

# четыре алгоритмических лица случайности

В. А. УСПЕНСКИЙ

Дубна, 23 июля 2005 г.

## ПРЕДИСЛОВИЕ

Если кто-либо скажет нам, что он подбросил «честную» монету двадцать раз и, обозначив герб единицей, а решётку нулём, получил такой результат:

100010111101111010000 (I)

или такой:

01111011001101110001, (II)

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

00000000000000000000 (III)

или таким:

01010101010101010101 (IV)

— мы будем поражены и вообще не поверим или же усомнимся в корректности эксперимента. Возникает вопрос, почему.

По-видимому, цепочки (I) и (II) воспринимаются как случайные, а цепочки (III) и (IV) — как неслучайные.

Но что означают слова «воспринимается как случайная»? Классическая теория вероятностей не даёт ответа на этот важный вопрос. Не столь редко можно услышать следующее объяснение: вероятность каждой из цепочек (III) и (IV) слишком мала, она равна  $2^{-20}$ , что меньше одной миллионной. Но ведь ровно такую же вероятность имеют цепочки (I) и (II).

Три важных замечания.

Во-первых, понимание случайности зависит от распределения вероятностей. Если одна сторона монеты тяжелее другой или если человек в процессе бросания научается подкидывать монету так, чтобы она падала на нужную сторону, то появление цепочек (III) или (IV) может стать вполне ожидаемым. Для простоты мы будем заниматься лишь независимыми бросаниями с вероятностью одна вторая для герба и решки.

Во-вторых, точной границы между случайными и неслучайными цепочками нет и не может быть. Ведь если в случайной цепочке заменить один знак, она должна остаться случайной. Но, заменяя много раз, мы от любой цепочки можем прийти к (III) или (IV).

В-третьих, говорить о случайности имеет смысл лишь в применении к очень длинным цепочкам. Бессмысленно спрашивать, которая из цепочек 00, 01, 10, 11 более случайна, чем другие.

Поэтому следует заниматься только очень длинными цепочками, а проще всего — бесконечными (т. е. **последовательностями**), что мы и будем делать. Вообще, бесконечность — это такое полезное приближение сверху к очень большому конечному.

Оказывается, бесконечные последовательности уже можно довольно осмысленно разделить на случайные и неслучайные. Иными словами, можно пытаться найти строгое математическое определение для понятия ‘случайная последовательностей нулей и единиц’.

Нам будут полезны некоторые термины и обозначения.

Множество всех бесконечных последовательностей нулей и единиц обозначаем  $\Omega$ . Всякая конечная цепочка  $\langle i_1, \dots, i_n \rangle$  нулей и единиц называется также *двоичным словом*, а число  $n$  — *длиной* этого слова. Длина слова может быть равна нулю; слово длины ноль ничего не содержит и называется *пустым*. Пустое слово обозначается буквой  $\Lambda$ . Множество всех двоичных слов обозначается буквой  $\Xi$ . Длина слова  $x$  обозначается так:  $|x|$ .

Для последовательности  $\langle a_1, a_2, a_3, a_4, \dots \rangle$  каждое из двоичных слов  $\langle a_1, a_2, \dots, a_n \rangle$  является её *началом*. Пусть  $x \in \Xi$  — двоичное слово. Множество всех тех последовательностей из  $\Omega$ , для которых  $x$  служит началом, обозначим через  $\Omega_x$ . Каждое множество вида  $\Omega_x$ , где  $x \in \Xi$ , принято называть *шаром*. Шару  $\Omega_x$  приписывается его *объём*  $\nu(x)$ , равный  $2^{-|x|}$ .

Содержательно, каждую последовательность из  $\Omega$  мы рассматриваем как серию результатов бросаний монеты. Ещё раз подчеркнём, что для наглядности мы ограничиваемся ситуацией, когда результаты всех бросаний равновероятны и независимы друг от друга. На математическом языке это означает, что на  $\Omega$  задано так называемое *равномерное распределение вероятностей* — а это, в свою очередь, означает следующее: для каждого шара вероятность того, что случайно выбранная последовательность из  $\Omega$  попадёт в этот шар, равна объёму этого шара.

Наша цель — попытаться выделить из  $\Omega$  некоторое точно описанное подмножество, претендующее на роль множества всех случайных последовательностей. Традиционная теория вероятностей не только не приближается к решению этой задачи, но даже не может её сформулировать в своих терминах. На помощь приходит теория алгоритмов. Может показаться парадоксальным, что понятие случайности уточняется на основе такого чуждого случайности понятия, как алгоритм, — тем не менее дело обстоит именно так: все известные до сих пор определения случайности индивидуального объекта (в нашем примере — индивидуальной последовательности нулей и единиц) опираются на понятие алгоритма.

Чтобы найти требуемое определение, поступают так. Формулируют некое характеристическое свойство, которым обладают случайные (в неформальном, интуитивном смысле) последовательности. А затем последовательности, обладающие этим свойством, и объявляют — по определению — случайными.

Какими же свойствами обладает случайная последовательность нулей и единиц?

Во-первых, она *частотойстойчива*. Вот что это означает для того простейшего случая, когда нули и единицы равновероятны — а только такой случай мы и будем рассматривать: частота нулей, как и частота единиц, стремится к одной второй. (*Частота нулей* — это их доля в начальном отрезке последовательности.) При этом указанная устойчивость частот выполняется не только для последовательности в целом, но и для любой её законной, разумной подпоследовательности.

Во-вторых, она *хаотична*. Это означает, что она сложно устроена и не может иметь

разумного описания. (Психологическому эксперименту, с которого мы начали, Колмогоров дал такое объяснение. Цепочки (I) и (II) потому воспринимаются как случайные, что они сложны, их устройство нельзя коротко описать. А вот цепочки (III) и (IV) имеют простое, легко описываемое устройство.)

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

В-четвёртых, она *непредсказуема*. Это означает, что играя против неё на деньги (то есть пытаясь угадать члены последовательности и делая ставки) последовательность невозможно обыграть, какой бы разумной стратегией не пользоваться.

Слово “разумный”, встречающееся в объяснениях перечисленных четырёх свойств, конечно, нуждается в уточнении. Теория алгоритмов как раз и предлагает такие уточнения, наполняя это слово точным смыслом — своим для каждого из наших четырёх свойств. Тем самым возникают четыре алгоритмических свойства: *частотная устойчивость*, *хаотичность*, *типичность*, *непредсказуемость*. Каждое из них представляет своё собственное алгоритмическое лицо случайности, и каждое из них с большими или меньшими основаниями может претендовать на роль строгого математического определения понятия случайности. Можно сказать и так: возникают четыре точно очерченных класса последовательностей, каждый из которых претендует на то, чтобы служить истинным классом случайных последовательностей; некоторые из этих претензий более оправданы, чем другие.

## Лицо Первое: ЧАСТОТУСТОЙЧИВОСТЬ

По-видимому, первым поставил вопрос о том, что такое отдельно взятая случайная последовательность замечательный немецкий математик Рихард фон Мизес в начале XX века — в 1919 г.. Во всяком случае, он первым попытался на этот вопрос ответить. Мизес исходил из того, что случайной последовательности должна быть присуща устойчивость частот. А именно, доля единиц (как и доля нулей) в начальном отрезке случайной последовательности должна стремиться к одной второй при неограниченном увеличении длины начального отрезка. Но этого недостаточно. Например, этим свойством устойчивости частот обладает заведомо неслучайная последовательность

$$0, 1, 0, 1, 0, 1, 0, 1, \dots$$

Очевидным образом необходимо, чтобы устойчивостью частот обладала не только вся последовательность целиком, но и её подпоследовательности. Однако устойчивостью частот не могут обладать все подпоследовательности. В самом деле, возьмём самую что ни на есть случайную последовательность и отберём в подпоследовательность те её члены, которые равны нулю... Значит, подпоследовательность, в которой должна соблюдаться устойчивость частот, должна быть **разумной**, или *допустимой*.

Сам Мизес объяснял, какие последовательности следует считать допустимыми, в расплывчатых терминах. Первое уточнение предложил в 1940 г. американский математик Алонзо Чёрч — один из создателей теории алгоритмов. Чёрч предложил, чтобы допустимые подпоследовательности строились на основе определённых алгоритмов. Последовательности, в которых устойчивость частот наблюдается во всех допустимых по Чёрчу подпоследовательностях, получили название *стохастических по Чёрчу*. Однако оказалось, что определение Чёрча слишком широко: существует стохастическая по Чёрчу последовательность, перестающая быть таковой после вычислимой перестановки её членов.

В 1963 г. Колмогоров усовершенствовал конструкцию Чёрча, предложив более обширный класс допустимых алгоритмов отбора членов исходной последовательности в подпоследовательность и тем самым более обширный класс допустимых подпоследовательностей; в частности, колмогоровский алгоритм разрешает членам подпоследовательности иметь иной порядок следования, чем в исходной последовательности. Поэтому класс всех последовательностей, стохастических по Колмогорову (т. е. таких, у которых устойчивость частот наблюдается во всех допустимых по Колмогорову подпоследовательностях), является подклассом класса последовательностей, стохастических по Чёрчу — на самом деле даже строгим подклассом, т. е. не совпадающим с объёмлющим классом. Но и этот подкласс оказался слишком широк: может быть, например, построена такая стохастическая по Колмогорову последовательность, в каждом начальном отрезке которой единиц больше, чем нулей; а это противоречит как интуиции, так и законам теории вероятностей (например, закону возвратности состояний при случайном блуждании). Для дальнейших ссылок класс всех последовательностей, стохастических по Колмогорову, обозначим буквой **S**.

Таким образом, наиболее продвинутое на сегодняшний день математически строгое уточнение идей Мизеса не даёт полного отражения интуитивного представления о случайности (хотя и это неполное отражение может оказаться полезным).

Для большей ясности приведём конструкции Чёрча и Колмогорова. Центральным для обеих конструкций является указание тех правил, согласно которым из членов рассматриваемой последовательности составляются допустимые подпоследовательности. Поскольку каждое такое правило производит отбор членов последовательности для включения их в допустимую подпоследовательность, сами эти правила принято называть *допустимыми правилами отбора*.

Для наглядности представим себе, что члены исследуемой последовательности написаны на картах. Карты лежат одна за другой, лицевой стороной вниз, так что мы не видим, что на них написано. Наша цель — отобрать некоторые из этих карт и составить из них другую последовательность — допустимую подпоследовательность. (Как мы увидим, в случае колмогоровской конструкции термин “подпоследовательность” имеет более широкий смысл, чем это обычно принято.) Допустимое правило отбора представляет собою алгоритм, который на каждом этапе построения подпоследовательности предписывает, во-первых, которую из карт надлежит открыть и, во-вторых, следует или нет включать эту уже открытую карту в подпоследовательность. Входом алгоритма служит информация о всех уже открытых к этому моменту картах. Может случиться, что алгоритм отберёт лишь конечное число членов исходной последовательности; в этом случае считается, что никакой допустимой подпоследовательности не образовалось. (Напомним, что исходная последовательность признаётся стохастической при наличии устойчивости частот в любой её допустимой подпоследовательности.)

Переходим к более точному описанию.

Функция называется *вычислимой*, коль скоро существует алгоритм, перерабатывающий каждый из её аргументов в соответствующее значение, — иными словами, **вычисляющий** её значение по заданному аргументу.

Исследуемую на устойчивость последовательность обозначим  $a_1 a_2 a_3 \dots$ , так что на  $n$ -ой карте записана цифра  $a_n$ , равная 0 или 1.

Допустимое по Чёрчу правило отбора состоит в предъявлении произвольной вычислимой функции  $G$ , определённой на множестве  $\Xi$  всех двоичных слов и принимающей одно из двух значений: *Да* и *Нет*. Карты открываются последовательно, одна за другой, начиная с первой, и каждый раз — до открытия очередной карты — решается вопрос, включать ли эту карту в подпоследовательность. Вопрос этот решается следующим образом. Пусть уже открыты  $n$  карт, на которых записаны, соответственно, цифры  $a_1, \dots, a_n$ . Если  $G(a_1 \dots a_n)$  есть *Нет*, то следующая,  $(n+1)$ -я карта не включается в подпоследовательность. Если  $G(a_1 \dots a_n)$  есть *Да*, то следующая,  $(n+1)$ -я карта включается в подпоследовательность в качестве очередного члена. Включать или не включать в подпоследовательность член  $a_1$ , зависит от значения  $G(\Lambda)$ . Таким образом, правилу  $G$  отвечает бесконечная подпоследовательность (в случае, если таковая образовалась)  $a_{n(1)} a_{n(2)} a_{n(3)} \dots$ , где числа  $n(1), n(2), n(3), \dots$  образуют возрастающую последовательность, составленную из всех чисел  $n$ , для которых  $G(a_1 \dots a_{n-1}) = \text{Да}$ .

Изложению конструкции Колмогорова предпослём следующее обобщение понятия последовательности. *Обобщённой подпоследовательностью* последовательности  $a_1 a_2 \dots a_n \dots$  назовём всякую последовательность вида

$$a_{\phi(1)} a_{\phi(2)} \dots a_{\phi(k)} \dots,$$

для которой выполнено условие

$$i < j \implies \phi(i) \neq \phi(j).$$

Для обычной подпоследовательности (которая является частным случаем обобщённой) это условие заменено на более сильное — с правой частью  $\phi(i) < \phi(j)$ .

Каждое допустимое по Колмогорову правило отбора образует некоторую обобщённую подпоследовательность исходной последовательности. (Для стохастичности по Колмогорову требуется устойчивость частот в каждой из допустимых, то есть образованных каким-либо допустимым по Колмогорову правилом, обобщённых подпоследовательностей.) Само правило состоит в предъявлении двух вычислимых функций:  $F$  и  $G$ . Первая служит для образования некой вспомогательной обобщённой подпоследовательности. Окончательная же обобщённая подпоследовательность исходной последовательности строится при помощи функции  $G$  в качестве обычной подпоследовательности вспомогательной обобщённой подпоследовательности. Аргументами функций  $F$  и  $G$  служат двоичные слова — но не обязательно все слова, так что каждая из этих функций определена на своём подмножестве множества  $\Xi$ . Значениями функции  $F$  служат целые положительные числа, у функции  $G$  два возможных значения: *Да* и *Нет*. Сперва строится последовательность натуральных чисел  $n(1) n(2) n(3) \dots$ :

$$n(1) = F(\Lambda), \quad n(2) = F(a_{n(1)}), \quad \dots, \quad n(k+1) = F(a_{n(1)} \dots a_{n(k)}).$$

Построение этой последовательности прекращается, как только наступает один из трёх случаев:

значение  $F(a_{n(1)} \dots a_{n(k)})$  не определено;

значение  $G(a_{n(1)} \dots a_{n(k)})$  не определено;

значение  $F(a_{n(1)} \dots a_{n(k)})$  совпадает с одним из чисел  $n(1), \dots, n(k)$ .

Если же остановки в построении не произошло и возникла бесконечная последовательность номеров  $n(1) n(2) n(3) \dots$ , то далее строится вспомогательная обобщённая подпоследовательность  $a_{n(1)} a_{n(2)} a_{n(3)} \dots$ . Из неё, наконец, отбираются — в порядке возрастания  $k$  — те её члены  $a_{n(k)}$ , для которых выполнено равенство  $G(a_{n(1)} \dots a_{n(k-1)}) = \text{Да}$ .

## Лицо Второе: ХАОТИЧНОСТЬ

Вернёмся к примерам конечных цепочек из Предисловия и вспомним объяснение, предложенное Колмогоровым: цепочки (I) и (II) воспринимаются как случайные, потому что они **сложны**; цепочки (III) и (IV) воспринимаются как неслучайные, потому что они **просты**. По-видимому, мы ожидаем, что результат случайного эксперимента окажется сложным и удивляемся, когда получаем что-то простое.

Некоторые объекты можно квалифицировать как большие или маленькие, другие — как тяжёлые или лёгкие, третьи — как сложные или простые. В начале 60-х годов Колмогоров наметил математическую теорию, позволяющую оценивать сложность объектов. Сейчас эта теория называется теорией колмогоровской сложности

В основе теории колмогоровской сложности лежит следующая простая и естественная идея:

**сложность объекта измеряется длиной его кратчайшего описания.**

В самом деле, каждый объект может иметь сколь угодно сложные описания, но сложный объект невозможно описать коротко.

Пусть  $Y$  — множество всех объектов, которые мы рассматриваем, а  $X$  — множество всех мыслимых описаний этих объектов. Через  $|x|$  обозначим длину описания  $x$ . Сложность объекта  $y$  обозначим  $\text{Comp}(y)$ . В соответствии со сказанным,

$$\text{Comp}(y) = \min_x \{|x| : x \text{ есть описание для } y\}.$$

Если объект  $y$  неопишуем, т. е. для него не существует описания, то, естественно, его сложность равна бесконечности.

Разумеется, длины надо мерить единообразным способом, чтобы не получилось, что описание одного и того же объекта имеет на китайском языке длину единица (один иероглиф), а на русском — сорок (сорок букв). Поэтому все описания кодируются в двоичном алфавите — в виде двоичных цепочек. Вот эти-то двоичные коды мы и будем отныне считать описаниями, так что отныне  $X = \Xi$ .

Множество всех таких пар  $\langle x, y \rangle$ , что  $x$  служит описанием для  $y$ , естественно называть *языком*.

Заметим, что и один и тот же объект может иметь в данном языке много описаний, и одно и то же описание может служить описанием многих объектов (таково, например, описание: «двоичное слово, состоящее только из нулей»).

Всё сказанное носило предварительный характер, чтобы оправдать нижеследующее формальное изложение.

В декартовом произведении  $\Xi \times Y$  рассматриваем произвольное подмножество  $E$ , которое называем *языком*. Если  $\langle x, y \rangle \in E$ , будем говорить, что  $x$  является *описанием* объекта  $y$ . Сложность  $\text{Comp}_E$  объекта  $y$  относительно языка  $E$  определяется так:

$$\text{Comp}_E(y) = \min_x \{|x| : \langle x, y \rangle\}.$$

Напомним, что минимум по пустому множеству принято считать равным бесконечности.

Для языке  $E = \Xi \times Y$ , в котором каждое  $x$  служит описанием для каждого  $y$ , сложность любого объекта  $y$  равна нулю, поскольку одним из описаний этого  $y$  является

пустое слово; такой язык мыслим, но не будет встречаться в тех семействах языков, которые мы будем рассматривать.

Представим себе два языка, причём описание какого-либо объекта на втором языке происходит путём удвоения описания этого же объекта на первом языке. Ясно, что второй язык «хуже» первого. Предпочтительнее те языки, которые в состоянии давать более короткие описания.

Будем говорить, что язык  $A$  не хуже языка  $B$  и писать  $A \leq B$ , если существует такая константа  $c$ , что для всякого  $y$  справедливо неравенство  $\text{Comp}_A(y) < \text{Comp}_B(y) + c$  — короче, если

$$\exists c \forall y \text{Comp}_A(y) < \text{Comp}_B(y) + c.$$

Если принять, что для так называемых естественных языков, то есть для тех, которыми пользуется человечество, могут быть сформулированы правила перевода с одного языка на другой, то становится очевидным, что каждый из этих языков не хуже другого. Ведь, скажем, турецкое описание какого-либо объекта можно составить так: взять японское описание этого объекта и дополнить его правилами японо-турецкого перевода. И турецкое, и японское описания, и правила перевода считаются закодированными в виде двоичных слов. Поэтому длина турецкого описания равна длине японского описания плюс не зависящая от выбора объекта длина правил перевода. В итоге получаем, что турецкий язык не хуже японского.

Встаёт вопрос о выборе оптимального языка — такого, который не хуже любого другого. Пусть  $\mathcal{L}$  — некоторое языковое семейство, т. е. попросту некоторое множество языков. Язык  $A$  из этого семейства  $\mathcal{L}$  называется *оптимальным* (для  $\mathcal{L}$ ), если он не хуже любого другого языка из этого семейства, т. е. если

$$\forall B \in \mathcal{L} A \leq B.$$

Если оптимальный язык существует, то именно с его помощью следует измерять сложность. Сложность объекта относительно оптимального языка называется *энтропией* этого объекта. Энтропию и рассматривают как окончательную искомую меру сложности — в рамках заданного языкового семейства.

Для некоторых важных языковых семейств имеет место теорема о существовании оптимального языка, а тем самым и энтропии. Эта теорема называется Теоремой Соломонова – Колмогорова.

Для данного семейства может существовать (и, как правило, существует) много оптимальных языков а, следовательно, много энтропий. Однако, в силу определения оптимальности, любые две энтропии (взяты для одного и того же фиксированного языкового семейства) различаются не более чем на аддитивную константу. Иными словами, если  $A$  и  $B$  суть два оптимальных для  $\mathcal{L}$  языка, то существует такая константа  $c$ , что

$$|\text{Comp}_A(y) - \text{Comp}_B(y)| < c.$$

**З а м е ч а н и е .** Конечно, неприятно, что энтропия, претендующая на роль «истинной сложности» определяется не однозначно, а всего лишь с точностью до аддитивной добавки, не превышающей константы. Однако попытки найти среди энтропий «наиболее оптимальную» пока что ни к чему хорошему не привели. С другой стороны,

ведь, скажем, и длина, и вес также определены неоднозначно, а с точностью до мультипликативной константы, зависящей от выбранной единицы измерения (10 см и 100 мм; 2 кг и 2000 г).

Из уважения к Колмогорову энтропию обозначают обычно буквой  $K$ ; иногда к этой букве  $K$  добавляют ещё вторую букву, указывающую то языковое семейство, к которому относится рассматриваемая энтропия. Если  $K'$  и  $K''$  суть две энтропии, относящиеся к одному и тому же языковому семейству, то, как отмечалось,

$$|K' - K''| < c.$$

Колмогоров не только сформулировал понятие энтропии объекта, но и приложил его к исследованию случайных последовательностей. Наблюдение Колмогорова состояло в том, что в случайной последовательности энтропия начального отрезка растёт достаточно быстро при неограниченном увеличении длины этого отрезка. (Ничего не поделаешь, случайная последовательность может начинаться с миллиона нулей — но и этот миллион ничто перед бесконечностью, и если взять все начальные отрезки в совокупности, то их энтропия неизбежно будет расти.)

Итак, в качестве описываемых объектов мы будем рассматривать двоичные цепочки — такие, как (I), (II), (III), (IV) и т. п. Поэтому отныне  $Y = \Xi$ .

Если язык содержит пару  $\langle z, z \rangle$ , то это означает попросту что цепочка  $z$  описывает саму себя. Язык  $D$ , состоящий в точности из пар такого вида, называется *диагональным* (термин из математики), или *автонимным* (термин из лингвистики). Для такого языка  $\text{Comp}_D(y) = |y|$ . Ограничимся языковыми семействами, содержащими автонимный язык (таким будет, в частности, семейство монотонных языков, описанное ниже). Тогда для каждой энтропии  $K$ , соответствующей этому семейству, и подходящей константы  $c$  будет выполнено неравенство:

$$K(y) < |y| + c.$$

Таким образом, если пренебречь аддитивной константой, максимально возможное значение для энтропии цепочки равно длине цепочки. Колмогоров предположил, что у случайной последовательности энтропия начального отрезка достигает этого максимума, т. е. равна длине этого отрезка, — опять-таки, при пренебрежении аддитивной константой. В этом состояла основная идея Колмогорова относительно хаотичности.

Итак, фиксируем некоторое языковое семейство и одну из соответствующих этому семейству энтропий  $K$ . Назовём последовательность

$$a_1, a_2, \dots, a_n, \dots$$

*хаотической*, если существует такая константа  $c$ , что

$$K(a_1, a_2, \dots, a_n) > n - c.$$

Очевидно, свойство хаотичности не зависит от выбранной энтропии, а только от выбранного семейства.

Оказалось, что при подходяще выбранном языковом семействе строгое понятие хаотичности хорошо отражает интуитивное представление о случайности.

Создавая теорию сложности объектов, Колмогоров придал соотношению ‘описание – объект’ алгоритмический характер. Следуя Колмогорову, ограничимся **перечислимыми** языками. Понятие *перечислимого множества* — это одно из основных понятий теории алгоритмов (да и математики в целом). Этому понятию можно дать такое наглядное объяснение. Представим себе безостановочно работающий принтер, последовательно печатающий слова. Перед напечатанием очередного слова принтер берёт время на размышление. Это время может оказаться бесконечным, и тогда слово не печатается вовсе; в таком случае напечатанным будет лишь конечное множество слов — в частности, пустое множество, если принтер в самом начале своей работы задумался навсегда. Так вот, множество всех когда-либо напечатанных таким принтером слов окажется перечислимым — и всякое перечислимое множество может быть получено таким образом при подсоединении принтера к «идеальному компьютеру». Перечислимо и множество теорем любой формальной теории. Здесь нет места разъяснять, ни что такое идеальный компьютер, ни что такое формальная теория. Но определение понятия ‘перечислимое множество’ мы сейчас приведём.

Для наглядности начнём с термина “счётное множество”. Термин этот имеет два варианта значения. В первом, более узком (и более распространённом) значении счётное множество — это такое множество, которое можно поставить во взаимно однозначное соответствие с натуральным рядом. Во втором, более широком значении счётное множество — это такое множество, которое можно поставить во взаимно однозначное соответствие с каким-либо начальным отрезком натурального ряда. Напомним, что *начальным отрезком натурального ряда* называется произвольное множество  $M$  натуральных чисел, содержащее вместе со всяким своим элементом и все меньшие числа. Таким образом, и весь натуральный ряд, и пустое множество являются начальными отрезками. Поэтому при втором понимании счётности все конечные множества, в том числе пустое множество, оказываются счётными. Именно это второе, более широкое понимание счётности удобно для наших целей; его мы и примем. А тогда счётному множеству можно дать и такое определение: **множество называется счётным, если оно либо пусто, либо может быть расположено в последовательность (то есть является множеством членов какой-либо последовательности)**. Например, конечное множество  $a, b, c$  можно расположить в последовательность  $a, b, c, c, c, c, \dots$ . Заменяя в приведённом определении счётного множества термин “последовательность” на термин “вычислимая последовательность”, мы получаем определение перечислимого множества. Что касается определения термина “вычислимая последовательность”, то оно сейчас будет дано.

Последовательность  $w_1, w_2, \dots, w_n, \dots$  называется *вычислимой*, коль скоро существует алгоритм, вычисляющий её  $n$ -ый член  $w_n$  по его номеру  $n$ . Поэтому понятие вычислимой последовательности называют *алгоритмическим аналогом* или *эффективным аналогом* понятия последовательности, а понятие перечислимого множества — алгоритмическим аналогом или эффективным аналогом понятия счётного множества. Перечислимые множества называются также *эффективно счётными*. Повторим определение: **множество называется перечислимым, или эффективно счётным, если оно либо пусто, либо может быть расположено в вычислимую последовательность (то есть является множеством членов какой-либо вычислимой**

последовательности).

Все рассматриваемые нами языки суть подмножества декартова произведения  $\Xi \times \Xi$  и, следовательно, счётны. Идеология Колмогорова состояла в том, чтобы рассматривать только эффективно счётные (они же — перечислимые) языки.

Окончательно надлежащий выбор языкового семейства произвёл колмогоровский ученик Леонид Левин. Именно, в 1973 г. он ввёл в рассмотрение семейство монотонных языков и изучил соответствующее этому семейству понятие хаотичности. Приведём необходимые определения.

Будем говорить, что двоичные слова  $u$  и  $v$  *согласованы* и писать  $u \approx v$ , если одно из этих слов есть начало другого.

Язык  $E$  называется *монотонным*, если он перечислим и выполнено условие:

$$[\langle x_1, y_1 \rangle \in E \ \& \ \langle x_2, y_2 \rangle \in E \ \& \ (x_1 \approx x_2)] \implies [y_1 \approx y_2].$$

Последовательность, удовлетворяющую условию хаотичности для семейства монотонных языков, то есть условию

$$\exists c \ \forall n \ \text{КР}(a_1, a_2, \dots, a_n) > n - c$$

условимся называть просто *хаотической*. Класс всех хаотических последовательностей обозначается буквой **C**.

Считается, что строгое понятие хаотической последовательности может служить хорошим отражением интуитивного понятия случайной последовательности. К тому есть два основания.

Во-первых, для каждой отдельно взятой хаотической последовательности выполняются основные законы теории вероятностей.

Во-вторых, класс хаотических последовательностей совпадает с другим «претендентом» на роль строгого аналога расплывчатого класса случайных последовательностей — а именно, с классом **T** всех типических последовательностей (о нём будет рассказано далее):

$$\mathbf{C} = \mathbf{T}.$$

Поэтому хаотические последовательности можно было бы называть *типическо-хаотическими* или *хаотическо-типическими*, а сам класс обозначать **ТС** или **СТ**. Этот класс уже класса **S** всех последовательностей, стохастических по Колмогорову (последний, как уже отмечалось, слишком широк):

$$\mathbf{ТС} \subset \mathbf{S}, \quad \mathbf{ТС} \neq \mathbf{S}.$$

## Лицо Третье: ТИПИЧНОСТЬ

Сказать про какой-либо объект, что он типичен — это значит сказать, что он принадлежит к любому подавляющему большинству. Например, типичный человек, встреченный на московской улице, имеет рост менее двух метров (т. е. принадлежит к подавляющему большинству людей, имеющих рост менее двух метров), возраст более трёх лет (т. е. принадлежит к подавляющему большинству людей, имеющих возраст более трёх лет) и т. д. Конечно, любое из большинств, о которых идёт речь, должно быть разумным: ведь какой объект ни возьми, он заведомо не принадлежит к подавляющему большинству, образованному всеми остальными, отличными от данного, объектами.

Мы исходим из интуитивного представления о том, что случайный объект должен обладать свойством типичности. Наша цель — дать строгое математическое определение этого свойства для последовательностей нулей и единиц при равномерном распределении вероятностей на пространстве  $\Omega$  таких последовательностей. Последовательно, удовлетворяющие этому строгому определению, мы будем называть *типическими*, оставляя слово *типичный* для употребления на интуитивном уровне. В силу сказанного выше, нам надо определить сперва, что есть подавляющее большинство последовательностей, а затем — какие множества следует считать разумными. А тогда класс типических последовательностей автоматически определится как пересечение всех разумных большинств,

Оставляя термин “подавляющее большинство” для интуитивного употребления, мы будем говорить о *больших* множествах последовательностей. Дополнение к большому множеству до всего  $\Omega$  будем называть *малым*. Очевидно, достаточно определить, что такое малое множество, — а тогда большие множества определятся как дополнения к малым.

Итак, пусть дано какое-то множество  $Q$  последовательностей,  $Q \subset \Omega$ . Мы хотим определить, когда его следует считать *малым*. В терминах теории вероятностей мы сказали бы, что  $Q$  малое, коль скоро вероятность попадания в  $Q$  равна нулю. В терминах теории меры мы сказали бы, что  $Q$  малое, коль скоро оно имеет меру нуль. Мы, однако, постараемся определить, что такое малое множество в более простых терминах.

Множество  $Q$  называется *малым*, если его можно покрыть шарами, сумма объёмов которых сколь угодно мала. Определению малого множества можно дать такую формулировку: множество  $Q$  называется *малым*, когда для каждого натурального числа  $m$  найдётся такая последовательность двоичных слов  $\langle x(1), x(2), \dots, x(n), \dots \rangle$ , что

$$Q \subset \bigcup_n \Omega_{x(n)};$$

$$\sum_n \mathbf{v}(\Omega_{x(n)}) = \sum_n 2^{-|x(n)|} < \frac{1}{m}.$$

Очевидно, каждая отдельная последовательность из  $\Omega$  образует малое множество, а потому понятие большого множества не может претендовать на роль уточнения понятия **разумного** большинства. Пересечение всех больших множеств пусто.

“Разумность” вводится путём следующей корректировки определения.

Во-первых, потребуем, чтобы упоминаемая в определении последовательность двоичных слов  $\langle x(1), x(2), \dots, x(n), \dots \rangle$  была вычислимой. Другими словами, мы требуем, чтобы существовал алгоритм, вычисляющий её  $n$ -ый член  $x_n$  по его номеру  $n$ .

Во-вторых, потребуем, чтобы такая вычислимая последовательность не просто существовала для каждого  $m$ , но строилась бы по этому  $m$  **эффективно**. Слово “эффективно” означает ‘с помощью алгоритма’. Здесь необходимы разъяснения. Дело в том, что алгоритм, получающий на вход  $m$  и выдающий на выходе требуемую последовательность, невозможен просто потому, что последовательность — это бесконечный объект, а алгоритмы оперируют лишь с конечными объектами. Однако в нашем случае, поскольку последовательность вычислима, то у неё есть алгоритм, который её вычисляет — даже очень много таких алгоритмов. Алгоритмы (в другой системе терминов — программы алгоритмов) являются конечными объектами и потому вполне осмысленно говорить об алгоритме, который по числу  $m$ , поступившему на его вход, выдаёт на выходе один из алгоритмов (в другой системе терминов — одну из программ), вычисляющих последовательность  $\langle x(1), x(2), \dots, x(n), \dots \rangle$ .

Внеся в определение малого множества эти два добавления, мы получаем определение *эффективно малого* множества — а тем самым и определение *эффективно большого* множества. Пересечение всех эффективно больших множеств оказывается непустым. Более того, оно само является эффективно большим. Это наименьшее среди эффективно больших множеств и есть искомое множество **T** всех *типических* последовательностей.

Типические последовательности называют также *случайными по Мартин-Лёфу* — по имени ученика Колмогорова, замечательного шведского математика Пера Мартин-Лёфа, который в 1966 г. сформулировал только что изложенное определение случайности.

Как уже говорилось, класс **T** всех типических последовательностей совпадает с классом **S** всех хаотических последовательностей:

$$\mathbf{T} = \mathbf{S}.$$

Поэтому типические последовательности можно именовать также *хаотическо-типическими* или *типическо-хаотическими*, а сам класс всех таких последовательностей обозначать **СТ** или **ТС**.

Как мы уже знаем,

$$\mathbf{CT} \subset \mathbf{S}, \quad \mathbf{CT} \neq \mathbf{S}.$$

## Лицо Четвёртое: НЕПРЕДСКАЗУЕМОСТЬ

Вот начало «Двенадцати стульев» Ильфа и Петрова: “В уездном городе N было так много парикмахерских заведений и бюро похоронных процессий, что казалось, жители города рождаются лишь затем, чтобы побриться, остричься, освежить голову вежеталем и умереть”. Суждение остаётся верным, если заменить N на Москву, парикмахерские на залы игровых автоматов, похоронные бюро на казино, а цель рождения на “играть и умереть”. Этот печальный факт, однако, позволяет наглядно объяснить то свойство случайных последовательностей, которое мы называем непредсказуемостью.

Интуитивно ясно, что всякая случайная последовательность является *непредсказуемой* в том смысле, что в каком бы порядке мы не выбирали её члены, знание значений уже выбранных членов не позволяет предсказать значение того следующего члена, который мы намереваемся выбрать. Таким образом, Казино, обладающее такой последовательностью и предлагающее Игроку угадывать её члены и делать при этом денежные ставки, не разорится; говоря более точно, Казино уверено, что никакой Игрок не может обладать такой стратегией игры, которая приведёт к разорению Казино.

Итак, представим себе что Игрок приходит в Казино. Каждая из сторон — и Казино, и Игрок — обладает своим начальным капиталом. Казино располагает некоторой фиксированной, но неизвестной Игроку последовательностью нулей и единиц и предлагает Игроку предсказывать её члены — необязательно в монотонном порядке их следования и даже необязательно все её члены.

Для наглядности представим себе, что члены последовательности написаны на картах, которые лежат рубашками вверх, так что Игрок не видит, что там написано. Последовательность предстает перед Игроком в виде бесконечного ряда таких карт. Игра состоит в том, что Игрок на своём ходу указывает ту карту, которая должна быть открыта, одновременно предсказывая значение, которое обнаружится на этой карте, и объявляя размер денежной ставки. Если предсказание окажется правильным, Казино выплачивает Игроку сумму ставки, если неправильным — Игрок выплачивает эту сумму Казино. Считается, что Игрок выиграл, если он сумел разорить Казино. Разумеется, если Игроку открыт кредит, он всегда может разорить Казино, удваивая ставки. Но Игра идёт на наличные, так что величина ставки ограничена текущим капиталом Игрока.

Последовательность называется *предсказуемой*, если Игрок обладает алгоритмом, позволяющим ему разорить Казино, каким бы начальным капиталом это Казино ни обладало. Последовательность называется *непредсказуемой*, если она не является предсказуемой.

На математическом языке ситуация описывается так.

Рассматривается бесконечная последовательность нулей и единиц

$$\mathbf{a} = \langle a_1, a_2, a_3, \dots \rangle.$$

При каждом ходе Игрока возникает тройка чисел

$$\langle n, i, v \rangle,$$

где

$$n \in \mathbf{N}, \quad i \in \{0, 1\}, \quad v \in \mathbf{Q}, \quad v \geq 0.$$

Содержательно: натуральное число  $n$  есть номер того члена последовательности, на который делается ставка;  $i$  есть предсказываемое значение этого члена; неотрицательное рациональное число  $v$  есть размер ставки. Ходы делаются друг за другом, начиная с первого; тройка, возникающая на  $k$ -том ходу, обозначается  $\langle n(k), i(k), v(k) \rangle$ . Более математически грамотно было бы сказать, что каждый ход **есть** тройка чисел, и что ходы не делаются, а **предъявляются**.

Капитал Игрока перед  $k$ -тым ходом обозначается  $V(k-1)$ ; без ограничения общности можно считать, что начальный капитал Игрока  $V(0)$  равен единице.

После каждого хода капитал Игрока меняется (если только ставка не была нулевой). А именно:

$$\begin{aligned} \text{если } i(k) = a_{n(k)} \text{ (Игрок угадал), то } V(k) &= V(k-1) + v(k) \\ \text{если } i(k) \neq a_{n(k)} \text{ (Игрок не угадал), то } V(k) &= V(k-1) - v(k). \end{aligned}$$

И ещё два прибавления к сказанному.

Во-первых, некоторые ходы объявляются незаконными. А именно,  $k$ -тый ход считается *незаконным* в одном из двух случаев:

1) если Игрок требует открыть карту, уже открытую ранее, т. е. если  $n(k)$  совпадает с одним из чисел  $n(1), \dots, n(k-1)$ ;

2) если делаемая Игроком ставка превышает имеющийся у него капитал, т. е. если  $v(k) > V(k-1)$ .

Как только Игрок делает незаконный ход, Игра останавливается, и Игрок остаётся при имеющемся у него к данному моменту капитале.

Во-вторых, не исключается возможность того, что Игрок вообще не делает очередного хода (даже самого первого хода!). В этом случае Игрок также остаётся при имеющемся у него к данному моменту капитале. Однако в этом случае — в отличие от случая незаконного хода — мы избегаем выражения «игра останавливается». Дело в том, что ситуацию можно наглядно представить себе следующим образом. Перед каждым своим ходом Игрок решает, делать ли ему ход и если делать, то какой именно. Решение требует обдумывания, и Игрок берёт время на обдумывание. Время обдумывания ничем не ограничено и может затянуться до бесконечности. Если Игрок за конечное время принимает решение не делать ход, то можно сказать, что игра останавливается и что капитал Игрока не меняется. Но если Игрок никогда не приходит ни к какому решению, его капитал тоже не меняется, но в этом случае не наступает момента остановки игры. Разумеется, всё сказанное носит иллюстративный характер, и математическое описание не включает в себя ссылку на такие понятия, как ‘решать’, ‘обдумывать’ и т. п.

По определению, Игрок *выигрывает*, если

$$\sup_k V(k) = \infty,$$

т. е. если

$$\forall W \exists k V(k) > W.$$

Содержательно это означает, что Игрок разоряет Казино, каким бы исходным капиталом  $W$  Казино ни обладало. Очевидно, что Игрок в состоянии выиграть лишь при условии, что всякий раз, когда ему предстоит делать ход, он его делает, причём всякий раз этот ход оказывается законным.

Как выглядит игра, мы описали. Перейдём теперь к понятию системы игры, или *стратегии*. Смысл стратегии в том, чтобы избавить Игрока от необходимости самостоятельно принимать решения: стратегия берёт эту функцию на себя. Стратегия есть правило, для каждого шага указывающее Игроку, какой на этом шагу он должен сделать ход (то есть какую тройку предьявить). Разумеется, стратегия выдаёт такое указание лишь в том случае, если ход должен быть сделан. Выше уже отмечалась возможность того, что никакого хода не делается; в этом случае, естественно, стратегия не выдаёт никакого указания. При указании хода стратегия опирается на всю предшествующую историю игры. История же игры состоит из всех уже сделанных к рассматриваемому моменту ходов и из всех ставших уже известными элементов последовательности. Мы лишь потому не включаем в историю игры информацию о капитале Игрока на каждый момент, что эта информация легко вычисляется из только что перечисленных сведений.

Таким образом, историю игры перед  $k$ -тым ходом можно записать в виде таблицы

$$\begin{array}{cccc} n(1) & n(2) & \dots & n(k-1) \\ i(1) & i(2) & \dots & i(k-1) \\ v(1) & v(2) & \dots & v(k-1) \\ a_{n(1)} & a_{n(2)} & \dots & a_{n(k-1)} \end{array}$$

(Если  $k = 0$ , таблица пуста.)

На математическом языке стратегия есть функция, которая каждой подобной таблице (в том числе пустой) либо ничего не ставит в соответствие, либо ставит в соответствие некоторый ход, т. е. тройку  $\langle n, i, v \rangle$ . Под “каждой подобной таблицей” мы понимаем отнюдь не только такую таблицу, которая отражает реальное течение игры, а произвольную таблицу, в которой в первой строке стоят положительные целые числа, в третьей — неотрицательные рациональные числа, а во второй и четвёртой — нули и единицы.

Если таблица реально встретилась в процессе игры (в качестве истории игры на каком-то этапе) и если задана стратегия, то первые три строки таблицы однозначно восстанавливаются по её четвёртой строке. В самом деле, применение стратегии к пустой таблице даёт нам первый ход  $\langle n(1), i(1), v(1) \rangle$ . Тем самым — поскольку четвёртая строка предполагается известной — мы получаем историю игры перед вторым ходом в виде таблицы

$$\begin{array}{c} n(1) \\ i(1) \\ v(1) \\ a_{n(1)} \end{array}$$

Теперь к этой таблице снова применяем стратегию, получаем второй ход  $\langle n(2), i(2), v(2) \rangle$  и таблицу

$$\begin{array}{cc} n(1) & n(2) \\ i(1) & i(2) \\ v(1) & v(2) \\ a_{n(1)} & a_{n(2)} \end{array}$$

И так далее.

Сказанное даёт нам право при определении стратегии брать в качестве аргумента не всю таблицу в целом, а лишь её четвёртую строку. Заметим, что в этой четвёртой строке стоит двоичное слово, т. е. элемент множества  $\Xi$ . Стратегия должна, имея этот элемент, или не выдавать ничего, или выдавать ход, который есть тройка, то есть элемент декартова произведения  $\mathbf{N} \times \{0, 1\} \times \mathbf{Q}^{\geq 0}$ . Здесь символом  $\mathbf{Q}^{\geq 0}$  обозначено множество всех неотрицательных рациональных чисел.

Мы приходим к окончательному определению понятия стратегии: *стратегия* есть отображение некоторого подмножества множества  $\Xi$  всех двоичных слов в декартово произведение  $\mathbf{N} \times \{0, 1\} \times \mathbf{Q}^{\geq 0}$ :

$$\Xi \longrightarrow \mathbf{N} \times \{0, 1\} \times \mathbf{Q}^{\geq 0}.$$

Для наших целей особый интерес представляет случай, когда указанное отображение задаётся каким-то алгоритмом. Поясним, что это значит. Пусть  $\mathbf{A}$  — алгоритм, на вход которого могут подаваться элементы из множества  $X$ , а на выходе получают элементы из множества  $Y$ . В множестве  $X$  выделяется *область результативности* алгоритма  $\mathbf{A}$ , состоящая из тех и только тех элементов, в применении к которым  $\mathbf{A}$  даёт результат. Алгоритм  $\mathbf{A}$  задаёт следующее отображение: область определения этого отображения совпадает с областью результативности алгоритма, и для каждого элемента этой области значение отображения на этом элементе совпадает с тем результатом, который получается при применении алгоритма к этому элементу.

Если стратегия задана каким-то алгоритмом, она называется *вычислимой*. (Если подать на вход задающего стратегию алгоритма такую историю игры, для которой очередной ход не определён, то алгоритм либо выдает сообщение об отсутствии результата, либо работает на этом входе бесконечно долго, не приходя ни к какому результату, но и не выдавая сообщение об отсутствии такового; эта бесконечная работа алгоритма и представляет собою то бесконечное время, взятое Игроком на обдумывание, о котором говорилось выше.)

Стратегия называется *выигрышной*, если при её использовании Игрок выигрывает.

Последовательность называется *предсказуемой*, если для неё существует выигрышная вычислимая стратегия, и *непредсказуемой* в противном случае.

Класс всех непредсказуемых последовательностей обозначается буквой  $\mathbf{U}$ .

Известно, что всякая непредсказуемая последовательность стохастична по Колмогорову (принадлежит классу  $\mathbf{S}$ ) и что всякая типическо-хаотическая последовательность (последовательность класса  $\mathbf{CT}$ ) непредсказуема:

$$\mathbf{CT} \subset \mathbf{U} \subset \mathbf{S}.$$

Известно также, что класс стохастических по Колмогорову последовательностей существенно шире класса непредсказуемых:

$$\mathbf{S} \neq \mathbf{U}.$$

Открытым остаётся вопрос о совпадении классов хаотических и непредсказуемых последовательностей:

$$\mathbf{CT} = \mathbf{U} ??$$

Эта важная проблема ждёт своего решения.

## ИСТОРИЯ И БИБЛИОГРАФИЯ

[КУ] А. Н. Колмогоров, В. А. Успенский. **Алгоритмы и случайность** // *Теория вероятностей и её применения*, 1987, т. 32, вып. 3, с. 425–455.

[УС] В. А. Успенский, А. Л. Семёнов. **Теория алгоритмов: основные открытия и приложения**. — М.: Физматлит, 1987. — 288 с.

[УСШ] В. А. Успенский, А. Л. Семенов, А. Х. Шень. **Может ли индивидуальная последовательность нулей и единиц быть случайной?** // *Успехи математических наук*, 1990, т. 45, вып. 1, с. 105–162.

[MSU] An. A. Mucnik, A. L. Semenov, V. A. Uspensky. **Mathematical metaphysics of randomness** // *Theoretical Computer Science*, 1998, vol. 207, pp. 263–317.

[Вь] В. В. Вьюгин. **Алгоритмическая энтропия (сложность) конечных объектов и её применение к определению случайности и количества информации** // *Семиотика и информатика. Выпуск 16*. — М: ВИНТИ, 1981. — С. 14–43.

Разумеется, этот список из пяти названий никоим образом не претендует на полноту. Однако в этих публикациях можно найти дальнейшие ссылки, и список этих ссылок уже претендует если не на полноту, то на некоторое приближение к ней.

В книге [УС] следует обратить внимание на § 2.6 «Приложения к теории вероятностей: определения случайной последовательности». Следует также иметь в виду, что терминология этой книги несколько архаична: хаотические последовательности называются там *случайными по Колмогорову*, а типические — *случайными по Мартин-Лёфу*. Частотоустойчивые, они же стохастические, последовательности называются там *случайными по Мизесу*, стохастические по Чёрчу — *случайными по Мизесу-Чёрчу*, стохастические по Колмогорову — *случайными по Мизесу-Колмогорову-Лавлэнду* (а в [УСШ] — *стохастическими по Колмогорову-Лавлэнду*; Д. Лавлэнд (D. Loveland) открыл эти последовательности независимо от Колмогорова, хотя и позже: соответствующая статья Колмогорова была опубликована в 1963 г., статья Лавлэнда — в 1966 г.) Что касается непредсказуемых последовательностей — в том точном понимании, как было сформулировано выше, — то они в указанной книге отсутствуют, поскольку появились в литературе лишь в 1998 г. в статье [MSU].

Предложение измерять сложность объекта длиной его кратчайшего описания и понятие оптимального языка были сформулированы Колмогоровым в статье 1965 г.; за год до того сходные идеи были опубликованы американским исследователем Рэем Соломоновым (Ray Solomonoff), о работах которого Колмогоров узнал лишь позже. Ввиду этого теорему о существовании оптимального языка мы называем Теоремой Соломонова-Колмогорова. В эти же годы (т. е. в середине 60-х годов XX в.) Колмогоров высказывает на своих семинарах предположение о том, что быстрота роста энтропии начальных отрезков последовательности может служить критерием случайности рассматриваемой последовательности. Однако введённое им в рассмотрение языковое семейство приводило к энтропии, непригодной для поставленной цели. Как уже говорилось выше, в главке о хаотичности, годные для этой цели семейства нашёл Леонид Левин соответственно в 1973 г. (монотонные языки) и в 1974 г. (префиксные языки).

Типические последовательности (под названием случайных, random) были, как уже отмечалось в главке о типичности, открыты в 1966 г. Пером Мартин-Лёфом (Per

Martin-Löf).

Утверждение о том, что класс последовательностей, стохастических по Колмогорову, шире класса типических (они же — хаотические) последовательности доказал Александр Шень (доказательство опубликовано в 1988 г.). При доказательстве он использовал последовательность, построенную для других целей голландским математиком ван Ламбальгеном, и потому и своё доказательство иногда приписывает Ламбальгеному. Здесь в Шене как бы проявляется ген его научного прадеда Н. Н. Лузина, относительно которого великий французский математик Анри Лебег писал: “Г-н Лузин лишь тогда бывает совершенно счастлив, когда ему удаётся приписать собственные открытия кому-либо другому”.

Пусть  $K$  — энтропия в каком-то из вариантов этого понятия (В литературе вопроса исследованы не менее шести таких вариантов: простая, априорная, монотонная, процессная, префиксная и энтропия разрешения; все эти варианты различны в том точном смысле, что разность энтропий, принадлежащих любым двум из перечисленных вариантов, неограниченна.) Скажем, что последовательность  $a_1 a_2 a_3 \dots a_n \dots$  является *хаотической относительно  $K$* , если

$$\exists c \forall n K(a_1 a_2 a_3 \dots a_n) > n - c.$$

(Заметим для ясности, что ни для простой энтропии, ни для энтропии разрешения последовательностей с таких последовательностей нет вовсе.) В той же статье Левина 1973 г., в которой была введена монотонная энтропия, содержалось и доказательство того факта, что понятие хаотичности относительно монотонной энтропии равнообъёмно понятию типичности. Независимо от Левина в том же 1973 г. К. П. Шнорр (С. Р. Schnorr) ввел в рассмотрение свой вариант энтропии, так называемую *процессную энтропию* (у Шнорра — *process complexity*), и доказал (независимо от Левина, но очень похожим способом), что понятие хаотичности относительно процессной энтропии также равнообъёмно понятию типичности. Хотя процессная энтропия и монотонная энтропия существенно различаются (как показал В. В. Вьюгин [Вь, с. 35, строки 6–4 снизу], их разность не ограничена никакой константой) и хотя впоследствии сам Шнорр перестал использовать свою процессную энтропию, вышеуказанную *Теорему Левина* иногда называют *Теоремой Левина–Шнорра*.

Префиксную энтропию ввёл в 1974 г. Левин и годом позже, но, надо полагать, независимо от Левина, Чэйтин (Gregory J. Chaitin) (см. *J. Assoc. Comput. Mach.*, 1975, v. 22, p. 329–340). В той же статье Чэйтин ввёл понятие хаотичности относительно префиксной энтропии и объявил, без доказательства, что этот вариант хаотичности равносильен типичности; первое опубликованное доказательство этого факта появилось в статье В. В. Вьюгина [Вь]: следствие 3.2 на с. 38. *Префиксная энтропия* — это, по определению, энтропия для семейства префиксных языков. Язык  $E$  называется *префиксным*, если он перечислим и выполнено условие:

$$[\langle x_1, y_1 \rangle \in E \ \& \ \langle x_2, y_2 \rangle \in E \ \& \ (x_1 \approx x_2)] \implies [y_1 = y_2].$$

Заметим ещё, что в литературе термин «сложность» («complexity») часто употребляется в смысле ‘энтропия’, то есть в смысле ‘сложность относительно оптимального языка’.

Непредсказуемые последовательности впервые возникли весной 1991 г. в совместном докладе под названием «Randomness and Lawlessness» («Случайность и Беззаконность»), который Ан. А. Мучник, А. Л. Семёнов и В. А. Успенский сделали на конференции в Калифорнии. Конференция была посвящена основаниям теории случайности и проходила с 4 по 7 марта в Станфордском университете. Содержание доклада было опубликовано в 1998 г. в виде статьи [MSU]. В этой статье приведены, в частности, теоремы о соотношении непредсказуемости с другими алгоритмическими лицами случайности. Сама идея о связи случайности с невозможностью гарантированного выигрыша достаточно очевидна и в тех или иных вариантах высказывалась и раньше, однако формулировка из [MSU] кажется более близкой к интуитивному представлению о случайности. В частности, предшествующие игровые определения заведомо не приводили к классу последовательностей, равнообъёмному классу типическо-хаотических последовательностей; для определения непредсказуемости из [MSU] этот вопрос, как уже отмечалось, открыт. Ясно, что чем бóльшим количеством лиц случайности характеризуется какой-то точно очерченный класс последовательностей, тем больше надежда, что этот класс может служить формальным аналогом расплывчатого интуитивного представления о случайности.