Отдельный оттиск

УСПЕХИ МАТЕМАТИЧЕСКИХ НАУК

ТОМ
XXV
ВЫПУСК
6(156)

1970

1970 г. ноябрь — декабрь т. XXV вып. 6 (1956)
УСПЕХИ МАТЕМАТИЧЕСКИХ НАУК

УДК 519.24+517.11+519.9

СЛОЖНОСТЬ КОНЕЧНЫХ ОБЪЕКТОВ И ОБОСНОВАНИЕ ПОНЯТИЙ ИНФОРМАЦИИ И СЛУЧАЙНОСТИ С ПОМОЩЬЮ ТЕОРИИ АЛГОРИТМОВ

А. К. Звонкин и Л. А. Левин

В 1964 г. А. Н. Колмогоров ввел понятие сложности конечного объекта (например, слова в некотором алфавите). Сложность он определял как минимальное число двоичных знаков, содержащих всю информацию о задаваемом объекте, достаточную для его восстановления (декодирования). Это определение существенно зависит от метода декодирования, однако с помощью общей теории алгоритмов А. Н. Колмогорову удалось дать инвариантное (универсальное) определение сложности. Близкие понятия рассматривались Р. Дж. Соломоновым (США) и А. А. Марковым. На базе понятия сложности А. Н. Колмогоров дал определение количества информации в конечных объектах и понятия случайной последовательности (уточненное потом в работах П. Мартин-Лёфа). Впоследствии этот круг вопросов быстро развивался. В частности, интересное развитие получили идеи А. А. Маркова о применении понятия сложности для изучения количественных вопросов теории алгоритмов. Настоящая статья представляет собой обзор основных результатов, связанных со всем изложенным.

СОДЕРЖАНИЕ

Предварительные замечания

При написании статьи мы, кроме цитированной литературы, существенно использовали материалы лекций А. Н. Колмогорова, спецкурса Н. В. Петри и М. И. Кановича, а также семинара В. А. Душского и Л. А. Левина. Мы горячо благодарны Андрею Николаевичу Колмогорову, который оказал нам большую помощь, редактируя все промежуточные варианты текста статьи; без его постоянной поддержки статья вообще не могла бы быть написана. Весьма ценным для нас был постоянный контакт и обсуждение результатов с М. И. Кановичем и Н. В. Петри, за что мы им крайне признательны. Мы

очень благодарны А. Б. Сосинскому, прочитавшему всю рукопись и сделавшему много ценных замечаний. Мы хотим также поблагодарить В. Н. Агафонова, Я. М. Барздиня, А. Н. Колодия, П. Мартин-Лёфа, Л. Б. Медведовского, В. А. Успенского. Дж. Т. Шварца и всех участников семинара А. А. Маркова за ценное обсуждение.

1. Некоторые определения и обозначения. Мы будем рассматривать *слова* в алфавите $\{0, 1\}$, т. е. конечные последовательности нулей и единиц. Установим взаимно однозначное соответствие между словами и *натуральными числами*:

$$\begin{gathered}\Lambda \leftrightarrow 0\\ 0 \leftrightarrow 1\\ 1 \leftrightarrow 2\\ 00 \leftrightarrow 3\\ 01 \leftrightarrow 4\\ 10 \leftrightarrow 5\\ 11 \leftrightarrow 6\\ 000 \leftrightarrow 7\\ 001 \leftrightarrow 8\\ \ldots\ldots\end{gathered}$$

(Λ — пустое слово), и в дальнейшем не будем различать эти объекты, употребляя произвольно любой из терминов «слово» или «число». Обозначать их мы будем, как правило, малыми латинскими буквами, множество всех слов-чисел будем обозначать S.

Если к слову x справа приписать слово y, получится слово, которое мы будем обозначать xy. Нам потребуется также уметь записывать одним словом упорядоченную пару слов (x, y). Для того чтобы не вводить специальных разделительных знаков (вроде запятой), условимся, что если $x = x_1x_2 \ldots x_n$ (x_i — 0 или 1), то

$$\bar{x} = x_1x_1x_2x_2 \ldots x_nx_n01. \tag{0.1}$$

Тогда по слову $\bar{x}y$ можно однозначно восстановить и x, и y. Обозначим $\pi_1(z)$ и $\pi_2(z)$ функции такие, что $\pi_1(\bar{x}y) = x$, $\pi_2(\bar{x}y) = y$; если слово z не представимо в виде $\bar{x}y$, то $\pi_1(z) = \Lambda$, $\pi_2(z) = \Lambda$ [1]).

Длиной $l(x)$ слова x будем называть количество знаков в нем; $l(\Lambda) = 0$. Очевидно,

$$l(xy) = l(x) + l(y), \tag{0.2}$$

$$l(\bar{x}) = 2l(x) + 2. \tag{0.3}$$

Обозначим $d(A)$ *количество элементов в множестве* A. Очевидно,

$$d\{x:\ l(x) = n\} = 2^n, \tag{0.4}$$

$$d\{x:\ l(x) \leqslant n\} = 2^{n+1} - 1. \tag{0.5}$$

---

[1]) Можно было бы устроить более стандартную нумерацию пар (x, y), однако для нас важно, чтобы выполнялось свойство (0.11) (см. ниже)

Объектом нашего рассмотрения будет также пространство Ω *бесконечных двоичных последовательностей* (их мы будем обозначать малыми греческими буквами). $\Omega^* = \Omega \cup S$ — множество всех *конечных и бесконечных последовательностей*. Пусть $\omega \in \Omega^*$; тогда будем называть *n-фрагментом* ω и обозначать $(\omega)_n$ слово, состоящее из первых n знаков ω (при этом если ω — слово, и $l(\omega) \leqslant n$, то, по определению, $(\omega)_n = \omega$). Последовательность $\omega \in \Omega$ будем называть *характеристической* для множества натуральных чисел $A = \{n_1, n_2, \ldots\}$, не содержащего 0, если в этой последовательности n_1-я, n_2-я, ... цифра — единицы, а все остальные цифры — нули. Множество A, для которого ω — характеристическая последовательность, будем обозначать также S_ω.

Обозначим Γ_x множество всех последовательностей (конечных и бесконечных или только бесконечных, в зависимости от того, рассматриваем мы пространство Ω^* или Ω; в каждом конкретном случае это будет ясно из контекста), начинающихся со слова x, т. е.

(0.6) $\quad \Gamma_x = \{\omega\colon (\omega)_{l(x)} = x\}.$

Будем обозначать $x \subset y$, если $\Gamma_x \supseteq \Gamma_y$ (т. е. слово x есть начало слова y). Отношение $\subset$ частично упорядочивает множество S (рис. 1).

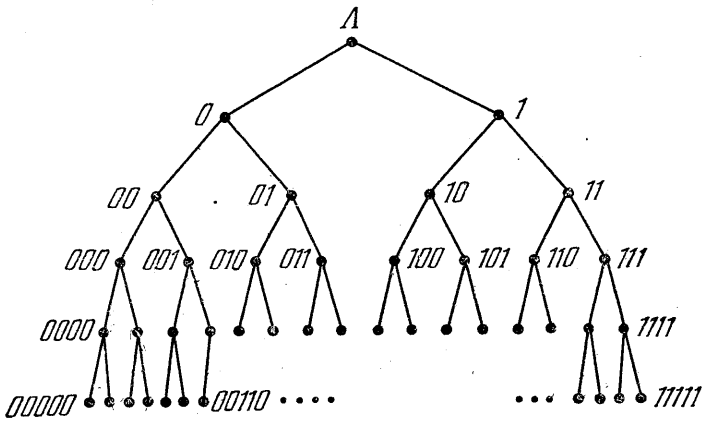

Рис. 1.

Функции, определенные на декартовом произведении $S^n = S \times S \times \ldots \times S$ (n раз), будем (за исключением, может быть, стандартных функций) обозначать большими латинскими буквами, иногда ставя вверху индекс n (обозначающий число переменных): $F^n = F^n(x_1, \ldots, x_n)$. Будем всегда стандартным образом заменять фразу: для любых допустимых значений переменных $y_1, \ldots, y_m$ найдется константа C такая, что для всех допустимых значений $x_1, \ldots, x_n$

(0.7) $\quad F^{n+m}(x_1, \ldots, x_n; y_1, \ldots, y_m) \leqslant G^{n+m}(x_1, \ldots, x_n; y_1, \ldots, y_m) + C,$

на более короткую фразу (использующую новое обозначение):

(0.8) $\quad F^{n+m}(x_1, \ldots, x_n; y_1, \ldots, y_m) \lesssim G^{n+m}(x_1, \ldots, x_n; y_1, \ldots, y_m)$

($y_1, \ldots, y_m$ входят как параметры).

Аналогично определяется отношение $\gtrsim$; $F \asymp G$ тогда и только тогда, когда $F \lesssim G$ и $G \lesssim F$. Очевидно, отношения $\lesssim$, $\gtrsim$ и $\asymp$ транзитивны. Очевидно также,

(0.9) $\quad l(x) \asymp \log_2 x$ для $x > 0$,

(0.10) $\quad l(\bar{x}) \asymp 2l(x)$,

(0.11) $\quad l(\bar{x}y) \asymp l(y)$ (x входит как параметр),

и т. д.

2. Необходимые сведения из теории алгоритмов. Приведем некоторые необходимые нам определения и теоремы из теории алгоритмов. Большинство из приведенных фактов доказывается в любом руководстве по теории алгоритмов (см., например, [1]—[4]), доказательство остальных не представит труда для читателя, знакомого с одним из таких руководств.

Пусть функции C^1, O^n, I_m^n принимают, по определению, следующие значения: $C^1(x)=x+1$, $O^n(x_1, \ldots, x_n)=0$, $I_m^n(x_1, \ldots, x_n)=x_m$. Говорят, что $(n+1)$-местная функция F возникает из n-местной функции G и $(n+2)$-местной функции H *примитивной рекурсией*, если для всех натуральных значений $x_1, \ldots, x_n, y$ имеем

$$F(x_1, \ldots, x_n, 0)=G(x_1, \ldots, x_n),$$

$$F(x_1, \ldots, x_n, y+1)=H(x_1, \ldots, x_n, y, F(x_1, \ldots, x_n, y)).$$

Обозначим через

$$\mu_y(F(x_1, \ldots, x_{n-1}, y)=x_n) \tag{0.12}$$

наименьшее значение a, для которого

$$F(x_1, \ldots, x_{n-1}, a)=x_n. \tag{0.13}$$

При этом будем считать, что значение (0.12) не определено в следующих случаях:

а) значения $F(x_1, \ldots, x_{n-1}, y)$ определены для всех $y<a$, но отличны от x_n, а значение $F(x_1, \ldots, x_{n-1}, a)$ не определено $(a=0, 1, 2, \ldots)$;

б) значения $F(x_1, \ldots, x_{n-1}, y)$ определены для всех $y=0, 1, 2, \ldots$ и отличны от x_n.

Значение выражения (0.12) при заданной функции F зависит от значений $x_1, \ldots, x_{n-1}, x_n$, т. е. является функцией от этих переменных. Говорят, что эта функция получена из функции F при помощи *операции минимизации*.

Определение 0.1. Функция F называется *частично рекурсивной*, если она может быть получена из функций C^1, O^n, I_m^n конечным числом операций *подстановки* (т. е. суперпозиции), примитивной рекурсии и минимизации. Всюду определенная частично рекурсивная функция называется *общерекурсивной*. Свойство n-ок чисел $\Pi^n(a_1, \ldots, a_n)$ называется *частично рекурсивным* (*общерекурсивным*) *предикатом*, если существует частично рекурсивная (общерекурсивная) функция, равная 0 на всех n-ках, удовлетворяющих этому свойству, и только на них.

Легко проверить, что функции $l(x)$, $\pi_1(z)$, $\pi_2(z)$, $F(x)=\bar{x}$, $G(x, y)=\overline{xy}$ общерекурсивны.

В настоящее время общепринятой является следующая естественнонаучная гипотеза:

Тезис Черча. *Класс алгоритмически вычислимых (в инстуитивно ясном смысле) числовых функций совпадает с классом всех частично рекурсивных фунций.*

В дальнейшем изложении мы неоднократно будем, приводя алгоритм, вычисляющий некоторую функцию, предполагать ее частичную рекурсив-

ность без доказательства, т. е. не выписывая (из-за громоздкости) построение, требуемое определением 0.1. Трудолюбивый читатель, не желающий в каждом таком случае принимать на веру тезис Черча, всегда сможет выписать такое построение самостоятельно.

Замечание 0.1. Легко видеть, что частично рекурсивные функции, строящиеся без употребления операции минимизации (такие функции называются *примитивно рекурсивными*), являются всюду определенными. Только операция минимизации может приводить к не всюду определенным функциям, так как процесс вычисления результата минимизации (состоящий в последовательной проверке справедливости равенства (0.13) для $a = 0, 1, 2, \ldots$) может никогда не кончиться. Будем говорить, что значение частично рекурсивной функции F^n на данном наборе $(x_1, \ldots, x_n)$ вычислено *не более чем за t шагов* (*операций*), если все процессы вычисления результатов минимизаций, входящих в построение функции F^n, закончились на значениях соответствующих параметров a, не превышающих t. Мы будем часто употреблять понятие количества шагов, совершенных алгоритмом, вычисляющим функцию F^n, в выше приведенном смысле [1]).

Теорема 0.1. *Какова бы ни была частично рекурсивная функция F^n, свойство набора $(t; x_1, \ldots, x_n)$, состоящее в том, что значение $F^n(x_1, \ldots, x_n)$ вычисляется не более чем за t шагов, является общерекурсивным предикатом.*

Определение 0.2. Частично рекурсивная функция $U^{n+1}(i; x_1, \ldots, x_n)$ называется *универсальной* для n-местных частично рекурсивных функций, если для любой частично рекурсивной функции $F^n(x_1, \ldots, x_n)$ найдется i такое, что

$$F^n(x_1, \ldots, x_n) \equiv U^{n+1}(i; x_1, \ldots, x_n). \tag{0.14}$$

Число i будем называть *номером* функции F^n относительно U^{n+1} (функция может иметь много номеров).

Теорема 0.2. *Для любого натурального n существует частично рекурсивная функция, универсальная для всех n-местных частично рекурсивных функций.*

Будем называть *нумерацией множества* S^n любую n-ку общерекурсивных функций F_i $(i = 1, 2, \ldots, n)$, отображающую S на S^n. Натуральное число k называется *номером* n-ки $(x_1, \ldots, x_n)$ в этой нумерации, если $F_i(k) = x_i$ для всех $i = 1, 2, \ldots, n$. Очевидно, пара функций $\pi_1(z)$, $\pi_2(z)$ является нумерацией S^2.

Следующее определение не зависит от нумерации.

Определение 0.3. Множество $X \subseteq S^n$ называется *перечислимым*, если множество номеров его элементов (в выбранной нумерации) является областью значений какой-либо частично рекурсивной функции (при этом говорят, что эта функция *перечисляет* множество X).

Замечание 0.2. Любое перечислимое множество перечисляется также и общерекурсивной функцией.

---

[1]) Определенное таким образом количество шагов является *сигнализирующей* функцией в смысле Трахтенброта [42].

Теорема 0.3. *Пусть есть частично рекурсивный предикат* Π^{n+k}. *Тогда множество* $\{(x_1, \ldots, x_n)$: $\exists a_1, \ldots, a_k\, \Pi^{n+k}(x_1, \ldots, x_n; a_1, \ldots, a_k)$ *истинно*$\}$ *перечислимо.*

Следующая теорема показывает, что семейство перечислимых множеств, зависящих от параметров $p_1, \ldots, p_k$, перечислимо *без повторений.*

Теорема 0.4. *Пусть дано перечислимое множество* $A \subseteq S^{n+k}$. *Тогда существует частично рекурсивная функция* $F(t; p_1, \ldots, p_k)$ *такая, что*

а) *при любых фиксированных* $p_1, \ldots, p_k$ *множество значений функции* $F(t; p_1, \ldots, p_k)$ *будет совпадать с множеством номеров наборов* $(x_1, \ldots, x_n)$ *таких, что* $(x_1, \ldots, x_n; p_1, \ldots, p_k) \in A$ *(номера берутся в некоторой фиксированной нумерации* S^n*);*

б) *если* $t_1 < t_2$ *и* $F(t_2; p_1, \ldots, p_k)$ *определено, то* $F(t_1; p_1, \ldots, p_k)$ *тоже определено и отлично от* $F(t_2; p_1, \ldots, p_k)$.

Определение 0.4. Множество $X \subseteq S^n$ называется *разрешимым*, если существует общерекурсивная функция, равная 0 на X и 1 на $S^n \setminus X$. Последовательность, характеристическую для разрешимого множества, будем называть *вычислимой.*

Очевидно, что всякое разрешимое множество перечислимо.

Теорема 0.5. *Всякое бесконечное перечислимое множество включает в себя бесконечное разрешимое подмножество.*

§ 1. Сложность

В этом параграфе вводится понятие сложности. Выводятся простейшие оценки величины сложности и изучаются алгоритмические свойства этой функции.

1. Определения. Теорема оптимальности. Одним из центральных понятий в этой статье будет понятие сложности некоторого текста (сообщения). Сложностью текста мы будем называть длину самого короткого двоичного слова, содержащего всю информацию, необходимую для восстановления рассматриваемого текста при помощи какого-нибудь фиксированного способа декодирования. Точнее:

Определение 1.1. (А. Н. Колмогоров). Пусть F^1 — произвольная частично рекурсивная функция. Тогда *сложность слова* x *по* F^1 *есть.*

$$K_{F^1}(x) = \begin{cases} \min l(p)\colon F^1(p) = x, \\ \infty, \text{ если } \forall p \in S\ F^1(p) \neq x. \end{cases} \tag{1.1}$$

Слово p такое, что $F^1(p) = x$, будем называть *кодом* или *программой*, по которой F^1 восстанавливает слово x.

Такое определение сложности очень сильно зависит от вида F^1. Однако следующая замечательная теорема позволяет дать инвариантное определение этого понятия (благодаря чему на базе понятия сложности смогла возникнуть теория, излагаемая в статье).

Теорема 1.1. (А. Н. Колмогоров, Р. Соломонов). *Существует частично рекурсивная функция* F^1_0 *(называемая оптимальной) такая, что для любой другой частично рекурсивной функции* G^1

$$K_{F^1_0}(x) \preccurlyeq K_{G^1}(x). \tag{1.2}$$

Доказательство. См. следствие 1.3.

Следствие 1.1. *Для любых двух оптимальных частично рекурсивных функций* F^1 *и* G^1

$$K_{F1}(x) \asymp K_{G1}(x). \tag{1.3}$$

Определение 1.2. *Слжностью* $K(x)$ слова x назовем сложность $K_{F_0^1}(x)$ по некоторой раз и навсегда фиксированной оптимальной частично рекурсивной функции F_0^1 (например, по той, которая будет определена в следствии 1.3).

Определение 1.3 (А. Н. Колмогоров). (*Условная*) *сложность слова* x *при известном* y *по частично рекурсивной функции* F^2 *есть*

$$K_{F^2}(x \mid y) = \begin{cases} \min l(p)\colon F^2(p, y) = x, \\ \infty, \text{ если } \forall p \in S\ F^2(p, y) \neq x. \end{cases} \tag{1.4}$$

Теорема 1.2 (А. Н. Колмогоров, Р. Соломонов). *Существует частично рекурсивная функция* F_0^2 (*называемая оптимальной*) *такая, что для любой частично рекурсивной функции* G^2

$$K_{F_0^2}(x \mid y) \preccurlyeq K_{G^2}(x \mid y). \tag{1.5}$$

Доказательство. Пусть $U^3(n;\ p,\ y)$ — частично рекурсивная функция, универсальная для всех двуместных частично рекурсивных функций (см. определение 0.2, теорему 0.2). Определим функцию

$$F_0^2(z, y) = U^3(\pi_1(z),\ \pi_2(z),\ y), \tag{1.6}$$

и докажем, что эта функция оптимальна. Действительно, пусть G^2 — частично рекурсивная функция, n_{G^2} — какой-нибудь ее номер (см. определение 0.2), и пусть

$$K_{G^2}(x \mid y) = l_0, \tag{1.7}$$

т. е. существует программа p_0 такая, что $G^2(p_0, y) = x$, $l(p_0) = l_0$, причем среди всех слов p таких, что $G^2(p, y) = x$, слово p_0 имеет минимальную длину. Тогда, если вместо z подставить $z = \bar{n}_{G^2} p_0$, согласно (1.6) мы получим

$$F_0^2(z, y) = F_0^2(\bar{n}_{G^2} p_0, y) = U^3(\pi_1(\bar{n}_{G^2} p_0),\ \pi_2(\bar{n}_{G^2} p_0),\ y) =$$
$$= U^3(n_{G^2};\ p_0, y) = G^2 p_0, y) = x,$$

поэтому из (1.4), (1.7) и (0.2) следует

$$K_{F_0^2}(x \mid y) \leqslant l(z) = l(\bar{n}_{G^2} p_0) = l(\bar{n}_{G^2}) + l(p_0) =$$
$$= l_0 + l(n_{G^2}) = K_{G^2}(x \mid y) + l(\bar{n}_{G^2}) \asymp K_{G^2}(x \mid y),$$

так как $l(\bar{n}_{G^2})$ не зависит от x и y, а зависит только от функции G^2.

Следствие 1.2. *Для любых двух оптимальных частично рекурсивных функций* F^2 *и* G^2

$$K_{F^2}(x \mid y) \asymp K_{G^2}(x \mid y). \tag{1.8}$$

Определение 1.4. (*Условной*) *сложностью слова* x *при известном* y $K(x \mid y)$ назовем сложность $K_{F_0^2}(x \mid y)$ по некоторой раз и навсегда фиксированной оптимальной частично рекурсивной функции F_0^2 (например, по функции, определяемой равенством (1.6)).

Следствие 1.3. *Частично рекурсивная функция*

$$F_0^1(p) = F_0^2(p, \Lambda) \tag{1.9}$$

будет оптимальной в смысле теоремы 1.1.

Доказательство. Покажем, что $K_{F_0^1}(x) \preccurlyeq K_{G^1}(x)$, где G^1 — произвольная частично рекурсивная функция. Действительно, определим $G^2(p, y) = G^1(p)$. Тогда из (1.5) и (1.9) $K_{G^1}(x) = K_{G^2}(x \mid \Lambda) \succcurlyeq K_{F_0^2}(x \mid \Lambda) = K_{F_0^1}(x)$, что и требовалось доказать.

В дальнейшем F_0^1 и F_0^2 будут обозначать раз и навсегда выбранные оптимальные функции.

2. Оценки величины сложности. В этом пункте мы докажем наиболее важные для дальнейшего оценки величин $K(x)$ и $K(x \mid y)$.

Теорема 1.3 (А. Н. Колмогоров). *Пусть A — перечислимое множество пар (x, a), и пусть $M_a = \{x: (x, a) \in A\}$. Тогда*

$$K(x \mid a) \preccurlyeq l(d(M_a)). \tag{1.10}$$

Доказательство. Пусть частично рекурсивная функция $F^2(p, a)$ вычисляется следующим алгоритмом: мы выбираем p-ю в порядке перечисления без повторений (см. теорему 0.4) пару вида (x, a) и выдаем в качестве значения функции F^2 первый элемент этой пары (т. е. слово x). Очевидно, если $x \in M_a$, то найдется $p \leqslant d(M_a)$ такое, что $F^2(p, a) = x$; отсюда согласно (1.5) $K(x \mid a) \preccurlyeq K_{F^2}(x \mid a) \leqslant l(d(M_a))$, что и требовалось доказать.

Замечание 1.1. Для произвольного слова y и конечного множества M доля тех $x \in M$, для которых

$$K(x \mid y) \leqslant l(d(M)) - m, \tag{1.11}$$

не превосходит 2^{-m+1}. Действительно, если $K(x \mid y) \leqslant n$, то найдется слово p длины, не превосходящей n, такое, что $F_0^2(p, y) = x$. Значит, количество таких слов x заведомо не превосходит количества всех программ p длины, не превосходящей n; количество таких программ p равно $2^{n+1} - 1$ (см. (0.5)). В свою очередь $d(M) \geqslant 2^{l(d(M))} - 1$. В итоге доля слов $x \in M$, удовлетворяющих условию (1.11), не превосходит $\dfrac{2^{l(d(M))-m+1} - 1}{2^{l(d(M))} - 1} < 2^{-m+1}$. Таким образом, оценка теоремы 1.3 для большинства слов точная; эта теорема часто дает возможность получать наилучшие (т. е. вообще говоря, неулучшаемые) оценки сложности многих типов слов и будет неоднократно использоваться нами в дальнейшем.

Докажем несколько свойств абсолютной (т. е. не условной) сложности.

Теорема 1.4 (А. Н. Колмогоров). *Справедливы утверждения:*

(1.12) а) *величина* $K(x) \preccurlyeq l(x)$

(*следовательно*, $K(x) < \infty$ *для всех* $x \in S$);

б) *доля слов* x, *для которых* $K(x) < l_0 - m$, *среди всех слов* x, $l(x) = l_0$, *не превосходит* 2^{-m+1} (*т. е. оценка* (1.12) *для большинства слов точная*);

в) *предел*

$$\lim_{x \to \infty} K(x) = \infty \tag{1.13}$$

(*следовательно, и* $\lim\limits_{x\to\infty} m(x)=\infty$, *где*

$$m(x)=\min_{y\geqslant x} K(y), \tag{1.14}$$

т. е. $m(x)$ *— наибольшая монотонно неубывающая функция, ограничивающая* $K(x)$ *снизу*);

г) *для любой монотонно стремящейся к бесконечности частично рекурсивной функции* $\Phi(x)$ *начиная с некоторого* x_0 $m(x)<\Phi(x)$ (*т. е.* $m(x)$ *хотя и стремится к бесконечности, но медленнее любой частично рекурсивной монотонно стремящейся к бесконечности функции*);

(1.15) д) *справедливо* $|K(x+h)-K(x)|\leqslant 2l(h)$

(*т. е. функция* $K(x)$ *хотя и колеблется все время между* $l(x)$ *и* $m(x)$, *но делает это довольно плавно*).

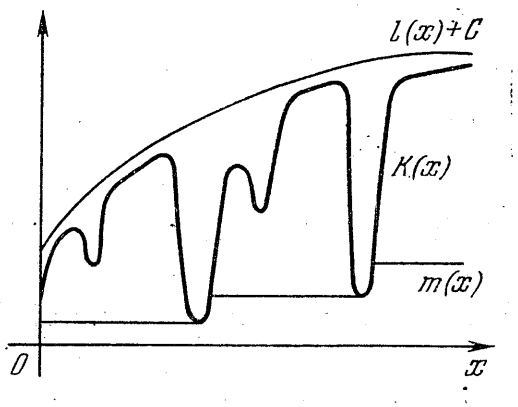

Рис. 2.

Доказательство (рис. 2). а) Пусть $G^1(x)=x$; тогда $K_{G1}(x)=l(x)$ и по теореме 1.1 $K(x)\leqslant K_{G1}(x)=l(x)$, что и требовалось доказать.

б) Это утверждение является тривиальным следствием замечания 1.1 (для $y=\Lambda$). Добавим к этому, что для любого l_0 найдется слово x длины l_0 такое, что $K(x)\geqslant l_0$ (так как количество текстов, имеющих длину l_0, равно 2^{l_0}, а количество программ, имеющих длину меньше l_0, равно $2^{l_0}-1$).

в) По аналогии с замечанием 1.1 количество слов x таких, что $K(x)\leqslant a$, не превосходит 2^{a+1}, т. е. конечно, значит, для любого a найдется x_0 ($x_0=\max\limits_{K(x)\leqslant a} x$) такое, что $K(x)>a$ для всех $x>x_0$, что и требовалось доказать.

г) Пусть утверждение теоремы неверно, т. е. существует частично рекурсивная монотонно стремящаяся к бесконечности функция $\Phi(x)$ такая, что $\Phi(x)\leqslant m(x)$ в бесконечном множестве точек x. Функция $\Phi(x)$ определена на бесконечном перечислимом множестве U. По теореме 0.5 U содержит бесконечное разрешимое подмножество V. Положим

$$\Psi(x)=\begin{cases}\Phi(x)\dot{-}1^{1}), & x\in V,\\ \Phi(\max\limits_{y\leqslant x,\ y\in V} y)\dot{-}1, & x\overline{\in}V.\end{cases}$$

Построенная функция $\Psi(x)$ общерекурсивна, монотонно стремится к бесконечности, и $\Psi(x)\leqslant m(x)$ на бесконечном множестве точек x. Обозначим $M(a)=\max\limits_{K(x)\leqslant a} x$. Легко проверить, что $M(a)+1=\min\limits_{m(x)>a} x$. Нетрудно показать, что $\max\limits_{\Psi(x)\leqslant a} x\geqslant \min\limits_{m(x)>a} x>M(a)$ на бесконечном множестве точек a, причем функция $F(a)=\max\limits_{\Psi(x)\leqslant a} x$, очевидно, общерекурсивна. Таким образом,

---

[1]) $a\dot{-}b=\max\{a-b;\ 0\}$; эта операция вводится для того, чтобы не выходить за пределы множества натуральных чисел.

$F(a) > M(a) = \max\limits_{K(x)\leqslant a} x$ на бесконечном множестве точек a, т. е. $K(F(a)) > a$. Но по теореме 1.1 $K(F(a)) \lesssim K_F(F(a)) \leqslant l(a)$. Отсюда найдется константа C такая, что $l(a) + C > a$ для бесконечно большого количества чисел a, что невозможно.

д) Пусть p_x — программа минимальной длины для слова x, т. е. $F_0^1(p_x) = x$ и $K(x) = l(p_x)$. Тогда слово $x + h$ можно получить из программы $\overline{h}p_x$, если применить к ней функцию $G^1(z) = F_0^1(\pi_2(z)) + \pi_1(z)$; поэтому из (0.2) и (0.10)

$$K_{G^1}(x+h) \leqslant l(\overline{h}p_x) = l(\overline{h}) + l(p_x) \asymp 2l(h) + l(p_x) = 2l(h) + K(x).$$

Но $K(x+h) \lesssim K_{G^1}(x+h)$, откуда $K(x+h) \lesssim K(x) + 2l(h)$, или $K(x+h) - K(x) \lesssim 2l(h)$. Аналогично, применяя функцию $H^1(z) = F_0^1(\pi_2(z)) \dot{-} \pi_1(z)$ [1]) к слову $\overline{h}p_{x+h}$, где p_{x+h} — программа слова $x+h$, получим неравенство

$$K(x) - K(x+h) \lesssim 2l(h).$$

3. Алгоритмические свойства сложности. Теорема 1.5 (А. Н. Колмогоров). а) *Функция $K(x)$ не частично рекурсивна, и, более того, никакая частично рекурсивная функция $\Phi(x)$, определенная на бесконечном множестве точек, не может во всей своей области определения совпадать с $K(x)$.*

б) *Существует общерекурсивная функция $H(t, x)$, монотонно не возрастающая по t, такая, что*

$$\lim_{t\to\infty} H(t, x) = K(x) \tag{1.16}$$

(т. е. хотя и нет способа вычислять $K(x)$, все же есть возможность получать сколь угодно хорошие оценки сверху этой величины).

Доказательство. а) Выделим в области определения U функции $\Phi(x)$ бесконечное разрешимое подмножество V (см. теорему 0.5). Функция $F(m) = \min\limits_{K(x)\geqslant m,\ x\in V} x$ общерекурсивная (так как $K(x) = \Phi(x)$ на V) и принимает сколь угодно большие значения, причем $K(F(m)) \geqslant m$ (по построению). Но, с другой стороны, $K(F(m)) \lesssim K_F(F(m)) \leqslant l(m)$, откуда $m \lesssim l(m)$, что неверно.

б) Пусть C — достаточно большая константа (такая, что $K(x) < l(x) + C$). Возьмем алгоритм, вычисляющий функцию F_0^1, и заставим его совершить по t шагов (см. замечание 0.1) на всех словах p длины, меньшей $l(x) + C$. Если слово x еще не получилось в качестве результата, положим $H(t, x) = l(x) + C$; если оно уже получилось (и, возможно, не один раз) в качестве результата, положим $H(t, x)$ равным минимальной длине программ p, из которых получено слово x. Ясно, что $H(t, x)$ общерекурсивна и монотонно не убывает по t. Если мы будем совершать все больше и больше шагов алгоритма, вычисляющего $F_0^1(p)$ (т. е. когда $t \to \infty$), мы, наконец, получим x из его «настоящей» программы p_0 минимальной длины, т. е. найдем сложность x ($K(x) = l(p_0)$) (правда, мы ни на каком шагу не сможем узнать, произошло это уже или нет).

---

[1]) См. сноску на стр. 93.

Теорема 1.6 (Я. М. Барздинь). *Пусть* $f(x)$ — *общерекурсивная функция и* $\lim\limits_{x\to\infty} f(x) = \infty$. *Тогда множество* $A = \{x : K(x) \leqslant f(x)\}$ *перечислимо* (*и вообще, предикат* $\Pi(x, a) \sim [K(x) \leqslant a]$ *частично рекурсивен*). *Дополнение к* A *бесконечно, но не содержит никакого бесконечного перечислимого подмножества* (*такие множества* A *называются простыми*).

Доказательство. Утверждение $[K(x) \leqslant a]$ эквивалентно утверждению $[\exists t : H(t, x) \leqslant a]$ (см. теорему 1.5б), что и доказывает первую часть теоремы.

Пусть D — бесконечное перечислимое множество, лежащее в дополнении к A, и пусть функция G^1 действует следующим образом: она берет первое в порядке перечисления без повторений (см. теорему 0.4) число $x \in D$ такое, что $f(x) \geqslant n$, и полагает $G^1(n) = x$. Ясно, что $K(x) \preccurlyeq K_{G^1}(x) \leqslant l(n)$. Но число x лежит в дополнении к A, т. е. по определению $K(x) \geqslant f(x)$, откуда $K(x) \geqslant n$ и $l(n) \succcurlyeq n$, что неверно.

4. Мажоранты сложности. Очевидно, если мы знаем само слово x и его сложность, то можно эффективно (например, перебором) найти одну из программ наименьшей длины, кодирующих слово x. Более того, если мы знаем слово x и какое-нибудь число $s \geqslant K(x)$, то можно эффективно найти одну из программ слова x, которая хотя, возможно, и не будет самой короткой, но все же будет иметь длину, не превосходящую s. Поскольку, как следует из теоремы 1.5, эффективно найти сложность нельзя, на практике приходится довольствоваться эффективно вычислимыми (точнее, частично рекурсивными) функциями, которые во всей своей области определения не меньше сложности, т. е. дают оценку длины кода, хотя и не самого короткого, но зато эффективно вычислимого.

Определение 1.5. Будем называть *мажорантой сложности* любую частично рекурсивную функцию $\Phi(x)$, для которой

$$K(x) \preccurlyeq \Phi(x). \tag{1.17}$$

Теорема 1.7 (Л. А. Левин). *Частично рекурсивная функция* $\Phi(x)$ *тогда и только тогда будет мажорантой сложности, когда*

$$l(d\{x : \Phi(x) = a\}) \preccurlyeq a. \tag{1.18}$$

Доказательство. Пусть Φ — мажоранта сложности, и x принадлежит области ее определения; $\Phi(x) = a$. Согласно (1.17) найдется константа C такая, что $K(x) \leqslant \Phi(x) + C$, откуда $d\{x : \Phi(x) = a\}$ не превосходит количества слов x таких, что $K(x) \leqslant a + C$, следовательно (аналогично замечанию 1.1),

$$d\{x : \Phi(x) = a\} \leqslant 2^{a+C+1} \quad \text{и} \quad l(d\{x : \Phi(x) = a\}) \leqslant a + C + 1,$$

что и доказывает теорему в одну сторону.

Пусть теперь для частично рекурсивной функции Φ выполнено условие (1.18), т. е. существует константа C такая, что $d\{x : \Phi(x) = a\} \leqslant 2^{a+C}$ для всех a. Если $\Phi(x) = a$, то слово x можно закодировать следующим образом: пусть $F(i, a)$ перечисляет без повторений все слова y такие, что $\Phi(y) = a$ (предикат $[\Phi(x) = a]$ частично рекурсивен, поэтому такая функция $F(i, a)$ существует; см. теоремы 0.4 и 0.3 и определение 0.1). Запишем слово i, для