Юрай Громкович,
Аблаев Фарид Мансурович

АЛГОРИТМЫ И СЛУЧАЙНОСТЬ
УРОК 1. ЧТО ТАКОЕ СЛУЧАЙНОСТЬ

В этой серии статей мы расскажем о том как алгоритмы помогают понять, что такое случайность, и о чудесной роли случайности в разработке алгоритмов. О том, как для решения задач применяется такое, казалось бы, не совместимое с построениями алгоритмов понятие, как случайность, и как при этом для некоторых задач значительно повышается эффективность их решения.

ДЕТЕРМИНИРОВАННОСТЬ И СЛУЧАЙНОСТЬ

Понятие случайности тесно связано с двумя другими важными понятиями: *детерминизм* и *недетерминизм.*

Детерминистический взгляд на устройство мира основан на принципе причинности, который можно описать следующим образом. У каждого явления или события есть причина, то есть события являются следствием причины. Каждое событие имеет какие-то эффекты (результаты), которые, в свою очередь, являются причинами для других событий и т. д. Из этого принципа следует, что если точно известно состояние Вселенной и действующие в ней законы, то можно точно предсказать будущее. И развитие рассматривается как цепь причин и следствий из них.

Недетерминизм, напротив, говорит, что причины могут иметь несколько возможных следствий, и нельзя точно предсказать, какое из них произойдет. Понятие случайности тесно связано с понятием недетерминизма.

В Википедии в статье «Случайность» (ru.wikipedia.org/wiki/случайность) авторы определяют случайность так:

«*Случайность* – проявление внешних неустойчивых связей в действительности, проявление результата пересечения (совпадения) независимых процессов или событий; проявление неотъемлемого дополнения к законам необходимости».

Естественно спросить: существует ли истинная (объективная) случайность или мы всего лишь используем этот термин для описания событий, происходящих по не известным нам законам. Философы и ученые спорили об этом с древнейших времен. Демокрит считал, что *случайное – это то же самое, что неизвестное, и детерминированность (определенность) Вселенной лежит в её основе.*

Таким образом, Демокрит утверждал, что миром правит порядок, но правит он по неизвестным нам законам, а «случайность» – это субъективное чувство, которым мы прикрываем невозможность полностью понять природу вещей и событий. Чтобы пояснить свою точку зрения, Демокрит приводит такой пример. Два человека решили послать своих рабов одновременно за водой, чтобы они встретились. Рабы же не знали об этом и, встретившись, сказали: «О, мы случайно тут встретились».

Эпикур придерживался противоположной точки зрения. Он утверждал, что *случай-*

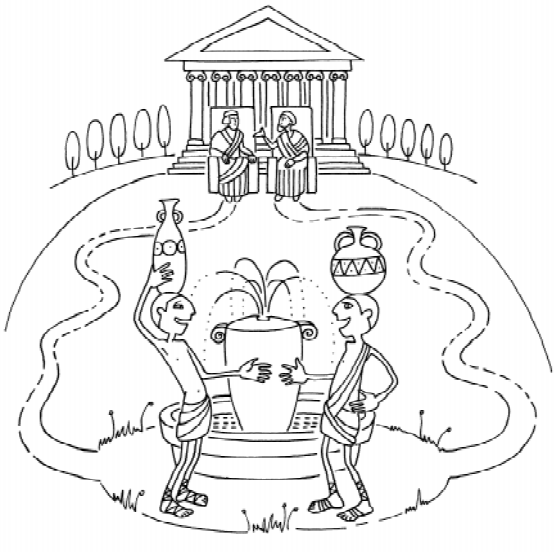
ность объективна, и она является свойством природы вещей.

Таким образом, Эпикур считал, что существует истинная случайность, не зависящая от наших знаний. Он полагал, что существуют процессы, развитие которых скорее многозначно, чем однозначно, и непредсказуемый выбор из существующих возможностей мы называем случайностью.

На первый взгляд кажется, что Эпикур был прав, так как существуют такие игры, как подкидывание монетки или рулетка, и результаты этих игр определяются случайно. Но не все так просто, и, обсуждая азартные игры, мы больше склоняемся к мнению Демокрита. Да, поведение подкинутой монетки сложное действие, но, измерив все параметры этого процесса, такие как сила, направление, поверхность, на которую монета подкидывается и т. д., можно подсчитать результат. Физики тоже часто используют понятие случайности, чтобы смоделировать и описать процессы, которые по сути не являются случайными (а зачастую являются полностью детерминированными), но которые слишком сложные, чтобы моделировать обычными способами.

До двадцатого века мировоззрение людей основывалось на причинности и детерминизме. Причинами этого были, во-первых, теологические воззрения того времени, которые не принимали существование случайности в мире, отождествляя случайность с хаосом. Во-вторых, успехи естественных наук XIX века, развитие механики, интегрального и дифференциального исчислений, заложенные И. Нютоном и другими великими мыслителями, бурное развитие основ электричества давали уверенность, что всё в мире можно изучить и объяснить принципами причинности и детерминированности: «если не будет причины, то не будет и явления».

Интересно, что великий физик, создатель теории относительности Альберт Эйнштейн рассматривал случайность только как относительность нашего незнания всех действующих факторов природы. Известна фраза Альберта Эйнштейна «Бог не играет в кости» («God does not roll dice») в споре с Нильсом Бором об основаниях квантовой механики. Так же знаменит ответ Нильса Бора: «Настоящий Бог не позволит людям устанавливать, что Он должен делать» («The true God does not allow anybody to prescribe what He has to do»).



Два человека решили послать своих рабов одновременно за водой, чтобы они встретились. Рабы же не знали об этом... «О, мы случайно тут встретились».

Открытия XX века (особенно в биологии и физике) вернули мир к эпикуровскому взгляду на случайность. Математическая модель эволюции показывает, что случайные мутации молекул ДНК являются ключевыми для эволюции. Математическая модель поведения элементарных частиц тесно связана с неоднозначностью, которая описывается понятиями теории вероятностей.

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

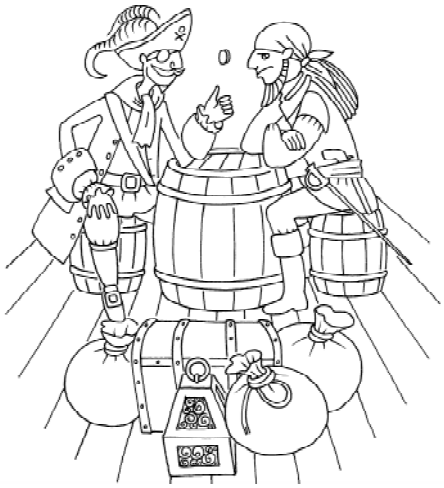
В этом параграфе мы напомним те положения теории вероятностей, на которых будет строиться наше дальнейшее изложение.

Теория вероятностей возникла как раздел математики из многолетних наблюдений, что в основе массовых случайных событий лежат устойчивые закономерности. Теория вероятности изучает данные закономерности. Например: определить однозначно результат выпадения «орла» или «решки»

в результате подбрасывания монеты (как мы это выяснили) практически нереально, но при многократном «честном» подбрасывании «правильной» монеты (многовековая практика игроков) выпадает примерно одинаковое число «орлов» и «решек». Стали полагать «вероятность выпадения орла и решки равна 1/2» (P(«орел») = P(«решка») = 1/2). Оказалось, что в случае подбрасывания «правильной» игральной кости вероятность выпадения конкретной грани – 1/6. При рассмотрении серии из бросаний одной монеты или одной кости (такие испытания называют независимыми) практика показывает, что если в опыте может произойти событие A с вероятностью $P(A)$ и опыт независимо повторили n раз, то событие B – «A произошло n раз» – происходит с вероятностью $P(B) = P^n(A)$, то есть вероятность события B равна произведению вероятностей событий $\underbrace{A, \ldots, A}_{n}$.

Рассмотрим серию «честного» подбрасывания «правильной» монеты. Результат выпадения «орла» или «решки» при подбрасывании будем обозначать «0» и «1» соответственно. Вероятность $P(0)$ события «выпал ноль» и вероятность $P(1)$ события «выпала единица» одинаковы: $P(0) = P(1) = 1/2$.

Пример 1. В серии из двух подбрасываний монеты, возможны четыре варианта результата {00, 01, 10, 11}. Говорят, что они образуют пространство Ω элементарных событий Ω = {00, 01, 10, 11} нашей серии.



Рассмотрим серию «честного» подбрасывания «правильной» монеты.

В данном случае все элементарные события равновероятны

$P(00) = P(01) = P(10) = P(11) = 1/4 = 2^{-2}$.

Действительно, вероятность $P(ab)$ каждого события $ab \in \Omega$ равна произведению вероятностей независимых событий a и b: $P(ab) = P(a)P(b) = 2^{-1} \cdot 2^{-1} = 2^{-2}$.

Пример 2. В условиях примера 1 рассмотрим событие A – «хотя бы один раз выпала единица». Событие A можно представить, как A = {01, 10, 11} – объединение трех элементарных событий или как $A = \Omega / \{00\}$ – множество всех событий исключая элементарное событие 00.

Вероятность $P(A)$ события A можно вычислить двумя способами в соответствии с его представлениями:

в виде суммы

$P(A) = P(01) + P(10) + P(11) = 3/4$

или в виде разности

$P(A) = P(\Omega) - P(00) = 1 - 1/4 = 3/4$.

Здесь и в дальнейшем мы используем понятие вероятности и закономерности поведения независимых испытаний, основанные на частотной точке зрения.

Отметим, что на частотном подходе строится изучение основ теории вероятностей и в школе.

А как обстоит дело со строгим математическим определением понятия вероятность? Этот вопрос находится за пределами школьной программы. Отметим лишь, что долгие годы поисков строгого математического понятия вероятности привели к построению современной теории вероятностей относительно недавно – в начале 30-х годов прошлого века. И сделал это известный российский математик Андрей Николаевич Колмогоров. Метод математического описания событий и вероятностей, предложенный им, придал теории вероятностей стиль, принятый в современной математике.

СЛУЧАЙНОСТЬ. АЛГОРИТМИЧЕСКИЙ ПОДХОД

Итак теория вероятностей – раздел математики, изучающий закономерности слу-

чайных явлений: случайные события, случайные величины, их свойства и операции над ними.

Наряду с изучением закономерностей случайных явлений в множествах элементарных событий, важным вопросом является вопрос проверки случайного характера индивидуальных последовательностей (мы будем говорить о двоичных последовательностях). Вопрос проверки, насколько предъявленную последовательность можно считать случайной, возникает в различных задачах, в частности, в области защиты информации. Этот вопрос имеет и чисто математический интерес. Может ли современная математика отвечать на такие вопросы?

К шестидесятым годам прошлого века сформировался алгоритмический подход к понятию случайности, который показал, как можно решать вопрос о случайности индивидуальных последовательностей. Значительная заслуга в этом принадлежит А.Н. Колмогорову и его ученикам. Ниже мы познакомимся с постановкой задачи и покажем пути, на которых математики построили свои ответы. Для более подробного знакомства с алгоритмическим подходом определения случайности мы советуем книгу [1]. В ней приводится такой пример.

Если кто-либо скажет нам, что он «честно» подбросил «правильную» монету двадцать раз и получил такой результат

10001011011111010000

или такой

01111011001101110001,

мы вряд ли будем удивлены. Однако, если нам скажут, что результат бросаний был таким:

00000000000000000000

или таким

01010101010101010101

мы будем поражены, или вообще не поверим, или же усомнимся в корректности эксперимента.

По-видимому, первая и вторая цепочки оцениваются нами как случайные, а третья и четвертая – как неслучайные. Но что значат слова «воспринимаются как случайные»? Классическая теория вероятностей не дает ответа на этот важный вопрос. Можно попробовать объяснить это, например, так: вероятности третьей или четвертой цепочек слишком малы и равны 2^{-20}. Но ведь при 20 независимых подбрасываниях нашей монеты каждая из 2^{20} возможных цепочек «ничем не лучше и не хуже других» – каждая может появиться с вероятностью 2^{-20}. То есть вероятности появления первой и второй цепочек тоже 2^{-20}.

Итак можно ли отличить случайные цепочки от неслучайных? Отметим сначала, что этот вопрос нельзя ставить о коротких цепочках. Например, вопрос о случайности цепочек длины два из нашего примера 1 вообще не имеет смысла. Следовательно, имеет смысл говорить только о достаточно длинных цепочках (для математики – лучше о последовательностях бесконечной длины). Длина цепочки в 20 символов – тоже немного. Но уже и в этом случае имеем большой (по человеческим меркам) объем материала. Итак, двоичных последовательностей длины 2 всего $2^2 = 4$, а длины 20 уже 2^{20}. Для математики это всего лишь добавление нуля, но число 2^{20} больше миллиона. А в жизни двадцать миллионов символов, напечатанных на бумаге, сколько это? Миллион букв заключает книга убористой печати в 600–800 страниц среднего формата, а двадцать миллионов букв – двадцать таких книг.

Но вернемся к вопросу, какие же последовательности из этого огромного множества последовательностей можно причислить к случайным? Для этого нужно ответить на вопрос, какими свойствами обладает случайная последовательность нулей и единиц? Математиками были сформулированы следующие четыре свойства, характеризующие случайность, с которыми нельзя не согласиться.

Во-первых, случайная последовательность *частотоустойчива*. Это означает, что доля нулей и единиц примерно одинаковы. Но хотя у четвертой последовательности это выполнено, для нас она не выглядит случайной. Поэтому это свойство уточняется так: в случайной последовательности указанная

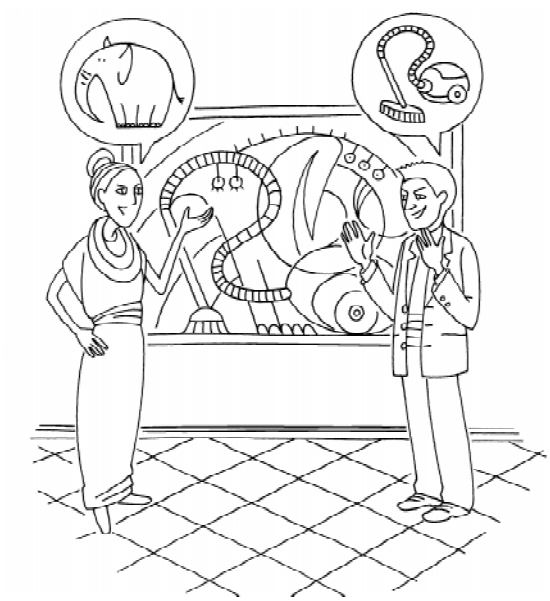
устойчивость частот выполняется для всей последовательности и для любой ее разумной подпоследовательности.

Во-вторых, случайная последовательность *хаотична*. Это означает, что она сложно устроена и не может иметь разумного описания. Например, первая и вторая цепочки воспринимаются как случайные, потому что их устройство сложно и в отличие от третьей и четвертой цепочек, их нельзя коротко описать.

В-третьих, случайная последовательность *типична*. Это означает, что она принадлежит любому разумному большинству.

В-четвертых, случайная последовательность *непредсказуема*. Это означает, что пытаясь угадывать члены последовательности в игре, когда на каждом шаге после нашего предположения открывается очередной элемент последовательности, ее (последовательность) невозможно обыграть, какой бы мы не придерживались стратегии.

Итак мы только что привели четыре свойства: *частотная устойчивость, хаотичность, типичность, непредсказуемость*. Теория алгоритмов обадает средствами выразить все эти интуитивные понятия в терминах своих понятий, наполняя каждое свойство точным смыслом – своим для каждого из четырех свойств. При этом возникают четыре точно очерченных класса, каждый из которых можно брать за образец случайных последовательностей.



Случайная последовательность хаотична. Это означает, что она сложно устроена и не может иметь разумного описания.

В последние десятилетия математики изучили их свойства, рассмотрели как эти четыре класса соотносятся между собой.

ВЕРОЯТНОСТНЫЕ АЛГОРИТМЫ

Использование случайности для разработки алгоритмов и на их основе компьютерных программ вместо полностью детерминированных может быть очень выгодным.

Представьте себе, что компьютерные программы со случайным управлением могут достигать своих целей в миллиарды раз быстрее, чем любые детерминированные компьютерные программы. И за такую огромную прибавку к скорости придется заплатить лишь вероятностью ошибки[1], меньшей, чем $1/10^{19}$?

Вероятность $1/10^{19}$ ошибки вычисления компьютера. Много это или мало? Предположим, что наш компьютер решает некую задачу за секунду, при этом вероятность его ошибки $1/10^{19}$. Если бы мы запускали решение такой задачи на нашем компьютере каждую секунду с момента зарождения вселенной (по подсчетам астрофизиков время жизни нашей вселенной с момента Большого Взрыва составляет 10^{18}–10^{19} секунд), то скорее всего ни разу не получили бы ошибочного результата.

На практике вероятностные (часто также применяют термин рандомизированные) алгоритмы с такой маленькой вероятностью ошибки надежнее, чем лучшие детерминированные алгоритмы. Почему так? Теоретически детерминированные алгоритмы никогда не ошибаются. Но на практике аппаратные сбои в работе компьютеров происходят, и ясно, что вероятность таких сбоев возрастает пропорционально времени выполнения программы. Поэтому быстрый вероятност-

[1] Вероятность получить неправильный результат называют вероятностью ошибки.

ный алгоритм может быть надежнее медленного детерминированного. Например, вероятностный алгоритм, вычисляющий результат за 10 секунд с вероятностью ошибки $1/10^{30}$ надежнее детерминированного алгоритма, работающего одну неделю. Таким образом, переход от обычных (детерминированных) алгоритмов к вероятностным необязательно ведет к потере надежности. И потеря абсолютной надежности не сильно нам мешает. Вот и оказывается, что эффективные вероятностные алгоритмы, которые делают ошибку один раз за миллиарды срабатываний очень надежны.

Чтобы показать читателю практическую пользу случайности, в следующей статье мы представим вероятностный протокол, который решает специальную задачу передачи информации в компьютерных сетях существенно эффективнее лучшего возможного детерминированного протокола.

ВМЕСТО ЗАКЛЮЧЕНИЯ

Понятие *случайности* является понятием современной математики, которое прошло долгий этап развития. Философы, математики, физики, биологи и химики спорили о существовании случайности и о её возможной роли в мире на протяжении многих лет. Для работы со случайными событиями математики создали теорию вероятностей. Были проведены исследования, где и когда выгодно применять случайность. Уже во второй половине двадцатого века были разработаны методы определения случайности на основе алгоритмических подходов. В последние десятилетия были разработаны и продолжают разрабатываться вероятностные алгоритмы и компьютерные программы на их основе, которые превышают детерминированные программы по своему быстродействию, не уступая им в надежности вычислений.

Литература

1. *Успенский В.А.* Четыре алгоритмических лица случайности / www.mccme.ru/dubna/books/pdf/vau-random.pdf.

Юрай Громкович,
Professor of Computer Science, Swiss Federal Institute of Technology, Zürich,

Аблаев Фарид Мансурович,
доктор физико-математических наук, профессор, заведующий кафедрой теоретической киберенетики Казанского федерального университета.