Гипергиперпростые множества, возникающие при вычислимой аппроксимации сверху префиксной сложности*

Андрей Альбертович Мучник†
Алексей Львович Семёнов‡

Вычислительный центр РАН, отделение кибернетики, 119991, Москва, ГСП-1

В статье [8] из сборника трудов международной конференции в честь 100-летия А. А. Маркова (которая прошла в августе 2003 г. в Санкт-Петербурге) авторы опубликовали ряд результатов, связывающих конструктивный математический анализ, алгоритмическую теорию информации (основанную А. Н. Колмогоровым) и понятия, введенные в [1] для решения проблемы Поста[1]. При обсуждении этих результатов на семинаре С. И. Адяна авторам были заданы вопросы, ответы на которые привели к настоящей работе, которая является продолжением статьи [8].

1. Сложность описаний

Простая энтропия натуральных чисел (введённая Колмогоровым в 1965 г.) — это минимальная длина кода относительно оптимального частично вычислимого кодирования двоичными словами.

Кодирование называется *префиксным*, если ни один код не может быть продолжением другого кода. Требование префиксности позволяет

*Работа выполнена при частичной поддержке Российского фонда фундаментальных исследований, проект N 04-01-00427, RFBR/NWO, проект N 047.017.014 и Совета поддержки научных школ при президенте РФ.

†e-mail: muchnik@lpcs.math.msu.su, fax: +(495)9156963.

‡e-mail: alsemenov@mtu-net.ru, fax: +(495)9156963.

[1]Введенные Постом понятия простого, гиперпростого и гипергиперпростого множества мы предполагаем известными. Их определения можно найти, например, в [9].

передавать последовательность закодированных сообщений без использования вспомогательного символа (например, пробела). *Префиксная энтропия* натуральных чисел (введённая Левиным в 1970 г.) — это минимальная длина кода относительно оптимального частично вычислимого префиксного кодирования двоичными словами.

Кроме того, имеются бескодовые определения (см. [6]).

Приведём формальные определения.

Простая энтропия — это минимальная с точностью до аддитивной константы перечислимая сверху функция $KS\colon \mathbb{N} \to \mathbb{N}$, для которой при любом n выполнено $|\{x\colon KS(x) < n\}| < 2^n$.

Префиксная энтропия — это минимальная с точностью до аддитивной константы перечислимая сверху функция $KP\colon \mathbb{N} \to \mathbb{N}$, для которой выполнено $\sum_x 2^{-KP(x)} \leqslant 1$.

Из определений очевидно, что функции KS и KP стремятся к бесконечности, и $\forall x\ KS(x) \leqslant KP(x) + O(1)$. Значительно сложнее доказывается

Теорема 1. *Разность* $(KP - KS)$ *стремится к бесконечности.*

Приводимое ниже доказательство было рассказано авторами на международной конференции в честь 100-летия А. Н. Колмогорова. Ли и Витаньи в своей книге [7] приводят формулировку этой теоремы без доказательства со ссылкой на неопубликованную рукопись [5] Соловэя.

Доказательство. Рассуждаем от противного. Предположим, что для некоторого C и бесконечно многих z выполнено $KP(z) - KS(z) < C$. Тогда очевидно, что такое C можно выбрать сколь угодно большим. Так как $\sum_x 2^{-KP(x)} \leqslant 1$, то для некоторого (сколь угодно большого) y при $n = KP(y)$ выполнено $\sum\limits_{x\colon KP(x) \geqslant n-2C} 2^{-KP(x)} < 2^{-2C-1}$ и $KP(y) - KS(y) < C$. Рассмотрим множество $U = \{u\colon KP(u) \leqslant m\}$. Ясно, что при известном m оно может быть перечислено. Фиксируем программу p, которая получив на вход число m, перечисляет множество U.

Докажем, что при $m = KP(y)$ имеет место $|U| < 2^{m-2C}$. Разобьем U на два подмножества U_1 и U_2. Подмножество U_1 состоит из тех $u \in U$, для которых $KP(u) < m - 2C$; подмножество U_2 состоит из тех $u \in U$, для которых $KP(u) \geqslant m - 2C$. С одной стороны, $\sum\limits_{u \in U_1} 2^{-KP(u)} \geqslant 2^{-m+2C+1} \cdot |U_1|$, с другой стороны, $\sum\limits_{u \in U_1} 2^{-KP(u)} < \sum\limits_x 2^{-KP(x)} \leqslant 1$. Поэтому $|U_1| < 2^{m-2C-1}$. Затем, с одной стороны, $\sum\limits_{u \in U_2} 2^{-KP(u)} \geqslant 2^{-m} \cdot |U_2|$, с другой

стороны, $\sum\limits_{u\in U_2} 2^{-KP(u)} < \sum\limits_{x:\ KP(x)\geqslant n-2C} 2^{-KP(x)} < 2^{-2C-1}$. Поэтому $|U_2| < 2^{m-2C-1}$. Следовательно, $|U| = |U_1|+|U_2| < 2^{m-2C-1}+2^{m-2C-1} = 2^{m-2C}$.

Положим $m = KP(y)$, тогда $y \in U$. Пусть w — двоичная запись номера, под которым программа p перечисляет y. Поскольку $|U| < 2^{m-2C}$, можно добавить к w нулевые старшие разряды и считать, что длина w равна $m - 2C$. Пусть слово v — кратчайший префиксный код числа C. Покажем, как по слову vw алгоритмически восстановить y. Во-первых, за счет префиксности кодирования отделяем начало слова vw, равное v. Во-вторых, находим C. В-третьих, сложив $2C$ с длиной w, получаем m. В-четвертых, запускаем программу p. Элемент, который будет перечислен под номером w, и есть искомое y.

Таким образом, $KS(y) < |vw| + O(1) = KP(C) + m - 2C + O(1) = KP(y) - 2C + KP(C) + O(1)$. При больших C имеем $O(1) < KP(C) < C/2$ и $KS(y) < KP(y) - C$, что противоречит предположению $KP(y) - KS(y) < C$. □

Г. Маранджан в 1969 г. доказал в статье [3], что функция KS не имеет нетривиальных частично вычислимых нижних оценок. Усиление этого утверждения доказано авторами в [8].

Теорема 2. *Для каждой частично вычислимой функции γ можно указать такую константу C, что*

$$\forall x \in Dom(\gamma)\ \ \gamma(x) \leqslant KP(x) \quad \Longrightarrow \quad \forall x \in Dom(\gamma)\ \ \gamma(x) < C.$$

Доказательство. Определим вспомогательную частично вычислимую функцию δ. На натуральном числе c значение δ определяется так. Запускаем процесс вычисления функции $\gamma(x)$ параллельно для всех x. Полагаем $\delta(c)$ равным первому x, для которого в запущенном процессе обнаружится $\gamma(x) \geqslant c$. Если такого x не обнаружится, то $\delta(c)$ остается неопределенным. Для каждого c, для которого δ определено, получаем[2]

$$KP(\delta(c)) < KP(c) + O(1) < 2\log c + O(1),$$

где $O(1)$ может быть найдено по программе вычисления δ (которая уже построена). Для найденного значения $O(1)$ можно указать такое c, для которого $c > 2\log c + O(1)$. Указанное c годится в качестве константы C из формулировки теоремы. Действительно, если $\exists x \quad \gamma(x) \geqslant C$, то $\delta(C)$ будет определено так, что $\gamma(\delta(C)) \geqslant C$. Последнее противоречит $\gamma(\delta(C)) \leqslant KP(\delta(C)) < 2\log C + O(1) < C$. □

[2]log будет обозначать логарифм по основанию 2.

В то же время функции KS и KP имеют нетривиальные вычислимые верхние оценки.

Теорема 3. *Существует всюду определённая вычислимая функция f, оценивающая сверху KS и на бесконечном множестве равная KS.*

Доказательство. Без ограничения общности будем считать областью определения функции KS множество всех двоичных слов. Через $\ell(x)$ будем обозначать длину слова x. Из мощностных соображений следует, что

$$\forall n\, \exists x\, \ell(x) = n \,\&\, KS(x) \geqslant n.$$

Поскольку функция ℓ вычислима и $|\{x\colon \ell(x) < n\}| < 2^n$, то по определению KS имеем $\exists c \forall x\, KS(x) < \ell(x) + c$.

Для каждого натурального C рассмотрим множество $M_C = \{x\colon \ell(x) + C \leqslant KS(x)\}$. Для $C = 0$ это множество бесконечно, для $C = c$ это множество пусто. С ростом C множество M_C уменьшается (нестрого). Рассмотрим наибольшее d, для которого M_d бесконечно. Ясно, что на дополнении до конечного множества M_{d+1} будет $\ell(x) + d \geqslant KS(x)$ и на бесконечном множестве $M_d \setminus M_{d+1}$ будет $\ell(x) + d = KS(x)$. Пусть функция f равна KS на конечном множестве M_{d+1} и равна $\ell(x) + d$ в остальных местах. Понятно, что f — вычислимая функция, оценивающая сверху KS и на бесконечном множестве равная KS. □

Теорема 4. *Если всюду определенная вычислимая функция f оценивает сверху KS и на бесконечном множестве равна KS, то множество $\{x\colon KS(x) < f(x)\}$ простое.*

Доказательство. Перечислимость множества $\{x\colon KS(x) < f(x)\}$ следует из перечислимости сверху функции KS и вычислимости функции f.

Предположим, что существует бесконечное перечислимое множество R, на котором $\forall x \in R$ выполняется $f(x) = KS(x)$. Пусть функция f' — не определена на дополнении до R, а на множестве R совпадает с f. Тогда f' — частично вычислимая нижняя оценка на KS. Последнее невозможно по теореме Маранджана, поскольку при каждом n множество $\{x\colon KS(x) < n\}$ конечно. □

Теорема 5. *Множество $\{x\colon KS(x) < f(x)\}$ из формулировки теоремы* 4

а) *может быть гиперпростым;*

б) *может быть не гиперпростым.*

Доказательство. Сначала докажем пункт б). Рассмотрим функцию $KS' := \min\{KS, \ell\} + 1$. Покажем, что KS' так же, как и KS, удовлетворяет определению простой энтропии. Операции «min» и «+» сохраняют свойство функции быть перечислимой сверху. Оценим мощность множества $\{x\colon KS'(x) < n\}$. Ясно, что это множество равно $\{x\colon \min\{KS(x), \ell(x)\} < n-1\} = \{x\colon KS(x) < n-1\} \cup \{x\colon \ell(x) < n-1\}$. Мощность последнего объединения меньше $2^{n-1} + 2^{n-1} = 2^n$. Наконец, $KS' \leqslant KS + 1$, за счет чего сохраняется минимальность с точностью до аддитивной константы.

Из мощностных соображений следует, что

$$\forall n\, \exists x\, \ell(x) = n \,\&\, KS(x) \geqslant n,$$

и, следовательно, $\forall n\, \exists x\, \ell(x) + 1 = n + 1 = KS'(x)$. С другой стороны, функция $f = \ell + 1$ является вычислимой верхней оценкой для KS'.

Обозначим через T_n набор всех слов длины n. Таким образом, мы получили вычислимую последовательность попарно непересекающихся наборов, каждый из которых содержит элемент, не принадлежащий множеству $\{x\colon KS'(x) < f(x)\}$.

Теперь докажем пункт а). Рассмотрим какое-нибудь гиперпростое множество R. Определим функцию KS'' так:

если $\ell(x) \notin R$, то $KS''(x) = KS'(x) + 1$;

если $\ell(x) \in R$, то $KS''(x) = KS'(x)$.

Легко проверить, что функция KS'' так же, как и KS', удовлетворяет определению простой энтропии (при этом используется только перечислимость R). Функция $g = \ell + 2$ очевидно является вычислимой верхней оценкой для KS''. Равенство $KS''(x) = g(x)$ имеет место на бесконечном множестве. А именно, слово x из этого множества существует в каждом наборе T_n при $n \notin R$.

Предположим, что t_n — вычислимая последовательность попарно непересекающихся наборов, каждый из которых содержит элемент, не принадлежащий множеству $\{x\colon KS''(x) < g(x)\}$. Конечность каждого набора T_n позволяет выбрать такую вычислимую подпоследовательность t_{n_i}, что при $i \neq j$ длина слова из t_{n_i} не может быть одновременно длиной слова из t_{n_j}. Пусть s_i — набор длин слов из t_{n_i}. Тогда s_i — вычислимая последовательность попарно непересекающихся наборов. Поскольку каждый набор t_n содержит элемент, не принадлежащий множеству $\{x\colon KS''(x) < g(x)\}$, то каждый набор s_i содержит длину, не принадлежащую множеству R. Последнее противоречит гиперпростоте R. □

Формулировки теорем 6 и 7 приведены без доказательства в книге [7] Ли и Витаньи со ссылкой на неопубликованную рукопись Соловэя [5].

Приводимые ниже доказательства были рассказаны авторами на международной конференции в честь 100-летия А. А. Маркова ([8]).

Теорема 6. *Существует всюду определённая вычислимая функция f, оценивающая сверху KP и на бесконечном множестве равная KP.*

Доказательство. Обозначим через KP_n аппроксимацию, полученную в процессе перечисления сверху функции KP. Построим вспомогательную вычислимую последовательность программ, перечисляющих множества $w_2, w_3, \ldots$ (для технического удобства мы начали нумерацию с 2). Перечисление этих множеств будет происходить так. На n-м шаге число n помещается ровно в одно множество w_i. В качестве i берется наименьший номер, для которого на всех числах x, помещенных в w_i до n-го шага, выполнено $KP_n(x) < i$. Интересующий нас номер найдется, поскольку указанному условию удовлетворяют w_i, в которые до n-го шага не было помещено ни одного элемента.

Из построения видно, что на всех элементах w_i, кроме последнего, значения KP меньше i. Из ограниченности суммы $\sum_x 2^{-KP(x)}$ следует конечность каждого w_i. Пусть y_i — последний элемент, помещенный в w_i, тогда $KP(y_i) \geqslant i$.

Рассмотрим вычислимую функцию g, сопоставляющую числу n номер того множества w_i, в которое n было помещено. Оценим сверху $\sum_n 2^{-g(n)}$. Для этого разобьем сумму на две части. В первую часть поместим слагаемые, соответствующие тем n, которые оказались последними элементами в каком-нибудь w_i. Во вторую часть поместим остальные слагаемые. Сумма слагаемых из первой части равна $\sum_{i=2}^{\infty} 2^{-i} = \frac{1}{2}$. Если слагаемое $2^{-g(n)}$ попало во вторую часть, то $2^{-g(n)} \leqslant 2^{-KP(n)-1}$. Таким образом, сумма слагаемых из второй части не превышает $\frac{1}{2}\sum_n 2^{-KP(n)} \leqslant \frac{1}{2}$. Следовательно, $\sum_n 2^{-g(n)} \leqslant 1$. Минимальность KP с точностью до аддитивной константы означает существование c, для которого $\forall n \;\; KP(n) < g(n) + c$.

Для каждого натурального числа C рассмотрим множество $M_C = \{n\colon KP(n) \geqslant g(n) + C\}$. Для $C = 0$ это множество бесконечно, для $C = c$ это множество пусто. С ростом C множество M_C уменьшается (нестрого). Рассмотрим наибольшее d, для которого M_d бесконечно. Так как $KP(n) < g(n) + d + 1 \iff KP(n) \leqslant g(n) + d$, то на дополнении до конечного множества M_{d+1} будет $KP(n) \leqslant g(n) + d$ и на бесконечном множестве $M_d \setminus M_{d+1}$ будет $KP(n) = g(n) + d$.

Искомая вычислимая верхняя оценка $f(n)$ определяется формулой $\max\{KP(n), g(n) + d\}$. □

Теорема 7. *Для произвольной всюду определённой вычислимой функции f, оценивающей сверху KP и на бесконечном множестве равной KP, множество $\{x\colon KP(x) < f(x)\}$ гиперпростое.*

Доказательство. Так как $\forall x\ \ f(x) \geqslant KP(x)$, то $\sum_x 2^{-f(x)} \leqslant 1$.

Пусть T_j — вычислимая последовательность непересекающихся отрезков натуральных чисел. Для каждого m найдём j_m, для которого $\sum_{x\in T_{j_m}} 2^{-f(x)} < 2^{-2m-1}$. Построим вспомогательную вычислимую функцию g:

на отрезках T_{j_m} положим $g(x) = f(x) - m$;

в остальных местах $g(x) = f(x) + 1$.

Оценим сверху $\sum_x 2^{-g(x)}$. Для этого разобьем сумму на две части. В первую часть поместим слагаемые с номерами из отрезков T_{j_m}. Во вторую часть поместим остальные слагаемые. Сумма слагаемых из первой части меньше $\sum_{m=1}^{\infty} 2^m \cdot 2^{-2m-1} = \frac{1}{2}$. Сумма слагаемых из второй части меньше $\frac{1}{2}\sum_x 2^{-f(x)} = \frac{1}{2}$. Таким образом, $\sum_x 2^{-g(x)} < 1$. Минимальность KP с точностью до аддитивной константы означает существование c, для которого $\forall x\ \ KP(x) < g(x) + c$.

Из определения функции g получается, что на отрезке T_{j_c} не может быть x, в котором $f(x) = KP(x)$. □

Теорема 8. *Множество $\{x\colon KP(x) < f(x)\}$ из формулировки теоремы* 7

а) *не является гипергиперпростым;*

б) *не может быть расширено до гипергиперпростого.*

Доказательство. Сначала объясним, каким образом пункт б) следует из пункта а). Пусть W — перечислимое расширение множества $\{x\colon KP(x) < f(x)\}$. Рассмотрим функцию KP', равную KP на множестве W и равную $KP + 1$ на дополнении до W. Просто проверяется, что KP' удовлетворяет требованиям на префиксную энтропию. Функция $f+1$ является вычислимой верхней оценкой KP'. Причем равенство $KP'(x) = f(x) + 1$ выполнено в точности на дополнении до W. Из пункта а) (который мы докажем для всех возможных KP и f) следует, что дополнение до W конечно.

Доказательство пункта а). Предположим, что множество $\{x\colon KP(x) = f(x)\}$ бесконечно. Обозначим через KP_n аппроксимацию, полученную в процессе перечисления сверху функции KP. Построим вычислимую последовательность программ, перечисляющих множества $w_1, w_2, \ldots$. Перечисление этих множеств будет происходить так. На n-м шаге число n

помещается ровно в одно множество w_i. В качестве i берется наименьший номер, удовлетворяющий следующему свойству:

для каждого числа $m < n$, помещенного в w_i,
если $KP_n(m) = f(m) < f(n)/3$, то $\sum\limits_{x<n, f(x)>3f(m)} 2^{-f(x)} > 2^{-2f(m)}$. $\quad (*)$

Интересующий нас номер найдется, поскольку указанному условию удовлетворяют w_i, в которые до n-го шага не было помещено ни одного элемента.

Индукцией по i докажем, что множество w_i конечно и в него попадет хотя бы одно x, для которого $KP(x) = f(x)$. Пусть для всех $j < i$ указанное свойство выполнено.

□

Теоремы 4 и 7 показывают, что для вычислимых верхних оценок f функций KS и KP множества $\{x\colon KS(x) = f(x)\}$ и $\{x\colon KP(x) = f(x)\}$ всегда «малы» с точки зрения теории алгоритмов. С точки зрения классической математики вопрос о «малости» указанных множеств естественно интепретируется как асимптотическое поведение при $k \to \infty$ функций

$$QS(k) := \frac{|\{x\colon KS(x) = f(x)\} \cap [1,k]|}{k},$$

$$QP(k) := \frac{|\{x\colon KP(x) = f(x)\} \cap [1,k]|}{k}.$$

Теорема 9. *Для простой энтропии*

нижний предел $QS(k)$ при $k \to \infty$ может быть положителен.

Теорема 10. *Для префиксной энтропии*

а) *нижний предел $QP(k)$ при $k \to \infty$ всегда равен нулю;*

б) *верхний предел $QP(k)$ при $k \to \infty$ может быть положителен.*

Доказательство. а) Рассмотрим вычислимое разбиение натурального ряда на последовательность отрезков T_j, длина которых достаточно быстро увеличивается с ростом j. Тогда по теореме 7 существует сколь угодно большое j, для которого $T_j \subset \{x\colon KP(x) < f(x)\}$. Полагая k равным правому концу T_j делаем $QP(k)$ сколь угодно малым.

б)

□

2. Заключение

Авторы очень благодарны руководителю семинара «Алгоритмические вопросы алгебры и логики» С. И. Адяну, а также В. Н. Крупскому и другим участникам указанного семинара за полезное обсуждение.

Литература

[1] *E. L. Post.* Recursively enumerable sets of positive integers and their decision problems // Bulletin of AMS. 1944. V. 50, № 5. P. 284–316.

[2] *A. N. Kolmogorov.* Three approaches to the quantative definition of information // Problems Inform. Transmission. 1965. V. 1, № 1. P. 1–7.

[3] *G. B. Marandzhyan.* On certain properties of asymptotically optimal recursive function // Izv. Akad. Nauk Armyan. SSR. 1969. V. 4. P. 3–22.

[4] *A. K. Zvonkin, L. A. Levin.* The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms // Russian Math. Surveys. 1970. V. 25, № 6. P. 83–124.

[5] *R. M. Solovay.* Lecture notes on algorithmic complexity. Unpublished, UCLA, 1975.

[6] *L. A. Levin.* Various measures of complexity for finite objects (axiomatic description) // Soviet Math. Dokl. 1976. V. 17. P. 522–526.

[7] *M. Li, P. M. B. Vitányi.* An introduction to Kolmogorov complexity and its applications. Springer, 2nd edition, 1997.

[8] *A. Semenov, An. Muchnik.* Effective Bounds for Convergence, Descriptive Complexity, and Natural Examples of Simple and Hypersimple sets // Annals of Pure and Applied Logic. 2006.

[9] *Роджерс Х.* Теория рекурсивных функций и эффективная вычислимость. М.: Мир, 1972.