Информационные процессы, Том 10, № 2, 2010, стр. 145–158.

Становление алгоритмической теории информации в России[1]

В.А. Успенский*, **В.В. Вьюгин****

**Московский государственный университет им. М.В. Ломоносова, механико-математический факультет, Москва, Россия*

***Институт проблем передачи информации им. А.А.Харкевича,Российская академия наук, Москва, Россия*

Поступила в редколлегию 28.04.2010

В начале 60-х гг. Колмогоров начал искать пути построения теории информации и теории вероятностей на принципиально новой, алгоритмической основе. Первой публикацией по алгоритмической теории информации является всемирно известная статья А.Н. Колмогорова «Три подхода к определению понятия «количество информации» [16], вышедшая в первом выпуске первого тома журнала «Проблемы передачи информации», который издается ИППИ им. А.А.Харкевича РАН с 1965г. В этой статье А.Н. Колмогоров указал способ измерения сложности конечного объекта (слова), для чего он ввел понятие, называемое сейчас колмогоровской сложностью (см. также [17], [18]). Свое новое понятие он применил для построения алгоритмического варианта теории информации, позволяющего измерять информацию в конечной строке знаков. Алгоритмическая теория информации, является естественным обобщением шенноновской вероятностной теории информации на конечные (дискретные) объекты. В области теории вероятностей целью Колмогорова было дать определение для понятия индивидуального случайного объекта (в классической теории вероятностей такое понятие отсутствует, хотя оно весьма важно для практических приложений). Первые попытки такого рода были начаты в 1919г. фон Мизесом [24], предложившим определение индивидуальной случайной последовательности на основе представлений об устойчивости частот, причем эта устойчивость должна была иметь место не только в самой последовательности, но и в любой ее «законной» подпоследовательности. Однако формулировки Мизеса относительно того, какие подпоследовательности надлежит считать «законными», были весьма расплывчаты. Их уточнение на основе понятия алгоритма предложил Черч в 1940г. [37]; формулировки Черча были точными, но приводили к недопустимо широкому понятию случайности. В области теории информации целью Колмогорова было предложить способ измерения информации в произвольной конечной строке знаков (в классической теории информации Шеннона носителями информации служат случайные величины). Основные научные события, происшедшие в 60-х и 70-х гг., в рамках программы Колмогорова таковы:

1. Развивая идеи фон Мизеса и усовершенствуя конструкцию Черча, Колмогоров в 1963г. [42] формулирует такое понятие индивидуальной случайной последовательности, которое является наиболее адекватным (т. е. наиболее близким к интуитивному понятию случайности) из тех, какие могут быть даны на основе представлений об устойчивости частот (1963г.); впоследствии такие последовательности были названы им стохастическими.

2. В 1966г. ученик Колмогорова шведский математик П. Мартин-Леф [46] (в 1964–1965гг. он проходил в Москве стажировку у Колмогорова) предлагает определение индивидуальной случайной последовательности на основе найденных им алгоритмических аналогов для понятий множеств нулевой и полной меры; впоследствии такие последовательности получили название типических.

[1] Работа выполнена при поддержке грантов Российского фонда фундаментальных исследований 09-07-00180-а и 09-01-00709-а

3. Колмогоров указывает способ измерения сложности объекта, для чего вводит понятие, называемое сейчас колмогоровской сложностью. Свое новое понятие он в 1965г. применяет для построения алгоритмического варианта теории информации, позволяющего измерять информацию в конечной строке знаков, а в 1969г. – для указания пути, на котором может быть найдено определение случайной последовательности в виде ее хаотичности (хаотичность состоит в том, что сложность начальных отрезков растет очень быстро) [17]. В эти годы Колмогоровым закладываются основы теории, получившей ныне название теории колмогоровской сложности.

Независимо от Колмогорова, и даже несколько ранее, идеи построения «универсального предсказателя» были предложены Р. Соломоновым. Соломонов хотел построить меру M с эффективно вычислимыми свойствами, которая бы предсказывала не хуже любой вычислимой меры P. Первоначальные идеи Соломонова были не ясны (например см. [53]), он уточнил их позже в статьях [54] и [55]. Его уточнения привели к построению универсальной предсказывающей меры, которая правда не обладала достаточными вычислимыми свойствами. Как выяснилось, построить предсказатель, который являлся бы одновременно вычислимым и универсальным невозможно; позже Л.А. Левин построил полувычислимый универсальный предсказатель. Здесь наиболее ценной является идея Соломонова об универсальности предсказателя, которая с самого начала присутствовала в его работах. Также как и колмогоровская сложность, универсальный предсказатель определялся с использованием универсальной машины Тьюринга, которая строится в теории рекурсивных функций. В этом заключается сходство подходов Колмогорова и Соломонова. Заметим, что Соломонов не рассматривал понятие алгоритмической сложности конечного объекта. Позже идея универсального предсказателя получила свое уточнение в виде понятия универсальной полумеры, введенной Л.А. Левиным в 1970г. [15] (см. ниже). Идеи универсального предсказания индивидуальной последовательности предвосхитили появившуюся позже в 1990 годах «теорию машинного обучения» (Machine Learning), которая имеет более прикладную направленность, чем алгоритмическая случайность [45].

Отметим также параллельные работы Г. Чейтина, в которых также были введены многие понятия алгоритмической сложности [32], [33], [34], [35], [36].

4. В 1973г. ученик Колмогорова Л.А. Левин [19], [20], [21], [22] открывает новую версию колмогоровской сложности, т. н. монотонную сложность, и доказывает совпадение класса хаотических (на основе этой версии) последовательностей с классом типических (т. е. случайных по Мартин-Лефу) последовательностей. Л.А. Левин также показал, что понятие перечислимой снизу полумеры играет важную роль при изучении понятия случайной последовательности. В частности, он доказал, что существует максимальная с точностью до мультипликативной константы перечислимая снизу (полувычислимая) полумера – универсальная или априорная полумера M; в терминах этой полумеры можно следующим образом дать эквивалентное определение понятия бесконечной случайной по Мартин-Лефу последовательности. Величина

$$t(\omega^n) = \frac{M(\omega^n)}{P(\omega^n)}$$

служит перечислимым снизу тестом случайности начального фрагмента $\omega^n = \omega_1 \dots \omega_n$ бесконечной последовательности $\omega = \omega_1\omega_2 \dots$ относительно вычислимой меры P. А именно, бесконечная последовательность ω случайна по Мартин–Лефу относительно вычислимой меры P тогда и только тогда, когда

$$\sup_n t(\omega^n) < \infty.$$

Заметим также, что величина $t(\omega^n)$ является ограниченным снизу супермартингалом относительно P, в частности, она удовлетворяет условиям

$$\begin{aligned} t(\Lambda) &\le 1, \\ t(x0)P(0|x) + t(x1)P(1|x) &\le t(x), \end{aligned}$$

для всех x, где x – произвольная конечная последовательность, Λ – пустая последовательность.

Определения случайности с использованием теории мартингалов позже привели к новому теоретико–игровому обоснованию теории вероятностей и финансовой математики, предложенному В.Г. Вовком и Г. Шейфером [50].

Аналогичное (хотя и несколько другое) понятие монотонной сложности рассматривал К.П. Шнорр [51], [52]. Он же применил понятие вычислимого мартингала для определения несколько более широкого понятия случайной (по Шнорру) последовательности.

В настоящее время теория колмогоровской сложности представляет собою один из разделов математики, по которому издаются монографии и проводятся международные конференции. Первая обзорная статья по колмогоровской сложности и случайности была опубликована в 1971г. А.К. Звонкиным и Л.А. Левиным [15]. Эта статья содержала также множество оригинальных результатов и идей, которые до сих пор не потеряли свою актуальность. Изложение колмогоровской концепции случайности и новые результаты в области алгоритмической теории информации были представлены в в 1981г. обзоре В.В. Вьюгина [6]. Здесь впервые были приведены доказательства новых результатов Л.А. Левина, опубликованных (без доказательства) в 1970-х годах в статьях [19], [20], [21]. Через 10 лет в 1990г. был опубликован обзор В.А. Успенского, А.Л. Семенова, А. Шеня [28], в котором были отражены последние достижения в области алгоритмической теории информации. Отметим также обзор [23], брошюру В.А. Успенского [29], а также монографию В.А. Успенского и А.Л. Семенова [27].

За рубежом в 1993г. была опубликована монография Ли и Витаньи [43], которая в настоящее время играет роль энциклопедии результатов в области колмогоровской сложности.

Определение свойства типичности и определение свойства хаотичности исходили из совершенно, казалось бы, различных представлений о случайности: первое – из того, что случайная последовательность должна принадлежать любому «допустимому» множеству единичной меры; второе – из того, что чередование членов случайной последовательности не может управляться простым законом. Демонстрация равносильности этих свойств привела к убеждению в адекватности каждого из этих определений. Адекватность понятия хаотичности была подтверждена и следующим. Опираясь только на учение о колмогоровской сложности, без привлечения вероятностных соображений, для ряда законов теории вероятностей удалось доказать, что каждая отдельно взятая хаотическая последовательность подчиняется этому закону. Так, в 1986г. ученик Колмогорова В.Г. Вовк предложил «сложностное» доказательство того, что для каждой хаотической последовательности справедлив закон повторного логарифма[3]; в 1996г. деятельность по перенесению вероятностных законов на индивидуальные последовательности была продолжена В.В. Вьюгиным [10], [12], [13]. В то же время в 1982 г. А. Шень [30] обнаружил, что класс стохастических последовательностей существенно шире класса типических (он же класс хаотических) последовательностей и тем самым определение на основе устойчивости частот не может полностью отражать понятие случайности. В начале 90-х гг. Ан.А. Мучник, А.Л. Семенов и В.А. Успенский [47] предложили определение непредсказуемой последовательности, формулируемое на основе представлений об алгоритмической игровой стратегии; непредсказумость трактовалась как необходимое (и, возможно, достаточное) свойство случайности. Класс непредсказуемых последовательностей оказался промежуточным между классами стохастических и типических-хаотических последовательностей (возможно, совпадающим с этим последним классом). Эти и другие результаты в области алгоритмической теории вероятностей были отражены в двух больших статьях [47], [28].

Определение случайности по Черчу (см. выше) потому, в частности, не является адекватным интуитивному пониманию случайности, что возможна случайная по Черчу последовательность, перестающая быть таковой после вычислимой перестановки ее членов (пример Лавлэнда 1966г.) [44]. Возникает вопрос о сравнительном изучении последовательностей с точки зрения возможности алгоритмического преобразования одной последовательности в другую. Такое изучение было предпринято Вьюгиным в десятилетие с 1976 по 1985г. [5], [7]. Оказалось, в частности, что последовательности, не являющиеся случайными ни по какой вычислимой мере, могут быть, тем не менее, с вероятностью, сколь угодно близкой к единице, порождены вероятностными алгоритмами. При рассмотрении последовательностей, т. е. бесконечных строк знаков, есть надежда четко разделить их на случайные и неслучайные. И действительно, эта надежда реализуется при любом из алгоритмических подходов к определению случайности. Со словами, т. е. с конечными строками знаков, дело обстоит иначе, здесь нет четкой границы между случайными и неслучайными словами: ведь заменяя, каждый раз по одной, единицу на ноль, любое двоичное (т. е. составленное из нулей и единиц) слово, объявленное случайным, можно превратить в заведомо неслучайное слово, состоящее из одних нолей. В применении к слову, поэтому, можно говорить лишь о той или иной оценке степени его случайности. Одной из таких оценок служит введенное Колмогоровым понятие дефекта случайности данного слова: чем этот дефект меньше, тем слово «случайнее». Ограниченность дефектов случайности начальных отрезков последовательности является необходимым и достаточным условием ее хаотичности. Еще в 1986г. В.Г. Вовк обнаружил, что многие важнейшие законы теории вероятностей (такие, как закон больших чисел или закон повторного логарифма) являются устойчивыми в следующем смысле: они остаются справедливыми не только для хаотических последовательностей, но и для последовательностей с небольшим ростом дефекта случайности начальных отрезков [3]. Однако знаменитую эргодическую теорему Биркгофа в течение десяти лет не удавалось перенести на случай произвольной хаотической последовательности. Наконец, в 1996г. В.В. Вьюгин сумел осуществить этот перенос [13], [57]. И он же выяснил причины того, почему этот перенос давался с таким трудом. Оказалось, во-первых, что эргодическая теорема неустойчива в том смысле, что для индивидуальной последовательности она перестает быть верна при сколь угодно медленном росте дефекта случайности начальных отрезков этой последовательности [14]. Оказалось, во-вторых, что предельный переход, присутствующий в эргодической теореме (а также в сходной с ней квазиэргодической теореме фон Неймана) неэффективен: для некоторых вероятностных мер, стационарных и вычислимых, не существует такого алгоритма, который позволял бы по заданной степени близости к предельному значению $\epsilon > 0$, указать тот момент n_0 процесса стремления к пределу, который обеспечивает близость с точностью до ϵ при всех $n \geq n_0$.

В статье [17] А.Н. Колмогоров рассматривал также асимптотические свойства понятия стохастической последовательности. В частности, он предположил, что существуют стохастические последовательности начальные отрезки которых имеют логарифмически возрастающую сложность $K(\omega^n) = O(\log n)$. Ан.А. Мучник показал, что это не так. Он доказал более общую теорему: если бесконечная двоичная последовательность ω такова, что сложность всех ее начальных отрезков не превосходит cn, для некоторого $c < 1$, то существует правило выбора по Колмогорову, которое выбирает бесконечную подпоследовательность последовательности ω, в которой предел частоты единиц не существует.

Отметим, что для понятия стохастичности по Черчу гипотеза Колмогорова верна.

В 1991г. В.Г. Вовк предложил алгоритмический подход к асимптотической теории точечного оценивания в математической статистике [4]. В частности, ему удалось усилить результат

Ле Кама о мере множества точек сверхэффективности для оценок метода максимального правдоподобия. Ле Кам доказал, что это множество имеет меру 0. В.Г. Вовк доказал, что оно – не более чем счетно.

А.Н. Колмогоров придавал большое значение изучению понятия алгоритмической случайности конечной последовательности, состоящей из 0-й и 1-ц. При этом понятие меры не должно входить в это определение. Основная идея Колмогорова заключалась в том, чтобы выводить стохастические свойства конечной последовательности из предположения о том, что ее сложность близка к максимальному значению. Приведем интерпретацию этой идеи в следующем виде. Вместо меры граничные условия на случайность задаются в виде разбиения множества всех конечных последовательностей длины n на конечные попарно непересекающиеся подмножества. Конечная последовательность x является случайной, если $K(x|D) \approx \log |D|$, где D – тот элемент разбиения, которому принадлежит x, $x \in D$, $|D|$ – число его элементов.[1] Из теории кодирования следует, что $K(x|D) \leq \log |D| + c$ для любого $x \in D$, где c – константа. Для характеризации «степени случайности» конечной последовательности вводится дефект случайности конечной последовательности x относительно конечного множества D :

$$d(x|D) = \log |D| - K(x|D),$$

где $K(x|D)$ – условная сложность конечной последовательности x относительно конечного объекта D.

Реализацией этой идеи занялся ученик Колмогорова Е.А. Асарин. В частности, он получил оценку для устойчивости частот для последовательности и ее подпоследовательностей, включающую в себя значение дефекта случайности этой последовательности. Приведем типичный результат из работы Асарина [1], характеризующий данную идею Колмогорова.

Рассматривается равномерный случай, т.е. разбиение состоит из одного элемента – это все конечные последовательности длины n. В этом случае дефект случайности конечной последовательности имеет вид

$$d(x|n) = n - K(x|n),$$

где $K(x|n)$ – условная сложность конечной последовательности x относительно ее длины n.

Теорема Асарина такова: для каждого $\epsilon > 0$ существуют целое число N и число $0 < \mu < 1$ такие, что для каждой последовательности x длины n и правила выбора R, если дефект случайности x и сложность правила выбора R достаточно малы по сравнению с длиной выбранной подпоследовательности:[2]

$$d(x|n) + K(R|n) + 2 \log K(R|n) < \mu n_R, \tag{1}$$

то имеет место следующее неравенство для частоты единиц в выбранной подпоследовательности:

$$\left| \frac{m_R}{n_R} - \frac{1}{2} \right| < \sqrt{\frac{d(x|n) + K(R|n) + 2 \log K(R|n) + (3 + \epsilon) \log n_R}{2 n_R (1 - \epsilon) \log e}},$$

где n_R – длина выбранной (посредством правила R) подпоследовательности, причем $n_R \geq N$, m_R – число единиц в ней. Здесь правило выбора это конечный объект, поэтому можно рассмотреть его сложность – $K(R|n)$.

[1] Далее рассматриваем логарифмы по основанию 2.

[2] Или же длина выбранной подпоследовательности достаточно велика по сравнению с дефектом случайности и сложностью правила выбора.

Таким образом, если последовательность x имеет достаточно малый дефект случайности, то частота единиц в любой «достаточно длинной» ее подпоследовательности, выбранной из x с помощью «достаточно простого» правила выбора, близка к $1/2$.

Идеи Колмогорова о случайности конечных последовательностей получили дальнейшее развитие в 1980 годах. В 1982г. А.Н. Колмогоров в своем докладе на семинаре в Московском университете предложил определение (α, β)-стохастической последовательности. Заданы произвольные неотрицательные числа α и β. Конечная двоичная последовательность x называется (α, β)-стохастической, если она является «β- случайным» элементом некоторого «α-простого» множества A. Точнее, конечная двоичная последовательность x называется (α, β)-стохастической, если $x \in A$ и $K(x|A) \geq \log |A| - \beta$ для некоторого конечного множества A сложности $K(A) \leq \alpha$; здесь $|A|$ – число элементов множества A. Те последовательности, для которых это неверно, называются (α, β)-нестохастическими. Для таких последовательностей невозможно подобрать простую стохастическую модель, объясняющую их поведение.

Если x – нестохастическая последовательность, то

$$d(x|A) = \log |A| - K(x|A) > \beta$$

для любого конечного множества A такого, что $x \in A$ и $K(A) \leq \alpha$. А.Н. Колмогоров поставил вопрос о существовании (α, β)-нестохастических последовательностей. А. Шень [31] привел условия на α и β, при которых такие последовательности существуют. В.В. Вьюгин [8] получил верхние и нижние оценки оценки априорной полувычислимой меры множества (α, β)-нестохастических последовательностей длины n в зависимости от α и β.

Асимптотический вариант этих результатов выглядит следующим образом. Можно оценивать в зависимости от n априорную полувычислимую меру всех (α, β)-нестохастических последовательностей длины n, где $\alpha = \alpha(n)$ и $\beta = \beta(n)$ – некоторые функции. Тогда априорная полувычислимая мера множества всех таких последовательностей длины n убывает как $2^{-\alpha(n)}$ в случае когда $\alpha(n)$ и $\beta(n)$ – неограниченные вычислимые функции. Однако, если отказаться от требования вычислимости этих функций, то как показал В.В. Вьюгин в работе [8] для некоторых (невычислимых) неограниченных неубывающих функций $\alpha(n)$ и $\beta(n)$ можно с близкой к единице вероятностью генерировать (с помощью вероятностной машины Тьюринга) бесконечные последовательности, у которых все начальные фрагменты длины n являются $(\alpha(n), \beta(n))$-нестохастическими последовательностями. [3]

Легко видеть, что такие бесконечные последовательности не могут быть случайными по Мартин–Лефу относительно любой вычислимой меры.

Понятие колмогоровской сложности и понятие (α, β)-стохастичности по Колмогорову нашли применение в математической статистике. Они позволили предложить как новую интерпретацию, так и новое обоснование для некоторых методов из этой области. Например, задачу найти правдоподобную гипотезу, объясняющую появление как результата эксперимента, можно понимать так: требуется найти такое конечное множество A («гипотезу»), что и его сложность $K(A)$, и величина дефекта случайности x относительно малы (напомним, что относительный дефект определяется как

$$d(x|A) = \log |A| - K(x|A),$$

где $K(x|A)$ есть условная сложность x относительно A). Функция

$$\beta_x(\alpha) = \min\{\delta(x|A) : K(x|A) \leq \alpha, x \in A\},$$

[3] Любую полумеру можно естественным образом рассматривать как меру на множестве всех конечных и бесконечных последовательностей [15]. Поэтому можно говорить об априорной полувычислимой мере множества бесконечных последовательностей. Если бесконечные последовательности из какого-либо множества можно генерировать с положительной вероятностью с помощью некоторой вероятностной машины Тьюринга, то априорная полувычислимая мера этого множества является положительной. Верно и обратное утверждение.

введенная В.В. Вьюгиным в 1987г. в работах [9], [58], отражает то, насколько хорошие объяснения возможны для x, если сложность гипотезы ограничена числом α . В 2002г. Н.К. Верещагин, совместно с П. Витаньи [59], дали исчерпывающее описание поведения функций $\beta_x(\alpha)$ и $d(x|A)$. Их результаты обобщили ранние результаты В.В. Вьюгина о возможных формах функции $\beta_x(\alpha)$ при малых значениях α и А. Шеня о существовании «абсолютно неслучайных объектов» (т.е. таких, которые имеют большой дефект случайности относительно всякого содержащего их конечного множества малой сложности) [31]. Они описали также поведение функции

$$h_x(\alpha) = \min\{\log |A| : K(A) \leq \alpha, x \in A\},$$

введенной Колмогоровым в 1974г. (т. н. структурной функции Колмогорова). Обнаруженные свойства функций и позволяют обосновать такие статистические методы, как метод наибольшего правдоподобия и метод MDL (Minimum Description Length). [4]

А.Н. Колмогоров ввел в 1969г. понятие конечной двоичной бернуллиевской последовательности [17]. Л.А. Левин определил понятие конечной двоичной последовательности случайной относительно класса всех бернуллиевских мер [19]. Точнее, он доказал существование универсального алгоритмического теста $d_L(x)$ корректного относительно всех бернуллиевских мер. В 1985г. В.Г. Вовк [2] нашел связь между определением бернуллиевской последовательности по Колмогорову и определением случайности относительно класса бернуллиевских мер, введенным Л.А. Левиным. Дефект бернуллиевости конечной последовательности x по Левину $d_L(x)$ равен дефекту бернуллиевости x по Колмогорову плюс дефект случайности числа единиц k в некотором элементе разбиения всех целых чисел от 0 до n (n – длина двоичной последовательности x, k – число единиц в x) на попарно непересекающиеся интервалы. Дефект бернуллиевости x по Колмогорову определяется

$$d_K(x) = \log \binom{n}{k} - K(x|n, k).$$

Кроме этого, было замечено, что дефект бернуллиевости по Колмогорову совпадает с точностью до константы с дефектом случайности относительно класса мер инвариантных относительно перестановки элементов последовательностей x (см. [58]). Все эти результаты можно интерпретировать как алгоритмический аналог известной теоремы де Финетти о связи мер инвариантных (относительно перестановок) и бернуллиевских мер.

Одним из последних значительных достижений алгоритмической теории вероятностей является решение Ан.А. Мучником и А.Л. Семеновым проблемы, поставленной Колмогоровым в его статье 1963г. «О таблицах случайных чисел» [42]. В этой статье речь шла о концепции случайности конечной строки, или слова, на основе частотного подхода. А.Н. Колмогоров предложил частотное определение случайности, которое является наиболее адекватным (т. е. наиболее близким к интуитивному понятию случайности) из тех, какие могут быть даны на основе представлений об устойчивости частот (1963г.); впоследствии такие последовательности были названы им стохастическими.

Центральную роль в определении стохастичности по Колмогорову играли верхняя и нижняя оценки для такого количества алгоритмов выбора подслова, которое гарантирует существование слова, являющегося случайным относительно этих алгоритмов при заданных числовых значениях параметров случайности (появление этих параметров неизбежно при переходе от бесконечных последовательностей к словам). Некоторые такие оценки были предложены в

[4] Согласно устному замечанию Л.А. Левина, он также в 1970 годах изучил все возможные формы кривых типа $\beta_x(\alpha)$ и $h_x(\alpha)$; однако он не опубликовал эти результаты. А.Н. Колмогоров ввел функцию $h_x(\alpha)$ на конференции по теории информации в Таллине в 1974г. и поставил вопрос об ее изучении – см. ссылку [38].

статье, но сам Колмогоров неоднократно указывал на значительный разрыв между оценками сверху и снизу. Ан.А.Мучник и А.Л. Семенов устранили этот разрыв. Они указали также очень точные оценки, при соблюдении которых в конечной строке из устойчивости частот вытекает типичность этой строки (т. е. ее принадлежность к любому «допустимому» множеству, имеющему меру, близкую к единице). Приводим более подробное изложение этих результатов. Работы Ан.А. Мучника и А.Л. Семенова развивают идеи А.Н. Колмогорова, сформулированные в работе [42]. Данная работа является первой публикацией А.Н. Колмогорова, посвященной разработке нового подхода к обоснованию приложений теории вероятностей и математической статистики, центральными понятиями которого позже стали понятия алгоритмической сложности и случайности индивидуальной последовательности событий. В данной работе Колмогоров существенно развил частотное определение случайности Мизеса: во-первых, он рассмотрел более широкий класс, так называемых, немонотонных правил выбора; во-вторых, он сформулировал частотное определение конечной последовательности исходов. Изучению случайности конечных объектов Колмогоров придавал особое значение, считая что соответствующие комбинаторные оценки будут иметь наибольшее теоретическое и практическое значение. Конечная двоичная последовательность называется (n, ϵ, p) -случайной (таблицей), если в каждой подпоследовательности длины не менее n, выбранной с помощью правила выбора из некоторого заранее фиксированного класса Θ , выполнено условие устойчивости частот, т.е. частота единиц в ней отличается от p не более чем на ϵ . Возникает естественная задача – насколько велик может быть этот заранее заданный класс правил выбора. А.Н. Колмогоров показал, что при $|\Theta| \le \frac{1}{2}e^{2n\epsilon^2(1-\epsilon)}$ случайные таблицы существуют. Заметим, что экспоненциальная оценка мощности множества правил выбора соответствует линейной оценке для сложности его элементов. Он также привел идею доказательства оценки противоположного типа: существует такой класс правил выбора Θ, для которого $|\Theta| \le e^{4n\epsilon(1+5\epsilon)}$ и не существует $(n, \epsilon, l/2)$-случайных таблиц при $\epsilon < 1/2$. При малых ϵ разрыв между верхней и нижней оценками становится существенным, поэтому А.Н. Колмогоров уже в статье 1963 поставил вопрос о сближении этих оценок. Впоследствии в предисловии к переводу своей статьи из Sankhya в 1982г. А.Н.Колмогоров вновь указал, что данная задача все еще ждет своего решения, и наконец, полное решение этой задачи было получено в работах [26], [25]. Они показали, что первая из приведенных оценок Колмогорова практически точна. Существование соответствующего множества правил выбора было доказано с помощью вероятностного метода и потребовало от авторов проведения тонких комбинаторных оценок. Во второй части исследований Ан.А. Мучника и А.Л. Семенова установлена точная количественная связь между колмогоровским частотным определением случайности и сложностным определением, введенным А.Н. Колмогоровым в последующих работах 1965 и 1969 годов. Согласно сложностному определению, последовательность является неслучайной, если она лежит в некотором достаточно малом множестве простой колмогоровской сложности (тесте случайности). При таком определении, тест случайности может учитывать закономерности любой природы, не обязательно только условия стабильности частот. Тем не менее, авторами получен чрезвычайно интересный результат – если проверять случайность с точностью до $o(n)$ (n-длина последовательности), то частотных тестов достаточно. Точнее, любой тест на бернуллиевость общего типа может быть покрыт небольшим набором частотных тестов. Таким образом, если рассматривать значения тестов с точностью до $o(n)$, то выполнение закона больших чисел для подпоследовательностей гарантирует бернуллиевость конечной последовательности при проверке тестами произвольного типа. Таким образом, с указанной точностью, колмогоровское определение частотной случайности 1963г. адекватно описывает случайность в целом.

В теории информации центральную роль играют величины $I(x)$, $I(x|y)$, $I(x : y)$, означающие, соответственно, информацию в x, информацию в при условии y, взаимную информацию в x и y. В вероятностной, шенноновской версии в качестве x и y, т. е. носителей информации, выступают случайные величины и составленные из них кортежи (т. е. конечные наборы), а в качестве I - шенноновская энтропия H. В алгоритмической, колмогоровской версии носителями информации служат слова и составленные из них кортежи, а I интерпретируется как колмогоровская сложность, или колмогоровская энтропия, K . Возникает естественный вопрос, сохраняются ли соотношения теории при переходе от шенноновской версии к колмогоровской. Первый нетривиальный и притом утвердительный ответ на этот вопрос был независимо получен Колмогоровым и Левиным в 1967г. [15] (теорема о симметрии взаимной информации): выяснилось, что соотношение $H(x : y) = H(y : x)$ сохраняется при замене H на K и знака равенства на знак $\cong$, означающий равенство с точностью до аддитивного члена, имеющего порядок логарифма от длин участвующих слов:

$$K(x : y) = K(y : x) + O(\log \max(|\nu_1|, \ldots, |\nu_n|)),$$

где $\nu_1, \ldots, \nu_n$ суть компоненты кортежей x и y. В 1996г. А.Е. Ромащенко [41] обобщил этот результат на случай произвольных линейных равенств и неравенств, связывающих между собой величины вида $H(x)$, $H(x|y)$ и $H(x : y)$; он показал, что такое соотношение является верным в шенноновской теории тогда и только тогда, когда в колмогоровской теории верно соотношение, получающееся из исходного заменой H на K, а также $=$ на $\cong$ и $\leq$ на $\preceq$ (где $\preceq$ означает меньше или равно с точностью до аддитивного члена, имеющего порядок логарифма от длин слов, участвующих в неравенстве). Колмогоровская программа построения теории информации на алгоритмической основе получила тем самым дополнительное подкрепление. Шенноновская теория информации содержит и утверждения, не укладывающиеся в рамки равенств и неравенств. Одно из них (теорема Гача и Кернера 1973г. [40]) утверждает существование таких объектов x и y, что их взаимная информация не может быть «материализована» в виде третьего объекта. В 1986г. Ан.А. Мучник [48] доказал аналогичную теорему для колмогоровской теории: он предъявил наглядные примеры таких пар слов, для которых их взаимная информация $I(x : y)$ не может быть выделена в виде третьего слова. Основополагающая работа Колмогорова 1965г. называлась «Три подхода к определению понятия «количество информации». Двумя из этих трех подходов были уже упомянутые вероятностный (энтропия Шеннона) и алгоритмический (алгоритмическая, или колмогоровская, сложность). Третий подход Колмогоров называл комбинаторным. В последние годы усилиями научной школы были обнаружены интересные связи между комбинаторным и алгоритмическим подходами. А именно, в 2002г. Н.К. Верещагин, А.Е. Ромащенко и А. Шень [49] показали, как преобразовать произвольное линейное неравенство для колмогоровской сложности в эквивалентное ему комбинаторное утверждение (о соотношении численных характеристик проекций конечного многомерного множества и его покрытий). Получаемое комбинаторное утверждение истинно если и только если соответствующее неравенство выполняется для любых кортежей слов. Таким образом, теперь известен параллелизм между тремя классами утверждений: неравенствами для шенноновской энтропии, неравенствами для колмогоровской сложности и классом комбинаторных утверждений. Неожиданное и вместе с тем естественное применение понятия колмогоровской сложности в математической логике предложил А. Шень. Еще в 30-е гг. Колмогоров предложил интерпретировать пропозициональную логику не как логику утверждений, а как логику задач и применил этот подход для описания интуиционистской логики (логики без закона исключенного третьего). Задачу можно трактовать как множество ее возможных решений, участвующие в построении логических формул логические связки - как операции над множествами, а сами решения – как слова. А. Шень предложил считать сложностью задачи наименьшую из сложностей ее решений. Для каждой логической формулы возникает проблема: по сложностям задач, подставляемых в формулу вместо ее переменных, вычислить сложность результирующей задачи. Первоначально такие вычисления ограничивались тем частным случаем, когда подставлялись задачи, имеющие единственное решение;

в публикациях Верещагина, Мучника, Шеня 2002г. такое вычисление осуществлено для ряда конкретных формул; вместе с тем Верещагин и Мучник опубликовали в 2001г. примеры формул, для которых такое вычисление в принципе невозможно. Если же разрешить подставлять в формулы произвольные задачи, то, как показали в 2002г. Н.К. Верещагин и его ученик А. Чернов, оказывается возможным дать исчерпывающее описание некоторых дедуктивных исчислений в терминах оценки сложности результирующих задач: выяснилось, что выводимость в этих исчислениях равносильна простоте результирующей задачи. Более конкретно, указанный факт был доказан для двух исчислений: для т. н. логики слабого закона исключенного третьего (она занимает промежуточное положение между интуиционистской и классической логиками) и для позитивного фрагмента интуиционистской логики (он состоит из всех формул этой логики, не содержащих отрицания). А. Шень и А.Г. Ромащенко [39] (совместно с Дюраном) решали задачу о замощениях плоскости с нарушением локальных правил в случайном редком множестве. Вопрос о замощении плоскости плитками (некоторого конечного набора типов плиток разной формы) связан с задачами из многих областей – от математической логики (классическая задача Entscheidungsproblem) до физики (квазикристаллы). Основная идея исследования замощений состоит в том, что выполнение некоторых глобальных свойств конфигурации (всего замощения плоскости) можно обеспечить с помощью локальных правил и ограничений (набор запретов для соседствующих плиток). При этом возникает следующий общий и фундаментальный вопрос: могут ли локальные правила обеспечить достаточно сложную (т.е. нерегулярную) структуру замощения? Минимальным требованием «нерегулярности» такого рода является свойство непериодичности. В 1960-х Р. Бергер построил такой набор плиток, которые допускают только непериодические замощения плоскости. Хотя замощения Бергера непериодичны, в интуитивном смысле они обладают вполне регулярной структурой. В 2000 г. Б. Дюран, Л.А. Левин и А. Шень доказали более сильный результат: они построили такой набор плиток, что всякое замощение имеет большую (в некотором смысле, максимальную) плотность информации. Разумеется, такое замощение не может быть периодическим (и даже вычислимым). А. Шень и А.Е. Ромащенко применили метод неподвижной точки (в духе теоремы Клини о рекурсии) и предложили новую конструкцию плиток, которые гарантируют высокую колмогоровскую сложность всех фрагментов. Новая конструкция очень устойчива и позволяет гарантировать нетривиальные свойства замощения (например, высокую колмогоровскую сложность фрагментов) даже если локальные правила разрешается нарушать в достаточно редком (случайно выбранном) множестве исключительных точек.

А.Н. Колмогоров определил сложность объекта как длину его кратчайшего описания. Разумеется, такое определение зависит от того, что понимается под описанием: для каждого способа описания – своя сложность. Сам Колмогоров указал важный класс способов описания и доказал существование в этом классе оптимальных способов, дающих – в разумном смысле – кратчайшие описания. Сложность относительно оптимального способа описания Колмогоров назвал энтропией. Часто энтропию называют просто сложностью, поскольку сложности относительно неоптимальных способов, как правило, не рассматривают. Другие классы способов описания приводят к другим видам колмогоровских энтропий. Определения этих различных видов энтропий по единой схеме с помощью непрерывных вычислимых отображений были предложены А.Шенем в 1984г.; вариант такой схемы без ссылок на непрерывность изложен в публикациях В.А. Успенского, А.Л. Семенова и А. Шеня [27], [28]. В статье В.А. Успенского и А. Шеня [56] приведены соотношения между различными видами колмогоровских энтропий. В 1998г. Успенский предложил рассматривать в качестве значения энтропии не число, а элемент упорядоченного векторного пространства; это предложение позволило представить шенноновскую и колмогоровскую теории информации как варианты единой теории.

Ан.А.Мучник доказал теорему об относительном описании, которая утверждает, что для любых двух слов a и b существует такая программа p, которая преобразует a в b, при этом имеет минимальную возможную длину (то есть ее длина равна $K(b|a)$) и при этом p имеет очень малую сложность относительно b. Другими словами, можно вычислить некоторое «хэш-значение» b длины $K(b|a)$, которого достаточно для восстановления b при заданном слове a. А.Е. Ромащенко и А. Шень обобщили теорему Мучника об относительном описания для программ с ограничением на используемые ресурсы (время и память). Они предложили два новых доказательства этой теоремы. Первое основано на on-line алгоритме поиска паросочетания в двудольном графе. Второе использует экстракторы. Поскольку известны явные конструкции графов-экстракторов, второе наше доказательство позволяет обобщить теорему Мучника для сложности с полиномиальным ограничением на память (используемую программой p). Также был доказан (с помощью конструкции экстракторов Тревисана) аналог теоремы Мучника для сложности, определенной в терминах вероятностно-недетерминированных программ p (в духе систем интерактивных доказательств). Еще в первой своей работе по алгоритмической теории информации А.Н. Колмогоров отмечал, что естественно рассмотреть определение алгоритмической сложности с ограничениями на время работы алгоритма. Однако аналоги важнейших классических теорем (например, теорема Колмогорова–Левина о симметрии взаимной информации) до сих пор не доказаны для колмогоровской сложности для полиномиальных алгоритмов. Трудности в решении этой задачи неслучайны: полиномиальные варианты утверждения о симметрии взаимной информации оказались связаны такими трудными открытыми вопросами как $P = NP$ и $BPP = EXP$. В целом, задача построения теории колмогоровской сложности и алгоритмической случайности с ограничениями на время работы соответствующих алгоритмов еще ждет своего решения.

В заключение заметим, что данный обзор отражает субъективный взгляд его авторов на данную область математики. В основном, в обзоре рассматривались вопросы применения понятия колмогоровской сложности к обоснованию теории вероятностей; эти вопросы изложены в исторической последовательности. В частности, за небольшим исключением, в обзор не включены многочисленные работы участников московской школы колмогоровской сложности по взаимной колмогоровской сложности строк, взаимной информации и применению сложности в комбинаторике и теории кодирования.

СПИСОК ЛИТЕРАТУРЫ

1. Асарин Е.А. О некоторых свойствах Δ-случайных по Колмогорову конечных последовательностей, Теор. вероятн. и ее применения, 1987, т.32, N3, с.556-Ц558.

2. Вовк В.Г. О понятии бернуллиевости. *Успехи математических наук*, 1986, т.41, N1, с.185–186.

3. Вовк В.Г. Закон повторного логарифма для случайных по Колмогорову, или хаотических последовательностей. *Теория вероятностей и ее применения*, 1987, т.32, N3, с.456–468.

4. Вовк В.Г. Асимптотическая эффективность оценок: алгоритмический подход. *Теория вероятностей и ее применения*, 1991, т.36, N2, с.247–261.

5. Вьюгин В.В. Об инвариантных по Тьюрингу множествах, *Доклады АН СССР*, 1976, т.229, N4, с.790–793.

6. Вьюгин В.В. Алгоритмическая энтропия (сложность) конечных объектов и ее применения к определению случайности и количества информации – в Сб.: Семиотика и информатика, В. 16. М.: ВИНИТИ, 1981, с.14–43. Перевод на англ. в *Selecta Mathematica* formerly *Sovietica*, 1994, v.13, N4, p.357–389.

7. Вьюгин В.В. Алгебра инвариантных свойств двоичных последовательностей. *Проблемы передачи информации*, 1982, т.18, N2, с.83–100.

8. Вьюгин В.В. О нестохастических объектах, *Проблемы передачи информации*, 1985, т.21, N2, с.3–9.

9. Вьюгин В.В. О дефекте случайности конечного объекта относительно мер с заданными границами их сложности. *Теория вероятностей и ее применения*, 1987, т.32, N3, с.558–563.

10. Вьюгин В.В. Эргодическая теорема для индивидуальной случайной последовательности. *Успехи математических наук*, 1996, т.51, N1, с.191–192.

11. Вьюгин В.В. Эффективная сходимость по вероятности и эргодическая теорема для индивидуальных случайных последовательностей. *Теория вероятностей и ее применения*, 1997, т.42, N1, с.35–50.

12. Вьюгин В.В. О длине максимальной серии «успехов» в индивидуальной случайной последовательности. *Теория вероятностей и ее применения*, 1997, т.42, N3, с.608–615.

13. Вьюгин В.В. Эффективная сходимость по вероятности и эргодическая теорема для индивидуальных случайных последовательностей. *Теория вероятностей и ее применения*, 1997, т.42, N1, с.35–50.

14. Вьюгин В.В. Проблемы устойчивости универсальных схем сжатия информации, Проблемы передачи информации, 2003, т.39, N1, с.36–52.

15. Звонкин А.К., Левин Л.А. Сложность конечных объектов и обоснование понятий информации и случайности с помощью теории алгоритмов. *Успехи математических наук*, 1970, т.25, N6, с.85–127.

16. Колмогоров А.Н. Три подхода к определению понятия «количество информации». *Проблемы передачи информации*, 1965, т.1, N1, с.3–11.

17. Колмогоров А.Н. К логическим основам теории информации и теории вероятностей. *Проблемы передачи информации*, 1969, т.5, N3, с.3–7.

18. Колмогоров А.Н. Комбинаторные основания теории информации и исчисления вероятностей. *Успехи математических наук*, 1983, т.38, N4, с.27–36.

19. Левин Л.А. О понятии случайной последовательности. *Доклады АН СССР*, 1973, т.212, N3, с.548–550.

20. Левин Л.А. Законы сохранения (невозрастания) информации и вопросы обоснования теории информации. *Проблемы передачи информации*, 1974, т.10, N3, с.30–35.

21. Левин Л.А. О различных мерах сложности конечных объектов (аксиоматическое описание). *Доклады АН СССР*, 1976, т.227, N4, с.804–807.

22. Левин Л.А. Об одном конкретном способе задания сложностных мер. *Доклады АН СССР*, 1977, т.234, N3, с.536–539.

23. Ли М., Витаньи П. Колмогоровская сложность двадцать лет спустя. *Успехи математических наук* т.43, N6, с.129–166.

24. Мизес Р. Вероятность и статистика. М. – Л.: Государственное издательство, 1930.

25. Мучник Ан. А., Семенов А Л. О роли закона больших чисел в теории случайности. *Пробл. передачи информ.*, 2003, т.39, N1, с.134–165.

26. Семенов А.Л., Мучник Ан.А. Об уточнении оценок Колмогорова, относящихся к датчикам случайных чисел и сложностному определению случайности. *Доклады РАН. 2003*, 2003, т.391, N1, с.1–3.

27. Успенский В.А., Семенов А.Л. Теория алгоритмов: основные открытия и приложения. М.:Наука. Гл. ред. физ.-мат. лит., 1987. –(Б-чка программиста).

28. Успенский В.А., Семенов А.Л., Шень А.Х. Может ли (индивидуальная) последовательность нулей и единиц быть случайной? *Успехи математических наук*, 1990, т.45, N1, с.105–162.

29. Успенский В.А. Четыре алгоритмических лица случайности. 2-е издание, исправленное, М.П: МЦНМО, 2009.

30. Шень А.Х. Частотный подход к определению понятия случайной последовательности. *Семиотика и информатика.* – М.: ВИНИТИ, 1982, Вып. 18. – (Второй выпуск за 1981 г.) с.14–42.

31. Шень А.Х. Понятие (α, β)-стохастичности по Колмогорову и его свойства. *Доклады АН СССР*, 1983, т.271, N6, с.1337–1340.

32. Chaitin G.J. On the length of programs for computing binary sequences, *Journal of the Association for Computing Machinery*, 1966, v.13, p.547–569.

33. Chaitin G.J. On the length of programs for computing binary sequences: Statistical considerations, *Journal of the Association for Computing Machinery*, 1969, v.16, p.145–159.

34. Chaitin G.J. A theory of program size formally identical to information theory, *Journal of the Association for Computing Machinery*, 1975, v.22, p.329–340.

35. Chaitin G.J. Algorithmic information theory, *IBM Journal of Research and Development*, 1977, v.21, p.350–359.

36. Chaitin G.J. Algorithmic information theory, Cambridge University Press, 1987.

37. Church A. On the concept of random sequence, *Bulletin of the American Mathematical Society*, 1940, v.46, N2, p.130–135.

38. Cover T.M., Gács P., Gray R.M. Kolmogorov's contributions to information theory and algorithmic complexity, *Ann. Probab.* 1989, v.17, N1, p.840–865.

39. Durand B., Romashchenko A., Shen A. Fixed Point and Aperiodic Tilings, Proceedings of 12th International Conference «Developments in Language Theory» (DLT), 2008, Lecture Notes in Computer in Science, v.5257, Springer-Verlag, 2008. p.276–288.

40. Gács P., Körner J. Common information is far less than mutual information, *Problems. of Control. and Inform. Theory*, 1973, v.2, p.149–162.

41. Hammer D., Romashchenko A., Shen A., Vereshchagin N. Inequalities for Shannon Entropy and Kolmogorov Complexity. *Journal of Computer and System Sciences*, 2000, v.60,p.442–464.

42. Kolmogorov A.N. On tables of random numbers. *Sankhya. The Indian Journal of Statistics, ser. A*, 1963, v.25, N4, p.369–376. (Русский перевод: О таблицах случайных чисел. Семиотика и информатика, М.: ВИНИТИ, 1982, Вып. 2 (второй выпуск за 1981г.), с.3–13).

43. Li M., Vitányi P. An Introduction to Kolmogorov Complexity and Its Applications, 2nd ed. New York: Springer–Verlag, 1997.

44. Loveland D. A new interpretation of the von Mises' concept of random sequence. *Zeitschrift fur Mathematische Logik und Grundlagen der Mathematik*, 1966, v.12, N3, p.279–294.

45. Lugosi G., Cesa-Bianchi N., Prediction, Learning and Games. Cambridge University Press, New York, 2006.

46. Martin-Löf P. The definition of random sequences. *Information and Control*, 1966, v.9, p.602–619.

47. Muchnic An.A., Semenov A.L., Uspensky V.A. Mathematical metaphysics of randomness. *Theoretical Computer Science*, 1998, v.207, p.263–317.

48. Muchnik An. A. On common information. *Theor. Comp. Sci.* 1998, v.207, p.319–328.

49. Romashchenko A., Shen A., Vereshchagin N. Combinatorial Interpretation of Kolmogorov Complexity. *Theoretical Computer Science.* 2002, v.271, p.111–123.

50. Shafer, G., Vovk, V., Probability and Finance. It's Only a Game! New York: Wiley, 2001.

51. Shnorr C.P. Process complexity and effective random tests. *Journal of Computer and System Sciences*, 1973, v.3, N4, p.376–378.

52. Schnor C.P. A survey of the theory of random sequences, Logic, foundations of mathematics and computability theory. Eds. R.E.Butts, J.Hintikka. – Dordrecht: D.Reidel. – 1977, p.193–211.

53. Solomonoff R.J. A preliminary report on a general theory of inductive inference. Technical Report, ZTB-138, Zator Company, Cambridge, Mass. (November, 1960).

54. Solomonoff R.J. A formal theory of inductive inference I, II. *Information and Control*, 1964, v.7, p.1–22, p.224–254.

55. Solomonoff R.J. Complexity-based induction systems: Comparisons and convergence theorems. *IEEE Transactions on Information Theory*, 1978, IT-24, p.422–432.

56. Uspensky V.A., Shen A. Relations between varieties of Kolmogorov complexities. *Mathematical Systems Theory*, 1996, v.29, p.271–292.

57. V'yugin V.V. Ergodic theorems for individual random sequences. *Theoretical Computer Science*, 1998, v.207, N4, p.343–361.

58. V'yugin V.V. Algorithmic complexity and stochastic properties of finite binary sequences. *The Computer Journal*, 1999, v.42, N4, p.294–317.

59. Vereshchagin N.K., Vitanyi P. KolmogorovTs Structure Functions and Model Selection. *IEEE Transactions on Information Theory*, 2004, v.50, N.12, p.3265–3289.