МАТЕМАТИКА ЗА ЧАШЕЧКОЙ КОФЕ

Маурицио Кодоньо

Maurizio Codogno

MATEMATICA IN PAUSA CAFFÈ

codice

EDIZIONI

Маурицио Кодоньо

МАТЕМАТИКА ЗА ЧАШЕЧКОЙ КОФЕ

Перевод с итальянского

Минск
«Дискурс»
2020

УДК 51
ББК 22.1
К57

Перевод с итальянского Олеси Пантелеенко

Кодоньо, М.

К57 Математика за чашечкой кофе = Matematica in pausa caffè / Маурицио Кодоньо ; перевод с итальянского Олеси Пантелеенко. — Минск : Дискурс, 2020. — 160 с.

ISBN 978-985-90515-7-9.

Парадоксов в математике предостаточно, но главный из них, пожалуй, состоит в том, что многие считают ее сложной, скучной и ненужной. И правда: зачем сегодня может понадобиться математика? Разве что прикинуть стоимость покупок в магазине или помочь ребенку решить задачу по математике... Стоит ли в таком случае тратить на нее время?

Безусловно! И прежде всего потому, что математика — это очень интересно. Если у вас есть 15 минут, налейте себе чашечку кофе и устраивайтесь поудобнее, а Маурицио Кодоньо объяснит, почему ваша очередь всегда движется медленнее соседней, существует ли надежный способ выиграть в казино и можно ли доверять большим данным.

УДК 51
ББК 22.1

Научно-популярное издание

Кодоньо Маурицио

МАТЕМАТИКА ЗА ЧАШЕЧКОЙ КОФЕ

Дизайн обложки *Т. Сиплевич*
Верстка *К. Подольцева*
Корректоры *Т. Радецкая, А. Павлович*

12+

Подписано в печать 21.09.20. Формат 84×108 1/32. Бумага офсетная.
Печать офсетная. Усл. печ. л. 8,4. Уч.-изд. л. 4,4. Доп. тираж 4000. Заказ № 2010860.

Частное унитарное предприятие «Издательство Дискурс».
Свидетельство о государственной регистрации издателя, изготовителя, распространителя печатных изданий № 1/519 от 11.08.2017.
Ул. Гусовского, д. 10, помещение № 9 (комн. 404), 220073, г. Минск.

Дата изготовления 21.10.20. Срок годности не ограничен.
Произведено в Российской Федерации

Отпечатано в полном соответствии с качеством предоставленного электронного оригинал-макета в ООО «Ярославский полиграфический комбинат»
150049, Россия, Ярославль, ул. Свободы, 97

ISBN 978-985-90515-7-9

Содержание

Посвящается Анне, моей спутнице жизни

Введение

Среди итальянцев 78,2 % считают, что математика — сложная наука. Более того, 42 % утверждают, что «боятся математики». Что можно сказать по этому поводу? Бесспорно, математика, которую преподают в университете, сложна. И что с того? Приготовить безукоризненный капучино тоже нелегко, и я изумляюсь всякий раз, когда смотрю, как бармен, наливая молоко в чашку, рисует листочек или сердечко. Я даже не знаю, с чего начать, чтобы получить нечто подобное. Тем не менее я могу включить кофемашину и приготовить кофе — пусть не такой идеальный с эстетической точки зрения, но более чем приличный по вкусу.

К сожалению, школьная программа на 99 % состоит из того, что не пригодится ученикам в дальнейшей жизни. И не так часто встречаются действительно хорошие педагоги, не только умеющие показать, что математика окружает нас в повседневной жизни (так говорят многие, и, если честно, это начинает надоедать), но способные объяснять математику без сложных вычислений. Здесь кто-нибудь может возразить, что математика применяется, только когда речь идет о вычислениях, но, на мой взгляд, это всего лишь предлог для того, чтобы держать нас подальше от настоящей математики. Не буду повторять избитые фразы о том, как прекрасна математика. Я убежден, что так оно и есть, хотя и не в состоянии разъяснить, почему это так. Я имею в виду, что бывает сложно объяснить окружающие нас явления, однако их можно понять интуитивно и рассказать о них в непринужденной беседе за чашечкой кофе. В конце концов, если эти несколько минут можно посвятить разговорам о политике, кино, экономике, хотя мы не являемся политиками, обладателями премии «Оскар» или Нобелевской премии, то почему же нам нельзя поболтать на любопытные математические темы?

Разделы книги объединяют ряд близких тем, каждой из которых посвящено несколько страниц. Она предназначена для людей с нематематическим складом ума. Да, иногда я использую примеры с числами, однако гарантирую доступность изложения. В тексте есть изображения, но их немного, потому что рисунок сам по себе бесполезен по сравнению с тем, что можно о нем рассказать. И я вас уверяю, что несколько формул, которые вы встретите, вставлены только из соображений их эстетической значимости. Так что если вы и пропустите несколько страниц, то ничего страшного не произойдет. Вы же не думаете, что мы будем рисовать графики и решать примеры, используя ложечку от кофемашины и пену от капучино?

Книга состоит из пяти разделов. Раздел «Арифметика» посвящен вопросам, которые вы, возможно, задавали себе в школьные годы, но о которых потом забыли. В раздел

«Парадокс, возможность, прогноз» включены на первый взгляд необъединяемые темы, однако при более внимательном рассмотрении такой подход оказывается логичным. Раздел «Игры» — ну, здесь название говорит само за себя: речь пойдет об играх азартных и не только. Раздел «На прогулке» рассказывает о проблемах, с которыми можно столкнуться, например, прогуливаясь по улице. И наконец, в разделе «Компьютеры и стандарты» затронуты вопросы, меньше всего связанные с математикой (хотя есть люди, которые уверены, что информатика — это та же математика, только в другом обличии). Каждую тему можно читать по отдельности, несмотря на отсылки к другим разделам, в которых она подается немного иначе. Книга включает и общеизвестные темы, однако я все-таки надеюсь, что мне удалось рассказать о них по-новому.

Развлекая вас любопытными фактами, я постараюсь донести их суть через интуицию. Я принадлежу к школе, полагающей, что в основе математики лежит создание моделей, которые практически в совершенстве находят применение в реальной жизни. Главная проблема состоит в том, чтобы создать логико-математическую структуру окружающего мира. А вот умение решать примеры или вычислять интегралы нам не очень нужно в повседневной жизни: если возникнет такая необходимость, всегда можно посчитать на компьютере.

Если у нас не сформирована качественная модель прогнозирования, мы подвергаемся реальной опасности: кто угодно может использовать огромное количество более или менее случайных чисел и формул, чтобы разыграть нас, воспользовавшись нашей безграмотностью.

С 2012 года по «Фейсбуку» гуляет текст следующего содержания: «Вчера Сенат республики утвердил 257 голосами "за" при 165 воздержавшихся проект закона сенатора Чиренга, предусматривающий создание фонда для "депутатов в кризисе", ввиду неизбежной кончины законодательной структуры. Этот фонд предусматривает заложить в бюджет 134 миллиарда евро для всех депутатов, которые не найдут работу в течение года, последующего за истечением мандата». Сколько тех, кто прочел пост и поделился им, однако не попытался суммировать количество указанных голосов, чтобы заметить, что их больше, чем членов Сената? А сколько тех, кто разделил указанную сумму на меньше чем тысячу депутатов, чтобы понять, что каждый из них должен был бы получить более 130 миллионов? Знаний по математике потребуется немного, здравого смысла — ненамного больше: и если первая вас не испугает, то я уверен, что второго у вас более чем достаточно.

В первую очередь я хотел бы выразить признательность Энрико Казадеи из издательского дома «Кодиче», который, читая мою болтовню в онлайн-газете «Пост», предложил написать об этом книгу. Также благодарю Стефано ди Милано из «Кодиче», который оказался втянут в выбор тем для написания. Спасибо (кого еще назвать?) Марко Фискетти и Массимо Манка за кропотливую работу над первой редакцией, которые не только выжили, но и указали, какие части подлежат корректировке. И наконец, спасибо Анне, Чечилии и Якопо, которые дали мне время написать и особенно переписать эту книгу.

Да, совсем забыл: статистические данные, приведенные выше, я выдумал. Прошу, не размещайте их в «Фейсбуке».

Раздел I

Арифметика

Что дает минус на минус (плюс или минус)?

Думаю, многие из вас заметили определенное сходство между математическими формулами и магическими заклинаниями. Не обязательно иметь волшебную палочку, чтобы привести в действие законы математики и получить желаемый результат, однако правила следует учить слово в слово, потому что малейшая неточность ведет к досадным ошибкам. Одни формулы мы воспринимаем как должное, например: объем шара равен три четвертых пи на радиус в кубе. Другие же сбивают нас с толку, например правило знаков, которое звучит следующим образом: плюс на плюс дает плюс; плюс на минус дает минус; минус на плюс дает минус... До сих пор правило не кажется сложным для усвоения. Но стоит произнести фразу «минус на минус дает плюс», в классе неизбежно раздается вопрос: «Как можно получить положительное число из отрицательных?»

На уроке учитель обычно торопится и лишь бросает в ответ: «Это нужно просто запомнить». Он не задумывается о том, насколько трудно школьнику разобраться в отрицательных числах. А ведь до XVII века их не изучали вовсе. В те времена математик избавлялся от неудобных минусов и записывал уравнение $x^2 = 5x + 6$ вместо $x^2 - 5x - 6 = 0$. Страх перед отрицательными числами сохраняется и в современном обществе. Он проявляется в том, что выражение «минимальная ночная температура составила минус 3 градуса» в нашей речи превращается в «три градуса ниже нуля». Этот комментарий не поможет понять правило знаков, но, к счастью, существует относительно простой способ разобраться в том, как оно действует, если воспользоваться воображением.

Мы можем, например, будущее обозначить положительными числами, а прошлое — отрицательными. Таким образом, позавчера будет равняться «минус два дня от сегодняшнего дня», а завтра — «плюс один день к сегодняшнему дню».

Еще проще представить, что доходы — это положительные числа, а долги — отрицательные. Если у меня совсем нет денег и при этом я должен отдать 1000 евро кредиторам, то я могу сказать, что у меня есть минус 1000 евро. Конечно, я так не говорю, но главное, что подход к объяснению правила становится понятным.

Объединим теперь обе идеи и посмотрим, что получится. Если я каждый год кладу на счет 100 евро (+100), то через десять лет (+10) у меня будет на 1000 евро больше, чем сейчас: плюс на плюс дает плюс. Представим, что эта ситуация длится уже не первый год. Если я начал пополнять счет в банке давно, то десять лет назад (–10) у меня было на 1000 евро меньше, чем сегодня: плюс на минус дает минус. А сейчас предположим, что давным-давно мои бабушка с дедушкой — пусть земля им будем пухом — завели чековую книжку в банке. И использовали они ее только для одной операции: снимали каждый год 100 евро, чтобы уплачивать налоги (−100). Через десять лет (+10) первоначальная сумма уменьшилась на 1000 евро: минус на плюс дает минус. И наконец, я с тоской вспоминаю, как ежегодно снимал со счета по 100 евро (−100). Если предположить, что я делал так каждый год, то какая сумма была на моем счету десять лет назад (−10)? Все согласятся, что на 1000 евро больше, чем сейчас: вот так минус на минус дает плюс.

Таким образом снимаются все вопросы и правило знаков становится понятным. Общеизвестно, что математические правила лучше усваиваются на примере денежных операций.

Осторожно: средние величины!

Я известный зануда. И когда я слышу заявление: «Демократичность общества определяется уровнем образованности избирателей, а он у нас ниже среднего», мне тотчас хочется исправить оратора, что не ниже среднего, а ниже медианы! Среднее арифметическое и медиана не являются взаимозаменяемыми синонимами ни в разговорной речи, ни в научном стиле, а представляют собой разные виды средних величин. Классификация усложняется, когда вводится еще один термин под названием «мода», который не имеет никакого отношения к весенне-летней коллекции. Давайте разберемся в разнице между этими понятиями и научимся использовать их в правильном контексте.

Итак, мы находимся в школьном классе. У всех детей есть конфеты, у кого-то больше, у кого-то меньше. Как узнать среднюю величину?

Сценарий 1. Учитель решает обеспечить социальное равенство: собирает все конфеты, ставит детей в ряд и начинает раздавать по одной каждому. Предположим, в результате все конфеты поровну и без остатка будут разделены между детьми. Таким образом, *среднее арифметическое* — это количество конфет, которым обладает каждый ребенок.

Сценарий 2. Учитель желает объяснить марксистскую теорию, для чего начинает формировать две группы: слева ставит ребенка с наименьшим количеством конфет, справа — с наибольшим. Так он распределяет детей до тех пор, пока не останутся один или два ребенка. Количество конфет, которыми обладает последний ребенок, и является *медианой*.

Сценарий 3. Учитель — заядлый любитель голосований в телешоу. Он делит класс таким образом, чтобы у всех детей из одной группы было одинаковое количество конфет. А затем определяет самую многочисленную группу. В итоге количество конфет, которыми обладает ребенок из данной группы, представляет собой *моду в распределении* конфет.

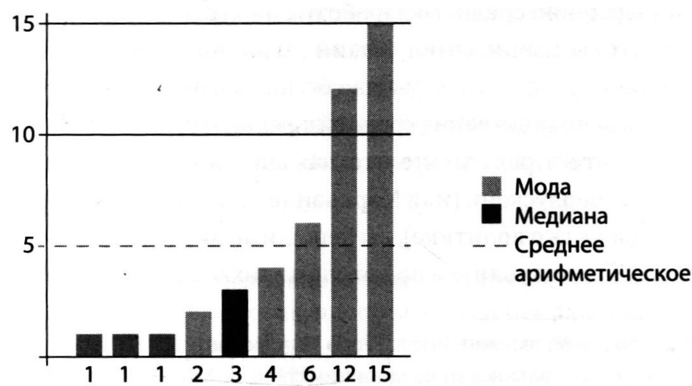
Сразу же отвечу на вопросы, которые, возможно, возникли у вас при чтении этих примеров.

- «Могло ли в первом сценарии у нескольких детей оказаться на одну конфету больше, чем у других?» Действительно, среднее арифметическое может иметь дробное значение, несмотря на то что все исходные числа являются целыми. Однако в любом случае данный пример помогает понять суть среднего арифметического, а заключается она в том, что у всех детей одинаковое количество конфет.
- «Что произойдет, если во втором сценарии останутся двое детей с разным количеством конфет?» В таком случае одного ребенка отправят налево, другого — направо, а среднее значение количества конфет

у этих двоих будет определять медиану. Важно усвоить, что медиана делит исходную группу на две одинаковые части.

- «Что будет, если в третьем сценарии окажется несколько групп с одинаковым количеством детей?» Такая ситуация в математике получила название *мультимодального распределения*. Мода — единственный вид средних величин, который может иметь сразу несколько значений.

Если в глубине души вы настоящий математик, то наверняка спросите: «Всегда ли среднее арифметическое, медиана и мода имеют разные значения? А могут ли они совпадать?» Вопрос серьезный, я не шучу. Существует масса примеров, когда и мода, и медиана, и среднее арифметическое приходятся на одну точку. Но также несложно выбрать совокупность чисел, в которой все три вида средних величин будут иметь разные значения: на рисунке вы видите как раз такую ситуацию.



Проще говоря, между средним арифметическим, медианой и модой нет никакой взаимосвязи типа «среднее арифметическое меньше медианы». Это всего лишь три независимых числа.

После того как разница между средним арифметическим, медианой и модой усвоена, становится понятен и контекст их использования. Поскольку среднее арифметическое является величиной, наиболее связанной с количеством, оно нередко пригождается в повседневной жизни. Такие вычисления полезны, когда тренируешься для участия в марафоне и каждый день записываешь в таблицу свой пробег в километрах. А если ежедневно в течение нескольких недель ходишь собирать дрова для камина, то в конце концов, вероятно, заинтересуешься средним объемом работы, проделываемой за день.

Если же мы желаем узнать значение некоего показателя по отношению к населению в целом, то логичнее было бы высчитать медиану, как в ситуации с конфетами. Я специально использовал условное наклонение, так как обычно поступают иначе. Классический пример — определение среднего заработка по стране, когда из-за простоты вычисления общий доход населения делят на число людей. Но есть еще более глупые примеры из разряда «определение средней температуры по палате». Не верите? Представьте, что, находясь в зале с дюжиной коллег Берлускони (или Маркьонне[1], если предпочитаете не говорить о политике), вы решили вычислить среднюю зарплату итальянцев по доходам этих богачей: вы сразу

---

[1] Серджо Маркьонне (итал. Sergio Marchionne; 1952–2018) — канадско-итальянский предприниматель в области автомобилестроения, возглавлявший компанию «Фиат» с 2004 по 2018 год. Под его руководством компания, находившаяся на пути к банкротству, вновь стала прибыльной. — *Здесь и далее прим. пер.*

заметите, что среднее арифметическое в этом случае не имеет практической значимости.

Другой пример, наглядно демонстрирующий суть медианы (или, лучше сказать, ее производной), — диаграммы роста младенцев. Если ваш ребенок оказался ниже 50-го перцентиля[1], это совсем не означает, что с ним что-то не так. По определению половина всех детей располагается до отметки 50 %, а другая половина — выше этого уровня. Короче, это мера относительная, а не абсолютная, поэтому она сама по себе ничего не говорит о здоровье ребенка.

А зачем необходимо знать моду? Ну, она может понадобиться, если вы следите за модой… или желаете просчитать результат лотереи, в которой не все числа имеют равную вероятность выигрыша. Так, при выбрасывании двух кубиков куда выше шанс получить 7, чем 12.

Мультимодальное распределение помогает обнаружить ошибки в расчетах. Если мы построим кривую для анализа роста итальянцев, то, скорее всего, заметим в ней два пика; так станет очевидно, что полезно составлять отдельные графики для мужчин и женщин.

Как видите, чтобы разобраться в этой теме и не затеряться в серединке, достаточно уделить ей немного внимания.

---

1 Перцентиль — значение антропометрического признака для сотой доли совокупности измеренных людей. Если кривую распределения всей совокупности измеренных людей разделить на 100 равных частей, то получим 99 площадей, в каждой из которых будет свое значение признака и частота ее встречаемости. Значение 50 окажется точно посередине.

Способ девятки

Прежде всего хочу сказать, что, по-моему, мы ожидаем от способа девятки слишком многого. Он известен с давних пор: упоминание о нем встречается уже в 1478 году в книге «Искусство абака», которая представляет собой первый текст на итальянском языке, посвященный арифметике. Однако я бы выбросил способ девятки из школьной программы (оказывается, есть еще педагоги, которые объясняют его ученикам, но затем никогда не применяют на уроках), так как он приносит больше проблем, чем пользы.

И все же я заметил, что проверка девяткой впечатляет взрослых людей. Меня часто спрашивают, как работает этот способ и — прежде всего — почему он иногда не срабатывает. Что ж, наконец-то вы сможете удовлетворить свое любопытство.

Первым делом, наверное, полезно вспомнить, в чем состоит суть этого способа.

Чтобы проверить умножение: $247 \cdot 53 = 13\,091$, нужно найти сумму цифр каждого числа. В случае если она оказывается двузначной, складываем цифры до тех пор, пока не получим однозначное число ($2 + 4 + 7 = 13 \rightarrow 1 + 3 = \mathbf{4}$;

5 + 3 = **8**; 1 + 3 + 0 + 9 + 1 = 14 → 1 + 4 = **5**)[1]. Затем перемножаем числа, полученные от исходных множителей, и при необходимости суммируем цифры итогового произведения, чтобы в конце концов вышло однозначное число (4 · 8 = 32 → 3 + 2 = **5**). Если финальные цифры, полученные после обработки множителей и ответа, различаются, это значит, что мы где-то ошиблись. Если же они совпадают, то, может быть (!), нам удалось посчитать правильно. Теперь четыре финальные цифры ставим внутри четырех полей. Не гарантирую, что положение цифр в полях, представленное на рисунке, соответствует объяснениям моего учителя: некоторые детали я уже подзабыл.

$$\begin{array}{r} 247\times \\ 53 \\ \hline 741 \\ 1235 \\ \hline 13091 \end{array} \qquad \begin{array}{r|l} 247 & 13091 \\ \mathbf{4} & \mathbf{5} \\ \hline \mathbf{8} & \mathbf{5} \\ 53 & 5\times 9 \end{array}$$

Пример проверки девяткой

1 В школах бывшего СССР проверка девяткой обычно не используется. Этот способ основан на правиле остатков: остаток от деления произведения на какое-либо число равен произведению остатков от деления каждого множителя на это же число. Если в качестве проверочного числа брать 9, задача упрощается, поскольку при делении числа на 9 получается тот же остаток, что и при делении на 9 суммы цифр этого числа. Например, число 247 при делении на 9 дает остаток 4, и то же получается в остатке от деления суммы цифр этого числа (2 + 4 + 7 = 13) на 9. В итальянских школах идут еще дальше: остаток от деления на 9 получают без деления, просто суммируя цифры, пока не получат однозначное число (1 + 3 = 4).

Способ девятки можно использовать для проверки сложения, вычитания и умножения; для деления выполняют обратные действия, то есть делимое рассматривается как произведение делителя на частное.

До сих пор мы рассматривали проверку девяткой на примере. Теория же гласит, что этот способ — применение модульной арифметики (которую в ряде учебников по математике называют арифметикой часов, совершенно забывая, что современная молодежь использует сотовый телефон, а не часы с циферблатом). Итак, способ девятки — это не что иное, как вычисление остатка по модулю 9, ведь на каждом этапе мы занимаемся тем, что вычитаем число, кратное 9, то есть модулю. Сложение, деление и умножение в модульной арифметике такие же, как и в классической, но, конечно, мы не заметим, что результат неверен, если разница между ним и правильным ответом составляет число, кратное модулю.

Но почему все-таки существует способ девятки, а не семерки или одиннадцати? Все просто. Из-за того что остаток от деления 10, 100, 1000 на 9 составляет 1, остаток от деления любого числа на 9 равняется сумме цифр этого числа. А теперь попробуйте сами, если сомневаетесь, произвести деление на 7 или на 15!

К сожалению, простота подсчета по модулю 9 сопровождается серьезной проблемой: если переместить между собой две цифры результата (10 391 вместо 13 091) либо ошибиться в распределении промежуточных цифр по полям, остатки по модулю 9 не меняют значения и, таким образом, мы не заметим ошибку. Как этого избежать? У меня есть решение данной проблемы — проверка одиннадцатью (см. рисунок на стр. 25).

$$
\begin{array}{r}
247\times\\
53\\
\hline
741\\
1235\\
\hline
13091
\end{array}
\qquad\qquad
\begin{array}{r|l}
247 & 13091\\
5 & 1\\
\hline
9 & 1\\
53 & 5\times 9
\end{array}
$$

Тот же пример с применением способа одиннадцати

Производить вычисления с модулем 11 в реальности не так сложно, как может показаться. Необходимо последовательно вычесть друг из друга цифры заданного числа справа налево. В случае если результат отрицательный, добавляем 11. В остатках мы получим цифры от 0 до 10. Если в ответе поменять местами две цифры или неверно расположить цифры в промежуточных вычислениях, то проверка сразу же выявит ошибку. Я справился с объяснением?

Цифры сомнительной известности

Изучая в школе различные виды чисел, вы, может быть, заметили, что названия многих из них производят не очень приятное впечатление. Конечно, с числами органическими… простите, натуральными[1], и даже с рациональными[2] дела обстоят довольно-таки благополучно. Но что насчет иррациональных и мнимых чисел? Откуда столько ненависти к ним? Этот вопрос сложнее, чем может показаться на первый взгляд. Иногда я и сам себя спрашиваю: а нет ли здесь подвоха?

Начнем с чисел *рациональных* и *иррациональных*. Изначально эти названия не имели позитивной или негативной окраски. В латинском языке слово *ratio* означает «отношение». Следовательно, число считается рациональным, если выражает отношение двух целых чисел, и иррациональным, если не выражает.

---

[1] Натуральные числа служат для счета. Например: 1, 2, 3.

[2] Рациональным называется число, которое можно представить обыкновенной дробью m/n, где числитель m — целое число, а знаменатель n — натуральное число.

Правда, греки расстроились, когда узнали, что не все числа являются рациональными. Шок был настолько сильным, что они решили в большей степени полагаться на геометрию, а не на арифметику. И в самом деле: $\sqrt{2}$ можно представить графически, несмотря на то что он не является отношением двух чисел. К сожалению, слово *ratio* означало также «разумный, правильный», и таким образом иррациональные числа стали «неразумными». Нечто подобное произошло и с несчастными отрицательными числами, которые сами по себе не сделали ничего плохого. Однако за ними тянется дурная слава, так же как и за словом «левый», которое нередко используется в отрицательном значении, в то время как быть правым — это хорошо.

Числа, которым математики присвоили по-настоящему скверное название, именуются *мнимыми*. Мы даже точно знаем, кто во всем виноват: Джироламо Кардано. Разрабатывая метод решения кубических уравнений, он обнаружил небольшую проблему: в определенный момент в вычислениях появлялись цифры, которых не могло существовать, а именно корни из отрицательных чисел. Ведь отрицательное число, умноженное на себя, дает положительный результат — естественно, в виде положительного числа. А главная беда заключалась в том, что эти числа появлялись в случае, когда кубическое уравнение имело три корня. К счастью, можно было притвориться, что ничего не случилось, и продолжить вычисления — несуществующие числа взаимно сокращались, и в итоге получался правильный ответ. Впрочем, будучи человеком прагматичным, Кардано решил, что раз уж результат верный, то можно представить, будто эти числа существуют. Отсюда и название «мнимые» в противовес «настоящим» вещественным числам[1].

Вещественные и мнимые числа можно рассматривать как частные случаи *комплексных*[2]. И что, их и впрямь так сложно понять? Да ладно! Более двух веков назад нашелся простой способ представлять положительные, отрицательные, мнимые и комплексные числа. И даже складывать их совсем несложно (вот умножать — да, нелегко; но часто ли их нужно умножать?). Однако предубеждения умирают медленно.

Забавно, что в XIX веке имена новым видам чисел давали люди набожные. По-моему, это смешно — называть числа типа греческого π, которые не являются корнями многочлена с целыми коэффициентами, *трансцендентными*. Ведь отношение между диаметром и окружностью можно представить на бумаге, оно не падает с небес.

Но самая курьезная история, на мой взгляд, связана с *трансфинитными* числами[3]. Георг Кантор решил отбросить две с половиной тысячи лет математики, в которой бесконечность была не значением, а только идеей, и дал определение бесконечного числа. Однако он обнаружил, что видов бесконечного числа было огромное множество и, как набожный католик, озадачился. Возможно ли

---

1 Вещественные, или действительные, числа служат для измерения величин окружающего мира (геометрических и физических).

2 Комплексные числа — числа вида $a + bi$, где a и b — вещественные числа, i — мнимая единица, то есть число, для которого выполняется равенство: $i = \sqrt{-1}$.

3 В русском языке в теории множеств принято название «порядковое число» или «ординал».

полагать, что бесконечный Бог является лишь одним из многих? Имелся лишь один способ узнать это, и Кантор, как и подобает порядочному немцу, написал в Ватикан — напрямую в Конгрегацию вероучения. Необходимо отдать должное римской курии, которая не выбросила письмо, а передала кардиналу Иоганну Баптисту Францелину. В ответном письме он разъяснил, что «два понятия — "абсолютная бесконечность" и "актуальная бесконечность", или "трансфинитность", — по сущности своей являются разными; если их сравнивать, то только первое может быть охарактеризовано как непосредственно бесконечность (eigentlich Unendliches), в то время как другое является бесконечностью несобственной (uneigentlich) и двусмысленной. Рассуждая таким образом, насколько я могу видеть, Ваш концепт "трансфинитность" не представляет никакой опасности для вероучения». Вот так трансфинитные числа получили благословение самой Святой Римской Церкви!

И последний забавный случай. В 60-х годах XX века Джон Хортон Конвей разработал «систему определений» для чисел, которые можно расположить в возрастающем порядке. Я вас уверяю: вы не хотите знать подробности. Скажу лишь, что эта система включает числа бесконечные, бесконечно малые, вещественные, гипервещественные и многие другие. Вот Конвей и назвал их *сюрреальными числами*. Дюшамп, Магритт и Дали были бы в восторге от этого.

Единица или нет

Равно ли единице число 0,999 999 — вот еще один вызывающий сомнения вопрос, который начальная школа оставляет без ответа и который имеет долгую историю, насчитывающую более двух тысяч лет. Ее начало восходит к апории «Ахиллес и черепаха» Зенона Элейского. Если бы этот древнегреческий мыслитель жил сейчас, то прослыл бы интернет-троллем. В свое время Зенон развлекался тем, что задавал невинные, на первый взгляд, вопросы, чтобы привести в замешательство собеседников-философов. Суть его знаменитого парадокса заключается в том, что быстрый Ахиллес состязается в беге с медленной черепахой. Поскольку скорость черепахи в десять раз меньше скорости Ахиллеса, он дает ей фору в девять десятых стадиона. Соперники подходят к стартовой черте. На старт! Внимание! Марш! Ахиллес стремительно преодолевает 0,9 стадиона и достигает точки, из которой стартовала черепаха, но та за это время проползла 0,09 стадиона. За один миг Ахиллес проходит и эту дистанцию. Меж тем черепаха продвинулась вперед, преодолев еще 0,009 стадиона, которые

Ахиллес пробегает в мгновение ока, но черепаха все ползет и ползет...

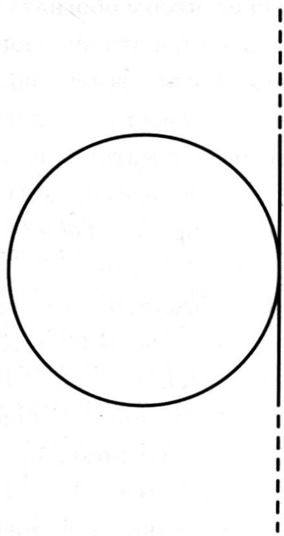
Подведем итоги. Все согласны с тем, что в настоящем соревновании Ахиллес рано или поздно догонит и обгонит черепаху, хотя по рассказу кажется, что этого не случится никогда.

Вопрос посложнее: чтобы обогнать черепаху, нужно сначала ее догнать, но где это произойдет? Преодолев 0,999 999... стадиона, Ахиллес все еще позади. Окажется ли точка обгона на беговой дорожке? Ответ положительный, но условный. Многие математики пытались решить эту задачу и найти неопровержимое доказательство, пока Рихард Дедекинд не разработал понятие «сечение» для определения вещественных чисел.

Если разделить множество чисел, размещая с одной стороны те, значение которых больше или равно единице, а с другой те, значение которых меньше единицы, то можно сказать, что посреди нет ничего. Однако признаемся сами себе: невозможно выполнять бесконечное число действий, как в рассказе. Физики утверждают, что существует минимальное значение длины, меньше которого нет смысла говорить о расстоянии, — это так называемая длина Планка. Есть еще один способ решения задачи. В уравнении 0,999 999... = 1 отнимите 0,999 999... с обеих сторон. Получится 0 = 0,000 000... Вы согласны, что представленная таким образом проблема принимает совсем другой вид и кажется разрешимой?

Конечно, можно предположить существование другого ответа. В 60-х годах XX века математик Абрахам Робинсон разработал теорию, известную как *нестандартный анализ*. Согласно этой теории 0,999 999... не равно 1.

Иными словами, существует число 0,000 000… отличное от нуля и меньше любого положительного числа. Могу привести пример таких чисел: угол между окружностью и касательной не равняется нулю, но и не может быть больше нуля, иначе касательная не будет касательной. Хотя, между нами, так ли все это важно?



Чему равен угол между прямой и окружностью?

Во что бы то ни стало нужно понимать, что 1,000 не равняется 1. Будем надеяться, что тот, кто написал число с тремя достаточно бесполезными нулями после запятой, хотел указать точное значение, которое больше 0,999 и меньше 1,001. Если я заявлю, что обхват моей талии составляет 1 метр, можно предположить, что у меня есть некоторый избыточный вес, причем никто не удивится, если на самом деле это 99,7 сантиметра. В то же время, если мне скажут, что длина трубки — 1,000 метра, я могу быть уверен, что возможная ошибка в длине составляет менее 1 миллиметра.

Ну что, вы заметили разницу между математической теорией (где 1 равняется 1,000) и практикой? Вот, запомните это теперь, когда вы уже не ходите в школу и никто не проверяет ваши вычисления!

Логарифмы

Существуют вещи, созданные для определенной цели, а потом со временем им находят другое применение. В этом нет ничего странного. Есть даже знаменитая поговорка: «Если нам нужно забить гвоздь, что бы ни было у нас в руках, может оказаться молотком». Вот только грустно, что мало кто руководствуется этим правилом в других ситуациях, и случай с логарифмами — прямое тому доказательство. В школах продолжают изучать логарифмы, не зная, для чего их применяли раньше. Правда, в наши дни первоначальная причина утратила значение, однако логарифмы становятся полезными для другой цели. Впрочем, начну с самого начала.

Если вы думаете, что «логарифм» — слово греческого происхождения, то ошибаетесь. Не является оно и арабским, несмотря на звуковое сходство со словом «алгоритм». В начале XVII века, когда Галилей еще только разрабатывал научный метод исследований природных явлений, в далекой Шотландии барон Джон Непер создал теорию логарифмов. Невероятное научное событие для этой географической зоны — тем более удивительное, что барон главным образом прославился толкованием

«Апокалипсиса». В частности, он доказал, что антихрист восседает на престоле Петра.

Формально логарифм — это операция, обратная возведению в степень, наряду с извлечением корня. Так, если 10 в кубе = 1000, то известно, что кубический корень из 1000 равен 10, а также что 3 является логарифмом числа 1000 по основанию 10. Кроме того, мы имеем две обратные операции, потому что 10^3 не равно 3^{10} степени.

Однако Непера интересовало не это. Он изобрел вычислительный инструмент, позволяющий переводить умножение в сложение и деление в вычитание, — логарифмические таблицы (они еще существуют?), чтобы определять логарифм числа и наоборот. Не знаю, помните ли вы со школьных времен, что произведение двух степеней числа равно этому числу в степени суммы показателей. Ну так вот, это и было ключом к решению задачи.

До середины прошлого века калькуляторами работали люди, вооруженные бумагой и ручкой, и для них логарифмы были чем-то вроде манны небесной. Лаплас сказал, что логарифмы удвоили продолжительность жизни астроно́мов, подразумевая, что процесс вычисления небесных орбит значительно ускорился. Инженеры тоже активно пользовались логарифмическими линейками. Так, для операций умножения особым образом совмещались шкалы линейки, способные передвигаться относительно друг друга.

В наши дни в распоряжении астрономов имеется масса компьютерных программ, а логарифм — всего лишь одна из кнопок калькулятора. Зачем же тогда учить что-либо, если оно даже бесполезнее, чем латинский язык, зная который можешь по крайней мере притворяться, будто читаешь древние надписи? Ответ прост: чтобы уметь обходиться с порядками величин. Раньше производить операции с повседневными числами не составляло труда. Хотя количество дней в году, представленное числом 365, было достаточно значительным, с ним легко справлялись, но уже предпочитали считать не днями, а месяцами. В настоящее время даже в обиходе мы говорим о миллионах и миллиардах. Когда мы слышим, что внешний долг Италии достиг 2000 миллиардов евро, то должны быть благодарны, что лира уже не в ходу, иначе речь шла бы о 4 миллионах миллиардов. Мы не имеем ни малейшего представления о том, насколько велико это число. Если бы нам сказали: «Это 2 фантастикльона», результат был бы таким же. Но запишем это число в виде $2 \cdot 10^{12}$ (или 2E12, как предлагает калькулятор). В данном примере 12 — это степень, в которую нужно возвести 10, чтобы получить 1000 миллиардов. Я уверен, что число, представленное в таком виде, лучше поддается осмыслению. Естественно, все не так просто: например, чтобы перейти от 12-й к 13-й степени, необходимо умножить число на 10, а не добавить к нему единицу. Короче, тренироваться надо. Однако, как только вы научитесь оперировать логарифмами, сможете легко производить вычисления, в частности умножение и деление, которые, как я уже упоминал, превращаются в сложение и вычитание.

Например, какая доля внешнего государственного долга приходится на одного итальянца? Нас 60 миллионов, то есть $6 \cdot 10^7$. Чтобы определить долю каждого, необходимо сначала 2 разделить на 6 (получится 1/3) и от 12 отнять 7 (получится 5). Затем полученные цифры надо подставить в следующее уравнение: 1/3 умножить на 10, возведенное в 5-ю степень ($1/3 \cdot 10^5$). Это равняется 1/3 от 100 000, или около 33 000 евро. Вот наконец понятная цифра, хотя и вызывающая тревогу!

Подведем итог: работая с логарифмами, мы вместо умножения складываем, а вместо деления вычитаем.

И напоследок. Если вы считаете противоестественным обращаться к логарифмическим шкалам, знайте, что наши внутренние сенсоры работают аналогично, опираясь на относительные показатели, а не на абсолютные. Например, звездная величина, то есть видимая яркость звезд, определяется так: Вега получает значение 0, а звезда в 100 раз тусклее Веги — значение 5; разница между двумя звуками, интенсивность которых составляет 1 и 10, равняется 10 децибелам; шкала Рихтера откалибрована так, что при землетрясении увеличение магнитуды на 1,0 соответствует увеличению амплитуды колебаний в 10 раз и увеличению энергии примерно в 31,6 раза ($\sqrt{1000}$). Короче говоря, знание того, как обращаться с логарифмами, тоже можно применять на практике!

Стремительный рост

Журналисты предпочитают использовать выражение «экспоненциальный рост», когда описывают сложные ситуации: если вы заметили, экспоненциально растут наши траты, преступность, налоги, а не зарплаты, зеленые зоны и число рабочих мест. Печально, но в большинстве приведенных примеров рост вовсе не экспоненциальный! Что не так уж и странно. Это правда, что в математике и физике экспоненциальная функция встречается довольно часто. Вы знали, например, что форма электропровода, провисшего между двумя лампами, называется цепной линией, которая, в свою очередь, представляет собой композицию из двух экспоненциальных кривых? Однако, чтобы распознать эту функцию, требуется тренировка. Наверное, лучше начать с основ. Итак, возьмем наиболее легкие для визуализации функции — те, значение которых зависит от времени, и начнем с *линейного возрастания* — самого простого типа.

Говорят, что функция линейно возрастает, если разница между значениями, которые она принимает через равные интервалы, тоже одинакова. Представьте, что вы едете с постоянной скоростью 50 километров в час:

за 1 час вы преодолеете 50 километров, за 2 часа — 100, за 3 часа — 150. Если вы нарисуете соответствующий график, то получите прямую. Приведу другой пример: пешеход движется со скоростью, скажем, 5 километров в час, а самолет — 500, но каждый из этих случаев можно представить в виде графика — прямой линии с большим или меньшим углом наклона.

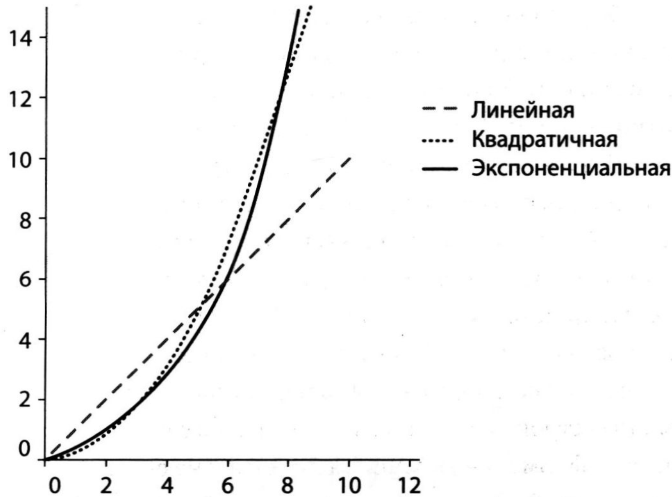
Однако есть и другие типы роста. Если скорость увеличивается линейно, после первой минуты путешествия достигая 10 километров в час, к концу второй минуты — 20, к концу третьей — 30 (и так до тех пор, пока ваша машина сможет разгоняться), то получается *квадратичная прогрессия*. Это, несомненно, быстрее линейной прогрессии, хотя в краткосрочной перспективе рост может быть даже медленнее; но для ученого не так уж они и отличаются, поскольку обе увеличиваются не слишком быстро.

Экспоненциальный рост — прогрессия другого типа. Ее самое простое определение схоже с определением линейной прогрессии, однако на равных отрезках остается неизменной не разница между значениями, а их отношение. Деление вместо вычитания. И это очень существенное различие! Есть достаточно известная задача, в которой амеба делилась надвое каждый день и заполнила чашку Петри за 10 дней. Если начать следующий эксперимент с двумя амебами вместо одной, сколько дней понадобится, чтобы заполнить чашку? Ответ, который дают те, кто думает о линейной прогрессии, — 5 дней. Увеличивая исходное количество амеб в два раза, мы во столько же раз уменьшаем количество дней. Но правильный ответ — 9 дней. Просто представьте, что если бы второй эксперимент начался на следующий день после первого, то в обеих чашках Петри содержалось бы по две амебы. Можно посмотреть на эту задачу и с другой стороны. Амеба тратит 9 дней для заполнения половины чашки Петри, и только 1 день, чтобы заполнить вторую половину. Вот что значит хороший прирост!

Еще один пример экспоненциального роста — на этот раз из области литературы — связан с именем Раймона Кено, который опубликовал «Сто тысяч миллиардов стихотворений». В книге всего 10 сонетов, каждый из которых разделен на 14 строк, напечатанных на отдельных полосках. Сборник построен таким образом, что, выбирая строки из любых сонетов, можно собрать новый сонет — формально правильный, хотя и не обязательно со смыслом. В 10 сонетах содержится 140 строк, что дает возможность составить сто тысяч миллиардов стихотворений. Выбрав строки случайным образом, вы получите стихотворение, которое еще никто никогда не читал!

Естественно, не всегда экспоненциальная прогрессия демонстрирует такой стремительный рост. Например, при увеличении на 0,1 % в год она достаточно долго не отличается от линейной. И все же, если набраться терпения, можно убедиться, что любая экспоненциальная прогрессия превзойдет любую и линейную, и квадратичную, — естественно, если это действительно прогрессия, то есть отношение, о котором я говорил выше, больше 1. В ином случае будет корректнее говорить о спаде, но это уже другая история (тоже очень красивая, однако эта глава посвящена совсем другой теме).

Надежда на то, что национальный ВВП вернется к росту 3 % в год, — большая утопия, по крайней мере в долгосрочной перспективе. При таком росте за 24 года



Три типа прогрессии

он удвоился бы, а за 100 лет увеличился бы в 16 раз. В реальности же ВВП изменяется скорее по графику линейной прогрессии — в 3 раза за 100 лет. Значительная разница, в общем. Добавлю также, что в большинстве случаев графики, которые иллюстрируют статьи в ежедневных газетах, содержат слишком мало значений, чтобы с их помощью можно было составить представление о том, что происходит на самом деле: вот и в графике, приведенном выше, парабола и экспоненциальная кривая кажутся достаточно похожими. Так что, возможно, говоря об «экспоненциальном» росте, журналисты просто хотят сказать, что речь идет об очень быстром росте, и при этом не слишком заботятся о выборе слов. Или же просто хотят показать, что знают такое классное словцо, как «экспоненциально».

Раздел II

Парадокс, возможность, прогноз

Один к тысяче не получится

После бактериологической атаки, совершенной сегодня неизвестной террористической группировкой, вас пригласили в клинику, чтобы определить, заразились ли вы болезнью, которая в течение недели приводит к летальному исходу. Известно, что вероятность контакта с вирусом — 1 к 1000. Точность теста составляет 99 %: это значит, что в 1 % случаев он дает ложноположительный результат (тест показывает, что человек заразился, хотя на самом деле он здоров) и в 1 % случаев — ложноотрицательный (тест показывает, что человек здоров, хотя на самом деле болен). Ваш результат оказывается положительным. Настало время срочно писать завещание? Совсем не обязательно!

Чтобы вычислить риск заражения болезнью, представим, что 1 миллион человек прошли этот тест. Тысяча из них действительно будут больны. Среди больных тест будет положительным для 990 человек и отрицательным для 10 (1 %); среди здоровых тест будет положительным для 9990 человек и отрицательным для остальных. Примем во внимание только положительные результаты теста: число здоровых в 10 раз превосходит число зараженных! Таким образом, заболеть можно в одном случае из одиннадцати, что составляет немногим более 9 %. Процент достаточно высокий, чтобы заставить человека провести несколько бессонных ночей, но с учетом 99%-ной точности теста положительные результаты не должны вызывать особое беспокойство.

Я вам только что привел типичный пример вычисления апостериорной вероятности, которую также называют байесовской по имени преподобного Томаса Байеса, который, возможно, что-то о ней писал. В действительности же первое исследование на эту тему опубликовал его друг Ричард Прайс в 1763 году — через два года после смерти Байеса.

Итак, первая оценка, которую мы делаем, является *априорной вероятностью*: «Какова вероятность того, что событие *X* (заражение болезнью) произойдет?» *Апостериорная вероятность* — это уточнение первоначальной оценки за счет введения деталей: «Какова вероятность того, что событие *X* произойдет, если мы знаем, что произошло событие *Y* (тест оказался положительным)?» Чтобы ее вычислить, необходимо знать *корреляцию* между *Y* и *X*, то есть как часто за первым событием следует второе. При условии совершенной корреляции, когда тест всегда дает безошибочный результат, вычисления просты: вероятность заболеть в реальности становится уверенностью. Если корреляция равна нулю, то есть тест настолько плох, что не дает никакой дополнительной информации (я мог бы не тратить время на его прохождение), вероятность заболеть остается априорной — 1 на 1000. В остальных случаях вам придется воспользоваться формулой Байеса. Если, конечно, вы ее помните. Если нет, поступайте как я:

представьте, что вы имеете дело с достаточно большим количеством примеров, и посмотрите, что произойдет.

Уметь отличать априорную вероятность от апостериорной очень важно, особенно в реальной жизни. Возьмем анализ ДНК, который определяет, является ли подозреваемый преступником: даже если точность этого теста достигает 99,99 %, в одной только Италии с ее населением в 60 миллионов результат был бы ложноположительным для 6000 человек. Безусловно, результат с такой погрешностью можно учитывать исключительно в качестве дополнения к другим, априорным уликам. Этот факт, к сожалению, часто игнорируют не только средства массовой информации (что уже серьезно влияет на ситуацию), но и судьи, которые, как правило, не имеют математических знаний, необходимых для понимания сути анализа ДНК.

Другой пример. Чтобы найти убийцу Яры Гамбиразио[1], было собрано 20 000 образцов ДНК взрослых мужчин, живших в районе Бергамо. Вероятность ошибки, равная 0,01 %, на практике означает, что двух человек могут ложно обвинить в убийстве. Как быть, если результаты анализа укажут на одного или трех человек? Что делать в таком случае?

---

1 Расследование убийства 13-летней Яры Гамбиразио, пропавшей 26 ноября 2010 года, считается одним из самых сложных за всю историю Италии. Вечером девочка вышла из спортивного центра в городке Брембате-ди-Сопра, но домой так и не вернулась. Тело нашли 26 февраля 2011 года в 10 километрах от Брембате. На нижнем белье девочки обнаружились следы крови преступника, но, несмотря на это, отыскать его оказалось совсем не просто.

Некоторые антропологические исследования показывают, что человек разумный эволюционировал под давлением обстоятельств, которые не требовали вычисления апостериорной вероятности. Например, при виде саблезубого тигра наши предки бежали прочь, не пытаясь предварительно узнать, ел он недавно или нет. Потому-то у нас и нет «байесовской интуиции». Но не беда: важно лишь научиться видеть, когда оказываешься в ситуации такого типа, и попытаться спокойно прикинуть ее исходы.

И наконец, простая задача, которая еще раз показывает, насколько важно обращаться с вероятностью осторожно. Арбуз весом 10 килограммов практически весь (если не учитывать корку) состоит из воды — на 99 %. После того как его оставили на солнце, он немного усох и теперь состоит из воды только на 98 %. Сколько весит арбуз сейчас? Девять с половиной килограммов? Больше? Попробуйте подсчитать, и, может быть, вы удивитесь. Хотите подсказку? Часть, которая не является водой, по-прежнему весит 100 граммов.

Парадокс двух конвертов

Карло и Аличе приняли участие в научном эксперименте. Их подвели к столу, на котором лежали два конверта, и предложили выбрать один из них. После этого друзей разделили, запретив общаться между собой, и попросили открыть конверты. В конверте лежали чек и записка. В записке объяснялось, что сумма чека в одном из конвертов в два раза больше, чем во втором, но не указывалось, в каком именно находится бо́льшая сумма. Испытуемым задали вопрос: «Согласны ли вы обменять свой конверт на конверт друга?» Какой стратегии им стоит придерживаться?

Предположим, что в конверте Аличе лежит чек на a евро. В таком случае можно рассуждать следующим образом: «Вероятность того, что у меня конверт с большей суммой денег, составляет 50 %. Если обменяться конвертами, у меня будет $a/2$ евро. Но вероятность того, что сейчас у меня чек на меньшую сумму, такая же. Значит, если обменяться, у меня будет $2a$ евро. Поскольку оба варианта имеют одинаковую вероятность, в среднем при обмене конвертами я заработаю $(5/4)a$ евро. Таким образом, мне лучше обменяться конвертами». Карло, у которого конверт с чеком на сумму *b*, будет размышлять аналогично; для него средняя сумма при обмене конвертами также будет равна $(5/4)b$ евро относительно *b*.

Возможно ли обменяться конвертами так, чтобы это было выгодно обоим участникам эксперимента? Для решения данного парадокса проще всего представить, что суммы чеков выбраны случайным образом, но ниже определенного предела. Для простоты допустим, что он составляет 100 евро. Это значит, что меньший чек будет достоинством от 1 цента до 50 евро, в то время как больший — от 2 центов до 100 евро. Случай с нулевой суммой можно исключить: в любом случае при удвоении нуля получается ноль, и таким образом, с одной стороны, нет смысла думать над обменом конвертов, с другой — исследователям не поздоровилось бы, вздумай они так пошутить. Но тогда, если кто-нибудь находит чек на 90 евро, он знает, что является обладателем чека на бóльшую сумму, и, конечно, откажется менять свой конверт. Итак, мы обнаружили, в чем разница между ситуациями, в которых оказались двое испытуемых: теперь уже не скажешь, что вероятность выиграть от обмена конвертов у обоих одинаковая. Въедливый читатель мог бы здесь добавить, что если кто-нибудь находит чек на 42,85 евро, то это явно меньшая сумма, ведь полуцентовых монет не бывает. Математик, с неодобрением глядя на такого читателя, ответил бы, что это банки занимаются округлением цифр, а не исследователи. И если вам это не нравится, представьте также, что исследователи были достаточно хитры, чтобы использовать исключительно чеки с четным числом центов (разумеется, без вашего ведома)...

В общем, у меня нет ни малейшей охоты приводить вычисления — при желании их можно найти в «Википедии». Однако хотелось бы показать математически, как лучше действовать в определенных условиях. Если заранее известно, что значения чеков были присвоены случайным образом в пределах между n и N евро, то необходимо обменяться конвертами, если сумма вашего чека меньше $2n$ евро, и оставить конверт себе, если она больше $N/2$ евро. При других условиях неважно, какое решение вы примете. На вопрос, допустимо ли свести n к нулю и принять, что N равняется бесконечности, если максимальная сумма неизвестна, ответом будет «нет». Половина бесконечности всегда бесконечность, и в таком случае неизвестно, начиная с какой точки можно быть уверенным в том, что обладаешь чеком на большую сумму.

Как же разрешить парадокс? Никто не знает. Приведенная выше аргументация Карло и Аличе, согласно которой чек на бо́льшую сумму может с одинаковой вероятностью оказаться в любом из конвертов, неприменима в случае потенциально неограниченной суммы: шанс выиграть сводится к нулю при бесконечном количестве равновероятных обменов. Возможно, лучший способ выйти из тупика — это «выйти из системы». Если посмотреть на ситуацию со стороны и определить общую сумму в двух конвертах как $3a$, то при обмене конвертами один из участников заберет a евро (таким образом, не увеличит выигрыш вдвое), в то время как другой потеряет *a евро* (таким образом, не уменьшит выигрыш вдвое). Видите, что стоит просчитывать?

Игра Пенни

При игре в орлянку испытываешь сильное, но быстро проходящее возбуждение: один бросок — и все окончено. Хотите продлить удовольствие? Предложите другу вариант орлянки с подбрасыванием монеты три раза подряд. Суть данной игры состоит в угадывании комбинации, которых существует всего восемь. Если мы условимся, что О — это орел, а Р — решка, то возможные комбинации можно представить следующим образом: ООО, ООР, ОРО, ОРР, РОО, РОР, РРО, РРР. Пусть ваш друг выберет какой-нибудь вариант, вы выбираете другой, и игра начинается. Бросайте монету и записывайте последовательность О и Р. Побеждает тот игрок, чей вариант появится первым. Вот и все. Самое сложное заключается в том, чтобы найти простофилю, который захочет бросить вам вызов.

Совершенно очевидно, что первый игрок при выборе комбинации должен руководствоваться логикой. Представим, например, что он ставит на ООО (обратите внимание на таблицу). Трижды выбросить орла при первых трех бросках может только невероятный везунчик или обладатель «крапленой» монеты. Такой шанс выпадает один раз из восьми при условии честной игры. Как будут

Вероятность победы в игре Пенни

А выбирает	В выбирает	В побеждает с вероятностью, %
ООО	РОО	87,5
ООР	РОО	75,0
ОРО	ООР	66,7
ОРР	ООР	66,7
РОО	РРО	66,7
РОР	РРО	66,7
РРО	ОРР	75,0
РРР	ОРР	87,5

развиваться события, если фортуна не на вашей стороне? Представим, что сочетание ООО выпало между шестым и восьмым бросками монеты, и вернемся к пятому броску. Мог выпасть орел? Нет, потому что в таком случае мы остановимся на седьмом броске с комбинацией из трех орлов. Таким образом, известно, что при пятом броске выпадет решка, — тогда при выборе комбинации РОО нас ожидает победа. Аналогично следует рассуждать в остальных случаях, кроме описанного выше, когда нам сказочно везет. Итак, выбрав комбинацию РОО, мы в среднем будем побеждать семь раз из восьми.

Хорошо, допустим. Но, может, все дело в том, что триплет ООО особенный? Вовсе нет. Как видно из таблицы, каким бы ни было сочетание, выбранное изначально вашим противником, вы всегда можете выбрать комбинацию, при которой вероятность победить будет больше 50 %. В целом не существует лучших или худших комбинаций. На рисунке стрелка обозначает, что первый триплет сильнее второго, и становится очевидно, что, кроме глупого выбора вроде ООО и РРР, а также комбинаций с чередованием орла и решки, не существует однозначно выигрышного варианта, но есть цикл.

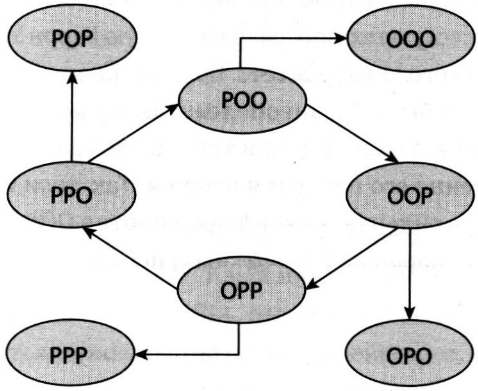


Схема выигрышных комбинаций

Математики в подобных случаях говорят о нетранзитивных связях. Как правило, нам привычно думать, что если *A* больше *B*, а *B* больше *C*, то *A* больше *C*. Но, как мы увидели, это не всегда так. Кстати, не такая уж это и редкость. Все мы знаем игру «Камень, ножницы, бумага»: камень побеждает ножницы, потому что разбивает их; ножницы побеждают бумагу, потому что разрезают ее; бумага побеждает камень, потому что в нее можно завернуть камень. С другой стороны, в чем удовольствие от игры, если заранее знаешь, кто победит?

Усложненный вариант орлянки называется игрой Пенни по имени ее создателя Уолтера Пенни, который в 1969 году написал соответствующую статью на целых десять строк в Journal of Recreational Mathematics. Но, как обычно случается, широкую известность игра получила лишь в 1974 году, когда Мартин Гарднер написал о ней в своей рубрике «Математические игры» в журнале Scientific American. Если вы не желаете заучивать наизусть выигрышные комбинации, можете использовать мнемоническую технику, разработанную Барри Уолком из Манитобского университета. Первый элемент комбинации должен быть противоположным второму элементу противника, а наш второй и третий элементы будут соответственно его первым и вторым. Так, если нам надо вычислить сильную комбинацию против OOP, то выбор будет POO. Хороших (возможных) побед!

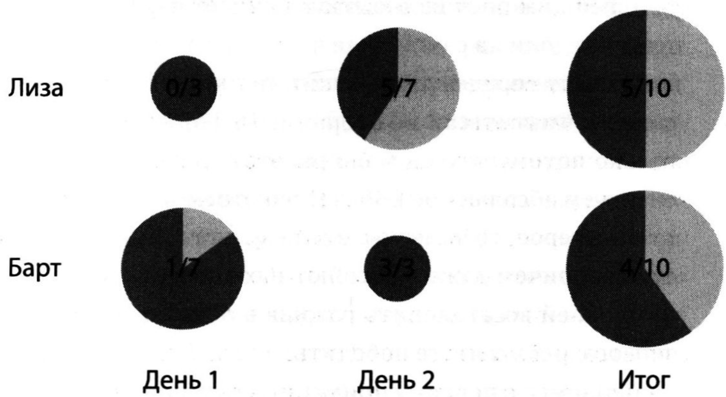
Парадокс Симпсона

Вы ведь не считаете, что «розовые привилегии» — итальянское изобретение?! Как и многие другие традиции, эта пришла к нам из США, где вот уже почти полвека пользуются термином «позитивное действие», которому в Италии соответствует выражение «позитивная дискриминация». В Штатах вовсю практикуется внутригрупповой фаворитизм по гендерному или расовому признаку, призванный оказывать помощь ущемленным категориям населения. Да уж, в таких вещах американцы разбираются по-настоящему. Только представьте себе: осенью 1973 года Калифорнийский университет в Беркли был обвинен в дискриминации женщин, подавших документы для поступления. Данные были очень точными: в университет поступили 44 % мужчин из 8442 подавших заявление и 35 % женщин из 4321. Разница оказалась слишком существенной, чтобы можно было сослаться на банальную статистическую погрешность. В ответ университет обнародовал информацию о зачислении студентов на различные факультеты. Согласно представленным данным, большинство факультетов приняло меньше мужчин, чем женщин, в процентном соотношении. Соответственно, ни о каком ущемлении прав женщин не могло быть и речи. Парадоксально, правда?

Приведу более шокирующий пример. В 2009 году уровень безработицы в США достиг 10,2 %, и многие эксперты задались вопросом, был ли этот кризис серьезнее того, что разразился в 1980-х. И сошлись во мнении, что нет, ведь в конце 1982 года без работы осталось 10,8 % населения. Однако при анализе безработицы по отдельным категориям работников (выпускники вузов, специалисты с опытом работы в университете, квалифицированные специалисты, неквалифицированные рабочие) обнаружилось, что в каждой из них уровень безработицы был выше, чем 27 лет назад.

Результаты поступления в Беркли и уровень безработицы — наиболее яркие примеры довольно распространенного явления, которое получило название «парадокс Симпсона». Эдвард Симпсон (а вовсе не мультяшный Гомер Симпсон) описал его еще в 1951 году. Несмотря на то что концепцию опубликовали на 50 лет позже первых исследований, проводившихся в этой области, Симпсону повезло и феномен носит его имя.

Однако дать имя понятию совсем не означает объяснить его! Размышляя над делом Беркли, можно прийти вот к какому заключению: парадокс вызван тем, что мужчины подали гораздо больше заявлений на поступление в университет, чем женщины. Но более глубокий анализ показывает, что причина не в этом, ведь мы оперируем процентами, а не абсолютными значениями. Впрочем, интуитивный ответ не такой уж и ошибочный, как показывает следующий простой пример.



Кто лучше?

На конкурсе профессионального мастерства Барт и Лиза должны отремонтировать по десять компьютеров за два дня. В первый день Лиза пытается починить три компьютера, но у нее ничего не выходит. Барт чинит один компьютер из семи. Результаты совсем невысокие, но Барт однозначно проявил себя лучше, по крайней мере неплохо. Во второй день показатели значительно улучшаются. Барту удается отремонтировать все три компьютера из тех, с которыми он еще не работал. Лиза, в свою очередь, чинит пять из семи. Таким образом, и во второй день у Барта выше процент отремонтированных компьютеров. Станет ли он в итоге победителем? Барт в общей сложности починил четыре компьютера из десяти, Лиза отремонтировала пять компьютеров. Здесь опять проявляется парадокс Симпсона. Итак, лучший ремонтник — Лиза, поздравим ее традиционным восклицанием мультяшной семейки Симпсонов: «Д'оу!»

Теперь наконец видно, что вызывает парадокс. Процентные доли на самом деле в пользу Барта, однако использование процентов слишком жестко гомогенизирует данные. Фактически во второй день Барт выигрывает только потому, что он мало работал: если мы проанализируем абсолютное количество отремонтированных компьютеров, то заметим, что результат Лизы гораздо лучше, причем даже в абсолютных значениях, что позволило ей восстановить разрыв в показателях после первого дня и в итоге победить.

Вернемся к первому примеру о поступлении в Беркли: более глубокое исследование показало интересные результаты. Женщины чаще подавали документы на гуманитарные факультеты, которые были намного популярнее, — следовательно, при высоком конкурсе количество отказов было больше. В то же время мужчины отдавали предпочтение естественно-научным факультетам, где конкурс был ниже.

Что касается примера с безработицей, то повышение уровня образованности между 1982 и 2009 годами привело к искажению общей статистики.

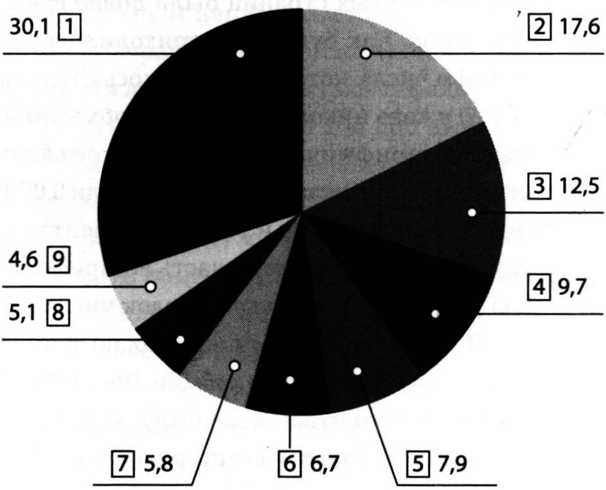
Ну а на вопрос о том, какие же данные являются корректными, ответом будет: «Все зависит от ситуации». В случае с Лизой и Бартом, естественно, важен конечный результат, а распределение объемов работы по дням не имеет значения. В деле Беркли университет не виноват, так как не удалось обнаружить предвзятости по отношению к женщинам. А на более глобальном уровне можно утверждать, что для борьбы с дискриминацией по гендерному признаку необходимо убедить женщин поступать на исконно «мужские» факультеты. Наконец, если говорить об американской безработице, я подозреваю, что каждый политик выберет данные, которые для него будут более удобными. А еще считается, что в математике нет мнений!

Закон Бенфорда

Будучи злостным неплательщиком, вы создали предприятие-призрак, чтобы обмануть налоговую. Сейчас вам остается лишь составить тысячу фальшивых счетов-фактур с суммами от 100 до 100 000 евро и отправить документы в налоговое управление. Но какие конкретно суммы указать в отдельных счетах? Вы прочли в научном приложении к одному журналу, что человек не способен генерировать по-настоящему случайные числа. Таким образом, вы идете на сайт random.org, и алгоритм генерирует для вас 976 (тоже случайное число) значений, равномерно распределенных в интервале от 100 до 100 000. Ну что, ваш план безупречен? В Италии, может быть, и да, а вот в США подобную попытку уйти от уплаты налогов раскрыло специально предназначенное для этого агентство. На самом деле если бы счета-фактуры были настоящими, то сумм, начинающихся с цифры 1, оказалось бы в семь раз больше, чем тех, которые начинаются с 9.

Что за всем этим скрывается? Закон Бенфорда. В математике слово «закон» означает, по сути, простую констатацию факта — в отличие от теоремы, доказательство которой рассматривается подробно. Кроме того, это слово нередко становится предметом шуток. Так, если закон носит имя Бенфорда, можно с полной уверенностью утверждать, что это не Бенфорд его открыл. И действительно, первая публикация на эту тему принадлежит некоему Саймону Ньюкому, занимавшемуся астрономией в конце XIX века. Как и все астрономы, Ньюком вынужден был производить уйму вычислений и поэтому использовал логарифмические таблицы (помните, я уже о них говорил?). Однажды на досуге он заметил кое-что необычное: поля первых страниц были более грязными, чем последних. Как будто ему приходилось чаще искать логарифм числа, которое начиналось с меньших цифр. (Для тех, у кого никогда не было необходимости использовать логарифмические таблицы, объясняю. Чтобы найти логарифм числа 42, или 42 000, или 0,00042, осуществляется поиск одного и того же элемента (4,2): таблицы дают только десятичную часть логарифма, так называемую мантиссу, поскольку порядок числа легко установить.) В 1881 году Ньюком написал по этому поводу статью, которую все быстро забыли. Понадобилось более полувека для того, чтобы кто-то другой попытался пролить свет на эту любопытную историю. Физик Фрэнк Бенфорд тогда работал в исследовательской лаборатории «Дженерал Электрик» и, вероятно, мог позволить себе проводить исследования такого рода. Он собрал множество данных всех типов и в 1938 году, изучив более 20 000 значений различного вида, изложил результаты исследования в статье, где описал вероятность появления той или иной первой цифры в распределениях величин, взятых из реальной жизни. Так родился закон Бенфорда.

В математической формуле закона, естественно, используются логарифмы. Я не стану ее приводить и ограничусь демонстрацией диаграммы, на которой видны проценты, соответствующие разным начальным цифрам числа. Они пропорциональны количеству страниц в таблице логарифмов, начинающихся с этого числа. На практике первые цифры чисел меняются, на первый взгляд, странным образом, но первые числа после запятой логарифмов чисел распределяются равномерно.



Вероятность распределения начальных цифр согласно закону Бенфорда

Можете самостоятельно проверить этот закон: возьмите достаточно большую подборку упорядоченных числовых данных и обратите внимание на первые цифры. Здесь надо учитывать только то, что все значения должны находиться в пределах нескольких порядков и не должны характеризовать некий специфический процесс или явление; они должны быть независимыми друг от друга. Скажем, для проверки закона Бенфорда не годятся значения роста восемнадцатилетних юношей и девушек в сантиметрах: могу поспорить, что более 95 % этих данных окажутся в пределах между 100 и 195 сантиметрами. Если подбросить монету 1000 раз, то количество выпавших орлов будет начинаться с 4 или 5. А вот число жителей итальянских городов, составляющее от нескольких десятков человек до миллионов, — отличный пример, для которого действует закон Бенфорда.

Так существует закон или нет? Ясный и четкий ответ: «данет». Можно лишь отметить, что если закон действительно работает для определенного набора данных, то он не должен зависеть от величин, в которых эти данные измеряются. Другими словами, распределение должно остаться неизменным — это так называемая *масштабная инвариантность*.

Предположим, что первые цифры совокупности данных, измеренных в килограммах, удовлетворяют какому-то определенному закону. Значит, этот же закон должен действовать и при пересчете веса в полукилограммы или в немецкие фунты. В таком случае все значения, которые до пересчета начинались с цифры от 5 до 9, теперь будут начинаться с 1. Следовательно, вероятность начала с 1 должна быть равна сумме вероятностей начала с 5, 6, 7, 8 или 9, а это именно то, что мы видим на диаграмме. Я знаю, что это одна из так называемых демонстраций на пальцах... А впрочем, какая разница? Я же говорил вам, что это не теорема, а закон. Разве нет?

Сколько весит «Википедия»?

Энрико Ферми можно без тени сомнений признать одним из величайших физиков XX века (и я так говорю вовсе не потому, что он тоже был итальянцем). Обычно его имя связывают с созданием атомной бомбы. И он действительно возглавлял команду исследователей, которые доказали необратимый и летальный характер цепной ядерной реакции. Ферми проявил себя и во многих других направлениях физики — существует как минимум два открытия, которые носят его имя.

Парадокс Ферми возник из фразы, которую он произнес за ужином с коллегами. Завязался разговор об НЛО, и Ферми прокомментировал: «Да, но где они все?» По его мнению, странно, что во Вселенной, существующей много миллиардов лет, человечество до сих пор не обнаружило признаков деятельности внеземного разума, даже таких, как электромагнитные сигналы. Для поиска следов инопланетян был создан проект SETI, который, впрочем, пока не принес особых результатов. На вопрос, заданный Ферми, можно дать много ответов, но в таком случае мы выйдем за рамки нашей темы.

Гораздо более интересными — по крайней мере с точки зрения математики — являются *задачи Ферми*. Они тоже родились в неформальных беседах с коллегами и друзьями. Говорят, Ферми преуспел в вычислении приблизительных ответов при наличии малого количества данных. Например, он смог оценить мощность ядерной бомбы в ходе первого ее испытания, измерив, как далеко ударная волна сместила клочки бумаги, которые он подбросил в воздух во время взрыва.

В одной из самых известных задач требуется определить число настройщиков пианино в Чикаго. На первый взгляд кажется, что ее невозможно решить, но Ферми показал, как получить правдоподобный результат, если воспользоваться простым алгоритмом: надо установить, сколько семей живет в Чикаго и его окрестностях; прикинуть, у скольких из них может быть пианино; узнать, сколько времени требуется, чтобы приехать к кому-то домой и настроить пианино; наконец, определить, сколько времени в год работает настройщик.

Но давайте обратимся к практическому примеру: сколько будет весить вся «Википедия» на итальянском языке, если напечатать ее на бумаге (с учетом двусторонней печати)? Чтобы ответить, необходимо произвести некоторые вычисления. В «Википедии» на момент создания книги, которую вы сейчас читаете, имелось чуть более миллиона статей. Можно предположить, что в среднем каждая из них занимает полторы страницы, — естественно, одни статьи меньше, другие больше. Короче, мы имеем чуть менее миллиона листов формата А4, скажем 800 000, заполненных с обеих сторон. Сейчас мы должны использовать малоизвестный, но любопытный факт: площадь листа А4 составляет шестнадцатую часть квадратного метра (чуть позже мы увидим почему), а бумага для принтера весит 80 граммов на квадратный метр. Следовательно, один лист весит 5 граммов. Таким образом, мы получаем 4 миллиона граммов, или 4 тонны. Не знаю, как вы, но я ожидал большего.

А если не использовать для решения этот малоизвестный и любопытный факт? Нет проблем, ищем другой способ. Пожалуй, нужно вспомнить, что площадь одного листа А4 равняется приблизительно 20 × 30 сантиметров, что в пачке содержится 500 листов, что в высоту она составляет около 5 сантиметров и что бумага весит более или менее как вода. Таким образом, 500 листов весят около 3 килограммов, а один лист — около 6 граммов, что дает общий вес «Википедии» почти 5 тонн. Упс, а результат-то другой. И что теперь делать? Ничего. Запомните: мы не ищем точный ответ, ведь мы не должны платить налоги на печать «Википедии» и нам не нужно знать, сколько марок понадобится приклеить. Нам достаточно приблизительной оценки.

Четыре или пять тонн — разница в данном случае несущественная. Конечно, если бы вторая оценка составила 1 центнер или 100 тонн, тогда да, где-то явно была бы ошибка. Ради любопытства я попросил нескольких друзей решить эту задачу — они дали ответы в пределах от 1,5 до 5 тонн. Я бы сказал, это отличный результат, потому что мы все назвали величины одного порядка, что является одним из практических применений логарифмов. Логарифм различных оценок веса в килограммах на самом деле заключается между 3,2 и 3,7.

Метод Ферми — типичный пример благородного искусства измерения на глаз, которое я называю королевой приблизительных наук. Это не что иное, как догадки — довольно полезное упражнение, помогающее привыкнуть к нестандартному мышлению. Тем не менее интуитивные измерения приносят и практическую пользу, когда необходимо оценить реальные вычисления. Во-первых, если результат, полученный другим способом, слишком отличается от найденного с помощью метода Ферми, причиной может служить банальная ошибка в вычислениях. Во-вторых, сам по себе процесс рассуждений позволяет понять, какой параметр следует уточнить, чтобы результат интуитивной оценки приблизился к реальному. Так, в случае веса «Википедии» самым слабым звеном в цепи рассуждений является оценка среднего объема статьи. Усилить эту позицию несложно: надо случайным образом выбрать достаточное количество статей, чтобы иметь значимую выборку, и благодаря обновленным данным получить гораздо более точный результат. Просто и эффективно, не правда ли?

Смещение в центр

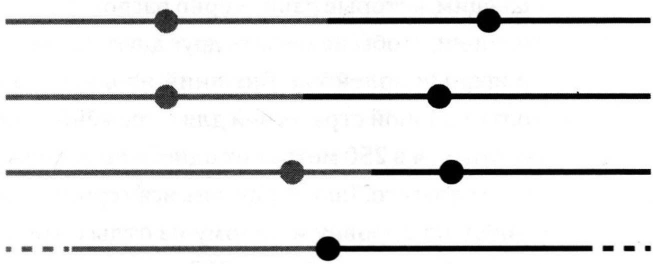
Коалиционная математика — сложная и разочаровывающая вещь для любителей детерминизма. Кеннет Эрроу, лауреат Нобелевской премии по экономике, внес значительный вклад во многие направления экономической теории. В том числе он разработал теорему, которая ответила на вопрос о рациональности коллективного выбора, волновавший умы исследователей на протяжении столетий. Смысл теоремы состоит в том, что не существует системы общего голосования, которая удовлетворяла бы ряду вполне справедливых условий и при этом всегда давала бы логически непротиворечивый результат в виде выбора оптимального кандидата из трех возможных. В определенном смысле теперь у нас есть математическое подтверждение того, что можно наблюдать после каждых выборов: все победили.

Но электоральные подсчеты вскрывают некоторые проблемы и в биполярной системе, где вроде бы должна быть очевидна победа большинством голосов. Вот вам математическое объяснение, почему два кандидата часто кажутся такими похожими друг на друга. Представьте великолепный пляж длиной километр, на котором два передвижных киоска держат монополию на продажу мороженого. Ограниченная площадь пляжа, похоже, не позволяет другим конкурентам предлагать аналогичный товар отдыхающим, которые равномерно располагаются по всей территории, чтобы не мешать друг другу загорать и играть в пляжный волейбол. Внешний наблюдатель скажет, что оптимальной стратегией для мороженщиков будет расположиться в 250 метрах от одного края пляжа и в 250 метрах от другого. Таким образом, вся территория окажется разделена поровну и никому из отдыхающих не придется преодолевать более 250 метров, чтобы утолить жажду.

Однако мороженщик, чей киоск стоит с правого края, рассуждает так: если я смещусь на 100 метров к центру, то сохраню всех своих клиентов. При этом клиент, который находится между 450-м и 500-м метрами пляжа, будет ближе ко мне, чем к моему конкуренту, и придет ко мне, так как я располагаюсь не на 750-м, а на 650-м метре пляжа! Но ведь и второй продавец не хочет терять клиентов, поэтому делает то же самое. Мало того, он передвигается на расстояние в два раза большее и располагается на 450-м метре. Так он становится единственным продавцом для всего левого участка пляжа. Он не только принимает своих старых клиентов, но также привлекает тех, кто находится между 500-м и 550-м метрами.

В конце всех этих танцев с перемещениями рождается новое решение: две тележки стоят бок о бок в центре пляжа, привлекая одних и тех же клиентов. Тем временем отдыхающие, которые находятся по краям пляжа, должны пройти 500 метров, чтобы добыть мороженое. Прекрасный результат, не так ли? Достигнута ситуация

равновесия — ни одному из мороженщиков не стоит удаляться от центра, иначе он потеряет часть своих голосующих... извините, отдыхающих.



Центр привлекает

Легко ли перенести этот подход на изучение выборных кампаний? Конечно, нет. Математические модели — это именно модели, они учитывают только некоторые аспекты действительности. Здесь проблема не в том, что мы предположили, что отдыхающие равномерно распределены по всему пляжу. Итоговый результат не изменится, если мерить мы будем не в метрах, а в купальщиках — по 500 человек с каждой стороны. Это, кстати, также важно в математике, нужно уметь правильно выбрать способ измерения.

Беда заключается в другом. Продолжим пример с пляжем: что произойдет, если «экстремисты» — то есть отдыхающие, которые находятся в крайних точках пляжа, — решат, что не стоит преодолевать весь этот путь ради жалкого мороженого? Их мороженщик потеряет «голоса», в которых был уверен. Таким образом, если оба продавца являются центристами, среди отдыхающих увеличивается количество «воздержавшихся». В этот момент периферийные зоны пляжа могут занять новые мороженщики, к которым переметнутся разочарованные клиенты. Скорее всего, таких будет немного, но они не вернутся к старым торговцам.

Наконец, как можно изменить модель, чтобы учесть сторону, которая явно отвергает лево-правую оппозицию? Это можно сделать, если ввести продавца нового товара, например бутербродов вместо мороженого.

Так или иначе, помните, что математические модели — это только модели: забавные, полезные, но не обязательно применимые к реальности!

Раздел III

Игры

Удваиваете ставку? Нет, оставляю

В XX веке название Санкт-Петербурга, второго по значимости города России, менялось несколько раз. Сначала большевики дали ему русифицированное имя Петроград, затем переименовали в честь «отца Родины» Владимира Ильича Ульянова, которого друзья звали Лениным. После падения коммунизма Ленинград вернул себе первоначальное имя. Только подумайте о той массе документов, которую приходилось каждый раз переделывать. Но математики всегда предпочитали связывать исконное название города с Леонардом Эйлером и с поистине невероятным парадоксом.

Представьте, что вы находитесь в казино Санкт-Петербурга... Там вообще есть казино? Ну, предположим, что да. Итак, пройдите в маленькую комнату, где крупье предложит вам «беспроигрышную» игру. Она очень проста: кто хочет поучаствовать, платит определенную сумму и потом начинает бросать монету, выбрав предварительно орел или решку. Если при первом броске выпадает решка, игра заканчивается и каждый игрок получает по рублю. Если выпадает орел, игрок снова бросает монету, и в этот раз на решке игра заканчивается, но участники получают по 2 рубля. Если же снова выпадает орел, то игра продолжается дальше. Если решка впервые выпадет на третьем броске, то игрок получает 4 рубля, на четвертом броске — 8 рублей, на пятом — 16 и так далее с постоянным увеличением суммы вдвое.

Представьте, что в банке неограниченное количество денег и игра, таким образом, может продолжаться до бесконечности: как и говорил крупье, она «беспроигрышная». Но сколько же вам необходимо заплатить за право поучаствовать в этой игре? Два рубля? Пять? Десять?

Произведем подсчеты. Это значит, необходимо взять все возможные случаи, умножить вероятность каждого случая на сумму выигрыша и суммировать все эти значения. Вероятность закончить партию на первом броске и выиграть 1 рубль составляет 1/2. Таким образом, данный случай стоит полрубля. Вероятность закончить партию на втором броске и выиграть 2 рубля составляет (1/2) × (1/2), то есть 1/4. Такой случай также стоит полрубля. Как вы, наверное, догадались, вероятность закончить партию на третьем броске составляет (1/8) × 4 = полрубля; на четвертом (1/16) × 8 = полрубля. И так далее. Сумма бесконечных полрублей бесконечна. Следовательно, какой бы ни была сумма, которую вы готовы заплатить за участие в игре, крупье, который не должен позволять банку проигрывать, вам вежливо откажет. Но все это кажется бессмысленным! Что за этим скрывается?

Ответ прост: чтобы теоретические подсчеты работали, не только игра должна продолжаться до бесконечности, но и банк должен иметь в распоряжении бесконечный резерв денег! Выигрыш увеличивается экспоненциально, а мы уже видели, как стремительно может возрастать экспоненциальная функция — даже слишком стремительно. Например, годовой ВВП всей Земли соответствует двум с половиной миллионам миллиардов рублей — приблизительно столько выиграл бы участник, которому удалось бы последовательно выбросить решку 51 раз. Правда, вероятность этого такая мизерная, что оказывается практически нулевой, но слово «практически» в математике не имеет прописки. При ограничении же максимального выигрыша стоимость одной партии в честной игре падает. При условии, что порог выигрыша составляет миллион рублей, средний выигрыш равен 11 рублям. Если максимальный выигрыш представлен миллиардом рублей (кругленькая сумма, равнозначная 20 миллионам евро), то средний выигрыш увеличивается лишь на 5 рублей, достигая в сумме 16 рублей, то есть немногим более 30 евроцентов.

Конечно, такого зала в санкт-петербургском казино не существует — как, может быть, не существует и самого казино. Но многие игроки убеждены, что стратегия, согласно которой надо начинать со ставки в один жетон на простых комбинациях (таких как «нечетное/четное» или «красное/черное») и удваивать ставку каждый раз, когда проигрываешь, несомненно, приведет к победе. Эта стратегия базируется на теории мартингалов; она считается выигрышной, поскольку алгебраическая сумма ставок меньше стоимости победной ставки, когда она наконец играет. Очень жаль, однако, что теория работает только при условии существования бесконечных денежных запасов у дилера и, что важно, огромного наследства у игрока и желания играть до бесконечности. Поверьте: дилер почти всегда остается в выигрыше и редко когда отказывается от игры. Конечно, этого недостаточно, чтобы вы вернулись домой с выигрышем (а вот о том, как этого добиться, я расскажу ниже).

Как выиграть в рулетку

Как вы, возможно, знаете, европейская рулетка представляет собой диск, разделенный на 37 равных секторов, которые пронумерованы от 0 до 36. Если вы поставите на определенное число и оно выпадет, то получаете в 37 раз больше суммы, которую поставили; если же оно не выпадет — вы потеряете деньги. Статистически выигрыш банка составляет 1/37, то есть примерно 2,7 %. Это значит, что, если 37 игроков ставят по жетону на все номера, банк забирает 37 жетонов, из которых 36 дает счастливому победителю.

Допустим, ваш бюджет — 105 евро. Вы решаете сделать 105 последовательных ставок по 1 евро на одно и то же счастливое поле. Как думаете, вероятность того, что вы покинете казино с выигрышем, составляет 10 %? А может, 30 %?

Вы не поверите, но она больше 50 %! Прежде всего легко увидеть, что достаточно выиграть трижды, чтобы получить 108 евро и, таким образом, иметь больше денег, чем вначале. Если произвести расчеты, то вероятность не выиграть ни одного раза составляет 5,63 %; вероятность выиграть один раз — 16,42 %; два раза — 23,72 %. В сумме получаем 45,77 %. Следовательно, вероятность того, что ваша ставка сыграет как минимум три раза и вы уйдете с выигрышем, равняется 54,23 %. Даже в американской рулетке, в которой добавлен второй ноль, чтобы обеспечить казино бо́льшие выигрыши, эта простая стратегия будет работать в 52,4 % случаев.

Где подвох? Я гарантирую, что подсчеты верны: с вероятностью более 54 % вы уйдете из казино, имея больше денег, чем когда входили в него. Жаль только, что речь идет о правильном ответе на неправильный вопрос. А правильный вопрос звучит так: «С каким количеством денег я могу выйти из казино в среднем?» Ответ: сумма составит примерно 102,16 евро, что связано с потерями, равными в среднем 1/37 от поставленных денег. Хорош парадокс, правда? Попытаюсь объяснить его на примере посерьезнее.

Русская рулетка. У нас пистолет с барабаном на шесть выстрелов, заряженный одним патроном. Крутим барабан, приставляем к виску и стреляем. Мне не нравятся кровавые примеры; чтобы не стрелять по-настоящему, я возьму пистолет, который выпускает флажок «Бах». Правила игры следующие: если пистолет стреляет впустую, банк дает вам 10 евро; если же вылетит «Бах», вы должны заплатить в банк 1000 евро. Пусть даже нет риска для жизни, я полагаю, вы, скорее всего, откажетесь играть. Вероятность выложить 1000 евро намного выше, чем шанс получить 10 при победе. Однако давайте призадумаемся на минутку: в конце концов, можно вернуться домой с деньгами в пяти случаях из шести, не так ли? Так почему же не попробовать?

Происходит то же самое, что и при 105 партиях игры в рулетку, пусть интуитивно это и сложно понять. Действительно, выигрываете вы чаще, чем проигрываете, но в большинстве случаев выигрыш слишком мал и, соответственно, вероятность вернуться домой с хорошим заработком ничтожна. В то же время есть отличная возможность потерять все деньги или почти все. Попробуем сравнить ситуацию с восхождением на вершину. Рывками по 1 метру вам удается подняться на 6–7 метров, но затем соскальзываете на 10 метров вниз. В итоге вы оказываетесь ниже исходной точки, несмотря на то что «практически всегда» поднимаетесь.

Теория мартингалов, о которой я говорил в предыдущей главе, сталкивается с такой же проблемой при условии ограничения наличных средств. Зайдите в казино с суммой 127 евро. Сделайте простую ставку («красное/черное», «чет/нечет», «большое/малое») в размере 1 евро. Если вам везет, забираете деньги и убегаете. Если же вы проигрываете, продолжайте играть, только теперь поставьте 2 евро. Если на этот раз вам удается выиграть, ваш чистый выигрыш составляет 1 евро: опять же забираете и сбегаете. Продолжайте в том же духе, каждый раз увеличивая ставку до тех пор, пока не выиграете. Либо после седьмой попытки вы потеряете все деньги и поймете, что азартные игры не для вас.

Но какова вероятность того, что вы везунчик? Предположим, нулевой сектор отсутствует, тогда вероятность поражения в каждой игре равна 1/2. Следовательно, вероятность проиграть всегда равняется 1/128. Ноль (зеро) благоприятствует казино; таким образом, вероятность все проиграть повышается, но тем не менее остается ниже 1 %. Это означает, что более чем в 99 % случаев вы можете сказать друзьям: «Съел? Я был в казино и выиграл!» Увы, существует вероятность все потерять.

В общем, те, кто заявляет: «Я знаю беспроигрышный метод почти всегда выигрывать в казино», определенно могут быть правы. Но, как я говорил, не всегда правильный ответ мы получаем на правильный вопрос.

Двойная ставка, двойная взятка

В нашу эпоху интернета и увлечения здоровым образом жизни люди не идут в киоск, чтобы купить почтовую марку. Если же все-таки найдется человек, который предпримет такой шаг, то он, вероятнее всего, услышит от продавца: «Извините, но марки нам давно не поставляют». Изменила свой статус и пачка сигарет: она уже не возглавляет список самых частых покупок. Вместо нее люди скорее приобретут лотерейный билет, чтобы испытать острые ощущения (которые обычно дарят нам азартные игры) в надежде изменить свою жизнь. Увы, жизнь вовсе не обязательно меняется к лучшему, даже если у вас билет с выигрышной комбинацией. В мгновенной лотерее, где для того, чтобы узнать результат, достаточно стереть соответствующее поле, вы играете против государства. В этом случае победа является возможным, хотя и редким исходом. Совсем другая ситуация с лотерейной игрой Win for Life, которая предлагает обладателям джекпота фиксированные ежемесячные платежи на значительную сумму в течение 20 лет.

В игре Win for Life ведущий выбирает 10 чисел от 1 до 20, каждый участник тоже выбирает свои 10 чисел. Соревнование идет между всеми игроками.

Например, в розыгрыше, который состоялся 13 марта 2010 года в 19:00, когда все 10 чисел попали в промежуток от 1 до 10, выигрышную комбинацию угадали более 59 участников, хотя каждый из них, естественно, надеялся быть единственным. В итоге их ежемесячный доход на ближайшие 20 лет составляет 67,80 евро. Мораль истории: во время игры важны не только математика и удача, но и психология. Мы живем в сложном мире.

Продолжим разговор о лотерее Win for Life. Необходимо отметить ее забавную особенность, которая на первый взгляд может показаться странной. В дополнение к описанному выше основному джекпоту и без учета дополнительной премии «нумерон» часть денежного фонда уходит на другие призы. Можно стать обладателем некоторой суммы, даже угадав 9, 8 и 7 чисел, хотя в последнем случае выигрыш составляет всего несколько евро. Небольшое отступление: это совсем не значит, что вы зарабатываете мало при отгадывании всего нескольких цифр. Дело в том, что существует масса способов ошибиться в игре и получить небольшой выигрыш — и всего один, чтобы стать обладателем джекпота. Лотерея Win for Life, как я уже говорил, стремится быть доброй даже с хроническими неудачниками. Она дает возможность поставить 2 евро вместо одного и выиграть, даже если все числа окажутся ошибочными или если не более трех из них будут правильными. Можно сделать ставку, при которой, если выпадает число 0, одновременно играет и число 10. Аналогично ставка на единицу связана с девяткой, на двойку — с восьмеркой, на тройку — с семеркой.

Стоит ли подстраховываться таким образом на случай неудачи? Но может быть, это просто способ вытянуть побольше денег из карманов наивного игрока? Ну а что касается выбора, удваивать ставку или нет (то есть поставить 2 евро или ограничиться одним), то над ним нет смысла задумываться! Доказать это очень легко, если знать, с чего начать. (Да, вы правы: всегда самое сложное — это узнать, что именно нужно сделать. Но подумайте, что интереснее: найти правильную идею или заняться вычислениями?) Как я объяснил, в лотерее Win for Life надо выбрать 10 чисел из набора, представленного 20 числами. После того как вы определились со своими 10 числами, вы думаете, что придет ваш клон и поставит на оставшиеся 10. Очевидно, что все выпавшие числа будут либо на одной карточке, либо на другой, но не на обеих: если вы ставите на n, то клон поставит на $10 - n$. Однако все возможные варианты одинаково вероятны, поэтому, если бы вы и ваш клон обменялись билетами до игры, не было бы никакой разницы (после того как числа озвучены, ситуация меняется!). Следовательно, вероятность иметь результат n или $10 - n$ одинакова и удвоение стоимости билета для победы на основании связанных ставок является справедливым предложением.

Не знаю, заметили ли вы, что я не рассчитал точную вероятность получения конкретного результата, но ограничился демонстрацией различных случаев. Математики по сути своей ленивы и никогда не производят вычисления, если их не принуждают... Что произошло бы, если бы, например, шар под номером 20 имел меньшую вероятность, чем другие, оказаться в розыгрыше? Столкнувшись с проблемой, не паникуйте — вместо этого начните искать путь к ее решению!

Пусть победит худший

Обычно в теннисном Уимблдонском турнире некоторые матчи переносятся из-за дождя. Однако в 2010 году матч между Джоном Изнером и Николя Маю длился целых три дня. Состязание, закончившееся со счетом 70:68 в пользу Изнера, побило все рекорды по продолжительности и набранным очкам, выведя из строя электронные табло, которые не были рассчитаны на такую ситуацию.

Существует множество видов спорта, в которых матч может длиться неопределенное время: подумайте о финальном футбольном матче, доведенном до серии пенальти, когда команды отбивают или забивают одинаковое количество мячей. Теннис, однако, имеет особенность: в наиболее тяжелых случаях победитель суммарно может выиграть меньше подач, чем его соперник. Действительно, в теннисе, как, например, в волейболе, играют не одну партию, а комплекс партий, то есть сеты. Сет, в свою очередь, представлен геймами, состоящими из нескольких розыгрышей. Это приводит к тому, что игрок может победить в матче, набрав гораздо меньше очков, чем противник. В частности, он может не иметь ни одного очка в геймах и сетах, которые выигрывает противник, и в то же время победить противника, позволяя ему при этом набрать максимальное количество очков в сетах.

Но каков минимальный процент для победы? Представьте, что вы ленивый победитель.

Несмотря на интересный способ обозначения баллов, на практике в теннисе победителем сета является игрок, который набрал по крайней мере 4 очка и у которого как минимум на 2 очка больше, чем у противника. Можете, таким образом, проиграть первые два сета со счетом 0:6 и 0:6, не заработав ни одного очка, и предоставить противнику преимущество в 48 очков. В этот момент наступает время играть по-настоящему! По-вашему, какой счет в гейме меньше всего способствует достижению победы: четыре к двум (40:30 и решающий мяч) или пять к трем (40:40 и две выигранные подачи)? Первая ситуация означает, что победитель выиграл 4 подачи, а его соперник — 2. Во второй ситуации гейм остается за тем, кто взял два последних розыгрыша подряд и, следовательно, набрал 5 очков против 3. Правильный ответ — первый. У вас есть три способа убедиться в этом: подсчитать самостоятельно, довериться мне или подметить, что чем больше розыгрышей в гейме, тем больше игра идет на равных. Что же касается итогов сета, то здесь надо рассуждать аналогично, чтобы проницательно выбрать лучший вариант между 6:4 и 7:6 (с тай-брейком на 7:5). В первом случае можно выиграть, взяв 24 очка против 28, во втором — 31 против 41, что, безусловно, лучше, если вы пытаетесь установить антирекорд.

Подведем итог: проиграв первый и второй сеты, вы должны закончить третий и четвертый со счетом 7:6 (7:5) и 7:6 (7:5). Если в розыгрышах вы проигрывали всухую, а соперник проигрывал, набирая 30 очков, то вы выиграли 62 подачи, а ваш оппонент — 130. Остается пятый сет. Поскольку, в отличие от других турниров, для победы в Уимблдоне необходимо взять в решающем сете по крайней мере на два гейма больше, чем противник, вам лучше остановиться на счете 6:4. Таким образом, в Уимблдонском турнире можно победить, выиграв 86 подач против 158, что составляет немногим более 35 % от их общего количества.

Не думаю, что подобная игра когда-либо имела место. Но если бы она все-таки состоялась, моей первой мыслью было бы, что это мошенничество. Впрочем, беззакония в теннисе немного.

Переложи карту

Возьмите 13 карт пиковой масти из колоды, смешайте их и выложите в ряд горизонтально. Крайняя левая карта будет иметь определенное числовое значение (обычно туз стоит 1, а валет, дама и король — 11, 12 и 13 соответственно). Отсчитайте количество карт, соответствующее этому значению, и переложите их в обратном порядке. Так, если исходная последовательность была **6**КХ594287ДТЗВ (где X соответствует 10), надо поменять порядок первых 6 карт — получится **4**95ХК6287ДТЗВ. Теперь повторите операцию, на этот раз перекладывая 4 карты. Продолжайте до тех пор, пока самая левая карта не станет тузом. На этом этапе игра, которая и раньше-то была не особо захватывающей, лишается всякого интереса: переложить одну карту — слишком незаметная операция. Но вопрос в другом: действительно ли мы уверены, что рано или поздно туз окажется на крайней левой позиции?

Ответ: да. Рано или поздно туз займет эту позицию, причем даже в том случае, если карт не 13. Важно, что все они пронумерованы в порядке возрастания. Понять, как все происходит, не слишком сложно, но от вас потребуется немного внимания (зато в итоге вы сможете удивить друзей, показав, насколько вы хороши в математических фокусах).

Первое соображение состоит в том, что существует определенное количество способов, которыми можно расположить *n* элементов: его можно узнать путем перемножения чисел от 1 до *n*, и обозначается оно символом *n*! (читается как «эн факториал»). Но нас интересует не этот результат. С другой стороны, если после определенного количества операций мы обнаруживаем комбинацию, которую уже видели, последующие шаги становятся очевидными в том смысле, что мы будем повторять тот же цикл комбинаций. Помните, что вы не тасуете карты, а просто перекладываете их в противоположном направлении, сохраняя порядок, и эта операция является предопределенной: если начальная точка одинакова, точка прибытия одинакова. Обратное неверно: последовательность 495XK6287QA3B также могла быть получена из 826KX5947QA3B. Это довольно обычно не только в математике, но и в повседневной жизни: нет уверенности в том, что мы можем вернуться пройденным путем в исходную точку. Однако, поскольку возможные комбинации не бесконечны, мы обязательно рано или поздно достигнем цикла, хотя никто не знает, как быстро. Мы хотим показать, что этот цикл длиной в единицу, то есть последовательность начинается с туза.

Я почти завершил объяснение. Последнее замечание, которое стоит сделать, заключается в том, что, если на последней позиции когда-нибудь окажется король, потому что на предыдущем шаге он был на первой позиции, мы можем быть уверены, что больше он оттуда никуда не сдвинется. Чтобы его переместить, нужно переложить 13 карт. Но для этого необходимо, чтобы король был экспертом по телекинезу, так как он должен находиться одновременно в двух местах: а именно на первой и на последней позиции. В более общем случае, если все m карт с наибольшим значением находятся на последних m позициях, они так и останутся в том порядке, в котором находятся.

Что ж, мы выяснили, что, меняя порядок карт на обратный, рано или поздно мы войдем в цикл. Если цикл имеет длину 1, то первая карта последовательности — туз: именно то, что и требовалось. А если это не так? Давайте построим противоречие и таким образом продемонстрируем, что невозможно достичь цикла длиной больше 1. Будем рассматривать все карты, которые появляются в первой позиции во время этого цикла, и выбирать ту, которая имеет наибольшее значение k. В тот момент, когда она фактически занимает первое место в ряду, мы знаем, что карты от $k + 1$ до короля находятся в позициях от $k + 1$ до 13. После перестановки первых k карт, как того требует алгоритм, k-карта окажется в положении k. Но, стало быть, с этого момента первые $k - 1$ позиции будут содержать карты от 1 до $k - 1$ и значение k больше не будет появляться в первой позиции, вопреки нашей исходной гипотезе. Следовательно, единственная возможность состоит в том, что $k = 1$. Именно это мы и хотели доказать.

Для любопытных: в ходе демонстрации я использовал концепцию моноварианта — название у нее заумное, но суть довольно простая. Возможно, вы слышали об инвариантах, то есть о выражениях, значение которых никогда не меняется. Например, в многограннике без отверстий значение $F + V - S$ (где F обозначает количество граней, V — вершин, а S — ребер) является инвариантом: оно всегда равно 2. Моновариант не такой жесткий, как инвариант, и значение может меняться, но только в одном направлении: или никогда не уменьшается, или никогда не возрастает. В нашем случае моновариант задается количеством карт в той части последовательности, которая остается фиксированной; постепенно он достигает 13.

Это тоже хитрость математиков: спрятать за кулисы легкость вычислений и показать только конечный результат!

Кубики честные и нечестные

Я никогда не понимал, почему в таких настольных играх, как «Монополия» или «Риск», право первого хода определяется броском кубика: кто выбросил большее число, тот и ходит первым. И проблема не в том, что приходится бросать кубик, который на самом деле дает случайный результат (кстати, вы знали, что в «правильных» игральных костях углубление, обозначающее единицу, по размеру намного больше и глубже остальных, чтобы обеспечить баланс кубика?). Проблема в том, что существует риск ничейного результата.

Начнем со случая, когда игроков всего двое. Вы не можете загадать орел или решку, потому что в кармане у вас только кредитная карта? Или вы не желаете этого делать, потому что, раз игра с кубиком, необходимо использовать именно кубик? Тогда, например, можно бросить один кубик на двоих: если выйдет четный номер, то начинает один из участников, если нечетный — то другой. Вас не очень устраивает этот вариант, потому что один из игроков не может реализовать свое конституционное право бросать кубик? Значит, самое время подключить математику и выяснить, как изготовить кубики, у которых нет обычных значений от 1 до 6, но для которых действовали бы следующие условия.

1. Ничья невозможна.
2. Каждый кубик обеспечивает равную возможность выиграть.

Играть одному не очень интересно: всегда выигрываешь. Если играют двое, то первое условие требует наличия разных чисел на кубиках: я подскажу, что это числа от 1 до 12. Несложно найти две совокупности значений, которые отвечают второму условию: например {1, 3, 5, 8, 10, 12} и {2, 4, 6, 7, 9, 11}. Трюк состоит вот в чем: шесть пар — (1, 2), (3, 4), (5, 6), (7, 8), (9, 10), (11, 12) — надо разделить между двумя кубиками таким образом, чтобы каждый имел три больших и три меньших значения. Не обязательно каждому кубику иметь все цифры от 1 до 12; можно использовать наборы значений {1, 1, 1, 4, 6, 6} и {2, 2, 2, 3, 5, 5}. Конечно, не получится использовать два стандартных кубика, но в крайнем случае возьмите один обычный, а второй — со значениями {0,5; 1,5; 2,5; 4,5; 5,5; 6,5}, если вы не боитесь десятых долей. Либо можно использовать два стандартных кубика, а на случай ничьей ввести правило: если выпадает 4, 5 или 6, выигрывает первый игрок, если же 1, 2 или 3 — второй. Ну что, осталось лишь объяснить правило обоим игрокам? Поверьте, лучше вырезать два кубика: будет меньше дискуссий.

Теперь предположим, что играют трое. Если мы хотим, чтобы все значения на кубиках были разными и последовательными, то они должны быть представлены числами от 1 до 18. Как их распределить? Можно разделить их, как и пары в предыдущем случае, одним из шести способов (по количеству сторон одного кубика), чтобы составить тройки (1, 2, 3). Необходимо, однако, аккуратно упорядочить триады, чтобы ни один из кубиков не обеспечил большего количества выигрышных комбинаций. Внизу слева вы видите подходящий порядок для перестановок в триадах, а справа — соответствующие значения для трех кубиков.

2 2 3 1 1 3	{2, 5, 9, 10, 13, 18}
1 3 2 2 3 1	{1, 6, 8, 11, 15, 16}
3 1 1 3 2 2	{3, 4, 7, 12, 14, 17}

Доказать, что эти кубики отвечают второму условию, достаточно просто. Как и в предыдущем случае, представьте, что на кубиках есть и второстепенные значения от 1 до 6 рядом с этими указанными. Как теперь определить приоритет хода? Если вспомогательные значения различаются — все хорошо; если они одинаковые, смотрим на схему перестановок в триадах: они абсолютно справедливые.

Беда, если игроков четверо: 6 сторон не делятся на 4. Лучше перейти к додекаэдрам, 12-сторонним кубикам, которые уже используют любители ролевых игр. Число сторон делится на 2, на 3 и на 4. Роберт Форд и Эрик Харшбаргер таким образом создали (и продают) кубики со следующими значениями:

{1, 8, 11, 14, 19, 22, 27, 30, 35, 38, 41, 48}
{2, 7, 10, 15, 18, 23, 26, 31, 34, 39, 42, 47}
{3, 6, 12, 13, 17, 24, 25, 32, 36, 37, 43, 46}
{4, 5, 9, 16, 20, 21, 28, 29, 33, 40, 44, 45}

Эти кубики отличаются важным свойством: при игре вдвоем, втроем и вчетвером каждый имеет одинаковую вероятность оказаться на каком-то месте относительно других игроков. Поэтому их можно использовать не только для того, чтобы решить, кто победит, но и для того, чтобы установить очередность игроков. Теперь (если оставить в стороне приобретение кубиков) остается лишь решить, нужна ли вам вся эта возня или нет, но это уже не моя проблема. А еще не хотел бы я начинать поиск значений для пяти икосаэдров — 20-гранных кубиков для игры впятером: подвох в том, что для троих они не будут работать. И еще одна проблема состоит в том, что никто мне их не изготовит.

Посмотрим правде в глаза: гораздо интереснее делать противоположное. Представьте, что у вас три кубика: зеленый, на гранях которого проставлены цифры {2, 2, 4, 4, 9, 9}, белый с числами {1, 1, 6, 6, 8, 8} и красный с числами {3, 3, 5, 5, 7, 7}. Если подсчитать, вы узнаете, что в среднем зеленый кубик побеждает белый, белый побеждает красный, а красный побеждает зеленый. Это немного напоминает игру «Камень, ножницы, бумага», где бумага сильнее камня, камень сильнее ножниц, а ножницы сильнее бумаги. Однако у кубиков есть преимущество: вы можете дать своему оппоненту возможность первым выбрать кубик, чтобы потом — уже с учетом его выбора — взять правильный кубик, то есть такой, который принесет вам выигрыш с большей вероятностью. Не кажется ли вам, что «несправедливость» этих кубиков куда привлекательнее «честности» обычных игральных костей?

Ищем секретаря

Мы живем в сексистском мире: существительное «секретарь» кажется нам гораздо более внушительным, чем «секретарша». Когда говорят «секретарь», обычно имеют в виду главу партии. В то же время при слове «секретарша» в памяти мгновенно всплывают забавные истории и анекдоты, которые рассказываются на протяжении многих десятилетий. Преимущество математических историй в том, что сексизм в них отсутствует и их персонажем может быть человек любого пола. А недостаток таких историй состоит в том, что на самом деле они являются зада́чами. Вот две задачи — совершенно разные, но между ними есть кое-что общее.

В первой секретарь подготовил 16 писем с конвертами. Однако он был так занят, комментируя последнее вирусное видео на «Фейсбуке», что вложил письма в конверты случайным образом. Какова вероятность того, что никто из адресатов не получит предназначенное ему письмо?

События второй истории разворачиваются в прошлом — в то счастливое время, когда найти работу было намного проще, чем сейчас. На должность секретаря претендуют 16 кандидатов. Отдел по работе с персоналом вызывает их по очереди на собеседование. Проблема в том, что невозможно сначала выслушать всех, а затем принять лучшего соискателя: необходимо сразу же по окончании собеседования или одобрить кандидатуру, или отказать, чтобы человек не откладывая отправился на поиски другой работы. (Я же говорил, что история не очень правдоподобная!) Какая стратегия позволит выбрать лучшего претендента на должность и не вынудит брать плохого только потому, что всем остальным мы уже отказали?

В первом случае, как вы могли понять, проблема принципиально комбинаторная: нужно найти формулу, позволившую бы определить количество перестановок в наборе из *n* элементов, которые не оставляют ни одного элемента в исходном месте. Однако, даже если вы изучали комбинаторику, очень маловероятно, что вам сказали, как называются перестановки этого типа. А все потому, что это один из самых тщательно хранимых секретов математики. Хорошо, знайте же: по крайней мере в итальянском языке мы говорим о дисмутациях[1], в то время как англичане используют термин «расстройство» (*derangement*), что, безусловно, дает очень яркое представление о происходящем.

Существует также символ, обозначающий количество беспорядков для *n* элементов. Так же как и для факториала, используется восклицательный знак, но он ставится слева от числа, а не справа. Итак, беспорядок для *n* элементов обозначается !*n*. Я не буду приводить

[1] По-русски такие перестановки называются беспорядками.

здесь строгую формулу для вычисления количества беспорядков — просто укажу, что для 10 элементов оно составляет около 37 % от общего числа перестановок. Иными словами, если мы смешаем колоду карт, а затем примемся переворачивать их по одной, произнося наугад: «Червовый туз, червовая двойка, червовый король, бубновый туз...» — то угадаем хотя бы одну карту в 63 %. Кто бы мог подумать?

В ситуации с собеседованием оптимальная стратегия состоит в том, чтобы оценить определенное число кандидатов, выявить, кто из них обладает самыми высокими показателями, а затем выбрать первого же кандидата, получившего более высокую оценку (если таких нет, очевидно, последний кандидат обязательно должен быть принят на работу). Остается определить, сколько, так сказать, заранее отвергнутых несчастных кандидатов нужно оценивать. В случае 16 претендентов ответ будет 6. Как я рассчитал это магическое число?

Я выбрал значение, близкое к 37 % от общего числа кандидатов. Опять это число! Совпадение? Я не смог найти никаких доказательств существования явной связи между этими двумя случаями, но 37 % — значение неслучайное. На самом деле это отношение $1/e$, где e (чуть более 2,718) является константой, которая встречается в математике едва ли не везде. В некотором смысле это даже более фундаментальное значение, чем греческое π. Кстати, есть люди, убежденные, что вместо π (3,14...) было бы намного лучше использовать в качестве фундаментальной константы 2π (6,28...), но идея с $1/e$ пока никому не пришла в голову.

Теория собеседований, подготовленных таким образом, «изящна», как любят говорить математики, когда в результате появляются числа, которые им нравятся. Но на практике все по-другому. Фактически процедура выбора, которую я описал, работает отлично, если вам приходится выбирать между определенным количеством абсолютно случайных значений без какого-либо верхнего ограничения. Но если вы поставили какому-либо соискателю оценку 9 из 10 возможных, было бы безрассудно и дальше рабски следовать алгоритму в надежде найти кого-нибудь с результатом 9,5.

Математика — полезная штука, но реальный мир не всегда такой математический, что бы там ни говорил Галилей.

Раздел IV

На прогулке

Остерегайтесь кольцевых дорог

Как вы, возможно, знаете, из сказочного Даксбурга в Мыштаун ведут два пути: первый проходит через город Мяувиль, а второй — через Гавбург. Каждое утро 4000 жителей рассаживаются по машинам, чтобы переместиться из одного города в другой. Участки дорог Мяувиль — Мыштаун и Гавбург — Даксбург достаточно широки, что позволяет водителям преодолевать расстояние, как правило, за 50 минут. Два других участка гористые и часто перегружены транспортом. На то, чтобы проехать каждый такой участок, уходит $N / 100$ минут, где N — количество автомобилей, которые находятся в пути; нижний предел составляет 15 минут, когда на маршруте менее 1500 автомобилей.

После череды невообразимых пробок ситуация стабилизировалась, потоки разделились между двумя направлениями поровну, и на каждом маршруте оказалось по 2000 машин. Теперь, когда дорожное движение нормализовалось, общее время в пути для каждого водителя составляет 2000 / 100 + 50 = 70 минут.

Мяувиль и Гавбург расположены очень близко друг к другу, и это позволило утке Фило Сганга убедить мэров построить кольцевую дорогу, которая соединит два пункта: всего за 5 минут можно будет переехать из одного города в другой. Скрудж Макдак хотел подать заявку на строительство кольцевой дороги, но, узнав, что комиссию возглавляет Бригитта Макбридж, сбежал, позволив Флинтхарту Гломгольду выиграть тендер.

Как вы думаете, что произошло после того, как дорогу построили?

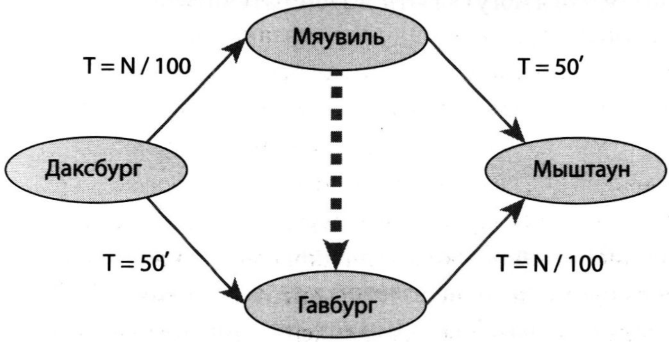
Ответ очень простой: все автомобилисты поняли, что ехать по новой дороге выгоднее.

Для прохождения участка Даксбург — Мяувиль — Гавбург требуется не более 4000 / 100 + 5 минут, то есть 45 минут, тогда как при выборе старой дороги Гавбург — Даксбург время в пути составляет 50 минут. Однако если все водители отдают предпочтение более быстрой трассе, то время, затрачиваемое на дорогу, увеличивается: 4000 / 100 + 5 + 4000 / 100 = 85 минут, несмотря на открытие новой дороги (более того, именно из-за этого!). В конце концов народ восстает, и Гломгольд вынужден разрушить дорогу и съесть свою шляпу, как обычно.

Ладно, я знаю, что автор комиксов Джорджо Каваццано никогда не нарисует историю, основанную на этом сюжете. Но приведенные выше математические вычисления абсолютно верны, и у них даже есть название — **парадокс Браеса**, по имени немецкого математика Дитриха Браеса, который впервые описал это явление. Чтобы понять суть парадокса, нужно взглянуть на него через призму теории игр: это направление, которое сочетает в себе элементы математики и экономики. Математики, работающие в этой области, могут претендовать на Нобелевскую премию по экономике, полное название которой «Премия Шведского национального банка по экономическим наукам памяти Альфреда Нобеля».

«Игры», охватываемые данной теорией, вовсе не шахматы, бридж или покер, а взаимодействие между двумя или более участниками, пытающимися получить максимальную выгоду от этих взаимодействий. Как правило, мы изучаем очень упрощенные модели того, что происходит в реальном мире. И это позволяет понять, как работает экономика и почему существует пропасть между теорией и практикой. Нередко, как и в нашем случае, речь идет о некооперативных играх: в таких ситуациях каждый частник пытается максимизировать свою выгоду, независимо от того, что происходит с другими. Грубо говоря, это все равно что убить бабушку, чтобы раздобыть немного деньжат... В некооперативных играх есть набор конкретных стратегий, называемых **равновесием Нэша** (да-да, того самого Джона Нэша из «Прекрасного разума»): это ситуация, при которой ни один из игроков не может увеличить свой выигрыш, меняя стратегию в одностороннем порядке.

Разделение водителей на две равные группы в начале нашей истории — это и есть равновесие Нэша. Теперь, если хотя бы один из них изменит маршрут, на дороге окажется больше машин и, следовательно, его скорость снизится. Проблема в том, что после открытия маршрута Мяувиль — Гавсбург равновесие Нэша стало неоптимальным. Самым разумным решением для водителей было бы игнорировать новую дорогу, но это должен быть совместный выбор: всем необходимо договориться, а это невозможно и потому неприменимо. Если только один человек решает не пользоваться кольцевой дорогой, это ничего не меняет: все обязаны сделать такой выбор, чтобы изменить ситуацию к лучшему.



Не стройте эту дорогу!

Есть и другие примеры того, как эгоистичное поведение в итоге приводит к поражению. Возьмем хотя бы прививки. При вакцинации всегда имеется небольшая вероятность осложнений, и поэтому, если один человек решает не прививать своего ребенка, у него есть несомненное преимущество, ведь, если все остальные вакцинированы, болезнь не может перерасти в эпидемию. Однако, если многие люди начнут рассуждать так же «умно», болезнь может распространиться и мы все в этой ситуации проиграем.

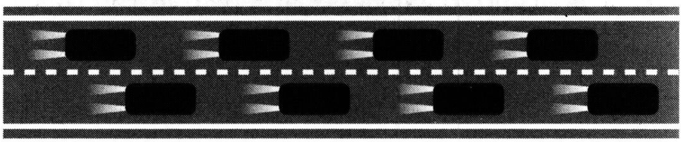
Но еще сильнее в ситуации с парадоксом Браеса впечатляет то, что создается иллюзия, будто маршрутов для выбора стало больше. Известны примеры этого парадокса в реальной жизни, причем не только в сфере дорожного движения, но и в других областях — скажем, при распределении электрического тока. Марксист мог бы указать, что свободный рынок, в конце концов, не всегда является панацеей, а у эколога наверняка имеются свои аргументы против строительства новой магистрали. Но все, что я могу сказать: в этом математика не виновата.

Соседняя очередь всегда движется быстрее

Один из известнейших примеров действия закона Мёрфи: соседняя очередь всегда движется быстрее. Допустим, вы собираетесь оплатить товары в супермаркете и раздумываете, какую кассу выбрать. По вашему мнению, ближайший к вам кассир обслуживает покупателей медленнее по сравнению с соседним. Но если вы предпочтете быструю, на ваш взгляд, кассу, то перед вами вполне может оказаться человек, который начнет скандалить по поводу цены и задержит всю очередь. Выбирая кассу, можно надеяться только на удачу: подходя к кассам, надо посмотреть, в какой из очередей покупатель, стоящий последним, первым оплатит товар.

Однако похожая задача — выбор полосы в автомобильной пробке — имеет совершенно другое решение, о чем Пол Кругман и Стивен Строгац несколько лет назад писали на страницах «Нью-Йорк таймс». Давайте посмотрим, почему в этом случае медлительность полосы имеет математическое объяснение, причем у тех, кто находится на соседней полосе, все происходит точно так же. Разве такое возможно? Математики утверждают, что да.

Представьте, что вы стоите на двухполосном шоссе в пробке, которая растянулась на 4 километра. На самом деле вы не стоите, а очень медленно движетесь: одну половину пути преодолеваете со скоростью 10 километров в час, а другую — со скоростью 30 километров в час.



Пробки на трассе

Для простоты предположим, что водители не перестраиваются и не меняют полосы движения. Знаю, что это совершенно абсурдная гипотеза для Италии, где автомобильный слалом — второй национальный вид спорта после футбола; но давайте представим, что водители продолжают движение по изначально выбранной полосе. Как будут разворачиваться события? Очевидно, что машины в обеих полосах будут преодолевать 4 километра за одинаковое время. Вы наверняка подумали, что это время составляет 12 минут, исходя из скорости 20 километров в час, однако это ошибочный ответ.

Машина преодолевает километр за 6 минут при скорости 10 километров в час и за 2 минуты при скорости 30 километров в час. Таким образом, общее время в пути составляет 16 минут. Но представьте, как все было! Из 16 минут вы потратили 4 минуты на движение по быстрой полосе и еще 12 минут (в три раза больше) — на ворчание, потому что соседние машины ехали быстрее вас. Напомню, что условие задачи составлено так, чтобы то же самое случилось и с автомобилистами в другом ряду.

Впрочем, результат может быть еще удивительнее. Представьте себе, что в другом ряду машины проходят занятые отрезки со скоростью 5 километров в час, а свободные — со скоростью 20 километров в час. Очевидно, что бедолагам потребуется больше времени, чем вам, чтобы преодолеть эти ужасные 4 километра. Однако анализ, проведенный выше, остается в силе, и поэтому вам кажется, что они едут быстрее вас!

В математике этот феномен известен как **парадокс Редельмейера**. Почему же он не работает в случае скандала у кассы супермаркета? Может, потому, что Мёрфи предпочитал тележки машинам?

Конечно, нет. В упрощенной модели с двумя дорожными полосами мы предположили, что расстояние, преодолеваемое быстрым и медленным ходом, одинаково. Таким образом мы поместили себя в размерность пространства, которая к очереди супермаркета неприменима, потому что нас не интересует пара метров до конвейерной ленты. Если же вместо этого мы поместим себя в размерность времени, воображая, будто каждая полоса движется быстрее в течение 20 % от общего времени и замедляется в оставшиеся 80 %, возникает парадокс… И закон Мёрфи может вступить в силу. А представьте, что может случиться при внезапных изменениях полосы движения и замедлениях волн трафика, но об этом я расскажу попозже.

У моих друзей больше друзей, чем у меня

В одном из выпусков комикса «Мелочь пузатая»[1] рассказывается о Дне святого Валентина. После вечеринки герои возвращаются домой с огромным ворохом полученных валентинок (в США их дарят не только возлюбленным, но зачастую также друзьям и родным). Все, кроме расстроенного Чарли Брауна, которому не досталось ни одной. Вы наверняка скажете: само собой, ведь Чарли Брауна никто не замечает и никто не удивляется, увидев, что у всех остальных больше друзей, чем у него.

В жизни, однако, большинство из нас и есть Чарли Брауны. Да, у нас есть друзья, но у наших друзей в среднем больше друзей, чем у нас. Это утверждение было доказано опытным путем: в мае 2011 года два аспиранта, Йохан Угандер и Брайан Каррер, проанализировали учетные

---

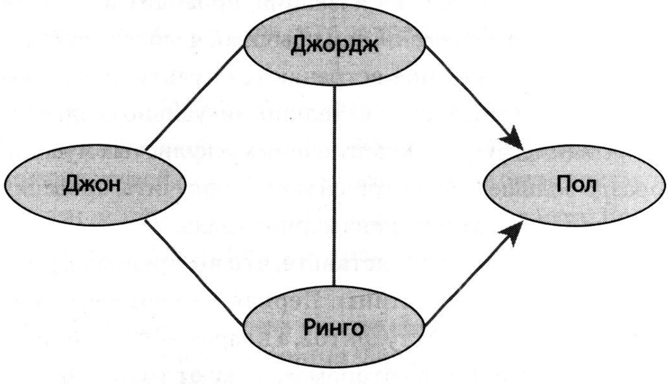
[1] Peanuts (англ. «Мелочь пузатая») — ежедневный американский комикс, созданный Чарльзом М. Шульцем и выходивший с 1950 по 2000 год; всего было опубликовано 17 897 выпусков.

записи в социальной сети «Фейсбук», в которой тогда был зарегистрирован всего 721 миллион пользователей. Они обнаружили, что в 93 % случаев друзья пользователей обладали большим количеством знакомых со статусом «друг», чем сами пользователи. В среднем у пользователя было 190 друзей, в то время как у его друзей — в три раза больше, а точнее 635. Каков сюжет! Поддельные данные? Как всегда, ответ намного проще; но лучше приступить к объяснению издалека.

Давайте сначала рассмотрим пример совершенно другого рода. Я добросовестно занимаюсь тяжелой атлетикой, просто чтобы поддерживать минимальную физическую форму. Я, конечно, не Мистер Мускул, но мне кажется, что у меня среднестатистическая физическая подготовка. Тем не менее, когда бы я ни пришел в спортзал (в любое время дня!), меня неизменно окружают одни культуристы. Заметьте, не то чтобы я слишком волновался по этому поводу, но на первый взгляд это сбивает с толку. Что может быть причиной такой аномалии? Что ж, качок, как следует из названия, проводит в качалке гораздо больше времени, чем обычный человек, поэтому в спортзале куда чаще встречаются культуристы, чем кто-либо другой. А следовательно, визуально создается впечатление, будто процент очень мускулистых мужчин гораздо больше того, который можно получить, проверив данные завсегдатаев тренажерного зала.

Второй пример: представьте, что вы преподаватель и ведете две дисциплины. Первая — вводный курс, рассчитанный на 90 студентов, а вторая — специализированный семинар, в котором участвуют только 10 студентов. Очевидно, что в среднем у вас по 50 студентов в группе (нас сейчас не интересует эффективность обучения). Но давайте посмотрим на это с точки зрения студентов: для 90 из них группа насчитывает 90 человек, а для оставшихся десятерых — 10 человек. В среднем мы имеем (90 × 90 + 10 × 10) / (90 + 10), то есть 82 студента в группе. Что-то здесь не так. Все же модель должна быть более точной: слишком уж велика разница между объективным взглядом (подсчет студенческих билетов и самих студентов) и субъективным (подсчет одногруппников). Когда существует неравенство между референтными группами: случайными посетителями тренажерного зала, которые редко тренируются, и завсегдатаями, которые всегда там; студентами из большой учебной группы и членами семинара для избранных — и среднее значение рассчитывается по всем элементам целого, у которого много связей, то мы получим искаженное среднее значение.

А сейчас попробуем разобраться с крошечной группой, представленной на рисунке ниже.



Четверо друзей из Ливерпуля

Эллипсы обозначают людей, а отрезки между ними — дружеские отношения. Суммируя число друзей каждого, мы получаем: 2 + 3 + 3 + 2 = 10 дружеских отношений (каждые отношения учитываются дважды) для 4 человек, то есть в среднем по 2,5 друга.

Теперь давайте узнаем среднее число друзей друзей следующим образом: сначала определим число друзей у каждого человека, затем — число друзей у каждого из его друзей и наконец выведем среднее значение. У Джона есть 2 друга: Джордж и Ринго; у каждого из них — по 3 друга. То же касается Пола. Соответственно, у их друзей в среднем по 3 друга. У Джорджа и Ринго есть 3 друга: у двух из них (Джон и Пол) по 2 друга, а еще у одного 3 друга. Среднее число друзей на одного друга составляет 7/3. Таким образом, общее среднее значение равняется 8/3, что немного больше, чем рассчитанное ранее.

Чем разветвленнее структура дружбы, тем больше разница между средним числом друзей. Если мы добавим в схему Элеонору, чей единственный друг — Ринго (а кому он не друг?!), то общее число дружеских отношений увеличится до 12, а среднее число друзей на всех пятерых уменьшится до 2,4. Однако у каждого из их друзей в среднем будет уже около 3,13 друга: больше, чем раньше! Нечто подобное происходит в «Фейсбуке».

Будьте осторожны! Этот очевидный парадокс работает с «Фейсбуком», но если вы проведете подобный анализ на основе подписчиков «Твиттера», то получите совершенно другие результаты. На самом деле подписка на кого-то в соцсетях не является симметричным действием: А может подписаться на В, но тот, возможно, даже не имеет представления о существовании А. Таким образом, мультипликативный эффект, который мы видели на примере двух студенческих групп (все учащиеся в одной группе считают всех остальных своими одногруппниками), больше не работает. Короче говоря, «Твиттер» — это не «Фейсбук», только с короткими сообщениями; между ними есть фундаментальная разница.

Исчезающие лифты

В прологе к «Занимательной математике» ее авторы, Георгий Гамов и Марвин Стерн, рассказывают, что летом 1956 года оба работали в здании авиационной компании «Конвэр». Кабинет Стерна, постоянного сотрудника, находился на шестом этаже, а Гамова, консультанта, — на втором. Гамов, часто поднимавшийся к Стерну, заметил, что в среднем пять раз из шести лифт, первым приезжавший на его этаж, двигался не вверх, а вниз. Он спросил у Стерна: может, «Конвэр» изготавливает лифты на крыше, а затем спускает их на склад в подвале? Стерн ответил: «Да нет же! Проверьте, что происходит, когда вы заходите в лифт, чтобы спуститься». Через некоторое время Гамов сказал: «Вы правы: когда я спускаюсь, только каждый шестой лифт, который останавливается на этаже, идет вниз. Значит ли это, что компания "Конвэр" строит лифты в подвале и посылает готовые лифты на крышу, откуда производимые компанией самолеты доставляют их к месту назначения?» На что Стерн ответил: «Конечно, нет! Но позвольте мне прежде заметить, что если бы я и не знал, сколько этажей в этом здании, то теперь, располагая той информацией, которую вы мне сообщили, смог бы сказать, что в здании семь этажей». (Я помню, что в Соединенных Штатах нет нулевого этажа, как в Италии, и что первый этаж обозначается номером 1.)

Так в чем же дело?

Объяснение несложное, если упростить вычисления. Представьте, что в здании есть только один очень медленный лифт, который останавливается на каждом этаже, причем перемещение с одного этажа на другой занимает 1 минуту. Если в 10:00 лифт находится на первом этаже, то в 10:01 он будет на втором этаже, в 10:02 — на третьем и так далее, пока не достигнет седьмого в 10:06. В этот момент он начинает снижаться: в 10:07 будет на шестом, в 10:08 — на пятом, а в 10:12 цикл закончится и лифт снова окажется на первом этаже. Если Гамов, покинув свой кабинет на втором этаже, подойдет к лифту между 10:00 и 10:01 или между 10:11 и 10:12, лифт прибудет на второй этаж, двигаясь вверх. Если же подойти к лифту в остальные 10 минут, он прибудет, направляясь вниз. Результат будет аналогичный, когда Гамов спускается с шестого этажа. Пресловутое выражение «каждый шестой» на практике представляет собой математическое объяснение того факта, что если верхних этажей больше, чем нижних, то, когда мы вызываем лифт, он с большей вероятностью находится выше нас, а не ниже. И наоборот. Любопытно: в своей книге Гамов и Стерн утверждают, что это рассуждение справедливо и при наличии нескольких лифтов.

Лет десять спустя Дональд Кнут написал статью, в которой доказал, что единственный лифт действительно чаще приходит с той стороны, где этажей больше, но при увеличении количества лифтов вероятность меняется и доходит до 50 %. Никогда не доверяйте интуиции!

Перейдем от математических лифтов к реальным. Вам может быть интересно узнать, что, когда в здании имеется более одного лифта, умное программное обеспечение нередко вынуждает нас тратить больше времени, когда мы нажимаем кнопку вызова. Представьте, например, что вы находитесь на подземной автостоянке, на этаже «–3», и видите, что два лифта стоят на нулевом и третьем этажах. Вы вызываете лифт, и спускается тот, что был на третьем этаже, хотя ему и пришлось проделать путь вдвое длиннее. Что происходит? Второй лифт не спускается так низко в целях безопасности? Причина гораздо прозаичнее: большинство людей входят в здание через главный вход.

Лифты запрограммированы так, чтобы как можно чаще останавливаться на первом этаже. Неизвестно ведь, когда подойдет группа людей, а их лучше не заставлять ждать слишком долго, тогда как один человек может позволить себе подождать еще несколько секунд.

Или представьте, что вы на шестом этаже вызываете лифт, чтобы спуститься. Лифтов два: один стоит на пятом этаже, а другой — на третьем. И вот первый начинает спускаться, а второй — подниматься: он без остановки проходит мимо вашего этажа и устремляется дальше вверх. А всё по той простой причине, что лифты также вызвали на втором этаже, чтобы подняться, и на восьмом, чтобы спуститься. Средняя продолжительность ожидания оптимизируется, даже если кому-то приходится ждать чуть дольше. Между нами говоря, я думаю, что это одна из причин, по которым табло с номером того этажа, где находится лифт, устанавливают только на первом этаже — чтобы у ожидающих не было информации, которая могла бы их разозлить.

Наконец, следует помнить, что с помощью математики можно сократить время ожидания, но психология нередко справляется с этой задачей гораздо эффективнее. Городская легенда гласит, что в здании одной компании были установлены очень медленные лифты и рядом с их дверями кто-то додумался повесить зеркала. В ожидании кабины люди начали прихорашиваться и перестали жаловаться на медлительность лифтов. Гениально, не правда ли?

Автобусные тройки

Лет двадцать назад я ездил на работу на автобусе номер 2, который теоретически должен был быть одной из сильных сторон туринской транспортной сети (а вот трамвайные пути так никогда и не открыли из-за нехватки денег). На линии было довольно много автобусов: жаль только, что они завели плохую традицию приезжать парами, а иногда даже колоннами по три — прямо друг за другом. Причина, конечно, же, не в том, что водители хотели избавиться от чувства одиночества, и не в том, что транспортная компания плохо составила график. Но эта ситуация наглядно демонстрирует банальное утверждение, что статистика работает только в среднем.

Давайте возьмем в качестве примера линию, от конечной станции которой автобус отправляется каждые 5 минут. В идеальном мире люди приходят на остановки через равные промежутки времени, движение не создает проблем и светофоры всегда горят зеленым. Чтобы попасть в зеленую волну в обоих направлениях, надо построить город подходящим образом, точно рассчитав расстояние между основными перекрестками: допустим, что все автобусы проходят одну и ту же последовательность светофоров. Но мы живем в неидеальном мире! Скажем, внедорожник может внезапно подрезать автобус, заставив водителя резко затормозить и задержаться перед светофором. Или на остановке собралась группа молодых людей, которые потратят больше времени на посадку, чем запланировано. В результате автобус начинает выбиваться из графика. Но чем сильнее он опаздывает, тем больше людей дожидается его прибытия на остановках. Таким образом, отставание от расписания продолжает расти.

Это, в свою очередь, подразумевает, что следующего автобуса будет ждать меньше людей. Возможно, он начнет прибывать на остановки раньше времени и даже догонит предыдущий автобус с учетом его задержки. С этого момента оба автобуса ведут себя как один — вы заметили, что люди группами перебегают из автобуса в автобус? — и задержка двух автобусов увеличится еще больше, иногда позволяя третьему автобусу присоединиться к колонне.

В теории все это может продолжаться вечно, но путь от одной конечной автостанции до другой редко занимает более часа, и поэтому на практике существует ограничение длины таких колонн. Короче говоря, сдваивание автобусов не связано ни с каким парадоксом — это простое следствие статистики. Тем не менее на практике вероятность удлинения автоколонны значительно возрастает с увеличением частоты автобусов. А вот если длительность маршрута — всего полчаса, приезд второго транспортного средства сразу после первого маловероятен. К счастью для бедных пассажиров!

Однако есть и реальный парадокс, связанный с временем ожидания: он проявляется именно тогда, когда собралась колонна из трех автобусов (если их два, ничего странного не происходит). Представьте, что ваш автобус теоретически должен ходить каждые 15 минут. Но в ситуации, когда образуются тройные автоколонны, второй и третий автобусы подъезжают к остановке с минутными интервалами. А затем вам нужно подождать 43 минуты, прежде чем появится четвертый автобус — он же первый из следующего триплета. Представьте также, что остановка находится недалеко от поворота, а значит, невозможно увидеть приближающийся транспорт на расстоянии. Как по-вашему, когда ожидание следующего автобуса будет короче: в случае если автобус уехал прямо перед вашим носом или если, подходя к остановке, вы не видите отъезжающего автобуса?

Для первого случая вычисления несложные. Вы могли увидеть любой из трех автобусов; если это был первый или второй, придется подождать всего 1 минуту, если же третий — 43 минуты; в среднем — четверть часа.

Во втором случае, однако, ситуация совсем иная. Существует 43 из 45 возможностей прийти в неудачный интервал, и тогда среднее время ожидания составит половину интервала, то есть 21,5 минуты. Конечно, остается еще две возможности, когда среднее время ожидания составляет полминуты. Однако они настолько маловероятны, что среднее время ожидания равняется почти 21 минуте.

Так что, получается, лучше опоздать на автобус, чем не опоздать? Ну не то чтобы… Но вы всегда можете убедить в этом и себя, и тех, кто вместе с вами дожидается автобуса на остановке!

Стоп — поехали

На мой взгляд, прослушивание новостей в эфире радиостанции Isoradio в перерыве между треками — чистой воды мазохизм даже для тех, кто, как и я, ценит старомодную музыку. Например, анонсированный часовой затор на дороге совсем не обязательно вас дождется, не говоря уже о том, что радиосообщения всегда отстают от реальных событий. Объявляют двухкилометровую пробку, а она, оказывается, растянулась на 10 километров. А может случиться, что вы и вовсе ее не увидите, как будто все машины испарились. Вот оно: иногда в машине — как слушая приемник, так и не слушая его — вы то ползете со скоростью пешехода, то внезапно набираете скорость, при этом по пути вам не встречаются ни аварии, ни ремонтные работы, из-за которых сужаются участки дороги. Естественно, что после сужения дороги или перед ним движение становится медленным. Но что, если нет никаких причин для замедления или ускорения? Нужно ли нам выдвигать гипотезу об испарении препятствий? Ответ опять же дает математика, и он намного проще, чем может показаться на первый взгляд.

Исходные данные в нашей задаче — однополосная дорога с интенсивным движением. Все автомобили едут со скоростью 80 километров в час и находятся так близко друг к другу, что в промежутке между двумя машинами в целях безопасности никакое другое транспортное средство не может вклиниться: короче, дорога максимально загружена. Теперь представьте, что кто-то внезапно замедляется по любой причине. Предположим оптимистичный вариант и допустим, что несчастных случаев нет. Все машины замедлятся: расстояние между двумя соседними машинами сократится до нескольких десятков сантиметров.

Если бы у нас была возможность подняться над дорогой и посмотреть на автомобили с высоты, мы заметили бы явление, которое физики называют **волной трафика** или стоп-волной, которая возникает, даже если нет никаких внешний воздействий: сокращение расстояний между транспортными средствами — двигаясь медленнее, вы можете оставаться ближе, — которое перемещается назад относительно движения потока. Когда первый автомобилист снова начинает ускоряться, другие также ускоряются и волна растворяется: поскольку автомобилисту требуется несколько долей секунды, чтобы понять, что у него есть возможность для движения вперед, то расстояния увеличиваются и кажется, что трафик не изменился.

Волны трафика хорошо заметны на кольцевой дороге, когда особо хитрый водитель перестраивается в соседнюю полосу, так как ему кажется, будто там машины движутся на 2 километра в час быстрее. Его действие немедленно создает стоп-волну, к «радости» сотен автомобилистов, которые замедлятся из-за этого.

Снижение пропускной способности дороги также приводит к пробкам в часы пик по очень простой причине. Представьте, что у нас есть светофор, в котором зеленый свет горит ровно минуту. За это время могут проехать 20 автомобилей, а затем придется подождать еще минуту, чтобы снова загорелся зеленый. Таким образом, пропускная способность дороги составляет 10 автомобилей в минуту, или, если хотите, 600 автомобилей в час. Если они прибывают постоянным потоком, то некоторые попадут на красный, но проедут на первый же зеленый сигнал. Но что произойдет, если поток прибывает с интенсивностью 11 машин в минуту? За первый цикл светофора не успеют проехать 2 автомобиля, и им придется ждать следующего зеленого сигнала. На втором цикле таких машин станет 4 и так далее. Через 10 минут всем понадобится ждать как минимум два цикла светофора для проезда.

Что еще хуже, неподалеку от нашего перекрестка есть пересечение с менее важной дорогой, где зеленый свет горит 75 секунд, и поэтому она может пропустить 25 машин, а не 20. Когда очередь машин растянется до этого светофора, тот, кто подъезжает к нему, вынужден будет останавливаться, даже если горит зеленый, — просто потому, что не сможет продолжать движение (за исключением Италии, где люди, как правило, останавливаются на середине перекрестка, блокируя поток в перпендикулярном направлении, но это уже выходит за рамки математики).

Таким образом, мы экспериментально доказали, что общая пропускная способность дороги определяется участком с наименьшей пропускной способностью. Всё как в цепи, где есть самое слабое звено, которое может разорваться. Но, даже если интенсивность дорожного движения ниже пропускной способности дороги, есть зеленая волна — тоже своего рода стоп-волна! Обеспечить ее — нетривиальная задача: во-первых, светофоры должны быть откалиброваны на скорости ниже максимально допустимого предела, во-вторых, необходимо учитывать время ускорения при срабатывании первого зеленого сигнала, а кроме того, надо помнить, что движение идет в двух противоположных направлениях и, создавая зеленую волну в одну сторону, вы рискуете создать красную волну в другую или в перпендикулярном направлении. В новом строящемся городе основные перекрестки можно расположить на таком расстоянии, чтобы оба противоположных потока попали в зеленую волну. Самое простое решение — одновременно переключать все светофоры, чтобы автомобили успевали преодолеть отрезок между перекрестками к следующему зеленому сигналу, но можно придумать и что-нибудь лучше.

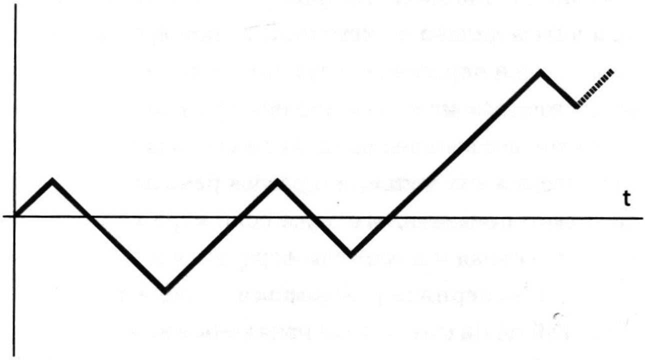
В наших городах, однако, это невозможно. Так что градостроители могут делать вид, что им ничего не известно, и регулировать светофоры наугад (как обычно и бывает в Италии), предлагать компромисс, который никого не устраивает, или двигаться в одном направлении — к худшему. Ситуация аховая, но математика может помочь... конечно, не извлечь кролика из цилиндра.

Случайные прогулки

Вечер пятницы, фактически субботнее утро. Вы уверены, что выпили не слишком много, но никак не можете вспомнить, каким образом очутились на очень узком пирсе, по которому можно двигаться только вперед или назад. Вы идете к суше, но ваш мозг настолько затуманен, что вы шагаете то вперед, то в обратную сторону. Как вы считаете, удастся ли вам добраться до материка и принять освежающую ванну, или вы так и будете ходить туда-сюда, пока не протрезвеете?

Пьяная походка — типичный пример одномерного случайного блуждания. Вы можете это смоделировать. Поставьте эксперимент: несколько раз подбросьте некрапленую монету, ставя +1, когда выпадает орел, и –1, когда выходит решка, а затем постройте график. На рисунке ниже представлен возможный результат эксперимента; горизонтальная ось указывает время, в течение которого проводился эксперимент, а вертикальная — положение, в котором находилась монета в разные моменты. По очевидным причинам симметрии (как говорят физики) среднее смещение относительно начала координат после определенного числа бросков равно нулю, это

эквивалентно тому, что, подбрасывая монету, с которой никто не мухлевал, в среднем вы ничего не выигрываете и не теряете.



Случайное блуждание

Однако более интересный вопрос — каким будет среднее расстояние от начальной точки после *n* шагов? Почему среднее расстояние будет для блуждающих неодинаковым? Ответ прост. Представим блуждающих на пирсе пьяниц. Как мы уже поняли, через некоторое время в среднем половина из них пойдет в одном направлении, а вторая половина — в другом, компенсируя друг друга; но это, конечно, не означает, что все они вернутся в исходную точку. Многие из них будут где-то поблизости от нее, кому-то удастся отойти достаточно далеко, однако нас интересует среднее значение модуля расстояний от начала координат.

И здесь нам помогут закон больших чисел или *центральная предельная теорема*, согласно которой после *n* подбрасываний среднее расстояние от начала координат составляет $\sqrt{2n/\pi}$. Отсюда возникает несколько следствий, на первый взгляд не совсем понятных. Прежде всего, если у вас достаточно времени, будут достигнуты все целые значения; действительно, это будет сделано бесконечное количество раз. Даже когда мы, кажется, находимся далеко от исходной точки, всегда есть шанс вернуться и пересечь ее: главное — не спешить, как говорят заядлые игроки, когда они проигрывают в казино (если же они выигрывают, то продолжают игру, потому что сегодня «их день»). Но раз уж речь зашла о казино, то можно показать, что даже если игра ведется совершенно честная и даже если играешь достаточно долго, то почти наверняка разоришься — если только ты не Билл Гейтс! На самом деле начальный капитал у игрока гораздо меньше, чем у дилера, и поэтому будет намного проще сохранить свои деньги, чем сорвать банк, даже если, возможно, в какой-то момент игры у вас на руках окажется намного больше вашего первоначального капитала.

Что происходит, когда мы перемещаемся в многомерное пространство? Допустим при этом, что наш пьяница может двигаться только ортогонально[1] — влево и вправо, вверх и вниз или даже (кто знает?) в четвертое измерение. Его маршрут становится гораздо разнообразнее. Если вы находитесь в двумерном пространстве и шаги очень малы, то ваши перемещения напоминают броуновское движение, что закономерно, так как оно возникает именно из-за случайных столкновений частиц. Но вот что забавно:

[1] Ортогональный — расположенный под прямым углом, перпендикулярный.

при двумерном перемещении, то есть на плоскости, рано или поздно мы определенно — в вероятностном смысле этого слова — вернемся в начальную точку. Однако, если мы перейдем к трем измерениям, ситуация изменится: мы будем возвращаться к исходной точке примерно один раз из трех (для педантов: в 34,05 % случаев).

Иными словами, совсем не удивительно, что герой фильма «Инопланетянин» был забыт на Земле, если он действительно бежал наугад. Но если в торговом центре вы потеряли из виду своего друга, у которого нет мобильного телефона, то лучшее решение — не бегать повсюду в надежде его найти, а спокойно подождать, пока он сам вас не отыщет. Если, конечно, он не намерен следовать той же стратегии.

Раздел V

Компьютеры и стандарты

Календарь в уме

Календарь, которым мы пользуемся, имеет один недостаток: нельзя предугадать, на какой день недели выпадет определенная дата. На то есть две причины. Во-первых, в каждом новом году все месяцы начинаются в дни недели, отличные от тех, что были в предыдущем. Во-вторых, в году насчитывается целых 52 недели и один день (в високосном — два дня). Вот почему из года в год мы и меняем настольный календарь к вящей радости производителей ежедневников, которые в зависимости от года предлагают до 14 вариантов календарей даже без учета подвижных праздников, таких как Пасха и Пасхальный понедельник.

Возможно, в конце концов кто-нибудь реформирует календарь и упростит нашу жизнь, а тем временем мы можем воспользоваться мнемотехниками — специальными приемами, позволяющими быстро определить день недели, который приходится на ту или иную дату. Математик Джон Хортон Конвей разработал один из таких мнемонических приемов, известный под названием «алгоритм судного дня». С его помощью можно определить день недели, соответствующий любой дате начиная с первого дня григорианского календаря, то есть с 15 октября 1582 года. Алгоритм работает как в традиционном итальянском формате «день-месяц», так и в американском «месяц-день».

У алгоритма существует три уровня сложности: определение дня недели по дате в текущем году, в текущем веке и наконец вечный календарь.

Конвей заметил, что в каждом году есть несколько дат (по одной в месяц), которые легко запоминаются и выпадают на один и тот же день недели. Например, даты, обозначенные парными числами: 4.04, 6.06, 8.08, 10.10 и 12.12 — другими словами, пары четных чисел, за исключением 2.02. Что касается нечетных месяцев (кроме января и марта), то для них такие дни — 9.05 и 5.09, а также 11.07 и 7.11. Их Конвей, как мнемоник, запоминает с помощью фразы: «Я работаю с девяти до пяти в торговом центре 7-Eleven»[1]. Все эти даты вместе с последним днем февраля (иногда его называют также нулевым днем марта) Конвей назвал *doomsdays*, что в переводе с английского означает «судные дни». Осталось лишь определить «судный день» января. В невисокосный год это будет 3.01, в високосный — 4.01.

Зная, какой день недели соответствует «судным дням» текущего года, при незначительной тренировке можно довольно легко установить день недели для любой заданной даты. Так, «судный день» в 2014 году — пятница, в 2015-м — суббота, в 2016-м — понедельник. При

[1] 7-Eleven — оператор крупнейшей сети небольших магазинов в 18 странах под управлением Seven-Eleven Japan Co., Ltd. Дословный перевод названия — «7-одиннадцать».

переходе от предыдущего года к следующему «судные дни» сдвигаются на один день в неделю, или на два дня, если последующий год является високосным.

Приведенных выше объяснений достаточно для практического использования алгоритма. Но если вы хотите рассчитать «судный день» для любого года текущего столетия — например, чтобы сказать ребенку, в какой день недели он родился, — система немного усложняется, так как необходимо производить деление двузначных чисел. Рассмотрим пример алгоритма для 2015 года. Сначала две последние цифры года (в нашем случае 15) разделите на 4. Затем суммируйте исходное число 15 с целым числом 3, которое получилось после деления. И наконец, добавьте магическую константу 2. После всех этих операций у нас выходит 20. При вычитании из него наибольшего числа, кратного 7 (то есть 14), получим 6. «Судный день» года рассчитывается следующим образом: 0 соответствует воскресенью, 1 — понедельнику и так далее. Значит, 6 — это суббота. Я гарантирую, что, немного попрактиковавшись, вы сможете без особого труда совершать такие вычисления в уме!

А что насчет вечного календаря? Нам понадобится знать магическое число для каждого столетия. Но прелесть в том, что каждые 400 лет цикл дней недели повторяется. А значит, достаточно запомнить только четыре магических числа. Итак, магическое число для 1600, 2000, 2400-го — 2; для 1900, 2300, 2700-го — 3; для 1800, 2200, 2600-го — 5; для 1700, 2100, 2500-го — 0. Если вы хотите упростить «алгоритм судного дня», особенно для дат прошлого века, то можете производить вычисления на временном отрезке в 20 лет. Дело в том, что на каждый такой период приходится пять високосных лет, поэтому в вычислениях значение, добавляемое за год, составляет 20 + 5 = 25. Вычитая наибольшее число, кратное 7, мы получаем 25 – 21 = 4. Например, Джон Леннон умер 8 декабря 1980 года: необходимые вычисления с учетом 20-летних отрезков дадут нам число 16, которое нужно суммировать с 3 — константой для XX века. Произведя все операции, получим 5. Это значит, что «судный день» попадает на пятницу. Следовательно, если 12 декабря — пятница, то 8 декабря было четырьмя днями раньше, то есть в понедельник.

Последняя любопытная задача, связанная с календарем: Сервантес и Шекспир скончались 23 апреля 1616 года. Кто умер раньше? Ответ: испанский писатель, потому что в Англии в то время еще использовали юлианский календарь, который отставал от григорианского на 10 дней.

Бумага формата А4

Для большинства итальянцев А4 — это название трассы между Турином и Триестом, но многие знают, что А4 также является форматом бумаги, которая используется в обычных принтерах. Есть ли связь между этими двумя обозначениями? Нет, за исключением того, что оба они предназначены для стандартизации. Идентификационные номера магистралей позволяют их различать, а буквы начинаются с А просто потому, что так принято. Из-за отсутствия определенной структуры дорог номера присваиваются случайным образом: например, исторически сложилось, что не существует трасс А2 и А17, в то время как столичную кольцевую А90 принято называть A-GRA (Grande Raccordo Anulare[1]). Однако с бумагой ситуация совершенно иная.

Лист, книга, газета имеют прямоугольную форму. В тех редких случаях, когда нам в руки попадает квадратный лист бумаги, он кажется странным. Но прямоугольник с какими пропорциями использовать? Сразу приходит

---

[1] Большая кольцевая магистраль (итал.).

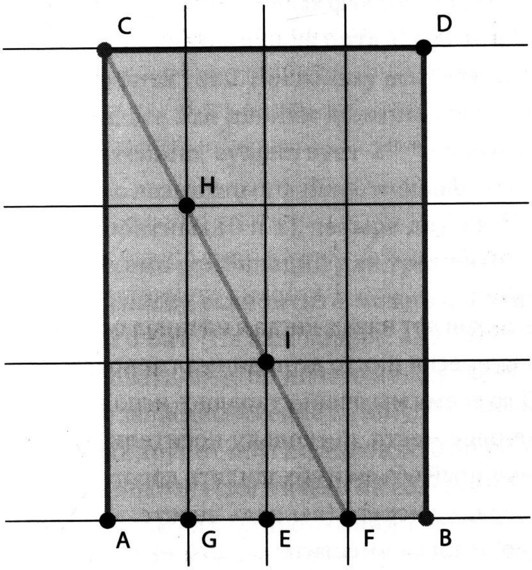
на ум идея о золотом прямоугольнике. Это такой прямоугольник, в котором отношение длин сторон равно математической константе φ (чуть больше 1,6), — его красота высоко ценится. Однако есть одно «но»: золотой прямоугольник, как правило, располагается горизонтально; если же развернуть его и расположить длинную сторону вертикально, то при письме кажется, будто лист не заканчивается. Фактически в Америке используется более или менее приближенный к золотому прямоугольнику формат, который называется *legal* (8,5 × 14 дюймов), что обусловлено, по моему мнению, его удобством для оформления сносок. Размеры стандартного американского формата *letter* — 8,5 × 11 дюймов, следовательно, он гораздо более сжатый.

Для того чтобы в этой истории всё расставить по своим местам, понадобилось вмешательство немцев. В 1786 году ученый Георг Кристоф Лихтенберг заметил, что, если бы отношение длины и ширины листа равнялось $\sqrt{2}$, это давало бы конкурентные преимущества: фактически, сложив лист пополам, можно получить две его половины, которые по пропорциям соответствуют оригиналу. К этой идее вернулись в начале XX века, а уже в 1922 году, прямо перед гиперинфляцией, Веймарская республика разработала стандарт DIN 476, который постепенно распространился практически повсеместно. Италия, например, приняла его в 1939 году, а в 1975-м он стал всеобщим под названием ISO 216. Сегодня этот стандарт принят во всем мире, за исключением Соединенных Штатов Америки, а также некоторых зон Канады (потому что официально у нее нет стандартов) и Мексики (очевидно, она покупает бумагу у северного соседа).

Базовый формат стандарта ISO 216 — A0, соотношение сторон которого равняется $\sqrt{2}$, а площадь составляет 1 квадратный метр. Площадь каждого последующего формата: A1, A2, A3 — равняется половине площади предыдущего формата. Именно поэтому площадь формата A4 (210 на 297 миллиметров) — 1/16 квадратного метра. Квадратный метр бумаги весит 80 граммов, соответственно, лист A4 весит 5 граммов. Об этом я уже писал в главе о задачах Ферми.

Но существуют не только форматы серии A — стандарт включает также серию B. Причем формат B0 фактически самый большой: его короткая сторона равняется 1 метру. Что касается формата C, то он является средним геометрическим между соседними форматами B и A и обычно используется для конвертов. При делении этого формата пополам пропорции сохраняются. Кстати, дальше всех пошли шведы, принявшие серии D, E, F и G!

Последний интересный вопрос: вам когда-нибудь приходилось складывать лист бумаги втрое для узкого конверта? Безусловно, самый простой вариант — это взять лист, по краям которого намечены линии сгиба. Но если вы оказались в стрессовой ситуации и при этом не желаете раздраженно комкать неудачно сложенный лист, то можете воспользоваться схемой, изображенной на рисунке. Разделите короткую сторону листа на четыре равные части, делая сгибы. Затем сложите бумагу по воображаемой линии между углом одного края и точкой третьего сгиба на другом краю. Образованная линия пересекает весь лист и имеет точки пересечения с двумя исходными сгибами на одной трети и двух третях страницы. Вы можете опробовать этот способ, чтобы проверить, а можете, как все, вспомнить свойства подобных треугольников.



Попробуй сложить!

Не доверяйте слишком сжатым файлам

Несколько десятков лет назад, когда я начинал осваивать интернет, практически никто не поднимал вопрос о сжатии файлов. В то время мы очень старались использовать как можно меньше места, поскольку носители данных имели ограниченный объем и обходились дорого. Проще говоря, моя первая дискета (вы ведь знаете, что такое дискета, верно?) могла вместить лишь 144 килобайта. Однако передача данных была еще более ограниченной и дорогой, потому-то и появились первые программы архивации. С тех пор прошло много лет. С одной стороны, после бума 1980-х годов теоретическое изучение вопроса прекратилось и, соответственно, технологии сжатия не развивались. С другой — в интернете то и дело предлагают новые алгоритмы сжатия с невероятной производительностью, способные даже уменьшить размер ранее сжатого файла. Меня неизменно удивляет, почему многие верят подобному обману и «практическим демонстрациям».

Очень легко доказать, что не может быть техники сжатия данных, которая работала бы всегда, то есть уменьшала бы размер любого файла! Фактически файл состоит из множества битов, которые кодируют данные через числовые значения 0 и 1. Выберем начальное значение любого вида, например 1000 бит. Сколько может быть разных 1000-битных файлов? Подсчитаем. Есть два возможных значения для бита 1, два — для бита 2 и так далее до бита 1000. Поскольку все биты не зависят друг от друга, все эти двойки должны быть перемножены: следовательно, существует 2^{1000} возможных файлов. А сколько возможных файлов меньше 1000 бит? Ну, есть два однобитных (0 и 1), четыре двухбитных (00, 01, 10 и 11), восемь комбинаций для трех битов и так далее... К этому списку, если хотите, можно добавить один файл объемом 0 бит. (На ваш вопрос: «Как такой файл может существовать, ведь он ничего не содержит?» — я отвечу: «Данные записаны именно потому, что не существует двух отличных друг от друга файлов по 0 бит». Как они могут быть разными, если ничего не содержат? Значит, существует только один такой файл.) Теперь посчитаем все эти файлы: сумма равна $2^{1000} - 1$, что на единицу меньше, чем у 1000 бит. У математиков есть убедительный способ доказать, что если нам нужно поместить $n + 1$ объектов в n ящиков, то хотя бы в одном ящике будет содержаться как минимум два объекта. Это так называемый принцип ящиков[1]. Перефразируя его для нашего случая, получим правило: каким бы образом мы ни связывали исходные длинные файлы со сжатыми, всегда останется один, и мы не будем знать, что с ним делать. В этом правиле мы не учитываем более короткие файлы, которые обязательно должны быть связаны с файлами весом не менее 1000 бит.

1 Другое название — принцип Дирихле.

После такого разочарования можно задаться вопросом: почему программы сжатия работают, причем довольно хорошо? Все просто: файлов, которые мы действительно хотим сжать, в итоге «мало». Давайте продолжим анализировать 1000-битные файлы, то есть 125 байт — мизерный объем. Мы знаем, что их можно записать в виде 2^{1000}: на бумаге это не впечатляет, однако на самом деле составляет огромный объем. Чтобы попытаться дать вам представление о том, насколько он велик, подумайте обо всей нашей Вселенной, о каждом ее протоне, электроне и нейтроне. Хорошо, представьте, что каждая из этих частиц — такая же вселенная, как наша; возьмите элементарные частицы всех этих вселенных и заставьте их тоже стать вселенными; наконец, подсчитайте частицы этих вселенных. Только теперь вы получили количество, сопоставимое с 2^{1000}. Короче, число возможных 1000-битных файлов просто немыслимо, в то время как те, что мы создаем, представляют собой лишь мизерную долю, и алгоритмы используют тот факт, что все наши файлы не случайны. На практике мы сжимаем нужные нам файлы и не рассматриваем те, которые были созданы случайным образом. Напротив, остерегайтесь тех, кто говорит вам, что они сжали файл, содержавший последовательность случайных чисел: если они действительно преуспели, это доказательство того, что числа были не такими уж случайными.

Для полноты картины напомню, что существуют разные методы сжатия. Очень хорошие результаты дают те, которые используются для аудио-, видео- и графических файлов. Просто задумайтесь: перейдя на цифровое наземное телевидение, мы смогли вместить восемь цифровых каналов в пространство одного аналогового канала. Где-то есть подвох? Ответ на этот вопрос очень простой: используются методы сжатия с потерями. Это значит, что при сжатии удаляются данные, которые считаются бесполезными. Таким образом, mp3-файл, когда вы его слушаете (то есть распаковываете, даже если не понимаете как, отправляя звук в колонки), является не совсем оригинальным файлом. Точно так же изображение формата JPEG не передает все начальные цвета. Для текстового файла такая технология не годится: подумайте, что произойдет, если вы сохраните шедевр литературный, а затем откроете его в более или менее хорошо перефразированном виде. Но с медиафайлами ситуация иная. Говорят, что мужчины распознают только 16 цветов... Это шутка, но никто не может по-настоящему распознать миллионы оттенков, которые способен воспроизводить экран компьютера. Несмотря на то что наш глаз специализируется в различении тонов, я все же думаю, что 50 оттенков серого видны только в названии бестселлера. Единственная важная вещь — не переусердствуйте с коэффициентом сжатия, иначе вы рискуете получить ложные цвета. Или несбалансированные гармонии... пусть даже у некоторых певцов это и сложно распознать!

Абсолютно безопасная криптография

Нет ничего необычного в том, что термин «криптография», будучи у всех на слуху, по сути, остается для многих неизвестным словом. Подумайте, ведь для управления автомобилем не обязательно знать, как работает двигатель внутреннего сгорания, что бы там ни говорили в автошколе. Точно так же мы можем совершать криптографически безопасные транзакции, доверяя тем, кто их создал. Однако, когда читаешь о методах шифрования, раскрытых хакерами (кстати, слово «хакер» имеет совершенно иное значение; тот, кто взламывает защиту, называется крякером), возникает вопрос: почему никому еще не удалось создать абсолютно безопасную криптографическую систему, которую нельзя было бы вскрыть никак иначе, кроме как украв кодовый ключ? Может, потому, что она в принципе невозможна? Да нет. Знайте, что совершенно безопасная криптографическая система известна уже более ста лет и было доказано, что она не поддается дешифрованию. Но почему тогда ее не используют?

Эта история началась добрых 2000 лет назад, в эпоху правления Цезаря. Говорят, император зашифровывал свои тексты, заменяя каждую букву той, которая располагалась тремя позициями позже в алфавите. Таким образом, «А» становится «Г», «Б» — «Д» и так далее; после «Я» вновь идет «А». Соответственно, ЦЕЗАРЬ превращается в ЩЗКГУЯ. Эта система, которую технически можно определить как *одноалфавитное шифрование по заданному алгоритму*, не обладает мощной защитой. Любой, кто увлекается головоломками, хорошо знает, что существуют и более сложные задачи. Среди них задача на определение числа, которому может соответствовать совершенно любая буква без соблюдения какого-либо алгоритма. Но я полагаю, что во времена Цезаря не было необходимости в исключительной изощренности — простого шифра было достаточно.

Первый качественный скачок в развитии криптографии произошел в 1586 году, когда появился *полиалфавитный шифр Виженера*. Принцип этого метода заключается в том, что создается 26 различных шифров Цезаря, каждый из которых связан с конкретной буквой алфавита[1]. Затем выбирается кодовое слово, например «КЛЮЧ». После этого мы шифруем текст, используя для первой буквы шифр, соответствующий «К», для второй буквы — «Л», для третьей — «Ю», для четвертой — «Ч». Кодовое слово используется циклически столько раз, сколько надо, как будто оно было «КЛЮЧКЛЮЧКЛЮЧКЛЮЧ...».

---

1 В латинском алфавите 26 букв. Для русского алфавита необходимо 33 таких шифра (строки).

Шифр Виженера имеет большое преимущество перед простой заменой: благодаря ему невозможно получить информацию о том, как часто встречается та или иная буква в исходном тексте. Известно, что одними из самых частых букв в итальянском языке являются гласные E, A, O, I, в то время как Q — наименее используемая согласная. С помощью частотного анализа можно легко восстановить оригинальное сообщение. Но при изменении кода для каждой буквы этот метод не сработает. Впрочем, если серьезному и мотивированному криптоаналитику предоставить достаточно длинный зашифрованный текст, он может подобрать ключи разной длины и посмотреть, какой из них совместим со статистическим распределением букв в итальянском языке.

Во время Первой мировой войны попытались использовать ключ длиной в весь текст — таким образом, становилось невозможно получить статистику распределения букв. Несмотря на то что идея была представлена тремя десятилетиями раньше, систему запатентовали лишь в 1919 году. Ее назвали шифром Вернама в честь одного из изобретателей, однако она более известна как разновидность *криптосистемы одноразовых блокнотов*. Именно тогда Клод Шеннон, автор теории информации, продемонстрировал, что новый шифр теоретически обладает абсолютной криптографической стойкостью при соблюдении двух условий: 1) ключ никогда не используется повторно; 2) он совершенно случайный. Первое условие делает бесполезным любой частотный анализ, а второе означает, что любой осмысленный текст такой же длины может быть зашифрован в имеющийся у нас шифрованный текст.

И где же подвох? Их как минимум два. Не только ключ должен быть достаточно длинным, чтобы иметь возможность шифровать все сообщения, но и адресат должен получить его раньше самих сообщений и через другой канал. Только подумайте: чтобы совершить безопасную транзакцию в Сети, вам сначала нужно обратиться к администратору сайта, чтобы он дал вам ключ! Вот почему в интернете используются системы шифрования с открытым ключом: хоть и нет теоретической уверенности в их надежности, но они не требуют предварительного обмена ключами.

Короче говоря, идеальный шифр неприменим на практике, за исключением особых случаев. Знаю: это плохая новость — все равно что сказать: «Операция прошла успешно, но пациент умер». Увы, математика не всегда помогает.

Почему компакт-диск не шуршит?

После череды исторических перипетий вернулись в моду виниловые пластинки, скорость вращения которых составляет 33 оборота в минуту. Те, кто их покупает, утверждают, что качество звука у пластинок намного лучше, чем у компакт-дисков. Вот только давайте не будем рассматривать сюжет о том, что mp3-файлы — результат заговора транснациональных корпораций, стремящихся не допустить, чтобы люди узнали, насколько прекрасна музыка на самом деле. Я не собираюсь участвовать в этой дискуссии; да, действительно, лет пятнадцать назад кое-кто из моих друзей, поклонников «Битлз», донимал меня песнями из альбома про сержанта Пеппера, и могу вас заверить, что на виниле и впрямь слышны звуки, которых нет на компакт-диске. Но причина скорее кроется не в разнице между аналоговым и цифровым носителями, а в неправильном обращении с ними. С другой стороны, даже поцарапанный CD не шуршит, в отличие от пластинки. Почему так происходит?

Прежде всего должно быть очевидно, что, в принципе, величина, измеренная цифровым способом, имеет конечное, заранее заданное число возможных значений и будет менее точной, чем величина, измеренная аналоговым способом, которая может принимать бесконечное количество разных значений. На практике все немного по-другому: разница между отрезками длиной 3,141 5926 и 3,141 5927 сантиметра настолько мала, что ее не разглядишь даже с увеличительным стеклом, и никто не сможет отличить эти два отрезка от отрезка длиной π сантиметров.

В случае звуков имеются одновременно и математическая, и физиологическая причины для записи цифровым способом. Существует теорема о дискретизации, согласно которой, если максимальная частота звука составляет N герц, то, обрабатывая не менее $2N$ сигналов в секунду, можно с точностью записать, а затем восстановить звук. В теории все просто: поскольку человек не способен слышать звуки, частота которых превышает 20 000 герц, если мы будем записывать более 40 000 значений в секунду (а стандарт CD составляет 44 100), этого будет достаточно. Жаль, на практике возникают проблемы. Прежде всего необходимо точно определить громкость сигнала в этих точках, поскольку она тоже будет оцифрована; еще необходимо отфильтровать звук, чтобы устранить слишком высокие частоты, которые мы не сможем записать и которые в противном случае будут заменены более низкими. Между нами, я думаю, что зачастую плохое качество звука с компакт-диска по сравнению с винилом является следствием неоптимального фильтра. Вместе с тем должен признать, что доказательств у меня нет. Скажем так: просто, я очень сильно верю в математику.

Довольно легко объяснить, почему царапины на CD, если их не слишком много, не влияют на звук. Хитрость в том, что на компакт-диск записываются не только значения, соответствующие звуку, но и контрольные значения, которые используются для выявления ошибок и позволяют исправлять их — в определенных пределах, разумеется. Математическая теория, лежащая в основе этих методов, начала стремительно развиваться в 1940-х годах, когда конструирование первых цифровых процессоров спровоцировало и потребность в их изучении. Выше я упоминал об оценке максимально возможной погрешности при аналоговых измерениях. Если кто-то поет ноту с частотой, немного отличающейся от правильной, только те, у кого действительно хороший слух, заметят, что певец чуточку не в ладу. Но кто знает, что произойдет, если из-за царапины 0 поменяется на 1 в кодировке звука?

Что делать? В таком случае следует добавить другие значения (помните, что и они могут быть ошибочными), предназначенные для того, чтобы максимально обезопасить исходные данные.

Таким образом, это похоже на ситуацию, когда вы пытаетесь доказать, что попали в мишень размером с почтовую марку, хотя на самом деле поразили область рядом с нею. Нам больше не нужно указывать значения с предельной точностью: достаточно приблизиться к желаемой точке и выбрать точку, ближайшую к указанной. Мы потратили впустую информационные объемы, но обеспечили определенный запас прочности.

На самом деле, конечно, все не так просто, как на словах. Существуют не только теоремы, которые объясняют, сколько и каких цифр нужно, чтобы гарантированно исправить определенное количество ошибок, — возвращаясь в мир физический, надо также учитывать, что царапины концентрируются в определенной области диска, и поэтому двоичные коды, которые соответствуют значениям, правильнее физически разнести, чтобы уменьшить общее количество ошибок на раздел. Однако принцип остается прежним. Хотите выиграть у иглы, идущей вдоль канавки?

Стеганография

Прежде всего хотелось бы уточнить, что стеганография не имеет ничего общего со стенографией. Многие из вас, скорее всего, попались на сходство этих двух слов. Вот почему о нем говорится даже в «Википедии» — в начале соответствующей статьи. Оба термина и впрямь обозначают способы письма, как можно догадаться по корню *-grafia*. Вот только стенография — это система, позволяющая быстро фиксировать устную речь с помощью особых знаков. Для того чтобы запись стала понятна всем, ее необходимо преобразовать в привычный нам текст. Стеганография же — явление абсолютно другого порядка: по-гречески *steganòs* означает «тайный, покрытый, непроницаемый» (этот же корень присутствует в греческом слове «крыша»). Сам термин означает искусство писать вещи, невидимые для тех, кто не должен их видеть.

Стеганография имеет очень древнюю историю. Еще в V веке до нашей эры об этом методе писал Геродот: мол, Демарат выгравировал сообщение на табличке, затем покрыл его воском и написал поверх другой текст. Геродот приводит еще один пример — сообщение в виде татуировки на бритой голове раба Гистея; когда волосы отросли, он ушел с сообщением. Лично я в последнюю историю не верю. Несмотря на то что древние эпохи не были такими беспокойными, как наш век, слишком уж расточительная трата времени — ожидать несколько недель, пока отрастут волосы. Кроме того, известно, что иногда Геродот не проверял свои источники информации.

В любом случае идея скрыть что-либо из поля зрения, чтобы снизить вероятность его обнаружения, была известна на протяжении тысячелетий. Тем не менее термин «стеганография» появился только в 1499 году: именно так назвал свой труд аббат бенедиктинского монастыря Иоганн Тритемий, который под видом трактата о магии фактически рассказывал о способах тайнописи. Книга вызвала такой ажиотаж, что Тритемий попытался уничтожить все копии рукописи, но безуспешно. Так «Стеганография» стала подпольным бестселлером — одним из тех незаменимых произведений, которые помогают зарекомендовать себя своим в кругу «тех, кто имеет значение».

Вариант применения стеганографии, подходящий для печатных книг, разработал Фрэнсис Бэкон. Он предложил использовать в тексте два типа символов (два шрифта, как сказали бы мы сегодня), немного отличающихся один от другого: читая символы, напечатанные вторым шрифтом, мы могли бы обнаружить истинное сообщение.

Перечисленные выше стеганографические методы были «любительскими», потому что не было конкретной основы для их классификации. Иначе говоря, каждый заново изобретал персональную систему шифрования, а затем применял ее. Однако с 1985 года ситуация внезапно изменилась, в чем большая заслуга — или, если хотите, вина — компьютера. С его помощью стало гораздо проще не только извлекать из текста скрытые в нем сообщения, но и прятать их! В качестве текста-основы можно брать что угодно — от поддельной рекламы на eBay до постов в блогах. Но идеальными вариантами для применения стеганографии являются аудиофайлы и особенно изображения, потому что никто не в состоянии заметить минимальные различия в цвете или в звуке, специально привнесенные для того, чтобы передать реальное сообщение. На практике это работает так: те биты изображения, которые кодируют нюансы, заменяются битами, которые после обработки будут соответствовать секретному сообщению. Конечный итог вполне предсказуем: даже если кто-то и знает, что изображение содержит скрытое сообщение, то его извлечение — всего лишь полдела. Сообщение, скорее всего, было к тому же зашифровано. Так зачем же его прятать, если прочесть все равно не удастся? Элементарно! Если человек видит зашифрованное сообщение, то сразу начинает что-то подозревать, пусть и знает, что там написано; если же он шифровку не видит, то остается спокойным.

Но действительно ли стеганография используется на практике? Позволю себе обратить ваше внимание только на один факт: Сеть буквально наводнена фотографиями котиков. С чего бы это вдруг?

Влияние больших данных

О больших данных говорят все. И некоторые даже понимают, о чем идет речь: используя постоянно растущую мощь компьютеров и бесконечное количество собираемых данных, можно делать прогнозы, недоступные для обычных статистических методов, когда надо тщательно составлять выборку — и все равно рискуешь потерей важной информации. Одной из мантр больших данных является фраза *N* = all: выборка соответствует всей совокупности, ничего не выбрасывается. Среди часто упоминаемых успехов этого подхода — проект Google Flu Trends: анализ запросов в поисковых системах, который позволяет оценить, например, масштабы эпидемии гриппа в Соединенных Штатах намного быстрее, чем это сделают Центры по контролю и профилактике заболеваний США, которые вынуждены ждать, пока сотрудники осуществят сбор и обработку медицинских отчетов. Вместо этого Google видит в реальном времени запросы, связанные с эпидемией гриппа или скорее в целом со здоровьем, и выдает свою оценку. Жаль только, что вот уже три года подряд Google Flu Trends ошибается в прогнозах: так, для эпидемии 2013 года система спрогнозировала почти вдвое больше случаев заболевания гриппом, чем было зарегистрировано в действительности. Если бы мы хотели позлорадствовать, то могли бы сказать, что прогнозы были верны только на время, необходимое для публикации первых статей и книг на эту тему, а затем воспользоваться копипастом, чтобы доказать, что это не работает.

Означает ли это, что большие данные — просто рекламный ход и практического значения они не имеют? Ну нет. Например, шахматные программы побеждают даже лучших гроссмейстеров, используя бесконечные библиотеки дебютов и эндшпилей партий. Машинный перевод и распознавание речи продвинулись далеко вперед, с тех пор как лингвистический подход, основанный на знании, был дополнен «тривиальным» механизмом чисто статистического анализа без учета семантических значений. Короче, положительные моменты есть, но, видимо, не все еще хорошо. Что могло пойти не так? Вот несколько возможных гипотез.

Прежде всего полученные результаты могли быть случайным результатом. Прогнозирующие методы Google в конечном счете являются чисто статистическими: парадигма больших данных заключается в том, что мы ищем не причинно-следственные связи, а простые корреляции между событиями. Если *A* и *B* очень часто происходят вместе, это может объясняться тем, что *A* является причиной *B*, что *A* является следствием *B*, что *A* и *B* являются результатом другого неизвестного события *C*... или что они произошли вместе случайно. И тогда возможно, что после первых нескольких лет, когда анализ поисковых запросов случайно совпадал с верным результатом, сейчас связи изменились. Но мне эта гипотеза кажется слишком упрощенной. Это правда, что некоторые из учитываемых Google Flu Trends поисковых запросов на первый взгляд кажутся притянутыми за уши, но многие другие запросы точно выглядят правильными, и совершенно непонятно, с какой стати совпадение совсем случайно.

Вторая возможная причина ошибок заключается в том, что данных недостаточно или по крайней мере их необходимо дополнять. В журнале Scientific American привели высказывание представителя Google: «Мы пересматриваем модель Google Flu Trends каждый год, чтобы посмотреть, как ее улучшить. Последнее обновление было проведено в октябре 2013 года для сезона-2013/14». Ясно, что для того, чтобы найти начальные корреляции между запросами в Google и эпидемией гриппа, необходимо было сопоставить наблюдения, то есть прошлые исследования, с официальными данными о вспышках эпидемии. Тем не менее только относительно недавно объем данных стал достаточным, чтобы рассматривать их как большие данные. То есть доступных данных стало так много, что корреляции могут быть найдены без использования статистических методов. Таким образом, все, что нужно сделать, — пересмотреть алгоритмы с учетом новых данных, полученных за это время. На первый взгляд эта гипотеза выглядит привлекательно, однако при ближайшем рассмотрении можно понять, что она никуда не годится. Если у нас нет возможности априори оценить, насколько большими должны быть данные, чтобы убедиться, что мы получаем правильные результаты, то какой смысл их использовать? Мы оказываемся в ситуации футбольного тренера, который говорит: «Если команда выиграет, то это благодаря моим схемам, если же проиграет, то в этом вина игроков, которые их не применили». Я тоже могу делать такие прогнозы.

Существует и третья гипотеза, в некотором смысле гораздо более тревожная для модели больших данных: наличие обратной связи. То, что широкая публика узнала об алгоритме Google Flu Trends — не сам алгоритм, а всего лишь факт его существования и его предназначение, — влияет, по крайней мере в широком смысле, на исследование этого знания. В двух словах: чем больше мы об этом говорим, тем хуже это работает. Вот тривиальный пример: услышав от многих людей о последнем вирусном видео на «Ютьюбе», я проникаюсь желанием пойти и посмотреть его, и таким образом видео становится еще более вирусным. Также не всегда возможно рассчитать эффект обратной связи и преобразовать алгоритмы, чтобы его учесть. Приведу пример: в 20-х годах прошлого века Альфред Джеймс Лотка и Вито Вольтерра изучали уравнения, которые описывают отношения между хищником и жертвой. Оказалось, что переход к экосистеме с пятью различными видами быстро приводит к хаосу в математическом смысле, то есть к непредсказуемости эволюции популяций. Айзек Азимов, который в цикле романов «Основание» о психоисториках с невероятной обоснованностью предсказал появление метода больших данных, хорошо объяснил это: чтобы Галактика следовала по пути, начертанному Гэри Селдоном, организация «Основание», обеспечивающая это развитие, должна была остаться неизвестной для всех, «чтобы не нарушать равновесие». Короче, для работы моделей, основанных на больших данных, недостаточно, чтобы данные были доступны каждому, — также необходимо, чтобы никто не знал, как они используются. Я не знаю как вас, а меня «настоящая сила — та, о которой никто не догадывается», волнует гораздо больше, чем тривиальные «сильные мира сего».

Моя точка зрения прагматична и может быть сформулирована как «достаточно хорошо». Большие данные работают довольно хорошо — даже намного лучше, чем мы наивно предполагали до их тестирования. Ключевое слово здесь «достаточно». Я могу ввести в программу «Google Переводчик» текст из газеты, написанной на китайском языке, и получить результат, который не совсем непонятен, после чего мне придется приложить много собственных усилий, чтобы более или менее полно уловить смысл. Я, конечно, не являюсь носителем английского языка, но мои переводы с итальянского лучше, чем переводы системы «Google Переводчик», и это что-то да значит. Я чувствую, что существует достаточно ограниченная область, где мы пока превосходим компьютеры, и жду момента, когда лучшим игроком в го станет компьютер. Но в большинстве случаев нам, людям, следует использовать наш естественный интеллект в сочетании с искусственным интеллектом компьютеров, чтобы получать действительно полезные результаты. Чисто статистический подход к искусственному интеллекту, появившийся в начале XXI века, намного лучше подхода «правил» прошлого века, но, на мой взгляд, и он уже исчерпал себя. Нас ждут годы, если не десятилетия, глобальных эффектов, связанных с большими данными, однако для реального прорыва придется подождать, пока какой-нибудь гений не изобретет совершенно иной метод.

Постскриптум: в отношении Айзека Азимова и цикла «Основание» я должен добавить, что Google Flu Trends не предвидел «ненормальную» эпидемию гриппа в 2009 году, связанную с вирусом H1N1, то есть птичьего гриппа. Однако кажется неправильным обвинять Google в полной неспособности прогнозировать: именно потому, что характеристики этой эпидемии не были стандартными, прогностическая модель была обречена на провал. Ситуация аналогична той, что сложилась в романе про «доброго доктора» с момента появления Мула: непредсказуемое одиночное событие, «черный лебедь» Нассима Талеба, приводит к еще более непредсказуемым результатам.

Маурицио Кодоньо — математик и специалист в области информатики, окончил университет Scuola Normale Superiore в Пизе. Кроме того, он переводчик, блогер — пишет для сайтов xmau.com и ilpost.it — и автор нескольких научно-популярных книг о математике. По его собственным словам, любит играть со словами и с числами, постоянно стремится узнавать что-нибудь новое и старается делиться полученными знаниями с другими.

НАУКА на досуге