Virtu's losing day was 1-in-1,238: odds say it shouldn't have happened at all // *Wall Street Journal*, 13 November 2014 // blogs.wsj.com/moneybeat/2014/11/13/virtus-losing-day-was-1-in-1238-odds-says-it-shouldnt-have-happened-at-all/.
83. Mamudi S. Virtu touting near-perfect record of profits backfired, CEO says // *Bloomberg News*, 4 June 2014 // bloomberg.com/news/2014-06-04/virtu-touting-near-perfect-record-of-profits-backfired-ceo-says.html.
84. $444\ 000 / 0,0027 = 164\ 444\ 444$.
85. В целях защиты личных данных имя изменено.
86. Krugman P. Three Expensive Milliseconds // *New York Times*, 13 April 2014 // nytimes.com/2014/04/14/opinion/krugman-three-expensive-milliseconds.html.

Глава 7

87. Подробности см.: medium.com/me/stats/post/2904fa0571bd.
88. Snapchat Marketing. The 17 types of Snapchat users, 7 June 2016 // snapchatmarketing.co/types-of-snapchat-users/.
89. Rosenberg N. A., Pritchard J. K., Weber J. L. et al. Genetic structure of human populations // *Science*. 2002. December. Vol. 298. No. 5602. Pp. 2381–2385.

90. Laughlin S. Gen Z goes beyond gender binaries in new Innovation Group data // J. Walter Thompson Intelligence, 11 March 2016 // jwttintelligence.com/2016/03/gen-z-goes-beyond-gender-binaries-in-new-innovation-group-data/.
91. См., например, Inglehart R., Baker W. E. Modernization, cultural change, and the persistence of traditional values // American Sociological Review. 2000. February. Vol. 65. No. 1. Pp. 19–51.
92. Inglehart R., Welzel C. Modernization, Cultural Change, and Democracy: The Human Development Sequence. Cambridge: Cambridge University Press, 2005.
93. Dillon M. Asynchrony in attitudes toward abortion and gay rights: the challenge to values alignment // Journal for the Scientific Study of Religion. 2014. March. Vol. 53. No. 1. Pp. 1–16.
94. Lambrecht A., Tucker C. E. On storks and babies: correlation, causality and field experiments // GfK Marketing Intelligence Review. 2016. November. Vol. 8. No. 2. Pp. 24–29.
95. Sumpter D. Outnumbered: From Facebook and Google to Fake News and Filter-Bubbles — The Algorithms that Control Our Lives. London: Bloomsbury Publishing, 2018.
96. О’Нил К. Убийственные большие данные. Как математика превратилась в оружие массового поражения. М. : ACT, 2018.
97. Cadwalladr C. Google, democracy and the truth about internet search // The Guardian, 4 December 2016 // theguardian.com/technology/2016/dec/04/google-democracy-truth-internet-search-facebook.
98. Caliskan A., Bryson J. J., Narayanan A. Semantics derived automatically from language corpora contain human-like biases // Science. 2017. Vol. 356. No. 6334. Pp. 183–186.
99. Angwin J., Tobin A., Varner M. Facebook (still) letting housing advertisers exclude users by race // ProPublica, 21 November 2017 // propublica.org/article/facebook-advertising-discrimination-housing-race-sex-national-origin.
100. Lambrecht A., Tucker C., Wiertz C. Advertising to early trend propagators: evidence from Twitter // Marketing Science. 2018, March. Vol. 37. No. 2. Pp. 177–199.

Глава 8

101. Robbins H., Monro S. A stochastic approximation method // Annals of Mathematical Statistics. 1951. September. Vol. 22. No. 3. Pp. 400–407.
102. Schultz W. Predictive reward signal of dopamine neurons // Journal of Neurophysiology. 1998. July. Vol. 80. No. 1. Pp. 1–27.
103. Przybylski A. K., Rigby C. S., Ryan R. M. A motivational model of video game engagement // Review of General Psychology. 2010. June. Vol. 14. No. 2. Pp. 154–166.
104. Dr Emily Collins speaking about digital games and mindfulness apps, EurekAlert!, University of Bath // eurekalert.org/multimedia/pub/207686.php.
105. Kálmán R. E. A new approach to linear filtering and prediction problems // Journal of Basic Engineering. 1960. Vol. 82. No. 1. Pp. 35–45.
106. Auger F., Hilairet M., Guerrero J. M. et al. Industrial applications of the Kálmán filter: a review // IEEE Transactions on Industrial Electronics. 2013. December. Vol. 60. No. 12. Pp. 5458–5471.
107. Flügge-Lotz I., Taylor C. F., Lindberg H. E. Investigation of a Nonlinear Control System, Report 1391 for the National Advisory Committee for Aeronautics. Washington DC: US Government Printing Office, 1958.
108. Один из самых авторитетных специалистов в этой области, формализовавший модель, — Жан-Луи Денебург. Исходной точкой была работа: Goss S., Aron S., Deneubourg J.-L., Pasteels J. M. Self-organized shortcuts in the Argentine ant // Naturwissenschaften. 1989. Vol. 76. No. 12. Pp. 579–581.
109. Мы можем переписать это уравнение для другой отслеживающей переменной. Оно будет выглядеть так и описывать вознаграждения для альтернативного варианта:

$$Q'_{t+1} = (1-\alpha)Q'_t + \alpha \left(\frac{(Q'_t + \beta)^2}{(Q'_t + \beta)^2 + (Q'_t + \beta)^2} \right) R'_t.$$

110. См., например, работы: Гладуэлл М. Переломный момент. Как незначительные изменения приводят к глобальным переменам.

- М. : Альпина Паблишер, 2016; Болл Ф. Критическая масса. Как одни явления порождают другие. М. : Гелиос, 2008.
111. Dussutour A., Nicolis S. C., Shephard G. et al. The role of multiple pheromones in food recruitment by ants // Journal of Experimental Biology. 2009. August. Vol. 212. No. 15. Pp. 2337–2348.
 112. Harris T. How technology is hijacking your mind — from a magician and Google design ethicist // Medium, 18 May 2016 // medium.com/thrive-global/how-technology-hijacks-peoples-minds-from-a-magician-and-google-s-design-ethicist-56d62ef5edf3.
 113. Krebs J. R., Kacelnik A., Taylor P. D. Test of optimal sampling by foraging great tits // Nature. 1978. September. Vol. 275. No. 5675. Pp. 27–31.
 114. Loader B. D., Vromen A., Xenos M. A. The networked young citizen: social media, political participation and civic engagement // Information, Communication & Society. 2014. January. Vol. 17. No. 2. Pp. 143–150.
 115. Этот вопрос активно изучала Анна Дорнхаус. Одна из работ: Charbonneau D., Hillis N., Dornhaus A. “Lazy” in nature: ant colony time budgets show high “inactivity” in the field as well as in the lab // Insectes Sociaux. 2014. February. Vol. 62. No. 1. Pp. 31–35.

Глава 9

116. Covington P., Adams J., Sargin E. Deep neural networks for YouTube recommendations. Conference paper, Proceedings of the 10th ACM Conference on Recommender Systems, September 2016. Pp. 191–198.
117. O’Neil-Hart C., Blumenstein H. The latest video trends: where your audience is watching // Google, Video, Consumer Insights // thinkwithgoogle.com/consumer-insights/video-trends-where-audience-watching/.
118. Stokel-Walker C. Algorithms won’t fix what’s wrong with You-Tube // The New York Times, 14 June 2019 // nytimes.com/2019/06/14/opinion/youtube-algorithm.html.
119. Orphanides K. G. Children’s YouTube is still churning out blood, suicide and cannibalism // Wired, 23 March 2018 // wired.co.uk/article/youtube-for-kids-videos-problems-algorithm-recommend.

120. Fisher M., Taub A. On YouTube's digital playground, an open gate for pedophiles // New York Times, 3 June 2019 // [nytimes.com/2019/06/03/world/americas/youtube-pedophiles.html?module=inline](https://www.nytimes.com/2019/06/03/world/americas/youtube-pedophiles.html?module=inline).
121. Silver D., Huang A., Maddison C. J. et al. Mastering the game of Go with deep neural networks and tree search // Nature. 2016. January. Vol. 529. No. 7587. Pp. 484–489.
122. Другое уравнение, которое называется Softmax, очень похоже на уравнение 1, но в некоторых ситуациях с ним работать проще. В большинстве случаев Softmax и уравнение 1 взаимозаменяемы. (Softmax — обобщение логистической функции для многомерного случая. *Прим. пер.*)
123. Mnih V., Kavukcuoglu K., Silver D. et al. Human-level control through deep reinforcement learning // Nature. 2015. February. Vol. 518. No. 7540. Pp. 529–533.
124. Mikolov T., Karafiat M., Burget L. et al. Recurrent neural network based language model. Conference paper, Interspeech 2010, Eleventh Annual Conference of the International Speech Communication Association, Japan, September 2010.

Глава 10

125. Misa T. J., Frana P. L. An interview with Edsger W. Dijkstra // Communications of the ACM. 2010. Vol. 53. No. 8. Pp. 41–47.
126. См. превосходную книгу: Cormen T. H., Leiserson C. E., Rivest R. L., Stein C. Introduction to Algorithms, 3rd edition. Cambridge, MA: MIT Press, 2009.
127. Loh P.-S. The Most Beautiful Equation in Math, video, Carnegie Mellon University, March 2016 // [youtube.com/watch?v=IUTGFQpKaPU](https://www.youtube.com/watch?v=IUTGFQpKaPU).
128. Подразумевается треугольник в евклидовой геометрии.
129. Rogers B. A. J. Ayer: A Life. London: Chatto and Windus, 1999.
130. Foot P. The problem of abortion and the doctrine of double effect // Oxford Review. 1967. Vol. 5. Pp. 5–15.
131. Ahlenius H., Tännsjö T. Chinese and Westerners respond differently to the trolley dilemmas // Journal of Cognition and Culture. 2012. January. Vol. 12. No. 3–4. Pp. 195–201.

132. Mikhail J. Universal moral grammar: theory, evidence and the future // Trends in Cognitive Sciences. 2007. April. Vol. 11. No. 4. Pp. 143–152.
133. Thomson J. J. Killing, letting die, and the trolley problem // The Monist. 1976. Vol. 59. No. 2. Pp. 204–217. Описание проблемы вагонетки в основном тексте взято из этой статьи.
134. Больше о философских аспектах проблемы вагонетки и идее моральной интуиции можно прочитать в статье: D’Olimpio L. The trolley dilemma: would you kill one person to save five? // The Conversation, 3 June 2016 // theconversation.com/the-trolley-dilemma-would-you-kill-one-person-to-save-five-57111.
135. В главах 3 и 5, излагая историю «Десятки», я не до конца объяснил, почему рассуждение Ричарда Прайса о чудесах было ложным. Научные доказательства против чудес, например Воскресения, исходят из базового понимания биологии, а не из того, что никто не совершил этого за 2000 лет после сообщений об Иисусе Христе. Нам следует рассматривать взгляд Прайса не как доказательство, что Воскресение могло быть, а как способ укрепить наши представления о доказательствах. Его рассуждение дает нам реальный и важный повседневный урок: тот факт, что мы не видели редкого события в прошлом, нельзя использовать как доказательство, что его не произойдет в будущем. Воскресение не выдерживает научного анализа, сочетающего данные и модель. Его можно объяснить только тем, что Иисус не был мертв либо о его смерти было неверно сообщено.
136. Spaiser V., Hedström P., Ranganathan S. et al. Identifying complex dynamics in social systems: a new methodological approach applied to study school segregation // Sociological Methods & Research. 2018. March. Vol. 47. No. 2. Pp. 103–135.
137. Вычисления Анны — часть следующей работы: UK’s carbon footprint 1997–2016: annual carbon dioxide emissions relating to UK consumption, 13 December 2012, Department for Environment, Food & Rural Affairs // gov.uk/government/statistics/uks-carbon-footprint.

Благодарности

Эта книга началась для меня с вызова Хелен Конфорд. Она предложила мне перестать писать для других и выразить то, что я действительно хочу сказать. Я сообщил ей, что я не особо интересный человек, а она ответила, что судить об этом не мне, а ей. И я выполнил ее просьбу.

Я все еще не уверен, что интересен, но знаю, что она, а потом и Касиана Ионита помогли мне выбрать то, что я действительно хотел сказать, и сделать это интересным. Заслуга здесь принадлежит в первую очередь Касиане, одновременно деликатная и жесткая редактура которой сделала эту книгу такой, какова она есть. Спасибо вам.

От моего суперагента Криса Уэллбела я узнал очень много об изложении и структурировании идей. Это сложно объяснить, но когда я пишу, то часто слышу, как Касиана, Крис и Хелен пропираются в моей голове. Спасибо вам за те дискуссии, которых вы никогда не вели.

Благодарю Джейн Робертсон за внимательное техническое редактирование, Бориса Грановского — за дополнительную проверку математических выкладок, а Рут Пьетрони и ее группу в издательстве Penguin — за то, что собрали всё воедино.

Большое спасибо Рольфу Ларссону за то, что внимательно прочитал книгу и обнаружил одну «серьезную ошибку» и несколько мелких. Также благодарю Оливера Джонсона за обратную связь и предложения по рисунку 2.

Лучше всего я пишу, когда вокруг меня бурлит жизнь, поэтому спасибо футбольному клубу «Хаммарбю», математическому

факультету Уппсальского университета, моей дочери Элизе и сыну Генри, а также друзьям, особенно Пеллингам, которые обеспечивали такую жизнь в течение последнего года.

Спасибо моему отцу за то, что познакомил меня с Алфредом Айером. Спасибо маме; жаль, что тебя «порезали», но все, что я написал о тебе, верно: ты вдохновляешь всех вокруг. Благодарю вас обоих за подробные вдумчивые комментарии.

Больше всего я хочу поблагодарить Ловису. Часто то, что я действительно хочу сказать, касается нашей жизни, разговоров, соглашений и споров. Надеюсь, что часть этого нашла отражение в книге.

Об авторе

Дэвид Самптер — профессор прикладной математики Уппсальского университета, Швеция. Автор книги «Футболоматика», переведенной на десять языков, и соавтор академического издания *Collective Animal Behaviour*. Консультировал ряд крупнейших футбольных клубов мира по вопросам аналитики, а также ставок.

Где купить наши книги

Специальное предложение для компаний

Если вы хотите купить сразу более 20 книг, например для своих сотрудников или в подарок партнерам, мы готовы обсудить с вами специальные условия работы. Для этого обращайтесь к нашему менеджеру по корпоративным продажам: +7 (495) 792-43-72, b2b@mann-ivanov-ferber.ru

Книготорговым организациям

Если вы оптовый покупатель, обратитесь, пожалуйста, к нашему партнеру — торговому дому «Эксмо», который осуществляет поставки во все книготорговые организации.

142701, Московская обл., г. Видное, Белокаменное ш., д. 1; +7 (495) 411-50-74;
reception@eksmo-sale.ru

Адрес издательства «Эксмо»
125252, Москва, ул. Зорге, д. 1;
+7 (495) 411-68-86;
info@eksmo.ru / www.eksmo.ru

Санкт-Петербург
СЗКО Санкт-Петербург,
192029, г. Санкт-Петербург,
пр-т Обуховской Обороны, д. 84е;
+7 (812) 365-46-03 / 04;
server@szko.ru

Нижний Новгород
Филиал «Эксмо» в Нижнем Новгороде,
603094, г. Нижний Новгород,
ул. Карпинского, д. 29; +7 (831) 216-15-91,
216-15-92, 216-15-93, 216-15-94;
reception@eksmonn.ru

Ростов-на-Дону
Филиал «Эксмо» в Ростове-на-Дону,
344023, г. Ростов-на-Дону,
ул. Страны Советов, д. 44а;
+7 (863) 303-62-10;
info@rnd.eksmo.ru

Самара
Филиал «Эксмо» в Самаре, 443052,
г. Самара, пр-т Кирова, д. 75/1, лит. «Е»;
+7 (846) 269-66-70 (71...73);
RDC-samara@mail.ru

Екатеринбург
Филиал «Эксмо» в Екатеринбурге,
620024, г. Екатеринбург,
ул. Новинская, д. 2щ;
+7 (343) 272-72-01 (02...08)

Новосибирск
Филиал «Эксмо» в Новосибирске,
630015, г. Новосибирск,
Комбинатский пер., д. 3;
+7 (383) 289-91-42;
eksmo-nsk@yandex.ru

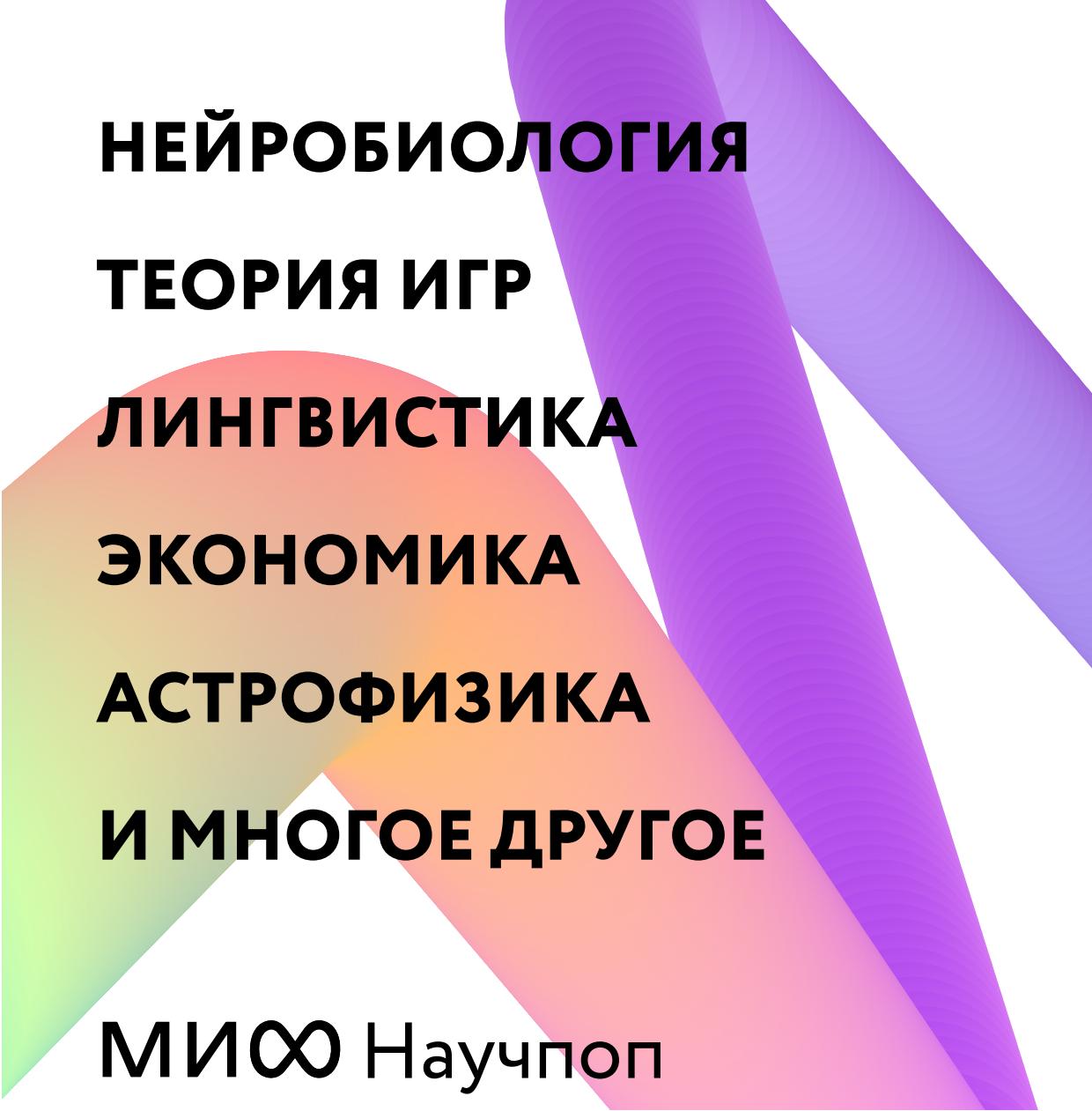
Хабаровск
Филиал «Эксмо Новосибирск» в Хабаровске,
680000, г. Хабаровск, пер. Дзержинского,
д. 24, лит. «Б», оф. 1; +7 (4212) 910-120;
eksmo-khv@mail.ru

Казахстан
«РДЦ Алматы», 050039,
г. Алматы, ул. Домбровского, д. За;
+7 (727) 251-59-89 (90, 91, 92);
RDC-almaty@eksmo.kz

Украина
«Эксмо-Украина», Киев,
ООО «Форс Украина», 04073, г. Киев,
Московский пр-т, д. 9;
+38 (044) 290-99-44;
sales@forsukraine.com



Если у вас есть замечания и комментарии к содержанию, переводу, редактуре и корректуре, то просим написать на be_better@m-i-f.ru, так мы быстрее сможем исправить недочеты.



НЕЙРОБИОЛОГИЯ

ТЕОРИЯ ИГР

ЛИНГВИСТИКА

ЭКОНОМИКА

АСТРОФИЗИКА

И МНОГОЕ ДРУГОЕ

МИФ Научпоп

Узнавай первым о новых книгах,
скидках и подарках
из нашей рассылки
mif.to/sci-letter

Весь научпоп
на одной странице:
mif.to/science

#mifbooks



*Научно-популярное издание
Серия «Принятие решений»*

Самптер Дэвид

Десять уравнений, которые правят миром
И как их можете использовать вы

Руководитель редакционной группы Светлана Мотылькова

Ответственный редактор Юлия Потемкина

Литературный редактор Ольга Свитова

Арт-директор Алексей Богомолов

Дизайн обложки Наталья Савиных

Верстка Екатерина Матусовская

Корректоры Людмила Широкова, Евлалия Мазаник

ООО «Манн, Иванов и Фербер»
123104, Россия, г. Москва, Б. Козихинский пер., д. 7, стр. 2
www.mann-ivanov-ferber.ru
www.facebook.com/mifbooks
www.vk.com/mifbooks
www.instagram.com/mifbooks



Есть ли секретная формула для того, чтобы разбогатеть? Или для того, чтобы рекламная кампания стала вирусной? Можно ли рассчитать, оставаться ли на нынешней работе или в отношениях?

Без всякого преувеличения, прикладная математика позволяет править миром. То, что кажется хаосом, на самом деле прошито теорией вероятностей и математической статистикой, теорией устойчивости и законами больших чисел. Фактически многие социальные и экономические процессы можно свести всего к десяти уравнениям.

Научившись вычленять из происходящего данные и математические модели, вы начнете видеть взаимосвязи, словно на рентгене.

И это даст возможность при необходимости переиграть букмекеров, не поддаться биржевой панике, отделить ценную информацию от неважной, или вычислить самую влиятельную персону.

Одним словом, математический подход дает ощущение твердой почвы там, где до того была неопределенность. Применяйте!

Автор книги, Дэвид Самптер, — профессор прикладной математики Уппсальского университета (Швеция). Автор книги «Футболоматика», переведенной на десять языков, и соавтор академического издания Collective Animal Behaviour, консультант ряда крупнейших футбольных клубов мира по вопросам аналитики и ставок.

