Обзор результатов А. Л. Семёнова и Ан. А. Мучника, относящиеся к колмогоровской теории сложности

I

В первой части обзора представлены результаты, относящиеся к проблемам, предложенным академиком Андреем Николаевичем Колмогоровым.

В статье «On tables of random numbers», опубликованной в 1963 году (см. [1]), Колмогоров ставит цель перенести с бесконечных последовательностей на достаточно длинные конечные последовательности систему определений, предназначенных для интерпретации понятия вероятности на языке частот.

Для бесконечных последовательностей первый вариант точного определения (основанный на понятии алгоритма) был предложен Чёрчем в 1940 году (см. [5]). В определении Чёрча цифры двоичной последовательности поступают на вход алгоритма (называемого *правилом выбора*) в том же порядке, в котором они расположены в последовательности. Перед поступлением очередной цифры алгоритм решает (используя информацию о предшествующих цифрах), должна ли очередная цифра быть включена в подпоследовательность, формируемую на выходе алгоритма. Пусть про бесконечную входную последовательность T предполагается, что её цифры порождаются независимо с вероятностями p для единицы и $1-p$ для нуля. Тогда естественно ожидать следующего:

> для каждого правила выбора, которое получив на вход T, на выходе формирует бесконечную подпоследовательность S, закон больших чисел выполняется в S (то есть доля единиц среди первых n цифр S стремится к p при $n \to \infty$). $(*)$

Достаточно просто доказывается, что вероятность нарушения свойства $(*)$ равна нулю. Таким образом получается точное описание некоторого класса бесконечных двоичных последовательностей, которые можно было бы назвать случайными относительно меры Бернулли с параметром p.

Вернёмся к упомянутой статье Колмогорова [1].

В ней, *во-первых*, Колмогоров обобщил понятие правила выбора. Для этого он рассмотрел алгоритмы, которые сами определяют номер во входной последовательности очередной выбираемой цифры. При этом

алгоритмы используют информацию о цифрах входной последовательности, выбранных ранее. Номер (во входной последовательности) следующей выбираемой цифры может быть как больше, так и меньше номера предыдущей цифры, но требуется, чтобы один номер не выбирался дважды в процессе работы алгоритма. Как и в определении Чёрча, перед поступлением очередной цифры алгоритм принимает решение, должна ли она быть включена в подпоследовательность, формируемую на выходе. Если правила выбора понимать по Колмогорову, то по-прежнему вероятность нарушения свойства $(*)$ равна нулю.

Во-вторых, Колмогоров рассмотрел не бесконечные, а конечные входные последовательности. В этом случае правило выбора может задаваться не программой алгоритма, а некоторой функцией. Её область определения и область значений эффективно указываются по длине входной последовательности. Тем самым все далее рассматриваемые вопросы о конечных последовательностях могут быть переведены с языка теории алгоритмов на язык комбинаторики.

Основные результаты статьи Колмогорова [1] состоят в следующем.

Пусть T — двоичная последовательность длины N; R — правило выбора (в смысле Колмогорова) на последовательностях длины N; A — множество номеров цифр из T, которые оказались в выходной подпоследовательности при применении R к T (пишем $A = R(T) \subset \{1, \ldots, N\}$). Рассмотрим некоторое множество правил выбора на последовательностях длины N и обозначим его $\mathcal{R}_N$. Двоичную последовательность $T = \langle t_1, \ldots, t_n \rangle$ Колмогоров называет (n, ε, p)-случайной относительно $\mathcal{R}_N$, когда для каждого правила $R \in \mathcal{R}_N$ выполнено

$$|R(T)| \geqslant n \;\Rightarrow\; \left| \frac{1}{|R(T)|} \sum_{k \in R(T)} t_k - p \right| \leqslant \varepsilon.$$

Основной вопрос, который интересовал Колмогорова: сколь велика может быть мощность $\mathcal{R}_N$ для того, чтобы гарантированно существовала (n, ε, p)-случайная относительно $\mathcal{R}_N$ двоичная последовательность T длины N (Колмогоров использует в отношении T термин «таблица случайных чисел»). Обозначим через $\tau(n, \varepsilon, N, p)$ верхнюю грань таких ρ, что для каждого множества правил $\mathcal{R}_N$ мощности ρ существует хотя бы одна таблица случайных чисел. Положим

$$l(n, \varepsilon) = \inf_{p,N} \log_2 \tau(n, \varepsilon, N, p).$$

Колмогоров доказал, что при достаточно малых $\varepsilon > 0$ и n, достаточно больших относительно $1/\varepsilon$, выполнено

$$2(\log_2 e)\, n\varepsilon^2(1 - \varepsilon) - 1 \leqslant l(n, \varepsilon) \leqslant 4n\varepsilon(1 + O(\varepsilon)).$$

Колмогоров пишет в [1], что ему не удалось избавиться от различия в показателях степеней, с которыми ε входит в нижнюю и верхние оценки на $l(n, \varepsilon)$. Через 20 лет в предисловии к переводу [1] на русский язык, опубликованному в [2], Колмогоров напоминает, что указанная проблема ждёт своего решения. Проблема оставалась открытой почти 40 лет.

Решение проблемы Колмогорова было найдено Семёновым и Мучником и опубликовано в [9, 10] ([10] — специальный выпуск журнала «Проблемы передачи информации» к 100-летию А. Н. Колмогорова). Оказалось, что нижняя оценка, полученная Колмогоровым, практически точна. А именно, при достаточно малых $\varepsilon > 0$ и n, достаточно больших относительно $1/\varepsilon$, выполнено

$$l(n, \varepsilon) \leqslant 2(\log_2 e)\, n\varepsilon^2(1 + 2\varepsilon)$$

(доказательство в теореме 3 из [9, 10]). Отметим, что приведённая верхняя оценка на $l(n, \varepsilon)$ достигается, например, когда вероятность появления единицы $p = 1/2$ и все правила из $\mathcal{R}_N$ неадаптивны. Класс неадаптивных правил выбора является даже более узким, чем у Чёрча. В неадаптивных правилах выбора множество номеров цифр входной последовательности, которые включаются в выходную подпоследовательность, указывается заранее (до начала просмотра входной последовательности).

II

Во второй части настоящего обзора рассматривается ряд других результатов авторов, связанных с работами А. Н. Колмогорова.

Сначала опишем, как проблема Колмогорова и некоторые разбираемые далее задачи вкладываются в контекст комбинаторной математики, где исследуются такие вопросы, как «*Чему равен минимальный диаметр шара, которым можно покрыть произвольное множество диаметра* 1?», «*Каково минимальное количество шаров диаметра* 1, *которыми можно покрыть произвольное множество диаметра* 1?».

В качестве объемлющего пространства будет выступать множество всех двоичных последовательностей длины N с заданным на нём равномерным распределением вероятностей. Следует выбрать удобную шкалу для измерения подмножеств этого пространства. (Подходящая шкала может делать математические утверждения более наглядными, а иногда и более информативными. Например, в центральной предельной теореме естественная единица измерения для отклонения частоты от вероятности — это $1/\sqrt{N}$.)

Удельным дефектом множества U, состоящего из последовательностей длины N, называется величина $1-\frac{\log_2|U|}{N}$. (Другими словами, удельный дефект множества равен δ, когда его мощность равна $2^{N(1-\delta)}$.) Заметим, что эта же единица измерения используется во многих результатах теории кодирования.

Каждому правилу выбора R на последовательностях длины N, натуральному числу n и числу $\varepsilon > 0$ сопоставляется подмножество $\{0,1\}^N$, называемое *частотным тестом*. Оно состоит из тех двоичных последовательностей T длины N, для которых $|R(T)| \geqslant n$ и доля единиц в выходной подпоследовательности отличается от $1/2$ не менее чем на ε. (Если числа n и ε фиксированы, то соответствующие частотные тесты будем называть (n,ε)-*тестами*.)

Проблема Колмогорова на языке частотных тестов формулируется так: *каким количеством (n,ε)-тестов можно покрыть множество всех двоичных последовательностей длины N?*

Колмогоров доказал, что при естественных ограничениях на N, ε, n *множество $\{0,1\}^N$ невозможно покрыть (n,ε)-тестами в количестве меньше $e^{2n\varepsilon^2(1-\varepsilon)}$.*

Семёнов и Мучник доказали, что при естественных ограничениях на N, ε, n *множество $\{0,1\}^N$ можно покрыть (n,ε)-тестами в количестве меньше $e^{2n\varepsilon^2\frac{1+\varepsilon}{1-n/(N-1)}}$.*

Перейдём к задаче покрытия частотными тестами произвольного множества $U \subset \{0,1\}^N$, имеющего удельный дефект больше δ. Семёнов и Мучник в [10, теоремы 4, 5] построили для каждого достаточно малого положительного δ натуральное число $\rho(\delta)$, для которого выполнено следующее. Пусть N достаточно велико относительно $1/\delta$.

I. Для любого $U \subset \{0,1\}^N$ с удельным дефектом больше δ существует $\rho(\delta)$ частотных тестов, объединение которых покрывает U и имеет удельный дефект больше $\dfrac{\delta}{2\ln(1/\delta)}$.

II. Для некоторого $U \subset \{0,1\}^N$ с удельным дефектом больше δ всякое покрытие множества U частотными тестами, удельный дефект которых больше $\dfrac{2\delta}{\ln(1/\delta)}$, состоит более чем из $e^{N/\rho(\delta)}$ тестов.

В 1965 году в первом выпуске журнала «Проблемы передачи информации» Колмогоров опубликовал знаменитую статью [3], в которой дал определения, связанные со сложностью описания двоичной последовательности. Для конечных двоичных последовательностей y, z сложность z при известном y принято обозначать $K(z \mid y)$. Пусть $0 < \delta < 1$,

y — двоичная последовательность, натуральное число N достаточно велико и множество U состоит из тех двоичных последовательностей z длины N, для которых $K(z \mid y) < N(1-\delta)$. Тогда удельный дефект U приблизительно равен δ. Это ключевое обстоятельство позволяет сформулировать аналоги вышеупомянутых результатов на языке колмогоровской сложности (см. [3, теоремы $1'$, $3'$, $4'$, $5'$]). При доказательстве некоторых из этих аналогов возникают дополнительные трудности, что можно усмотреть из следующего обстоятельства. *Все вышеприведённые комбинаторные результаты в равной степени верны для правил выбора в смысле Чёрча и в смысле Колмогорова. Некоторые их алгоритмические аналоги верны для правил выбора в смысле Колмогорова, но не в смысле Чёрча* (*доказано Семёновым и Мучником в* [3, теоремы 4, $4'$, $6'$]). Таким образом, уже в 1962 году Колмогоров замечательно предвидел, что его обобщение правил выбора Чёрча в дальнейшем окажется необходимым.

Обратимся к рассмотрению бесконечных двоичных последовательностей. Назовём бесконечные двоичные последовательности, случайные относительно всех правил выбора в смысле Чёрча, стохастическими по Чёрчу. Назовём бесконечные двоичные последовательности, случайные относительно всех правил выбора в смысле Колмогорова, стохастическими по Колмогорову. В 1969 году в статье [4] Колмогоров доказал, что стохастическая по Чёрчу последовательность может быть очень неслучайной относительно колмогоровской сложности; а именно, сложность её начал длины N может расти медленнее $\log^2 N$. В той же работе Колмогоров выдвинул гипотезу, что тем же свойством обладают некоторые последовательности, стохастические по Колмогорову. Эта гипотеза была опровергнута Мучником. Пусть $\delta > 0$. Оказалось, что последовательность, сложность начал которой длины N растёт медленнее $N(1-\delta)$, не может быть стохастической по Колмогорову. Наиболее подробное изложение доказательства опубликовано в [8].

Интересно отметить ещё два результата, доложенных на Колмогоровском семинаре и опубликованных в [6, 7]. Они относятся к конечным последовательностям. Начнём с неформальных мотивировок. Двоичное слово можно представлять как код некоторой информации. Тогда совокупность всех двоичных слов образует что-то вроде верхней полурешётки. Скажем, что информация слова z включена в информацию слова y, если $K(z \mid y) \approx 0$ (пишем $z \preceq y$). Очевидно, отношение $\preceq$ является частичным порядком и у каждых двух слов есть точная верхняя грань, а именно, код их пары. Возникают естественные вопросы:

- Можно ли в этой полурешётке задать пересечение?
- Можно ли в этой полурешётке задать разность?

В [6] показано, что ответ на первый вопрос отрицателен; в [7] показано, что ответ на второй вопрос положителен. Приведём точные утверждения.

Для любого достаточно большого числа m существуют такие слова x_1, x_2, что $K(x_1) > m$, $K(x_2) > m$, $K(x_1, x_2) < \frac{4}{3}m$ и для всякого слова w, если $K(w \mid x_1) < \frac{1}{3}m$, $K(w \mid x_2) < \frac{1}{3}m$, то $K(w) < \frac{1}{2}m$.

То есть размер «общей информации, содержащейся в x_1 и x_2» больше $\frac{2}{3}m - o(m)$, но эту общую информацию невозможно представить никаким двоичным словом w (доказательство в [6]).

Для любых слов u, v существует такое слово w, что

$$K(w) < K(u \mid v),$$
$$K(w \mid u) < O(\log K(u)),$$
$$K(u \mid v, w) < O(\log K(u)).$$

То есть w можно рассматривать как разность u и v. В [7] доказано, что эта разность определена неоднозначно.

Литература

[1] K o l m o g o r o v A. N. On tables of random numbers // The Indian Journal of Statistics. Series A. 1963. V. 25, part 4. (Reprinted in Theoretical Computer Science. 1998. V. 207. P. 387–395.)

[2] К о л м о г о р о в А. Н. О таблицах случайных чисел // Семиотика и информатика. 1982. Вып. 18. С. 3–13.

[3] К о л м о г о р о в А. Н. Три подхода к определению понятия «количество информации» // Проблемы передачи информации. 1965. Т. 1, № 1. С. 3–11.

[4] К о л м о г о р о в А. Н. К логическим основам теории информации и теории вероятностей // Проблемы передачи информации. 1969. Т. 5, № 3. С. 1–4.

[5] C h u r c h. On the concept of a random sequence // Bull. Amer. Math. Soc. 1940. V. 46. P. 130–135.

[6] M u c h n i k An. A. On common information // Theoretical Computer Science. 1998. V. 207. P. 319–328.

[7] M u c h n i k An. A. Conditional complexity and codes // Theoretical Computer Science. 2002. V. 271, № 1–2. P. 97–109.

[8] M u c h n i k An. A., S e m e n o v A. L., U s p e n s k y V. A. Mathematical Metaphysics of Randomness // Theoretical Computer Science. 1998. V. 207. P. 263–317.

[9] С е м ё н о в А. Л., М у ч н и к Ан. А. Об уточнении оценок Колмогорова, относящихся к датчикам случайных чисел и сложностному определению случайности // Доклады Академии Наук. 2003. Т. 391. С. 1–3.

[10] М у ч н и к Ан. А., С е м ё н о в А. Л. О роли закона больших чисел в теории случайности // Проблемы передачи информации. 2003. Т. 39, вып. 1. С. 134–165.