Институт прикладной математики им. М. В. Келдыша
Российской академии наук
Московский государственный университет им. М.В. Ломоносова
Механико-математический факультет

ДИСКРЕТНАЯ МАТЕМАТИКА И ЕЕ ПРИЛОЖЕНИЯ

СБОРНИК ЛЕКЦИЙ МОЛОДЕЖНЫХ НАУЧНЫХ ШКОЛ ПО ДИСКРЕТНОЙ МАТЕМАТИКЕ И ЕЕ ПРИЛОЖЕНИЯМ

IV

Москва · 2007

М34
УДК 519.7

Издание осуществлено при поддержке Российского фонда фундаментальных исследований по проекту 07–01–06018

М34 Дискретная математика и ее приложения: Сборник лекций молодежных научных школ по дискретной математике и ее приложениям. Выпуск IV. Под редакцией А. В. Чашкина. — М.: ИПМ им. М.В. Келдыша РАН, 2007. — 71 с.

Четвертый выпуск лекций содержит содержит лекции, прочитанные на VI молодежной научной школе по дискретной математике и ее приложениям, проходившей в Москве в ИПМ им. М.В. Келдыша РАН с 16 по 21 апреля 2007 г. при поддержке Российского фонда фундаментальных исследований (проект 07–01–60018). Для студентов, аспирантов и научных работников в области дискретной математики и математической кибернетики.

Научное издание

ДИСКРЕТНАЯ МАТЕМАТИКА
И ЕЕ ПРИЛОЖЕНИЯ

Сборник лекций
Выпуск IV

Под общей редакцией А. В. ЧАШКИНА

Ответственный за выпуск *Ф. М. Ковалев*

МЕТОДЫ ПОСТРОЕНИЯ НЕЛИНЕЙНЫХ СОВЕРШЕННЫХ ДВОИЧНЫХ КОДОВ (обзор работ)

А. М. Романов

Институт математики им. С. Л. Соболева СО РАН,
630090 г. Новосибирск, пр. ак. Коптюга, 4

e-mail: rom@math.nsc.ru

Теория совершенных кодов — область, которая находится на стыке теории кодирования и теории дизайнов или t-схем и является трудной для исследования. Линейные совершенные коды были построены М. Голеем и Р. Хеммингом в конце 40-х годов прошлого века. Нелинейные совершенные коды были открыты Ю. Л. Васильевым в 1961 году. В настоящее время известно достаточно много различных методов построения совершенных кодов. В работе представлен обзор методов построения нелинейных совершенных двоичных кодов.

Пусть $\mathbb{F}_2^n$ — векторное пространство размерности n над полем Галуа $GF(2)$. Произвольное подмножество $\mathbb{C} \subseteq \mathbb{F}_2^n$ называется *двоичным кодом* длины n. Векторы, принадлежащие коду, называются *кодовыми словами*. *Расстоянием Хемминга* $d(\mathbf{x}, \mathbf{y})$ между векторами $\mathbf{x} \in \mathbb{F}_2^n$ и $\mathbf{y} \in \mathbb{F}_2^n$ называется число координат, в которых векторы $\mathbf{x}$ и $\mathbf{y}$ различаются. Минимально возможное расстояние d между двумя различными кодовыми словами называется *минимальным расстоянием* кода. *Радиусом упаковки* ρ кода $\mathbb{C}$ называется величина $\rho(\mathbb{C}) = \frac{d-1}{2}$. *Радиус покрытия* кода $\mathbb{C}$ равен $r(\mathbb{C}) = \max\limits_{\mathbf{x} \in \mathbb{F}_2^n} \min\limits_{\mathbf{c} \in \mathbb{C}} d(\mathbf{x}, \mathbf{c})$. Код $\mathbb{C}$ называется *совершенным*, если $r(\mathbb{C}) = \rho(\mathbb{C})$. Совершенный код образует совершенную упаковку (также совершенное покрытие) шарами Хемминга радиуса ρ; при этом центрами этих шаров являются кодовые слова. Следовательно, множество $\mathbb{C}$ является совершенным кодом с минимальным расстоянием $d = 2\rho + 1$ тогда и только тогда, когда

для каждого вектора $\mathbf{x} \in \mathbb{F}_2^n$ существует единственное кодовое слово $\mathbf{c} \in \mathbb{C}$ такое, что $d(\mathbf{x}, \mathbf{c}) \leqslant \rho$.

Коды $\mathbb{C}_1, \mathbb{C}_2 \subseteq \mathbb{F}_2^n$ называются *изоморфными*, если существует перестановка координат π такая, что $\mathbb{C}_2 = \{\pi(\mathbf{c}) \mid \mathbf{c} \in \mathbb{C}_1\}$. Коды $\mathbb{C}_1$ и $\mathbb{C}_2$ называются *эквивалентными*, если существует вектор $\mathbf{x} \in \mathbb{F}_2^n$ и перестановка π такие, что $\mathbb{C}_2 = \{\pi(\mathbf{c}) + \mathbf{x} \mid \mathbf{c} \in \mathbb{C}_1\}$. Код называется *линейным*, если его слова образуют линейное подпространство в $\mathbb{F}_2^n$. Линейные совершенные коды с минимальным расстоянием $d = 3$ называются *кодами Хемминга* [38]. С точностью до эквивалентности существует единственный двоичный код Хемминга длины n.

Числа n, M, d называются *параметрами кода*, если его длина равна n, мощность — M, минимальное расстояние — d. Известно, что совершенные двоичные коды с параметрами кодов Хемминга существуют только при $n = 2^s - 1, \quad s = 2, 3, \ldots$. В данной статье будут рассматриваться именно такие коды. Будем предполагать (если не оговорено обратное), что нулевой вектор всегда принадлежит коду. Все нелинейные совершенные двоичные коды имеют параметры кодов Хемминга и существуют при $n \geqslant 15$.

В [37, 36], а также независимо в [7] установлено, что кроме совершенных двоичных кодов с параметрами кодов Хемминга, двоичного кода Голея длины $n = 23$ с минимальным расстоянием 7, троичного кода Голея длины $n = 11$ с минимальным расстоянием 5 и тривиальных кодов (код из одного слова, полный код и двоичный код с повторением нёчетной длины) никакие другие совершенные коды над полями Галуа не существуют. В [35, 18] с точностью до эквивалентности доказана единственность кодов Голея.

У. Хеден [21] обнаружил, что среди совершенных кодов, построенных методом конкатенации (которая определяется перестановкой и тривиальными разбиениями $\mathbb{F}_2^n$ на совершенные коды длины n), существует совершенный код длины $n = 15$, неэквивалентный ни одному из кодов Васильева. Ф. И. Соловьёва [16] привела примеры нетривиальных разбиений $\mathbb{F}_2^n$ на совершенные коды длины n и показала, что из этих нетривиальных разбиений методом конкатенации можно построить совершенные коды, неэквивалентные кодам Васильева и кодам Хедена [21]. К. Фелпс [31] также привёл примеры нетривиальных разбиений $\mathbb{F}_2^n$ на совершенные коды. В [32] он описал конструкцию совершенных двоичных кодов, в которой конкатенация определяется m-арной квазигруппой (перестановку можно рассматривать как некоторую унарную квазигруппу). В [17] построены три совершенных кода длины $n = 15$, которые неэквивалентны кодам Васильева. М. Моллар [29] обобщил конструкцию Васильева. В. А. Зиновьев и А. Лобстейн [28, 8] предложили каскадные конструкции совершенных кодов. Вариации кас-

кадных конструкций можно найти в более ранних работах В. А. Зиновьева, например, в [3]. У. Хеден в [22] построил совершенные коды длины $n = 15$ с ядрами размерности $1, 2, 3$ и в [23] — совершенные коды полного ранга длины $n = 31$ с ядром размерности 21.

Т. Этцион и А. Варди [25] предложили некоторые упорядоченные семейства подмножеств из $\mathbb{F}_2^n$ и назвали их совершенными сегментациями. Используя эти сегментации, они методом конкатенации построили совершенные двоичные коды и показали, что среди этих кодов содержатся новые коды длины $n = 15$. В этой же работе методом сдвига компонент кода Хемминга или свитчингами они также построили совершенные двоичные коды полного ранга с тривиальным ядром и показали, что совершенные двоичные коды полного ранга не могут быть построены конкатенацией.

Свитчинги в совершенных двоичных кодах были открыты Ю. Л. Васильевым [2]. Свитчинговые методы широко известны в комбинаторике. Так, например, в системах Штейнера известны свитчинги Паша (Pash). Существует всего 80 попарно неизоморфных систем троек Штейнера порядка 15, все они занумерованы некоторым фиксированным образом [38] (таблица, в которой перечисляются системы троек Штейнера, также содержится в более доступной работе [27]). Как установлено в [19], эти 80 систем троек Штейнера разбиваются на два свитчинговых класса. Один класс содержит системы от № 1 до № 79. Другой класс состоит из одной системы № 80. Что касается числа свитчинговых классов, на которые разбивается множество совершенных двоичных кодов длины n, в настоящее время наиболее изученным является свитчинговый класс кода Хемминга. Кроме того, в [34] приведён пример двух совершенных кодов длины $n = 15$, которые не принадлежат свитчинговому классу кода Хемминга и образуют собственный свитчинговый класс, состоящий из двух кодов. На самом деле, по-видимому, множество совершенных двоичных кодов дины $n = 15$ разбивается на несколько тысяч свитчинговых классов.

Следует заметить, что в [30] недавно решена известная проблема о пополнении характеристических векторов, соответствующих системам троек Штейнера, до совершенных кодов. В ней показано, что характеристические векторы, соответствующие системам троек Штейнера № 79 и № 80, не могут принадлежать ни одному совершенному коду длины $n = 15$.

К. Фелпс и М. Ле Ван [33] заметили, что подмножества кода Хемминга, которые сдвигали Т. Этцион и А. Варди [25] и которые мы называем компонентами, являются смежными классами некоторых подпространств. Используя групповые свойства компонент кода Хемминга, они неконструктивными методами доказали существование в коде Хемминга непересека-

ющихся компонент, отвечающих различным координатам, и тем самым доказали существование совершенных двоичных кодов со всеми допустимыми размерностями ядер. С. В. Августинович и Ф. И. Соловьёва [1] обратили внимание на то, что если в коде Хемминга сдвинуть n непересекающихся компонент по n различным направлениям, то код Хемминга превратится в несистематический код. С. А. Малюгин [9] показал, что для превращения кода Хемминга в несистематический код в нём достаточно сдвинуть 7 компонент. В [13] получены достаточные условия непересекаемости компонент кода Хемминга и построены регулярные разбиения кода Хемминга на компоненты. С использованием этих условий в [14] построены несистематические коды длины $n = 15$, а в [15] — семейство непересекающихся компонент, которое даёт коды полного ранга с тривиальным ядром. В [12] получен критерий непересекаемости компонент двоичного кода Хемминга и построены регулярные разбиения кодов Хемминга на компоненты с новыми параметрами. В [26] построены совершенные двоичные коды полного ранга с большими размерностями ядер исходя из мозаичных замощений $\mathbb{F}_2^n$.

Пусть $N(n)$ — число попарно неэквивалентных совершенных двоичных кодов длины n. Тогда $2^{2^{(0,5+o(1))n}} \leqslant N(n) \leqslant 2^{2^{(1+o(1))n}}$. Нижняя оценка получена Ю. Л. Васильевым [2] и неоднократно передоказана многими авторами. Верхняя оценка является тривиальной. Далее говоря о числе неэквивалентных кодов, мы будем иметь ввиду попарную неэквивалентность.

В настоящее время неизвестно даже число неэквивалентных совершенных двоичных кодов длины $n = 15$. Известные оценки числа неэквивалентных совершенных двоичных кодов длины $n = 15$ и числа неэквивалентных расширенных совершенных двоичных кодов длины $n = 16$ приведены в таблице 1.

Т а б л и ц а 1

	$n = 15$	$n = 15$	$n = 16$	$n = 16$
Ранг 11	1	1	1	1
Ранг 12	18	18	12	12
Ранг 13	758	758	272	272
Ранг 14	?	?	1719	?
Ранг 15	?	51	?	51

Во второй колонке таблицы 1 перечислены неэквивалентные совершенные двоичные коды длины $n = 15$ и ранга 11, 12, 13. Число неэквивалентных совершенных двоичных кодов ранга 14 и 15 неизвестно. Кроме того, в

настоящее время не известны никакие теоретические методы, с помощью которых можно было бы построить коды полного ранга длины $n = 15$, не принадлежащие свитчинговому классу кода Хемминга. В третьей колонке таблицы 1 перечислены неэквивалентные совершенные двоичные коды, принадлежащие свитчинговому классу кода Хемминга длины $n = 15$. В четвёртой и пятой колонках соответственно перечислены неэквивалентные расширенные совершенные двоичные коды длины $n = 16$ и неэквивалентные расширенные совершенные двоичные коды, принадлежащие свитчинговому классу расширенного кода Хемминга длины $n = 16$. Приведённые оценки заимствованы из работ [24, 4–6, 10, 11]. В работе [11] приводится нижняя оценка числа неэквивалентных кодов полного ранга, которая по утверждению С. А. Малюгина является точной. Как видно из таблицы 1 число неэквивалентных расширенных совершенных двоичных кодов ранга 14 равно 1719. Все такие коды строятся методом конкатенации; из них 844 кода — исходя из разбиений, 875 кода — исходя из совершенных сегментаций и обобщений [6]. Ранг кода Хемминга длины n равен $n - s$. Можно показать, что все совершенные коды длины n и ранга $n - s + 1$, $n - s + 2$ принадлежат свитчинговому классу кода Хемминга длины n. Вполне вероятно, что в ближайшее время все совершенные двоичные коды длины $n = 15$ будут перечислены при помощи компьютера.

Литература

1. Августинович С. В., Соловьёва Ф. И., О несистематических совершенных двоичных кодах // Проблемы передачи информации. 1996. Т. 32, вып. 3. С. 47–50.
2. Васильев Ю. Л., О негрупповых плотно упакованных кодах // Проблемы кибернетики. Вып. 8. М.: Физматгиз, 1962. С. 337–339.
3. Зиновьев В. А., Коды для корреляционной многоадресной селекции. Дис. ... канд. техн. наук. М., 1970.
4. Зиновьев В. А., Зиновьев Д. В., Двоичные расширенные совершенные коды длины 16, построенные обобщённой каскадной конструкцией // Проблемы передачи информации. 2002. Т. 38, вып. 4. С. 56–84.
5. Зиновьев В. А., Зиновьев Д. В., Двоичные совершенные коды длины 15, построенные обобщённой каскадной конструкцией // Проблемы передачи информации. 2004. Т. 40, вып. 1. С. 27–39.
6. Зиновьев В. А., Зиновьев Д. В., Двоичные расширенные совершенные коды длины 16 ранга 14 // Проблемы передачи информации. 2006. Т. 42, вып. 2. С. 63–80.

7. Зиновьев В. А., Леонтьев В. К., Несуществование совершенных кодов над полями Галуа // Проблемы управления и теории информации. 1973. Т. 2, № 2. С. 123–132.

8. Зиновьев В. А., Лобстейн А. С., Об обобщённых каскадных конструкциях совершенных двоичных нелинейных кодов // Проблемы передачи информации. 2000. Т. 36, вып. 4. С. 59–73.

9. Малюгин С. А., Несистематические совершенные двоичные коды // Дискрет. анализ. и исслед. операций. Сер. 1. 2001. Т. 8, № 1. С. 55–76.

10. Малюгин С. А., О классах эквивалентности совершенных двоичных кодов длины 15 // Новосибирск, 2004. 34 с. (Препринт / РАН, Сиб. отд-ние. Институт математики; № 138).

11. Малюгин С. А., О перечислении неэквивалентных совершенных двоичных кодов длины 15 и ранга 15 // Дискрет. анализ. и исслед. операций. Сер. 1. 2006. Т. 13, № 1. С. 77–98.

12. Малюгин С. А., Романов А. М., О разбиениях кодов Хемминга на непересекающиеся компоненты // Дискрет. анализ. и исслед. операций. Сер. 1. 2002. Т. 9, № 1. С. 42–48.

13. Романов А. М., О построении совершенных нелинейных двоичных кодов инверсией символов // Дискрет. анализ. и исслед. операций. Сер. 1. 1997. Т. 4, № 1. С. 46–52.

14. Романов А. М., О несистематических совершенных кодах длины 15 // Дискрет. анализ. и исслед. операций. Сер. 1. 1997. Т. 4, № 4. С. 75–78.

15. Романов А. М., Совершенные двоичные коды с тривиальным ядром // Дискрет. анализ. и исслед. операций. Сер. 1. 2000. Т. 7, № 2. С. 71–74.

16. Соловьёва Ф. И., О двоичных негрупповых кодах // Методы дискретного анализа в изучении булевых функций и графов. Сб. науч. тр. Вып. 37. Новосибирск: Ин-т математики, 1981. С. 65–76.

17. Bauer H., Ganter B., Hergert F., Algebraic techniques for nonlinear codes // Combinatorica. 1983. V. 3, N 1. P. 21–33.

18. Delsarte P., Goethals J. M., Unrestricted codes with the Golay parameters are unique // Discrete Math. 1975. V. 12, N 3. P. 211–224.

19. Gibbons P. B., Computing techniques for the construction and analysis of block designs. Ph.D. Thesis, Department of Computer Science, University of Toronto, 1976.

20. Hamming R. W., Error detecting and error correcting codes // Bell System Tech. J. 1950. V. 29, N 2. P. 147–160.

21. Heden O., A new construction of group and nongroup perfect codes // Inform. and Control. 1977. V. 34, N 4. P. 314–323.

22. Heden O., A binary perfect code of length 15 and codimension 0 // Designs, Codes and Cryptog. 1994. V. 4, N 3. P. 213–220.

23. Heden O., A full rank perfect code on length 31 // Designs, Codes and Cryptog. 2006. V. 38, N 1. P. 125–129.

24. Hergert F., The equivalence classes of the Vasil'ev codes of length 15 // Combinatorial Theory. Berlin: Springer, 1982. P. 176–186. (Lectures Notes in Math. V. 969).

25. Etzion T., Vardy A., Perfect binary codes: constructions, properties, and enumeration // IEEE Trans. on Inform. Theory. 1994. V. 40, N 3. P. 754–763.

26. Etzion T., Vardy A., On perfect codes and tilings: problems and solution // SIAM J. Discrete Math. 1998. V. 11, N 2. P. 205–253.

27. Limbos M., Projective embeddings of small "Steiner triple systems"// Ann. Discrete Math. 1980. V. 7. P. 151–173.

28. Lobstein A. S., Zinoviev V. A., On new perfect binary nonlinear codes // Applicable Algebra in Engineering, Communication and Computing. 1997. V. 8, N 5. P. 415–420.

29. Mollard M., A generalized parity function and its use in construction of perfect codes // SIAM J. Algebraic Discrete Methods. 1986. V. 7, N 1. P. 113–115.

30. Östergård P. R. J., Pottonen O., The exist Steiner triple systems of order 15 that do not occur in a perfect binary one-error-correcting code // J. of Combinatorial Designs, to appear.

31. Phelps. K. T., A combinatorial construction of perfect codes // SIAM J. Algebraic Discrete Methods. 1983. V. 4, N 3. P. 398–403.

32. Phelps K. T., A general product construction for error correcting codes // SIAM J. Algebraic Discrete Methods 1984. V. 5, N 2. P. 224–228.

33. Phelps K. T., LeVan M., Kernels of nonlinear Hamming codes // Designs Codes and Cryptogr. 1995. V. 6, N 3. P. 247–257.

34. Phelps K. T., LeVan M., Switching equivalence classes of perfect codes // Designs Codes and Cryptogr. 1999. V. 16, N 2. P. 179–184.

35. Pless V., On the uniqueness of the Golay codes // J. Combin. Theory. 1968. V. 5, N 3. P. 215–228.

36. Tietäväinen A., On the nonexistence of perfect codes over finite fields // SIAM J. Applied Math. 1973. V. 24, N 1. P. 88–96.

37. van Lint J. H., Introduction to coding theory. New York–Berlin: Springer-Verlag, 1982.

38. White H. S., Cole F. N. Cummings L. D., Complete classification of triad systems on fifteen elements // Memoirs Nat. Acad. Sci. USA. 1919. V. 14. P. 1–89.

О МЕТОДЕ КОНТЕЙНЕРОВ

А. А. Сапоженко

**Московский государственный университет
им. М. В. Ломоносова,
факультет вычислительной математики и кибернетики,
119992 Москва, Ленинские горы**

e-mail: sapozhenko@mail.ru

1. Введение

Предлагается метод решения перечислительных задач. Рассматриваемый метод не предполагает наличия разрешимых рекуррентных соотношений, которые оказываются необходимыми при использовании традиционного метода производящих функций. Целью метода является получение асимптотически совпадающих верхних и нижних оценок. Идея метода состоит в том, чтобы построить семейство подмножеств, называемых контейнерами, такое, что каждый из перечисляемых объектов содержится хотя бы в одном из контейнеров. Если при этом некоторое подсемейство содержит почти все объекты, то оценивая число объектов в этой подсистеме, удается получить достаточно близкие верхние и нижние оценки. Объектами, для которых удается получить такие оценки, являются независимые множества в графах, антицепи в частично упорядоченных множествах, множества, свободные от сумм, в группах и множестве натуральных чисел и др. Мы демонстрируем метод контейнеров на примере получения верхних оценок для числа независимых множеств в графах. Полученные оценки используются в дальнейшем для получения асимптотик при рассмотрении других упомянутых выше объектов.

2. Определения

Рассматриваемые в статье графы являются конечными неориентированными и простыми (без петель и кратных ребер). Вершины графа считаются занумерованными. Степень вершины v обозначается через $\sigma(v)$.

Подмножество вершин графа G называется *независимым*, если никакие две его вершины не соединены ребром в G. Семейство всех независимых множеств графа G обозначим через $\mathcal{I}(G)$. Положим $I(G) = |\mathcal{I}(G)|$. Пусть $G = (V; E)$ — граф с множеством вершин V и множеством ребер E, и $v \in V$. Множество $\partial v = \{u : (u, v) \in E\}$ назовем *границей* вершины v. Ясно, что $\sigma(v) = |\partial v|$. *Границей* множества $A \subseteq V$ в графе $G = (V; E)$ назовем множество $\partial A = \left(\bigcup_{v \in A} \partial v\right) \setminus A$. Подмножество A вершин графа $G = (V, E)$ называется *доминирующим*, если $V = A \cup \partial A$. Семейство всех доминирующих независимых множеств обозначим через $\mathcal{D}(G)$. Пусть $0 \leq \delta < 1$. Граф G называется *δ-расширителем*, если $|A| \leq |\partial A|(1-\delta)$ для любого его независимого множества A.

3. Метод контейнеров

Метод контейнеров предназначен для получения верхних оценок и асимптотик мощности семейств множеств с заданными свойствами. Примерами являются семейства независимых множеств и клик в графах, множеств, свободных от сумм, в группах, антицепей в частично упорядоченных множествах, кодов с заданным минимальным расстоянием и т. п.

Пусть задано некоторое свойство, например, свойство независимости множеств вершин графа. Оно определяет семейство $\mathcal{I}(G)$ независимых множеств этого графа. Как оценить сверху $I(G) = |\mathcal{I}(G)|$? Идея состоит в том, чтобы найти другое семейство $\mathcal{F}$ вершин графа G, такое, что для каждого $A \in \mathcal{I}(G)$ существует $B \in \mathcal{F}$, такое, что $A \subseteq B$. Элементы семейства $\mathcal{F}$ назовем *контейнерами*, а само $\mathcal{F}$ – *системой контейнеров* для $\mathcal{I}(G)$. Положим $L(B) = |\{A \subseteq B : A \in \mathcal{I}(G)\}|$. Очевидны следующие неравенства

$$|I(G)| \leq \sum_{B \in \mathcal{F}} L(B) \leq \sum_{B \in \mathcal{F}} 2^{|B|}. \tag{1}$$

Неравенства представляются довольно грубыми. Однако в ряде случаев они оказываются достаточно точными. Это, чаще всего бывает тогда, когда свойство, определяющее семейство, мощность которого оценивается (обозначим его через $\mathcal{A}$), оказывается наследственным, т. е. когда из того, что $A \in \mathcal{A}$ и $C \subseteq A$ вытекает включение $C \in \mathcal{A}$. Заметим, что все упомянутые выше свойства являются наследственными. Например, независимое множество или клика сохраняют свойства быть независимым множество или кликой после удаления элементов.

Таким образом для получения точных верхних оценок некоторого семейства $A \in \mathcal{A}$ требуется найти систему контейнеров $\mathcal{F}$ как можно меньшей мощности, такую чтобы контейнеры "слабо" пересекались, а *плот-*

ность $\lambda(B)$ каждого контейнера $B \in \mathcal{F}$, определяемая равенством $\lambda(B) = L(B)2^{-|B|}$, была как можно ближе к 1.

Ниже мы показываем, пример построения подходящей системы контейнеров для семейства $\mathcal{I}(G)$ независимых множеств регулярного графа G. Далее даем краткий обзор верхних оценок числа независимых множеств для некоторых классов графов. Почти все эти оценки получены методом контейнеров. Кроме того мы показываем, как с помощью этого метода можно получить не только верхние оценки, но и асимптотики числа множеств, свободных от сумм. Эти результаты получены сведением задачи к оценке $I(G)$ для подходящих графов Кэли. Упомянутые сведения оказываются достаточно нетривиальными. Примеры простых сведений демонстрируются в следующем параграфе.

4. Сведения задач к подсчету числа независимых множеств

Простейшее из рассматриваемых сведений касается подсчета числа клик (т. е. полных подграфов) в графе. Для того чтобы найти число клик в графе G достаточно посчитать $I(\bar{G})$ для графа $\bar{G}$, дополнительного к G.

Подсчет числа двоичных кодов с расстоянием r сводится к подсчету числа независимых множеств в графе, вершинами которого являются двоичные векторы длины n, а ребрами — пары вершин, находящихся на расстоянии, меньшем чем r.

Асимптотика числа антицепей в унимодальных частично упорядоченных множествах была найдена автором в [11]. Ранее подобная задача была решена для частного случая n-мерного куба А. Д. Коршуновым в [19]. Задача подсчета числа антицепей в произвольном частично упорядоченном множестве $P = (X, Z, Y)$ сводится подсчету числа независимых множеств в графе G_Q, в котором множеством вершин является P, а вершины соединены ребром, если одна из них предшествует другой.

5. О числе независимых множеств в графах

В этом параграфе мы получим верхние оценки для числа независимых множеств в графах. Доказательство основывается на построении системы контейнеров для семейства $\mathcal{I}(G)$ независимых множеств графа G. Очевидно, что системой контейнеров для $\mathcal{I}(G)$ является семейство $\mathcal{D}(G)$ всех доминирующих независимых множеств. Однако такая система неудобна по ряду причин. Построение семейства $\mathcal{D}(G)$ и получение верхней оценки его мощности обычно не проще, чем решение аналогичных проблем для $\mathcal{I}(G)$. Кроме того, размер семейства $\mathcal{D}(G)$ может оказаться сравнимым с

размером семейства $\mathcal{D}(G)$, что неприемлемо при нашем подходе. Следующая лемма дает способ построения подходящей системы контейнеров для семейства $\mathcal{I}(G)$ в случае регулярных графов. Регулярный граф степени k на n вершинах назовем (n,k)-*графом*.

Лемма 1. *Пусть $G=(V;E)$ является $(n,k)-$графом, $k>1$, $A\in\mathcal{I}(G)$ и $0<\varphi<k$. Тогда существует $T\subseteq A$, такое, что*:

$$|T|\le|\partial A|/\varphi, \tag{2}$$

$$A\subseteq D,\quad \textit{где}\quad D=D(T,\varphi)=\{v\in V\setminus\partial T:\ |\partial v\setminus\partial T|<\varphi\}, \tag{3}$$

$$|D|\le|\partial T|\,\frac{k}{k-\varphi}. \tag{4}$$

Доказательство. Искомое множество T построим с помощью следующей пошаговой процедуры

Шаг 1. Пусть u_1 — произвольная вершина из A. Положим $T_1=\{u_1\}$. Предположим, что m шагов сделаны и построено множество $T_m=\{u_1,\dots,u_m\}$.

Шаг $m+1$. Если существует $u_{m+1}\in A$ такая, что $|\partial u_{m+1}\setminus\partial T_m|\ge\varphi$, полагаем $T_{m+1}=T_m\cup\{u_{m+1}\}$ и продолжаем процесс.

В противном случае процедура закончена, а ее результатом является множество $T=T_m$. Неравенство (2) и включение (3), вытекают из построения множества T. Неравенство (4) — из того, что $|\partial v\cap\partial T|\ge k-\varphi$ для всякого $v\in D$ и $|\partial u|\le k$ для всякого $u\in\partial T$. □

Следствие 1. *Пусть G является $(n,k)-$графом, $1\le\varphi<k$. Тогда существует система контейнеров $\mathcal{F}$ для $\mathcal{I}(G)$, такая, что*:

(i)

$$|\mathcal{F}|\le\sum_{i\le n/\varphi}\binom{n}{i}, \tag{5}$$

(ii) для любого $D\in F$

$$|D|\le nk/(2k-\varphi). \tag{6}$$

Доказательство. По лемме каждое $A\in\mathcal{I}(G)$ содержит подмножество T мощности, не превышающей $|\partial A|/\varphi$, причем $A\subseteq D(T,\varphi)$. Заметим, что контейнер $D(T,\varphi)$ однозначно определяется множеством T при фиксированном φ. таким образом число контейнеров не превосходит $|\partial A|/\varphi\le n/\varphi$. Отсюда вытекает (i).

Неравенство (ii) следует из (4) ввиду того, что $|D|\le n-|\partial T|$. □

Следующая теорема улучшает остаточный член в аналогичном результате Н. Алона [16] и является иллюстрацией нашего подхода к получению верхних оценок.

Теорема 1. *Для любого (n,k)-графа*

$$I(G) \le 2^{\frac{n}{2}\left(1+O\left(\sqrt{(\log k)/k}\right)\right)}. \tag{7}$$

Доказательство. Неравенство (7) вытекает из следствия. Независимые множества могут быть перечислены следующим образом. Зафиксируем $\varphi = \sqrt{k\log k}$. Выберем произвольное множество $T \subseteq V$ и построим $D = D(T,\varphi)$. Далее выберем $A \subseteq D$. Заметим, что

$$|D| \le nk/(2k-\varphi) \tag{8}$$

ввиду (6). Следовательно, положив $\varphi = \sqrt{k\log k}$, имеем

$$\begin{aligned} I(G) &\le \sum_{T\subseteq V, |T|\le n/\varphi} 2^{|D(T,\varphi)|} \le \sum_{i\le n/\varphi} \binom{n}{i} 2^{nk/(2k-\varphi)} \\ &\le 2^{\frac{n}{2}\left(1+O\left(\sqrt{(\log k)/k}\right)\right)}. \end{aligned} \tag{9}$$

Теорема 1 обобщается на случай нерегулярных графов (см. [7], [9]). Необходимость подобных обобщений продиктована приложениями к комбинаторным проблемам алгебры и теории чисел (см. [13], [14], [18], [20]).

G называется (n,k,θ)-*графом*, если в нем n вершин и при этом $k \le \sigma(v) \le k+\theta$ для всякой вершины v. По определению $(n,k,0)$-граф является (n,k)-графом.

Теорема 2. *Для всякого (n,k,θ)-графа* Γ

$$I(\Gamma) \le 2^{\frac{n}{2}\left(1+O\left(\theta/k+\sqrt{(\log k)/k}\right)\right)}. \tag{10}$$

Теорема 2 обобщает теорему 1 на случай (n,k,θ)-графов. Обозначим через $I_\beta(\Gamma)$ число множеств $A \in \mathbf{I}(\Gamma)$, удовлетворяющих неравенству $\big||A| - n/4\big| \ge \beta n/4$.

Теорема 3. *Пусть $\Gamma = (V;E)$ является (n,k,θ)-графом и $0 < \beta < 1$. Тогда*

$$I_\beta(\Gamma) \le 2^{\frac{n}{2}\left(1-\frac{\beta^2}{2\ln 2}+O\left(\frac{\theta}{k}+\sqrt{\frac{\log k}{k}}\right)\right)}. \tag{11}$$

Теорема 4. *Пусть (n,k,θ)-граф $\Gamma=(V;\,E)$ является δ-расширителем и $0\le\delta<1$. Тогда*

$$I(\Gamma)\le 2^{\frac{n}{2}\left(1-\delta/7+O\left(\theta/k+\sqrt{(\log k)/k}\right)\right)}. \tag{12}$$

Теоремы 3 и 4 позволяет получать верхние оценки вида $I(G)\le 2^{n(1/2-c)}$, где $c>0$ некоторая константа. В ряде случаев такие оценки оказываются полезными (см., например, [13]).

Дальнейшие результаты в этом направлении получены в [9]. Пусть l,k,θ,n удовлетворяют неравенствам $l\le k-\theta\le k+\theta\le n$. Граф с n вершинами называется $(n,l,k,m,\delta,\varepsilon,\theta)$-*графом*, если удовлетворяются следующие условия: минимальная степень вершины не меньше l, максимальная степень вершины не больше m, доля вершин степени превышающей $k+\theta$, не больше ε, доля вершин степени меньшей чем $k-\theta$, не больше δ. В [14] доказана следующая

Теорема 5. *Пусть $G=(V,E)$ является $(n,l,k,m,\delta,\varepsilon,\theta)$-графом. Тогда существует система $\mathcal{B}$ контейнеров для $\mathcal{I}(G)$ удовлетворяющая следующим условиям*:

1) *Для всякого $B\in\mathcal{B}$*

$$|B|\le n\frac{k+\delta(k-l)+\varepsilon(m-k)+\theta}{2k-\sqrt{k\log k}}; \tag{13}$$

2) *Для $k>3$ и для достаточно больших n*

$$|\mathcal{B}|\le 2^{n\sqrt{\frac{\log k}{k}}}. \tag{14}$$

Кроме того для достаточно больших n

$$I(G)\le 2^{\frac{n}{2}\left(1+\delta\left(1-\frac{l}{k}\right)+\varepsilon\left(\frac{m}{k}-1\right)+O\left(\frac{\theta}{k}+\sqrt{\frac{\log k}{k}}\right)\right)}. \tag{15}$$

Следующие две теоремы, доказанные в [9], являются обобщениями теорем 3 и 4.

Теорема 6. *Пусть G является $(n,l,k,m,\delta,\Delta,\theta)$-графом и $\gamma=\delta\left(1-l/k\right)+\Delta\left(m/k-1\right)+O\left((\theta+\sqrt{k\log k})/k\right)$. Тогда для достаточно больших n*

$$I_\beta(G)\le 2^{\frac{n}{2}\left(1-(\beta-\gamma)^2/(2(1+\gamma)\ln 2)\right)}. \tag{16}$$

Теорема 7. *Пусть $(n, l, k, m, \delta, \Delta, \theta)$-граф G является ϵ-расширителем. Тогда для некоторого $c > 0$ и достаточно больших n*

$$I(G) \le 2^{\frac{n}{2}\left(1-c\epsilon+\delta(1-l/k)+\Delta\left(m/k-1\right)+O\left(\left(\theta+\sqrt{k\log k}\right)/k\right)\right)} \tag{17}$$

Теоремы 2–4 использовались при оценке числа множеств, свободных от сумм, в группах (см. [6] и [13]). Теорема 5 применялась для доказательства гипотезы Камерона-Эрдёша [14]. Теоремы 2–7 используют только идею контейнеров в том виде, в котором она работала в доказательстве Теоремы 1.

Дальнейшие результаты в этом направлении получены с помощью дополнительных соображений. Обозначим через $\alpha(G)$ максимальный размер независимого множества в графе G. В. Е. Алексеев [2] доказал следующую оценку.

Теорема 8. *Для всякого графа G на n вершинах, такого, что $\alpha(G) = \mu$,*

$$i(G) \le \left(\frac{n}{\mu}+1\right)^{\mu}. \tag{18}$$

Эта верхняя оценка достигается на полном и пустом графах. Во многих других случаях она оказывается довольно грубой, если на рассматриваемый класс графов накладываются ограничения. Так, например, для (n, k)-графов с помощью соображений из доказательства теоремы леммы 1 удается получить более точные оценки. В [9] доказаны следующие утверждения.

Теорема 9. *Пусть граф G на n вершинах является регулярным степени k, $\alpha(G) = \mu$. Тогда*

$$i(G) \le 2^{\mu\log\left(1+\frac{n}{2\mu}\right)+O\left(n\sqrt{k^{-1}\log k}\right)}. \tag{19}$$

Теорема 10. *Пусть G является $(n, l, k, m, \delta, \Delta, \theta)$-графом, $\alpha(G) = \mu$, а n и k — достаточно велики. Тогда*

$$i(G) \le 2^{\mu\log\left(1+\frac{n}{2\mu}\right)+n\left(\delta(1-l/k)+\Delta\left(m/k-1\right)+O\left(\left(\theta+\sqrt{k\log k}\right)/k\right)\right)}. \tag{20}$$

Теорема 11. *Пусть G является $(n, l, k, m, \delta, \Delta, \theta)$-графом и ε-расширителем, а n и k — достаточно велики. Тогда*

$$i(G) \le 2^{\frac{n}{2}\left(\frac{2-2\varepsilon}{2-\varepsilon}\log\frac{4-3\varepsilon}{2-2\varepsilon}+\delta(1-l/k)+\Delta\left(m/k-1\right)+O\left(\left(\theta+\sqrt{k\log k}\right)/k\right)\right)}. \tag{21}$$

6. Двудольные графы

Двудольный граф $G = (X, Z; E)$ с долями вершин X, Z и множеством ребер E назовем (ϵ, δ)-*расширителем*, если $|A| \le (1-\delta)|\partial A|$ для всякого $A \subseteq X$ такого, что $|A| \le \epsilon\,|X|$ и для всякого $A \subseteq Z$ такого, что $|A| \le \epsilon\,|Z|$. В случае двудольных графов метод контейнеров позволяет получить асимптотику, если граф является расширителем.

Первый результат в этом направлении, полученный с помощью идеи контейнеров, касался n-мерного куба. Через B^n обозначим граф n-мерного куба, вершинами которого являются двоичные наборы длины n, а ребрами — пары наборов, отличающихся ровно в одном разряде. В [4] доказана следующая

Теорема 12.

$$I(B^n) \sim 2\sqrt{e}\,2^{2^{n-1}} = 2\sqrt{e}\,2^{N/2}. \tag{22}$$

В [22] доказаны следующие оценки.

Теорема 13. *Пусть двудольный (n, k, θ)-граф $G = (X, Z; E)$ является $(1/2, \delta)$-расширителем и z — наибольшее из решений уравнения $x = \log_2(2ex/c\delta)$. Тогда для достаточно больших n и k*

$$2^{|X|} + 2^{|Z|} - 1 \le I(G) \le \left(2^{|X|} + 2^{|Z|}\right)\left(1 + 2^{-k\delta/z + O\left(\sqrt{k}\log k + \theta\right)}\right). \tag{23}$$

Доказательство в (23) основано на той идее, что X и Z образуют главную систему контейнеров, содержащую в совокупности почти все независимые множества графа G. Отсюда вытекает асимптотика.

Следствие 2. *Рассмотрим последовательность $G_n = (X_n, Z_n; E_n)$ двудольных $(n, k(n), \theta_n)$-графов, являющихся $(1/2, \delta_n)$-расширителями. Пусть*

$$k(n)\delta_n \to \infty \quad \text{и} \quad \theta_n/k(n) + k(n)^{-1/2}\log k(n) \to 0 \tag{24}$$

при $n \to \infty$. Тогда

$$I(G_n) \sim 2^{|X_n|} + 2^{|Z_n|}. \tag{25}$$

7. Множества чисел, свободные от сумм

Здесь на примере доказательства гипотезы Камерона-Эрдёша мы покажем как метод контейнеров применяется для решения комбинаторных задач теории чисел. Множество A называется *свободным от сумм* (сокращенно, МСС), если $a + b \notin A$ любых $a, b \in A$. Для действительных

чисел $p \le q$ обозначим через $[p, q]$ множество натуральных x, таких, что $p \le x \le q$. Семейство всех подмножеств из отрезка $[t, n]$, свободных от сумм, обозначим через $S(t, n)$. Положим $s(t, n) = |S(t, n)|$, $S(n) = S(1, n)$ и $s(n) = |S(n)|$. В 1988 г. Камерон и Эрдёш предположили [17], что $s(n) = O(2^{n/2})$. Это предположение было доказано независимо Б. Грином [18] и автором [14]. Б. Грин использовал технику преобразований Фурье. Доказательство в [14] — чисто комбинаторное. Однако в том и другом случае использовалась идея контейнеров.

Легко видеть, что всякое подмножество нечетных чисел свободно от сумм. Обозначим семейство всех таких подмножеств из отрезка $[1, n]$ через $S_1(n)$. Ясно, что $|S_1(n)| = 2^{\lceil n/2 \rceil}$. Другим "большим" семейством множеств, свободных от сумм, является семейство всех подмножеств чисел из отрезка $[\lfloor n/2 \rfloor + 1, n]$. Положим $\hat{q} = n^{3/4} \log n$ и обозначим через $S_2(n)$ семейство всех МСС, содержащихся в интервале $[n/2 - \hat{q}, n]$. Идея доказательства из [14] состоит в том, чтобы доказать асимптотику вида

$$|S(n)| \sim |S_1(n) \cup S_2(n)|,$$

т. е., что множество N^1 нечетных чисел из отрезка $[1, n]$ и интервал $[n/2 - \hat{q}, n]$ составляют главную систему контейнеров для $S(n)$. Задача свелась к тому, чтобы доказать, что

$$|S(n) \setminus (S_1(n) \cup S_2(n))| = o(2^{n/2}). \tag{26}$$

Для получения верхней оценки мощности семейства $\widetilde{S}(n) = S(n) \setminus (S_1(n) \cup S_2(n))$ доказывается существование так называемой правильной системы контейнеров для семейства $\widetilde{S}(n)$.

Положим $\widetilde{q} = \hat{q} \log n$, $N^\sigma = \{i \in [1, n] : i \equiv \sigma \pmod 2\}$ и $B_{i,p} = B \cap [i, p]$. Семейство $\mathcal{B}$ подмножеств из отрезка $[1, n]$ назовем *правильной* системой контейнеров для $\widetilde{S}(n)$, если выполнены следующие условия:

1) Для достаточно больших n и любого $B \in \mathcal{B}$

$$|B| \le n/2 + O(\hat{q}). \tag{27}$$

2) Для достаточно больших n

$$|\mathcal{B}| \le 2^{o(\hat{q})}. \tag{28}$$

3) Для любого $i \in [\widetilde{q}, n - \widetilde{q}]$ и $p \in [\widetilde{q}, n - i]$

$$||B_{i,p}| - p/2| \le \hat{q}. \tag{29}$$

4) Для любого $\sigma \in \{0,1\}$, $i \in [\widetilde{q}, n-\widetilde{q}]$ и $\in [wq, n-i]$

$$||B_{i,p} \cap N^{\sigma}| - p/4| \leq \hat{q}. \tag{30}$$

Первые два пункта определения накладывают ограничения сверху на размер контейнеров и их число. Третий и четвертый пункты говорят о том, что элементы каждого контейнера распределены равномерно (с плотностью 0,5) по прогрессиям разности 1 и 2 в любом достаточно большом отрезке, достаточно удаленном от концов отрезка $[1, n]$.

Семейство множеств $\mathcal{B}$ называется *почти правильной* системой контейнеров для семейства $\widetilde{S}(n)$, если она является правильной для некоторого подсемейства $\mathcal{A}' \subseteq \widetilde{S}(n)$, такого, что $|\widetilde{S}(n) \setminus \mathcal{A}'| = o(2^{n/2})$.

Существование почти правильной системы контейнеров для семейства $\widetilde{S}(n)$ доказывается сведением к оценке числа независимых множеств в соответствующем графе Кэли с последующим применением теоремы 5. Однако существование такой системы еще недостаточно для достижения поставленной цели. Дело в том, что ограничение (27) на размер контейнера не позволяет непосредственно получить оценку вида $|\widetilde{S}(n)| = o(2^{n/2})$. Идея состоит в том, чтобы показать, что в каждом контейнере содержится не более 2^{cn}, где $c < 0.5$, МСС. С учетом (28) это позволит доказать неравенство (26).

С этой целью выделим в каждом контейнере B два фрагмента $D = B \cap [n/4+1, n/2]$ и $H = B \cap [n/2+1, n]$ докажем, что число МСС в этих фрагментах существенно меньше числа всех подмножеств множества $D \cup H$.

Пусть $Q = D + D$ (по определению $D + D = \{i + j : \{i, j\} \subseteq D\}$) и $\widetilde{Q} = Q \cap H$. Рассмотрим граф $\Gamma = (D, E)$ с множеством вершин D и множеством E ребер вида $\{\{i, j\} : i + j \in \widetilde{Q}\}$. Подмножество ребер графа называется *паросочетанием*, если в нем никакие два ребра не смежны. На каждом ребре паросочетания графа Γ не более одной вершины может принадлежать МСС. Если в n-вершинном графе существует паросочетание из p ребер, то, очевидно, число независимых множеств в нем не больше $2^{n-2p}3^p$. При условии, что $p \geq cn$, где $0 < c \geq 0.5$ имеем $I(\Gamma) \leq 2^{|D|(1-\varepsilon(c))}$, где $\varepsilon(c) > 0$. Отсюда следует, что число МСС, содержащихся в произвольном контейнере B из правильной системы, не превосходит $2^{d|B|}$, для некоторого $d < 1$, что с учетом (27) и (28) влечет (26).

Таким образом достаточно лишь доказать, что в графе Γ существует достаточно большое паросочетание. Этот факт доказывается использованием неравенств (29), (30) и следующего факта из теории сложения множеств.

Теорема 14. (Г. А. Фрейман [15]). *Пусть множество K целых чисел удовлетворяет условию $|K+K| \le 2|K|-1+b$, где $0 \le b \le |K|-3$. Тогда K содержится в некоторой арифметической прогрессии длины $|K|+b$.*

При доказательстве гипотезы Камерона-Эрдёша мы использовали главную систему контейнеров для семейства $\widetilde{S}(n)$, состоящую из двух контейнеров: множества N^1 нечетных чисел, не превосходящих n, и отрезка $[n/2-\hat{q}, n]$. Заметим, что контейнер N^1 является *полным* в том смысле, что каждое его подмножество принадлежит семейству $\widetilde{S}(n)$. В то же время контейнер $[n/2-\hat{q}, n]$ является *почти пустым* в том смысле, что доля его подмножеств, принадлежащих семейству $\widetilde{S}(n)$, стремится к 0 при $n \to \infty$.

Подсчет числа МСС в первом из контейнеров, очевидно не представляет труда. Нахождение асимптотики числа МСС во втором контейнере является отдельной и достаточно трудной задачей. Ее существование следует уже из статьи [17]. К. Г. Омельянов в [5] показал, что $|S_2(n)| \sim c_0 2^{n/2}$ при четных n и $|S_2(n)| \sim c_1 2^{\lceil n/2 \rceil}$ при нечетных n, где $6.070 \le c_0 \le 6.099$ и $4.810 \le c_1 \le 4.837$. Таким образом он вычислил так называемые константы Камерона-Эрдёша c_0 и c_1 с точностью двух знаков. Кроме того в [5] указан способ вычисления этих констант с любой наперед заданной точностью.

В доказательстве Б. Грина [18] аналог понятия "контейнер" также присутствует. Назовем тройку чисел *суммируемой*, если ее элементы, взятые в некотором порядке, удовлетворяют уравнению $x+y=z$. Предложение 6 из [18] указывает семейство $\mathcal{F}$, удовлетворяющее следующим трем свойствам:

1. Каждое МСС в $[1,n]$ содержится в некотором $A \in \mathcal{F}$,
2. $|\mathcal{F}| = 2^{o(n)}$,
3. Каждый элемент из $\mathcal{F}$ содержит $o(n^2)$ суммируемых троек.

В дальнейшем в статье [18] мощность каждого элемента A из $\mathcal{F}$ ограничивается величиной $n(1/2+1/120+o(1))$. Грубо говоря, семейство $\mathcal{F}$ отличается от правильной системы контейнеров тем, что условия (29) и (30) заменены ограничением на число суммируемых троек. Б. Грин называет способ построения такого семейства $\mathcal{F}$ *гранулярризацией*. Идейная близость понятий "семейства $\mathcal{F}$" и "системы контейнеров" очевидна, но подходы к построению весьма различны.

8. Множества, свободные от сумм в группах

Здесь мы рассмотрим применение метода контейнеров к оценке числа МСС в группах. Как и при доказательстве гипотезы Камерона-Эрдёша метод приводит к асимптотически точному результату в случае, когда в

группе существует подгруппа индекса 2. Для абелевых групп асимптотика числа МСС была получена в [20] и независимо в [13].

Теорема 15. *Для достаточно больших четных n и любой абелевой группы G порядка n с t подгруппами индекса* 2

$$t \cdot 2^{n/2} - 2^{(n/4)\left(1+o(1)\right)} \le s(G) \le t \cdot 2^{n/2} + 2^{n(1/2-c)}, \qquad (31)$$

где $c > 0.01$.

Доказательство основано на том, что каждый класс смежности подгруппы индекса 2 представляет собой МСС, в t таких классах в совокупности содержится не менее $t \cdot 2^{n/2} - 2^{(n/4)\left(1+o(1)\right)}$ МСС. Назовем те МСС, которые не содержатся ни в одном смежном классе подгруппы индекса 2 *нерегулярными*. Основная задача состоит в том, чтобы показать, что семейство всех смежных классов подгрупп индекса 2 представляет собой главную систему контейнеров для семейства МСС группы, т. е., что доля нерегулярных МСС мала. Для каждой подгруппы H индекса 2 и любого нерегулярного МСС A выполняется условие $A \cap H \neq \oslash$. Граф $C_D(V) = (V, E)$ на множестве вершин V, порожденный множеством D, в котором пара $\{u, v\}$ вершин является ребром тогда и только тогда, когда $u + v$, $u - v$ или $v - u$ принадлежат множеству D, называется графом Кэли, порожденным множеством D на множестве V. При оценке числа нерегулярных множеств мы пользуемся тем, что для каждого такого множества A и любой подгруппы H индекса 2 группы G множество $V = A \cap G \setminus H$ является независимым в графе $C_D(V)$ с $D = A \cap H$. Таким образом задача сводится к оценке числа независимых множеств в графах Кэли, порожденных подмножествами элементов произвольной подгруппы индекса 2.

Для получения требуемых оценок недостаточно оценок типа (2). Нужны оценки вида (7) и (11), в которых показатель имеет вид $n(1/2-c)$. Такие оценки накладывают на граф Кэли дополнительное условие расширительности. Выполнение этого условия доказывается применением результатов из теории сложения множеств, к которым относится следующая

Теорема 16. [19] *Пусть A и B — два непустых подмножества элементов абелевой группы G. Пусть $|A + B| \le |A| + |B| - 1$. Тогда существует такая подгруппа H группы G такая, что*

$$A + B + H = A + B \qquad \text{и} \qquad |A + B| \ge |A + H| + |B + H| - |H|. \qquad (32)$$

Аналогичный теореме 15 результат для некоммутативного случая был получен с помощью метода контейнеров Т. Г. Петросяном [6]. Пусть $PF(G)$ — число множеств, свободных от произведения, в подгруппе G.

Теорема 17. *Для любой группы G порядка n с числом подгрупп индекса 2, равным t, $t \geq 1$,*

$$t \cdot 2^{n/2} - 2^{n\left(1+o(1)\right)/4} \leq |\mathrm{PF}(G)| \leq t \cdot 2^{n/2} + 2^{n(1/2-\epsilon)}, \tag{33}$$

где $\epsilon > 0$ при $n \to \infty$.

Для доказательства расширительности соответствующих графов Кэли использовалась следующая

Теорема 18. [21] *Пусть A и B являются конечными непустыми подмножествами группы G, тогда существует подмножество S множества AB и подгруппа H группы G, такие, что $|S| \geq |A|+|B|-|H|$, и либо $SH = S$, либо $HS = S$.*

9. Некоторые нерешенные задачи

Здесь формулируются нерешенные задачи, в решении которых метод контейнеров может оказаться полезным.

1) Коды исправляющие ошибки.

Асимптотика числа двоичных кодов с расстоянием 2 найдена в [4]. Открытым остается вопрос о числе двоичных кодов с расстоянием, превышающем 2. Аналогичную задачу можно рассматривать для q-ичных кодов.

2) Антицепи в декартовой степени возрастающей цепи из трех 3 элементов.

Вопрос о числе антицепей в декартовой степени цепи из двух элементов эквивалентен проблеме Дедекинда о числе антицепей в n-мерном кубе, решенной в 1981 г. (см. [19]). В [12] найдена асимптотика числа антицепей в унимодальных (кубоподобных) частично упорядоченных множеств. В частности, это касается декартовых степеней звезд. Вопрос о числе антицепей в декартовой степени цепи из трех элементов остается открытым, хотя асимптотика логарифма получена В. Б. Алексеев [1] для декартовой степени произвольных частичных порядков.

3) Множества, свободные от сумм, в группах.

Асимптотика логарифма числа МСС найдена Б. Грином и И. Ружа для всех абелевых групп. Асимптотика числа МСС найдена лишь для некоторых классов абелевых групп. Асимптотик для некоммутативных групп не известно кроме упомянутых выше результатов Т. Г. Петросяна.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект 07-01-00444).

Литература

1. Алексеев В. Б., О числе монотонных k-значных функций // Проблемы кибернетики, М., Наука, Вып. 28, 1973, 5–24.

2. Алексеев В. Е., Верхняя оценка числа максимальных независимых множеств // Дискретная математика, **19** (2007), №2, 84–89.

3. Коршунов А. Д., О числе монотонных булевых функций // Проблемы кибернетики, вып. 38, М. Наука, 1981, 5–108.

4. Коршунов А. Д., Сапоженко А. А., О числе двоичных кодов с расстоянием 2 // Проблемы кибернетики, М., Наука, Вып. 40, 1983, 111–140.

5. Омельянов К. Г., Оценки констант Камерона-Эрдёша // Дискретная математика (2006) 18, №2, 2006, с. 55–70.

6. Петросян Т. Г., *О числе множеств, свободных от произведений, в группах четного порядка*, Дискретная математика, **17** (2005), №1, 89-101.

7. Сапоженко А. А., О числе независимых множеств в расширителях // Дискретная математика, т. 13, вып. 1, 2001, С. 56-62.

8. Сапоженко А. А., О числе множеств, свободных от сумм в абелевых группах // Вестник московского университета, Серия 1, Математика, Механика, №4, 2002, С. 14-18. 2001, С. 56-62.

9. Сапоженко А. А., О числе независимых множеств в графах. // Вестник московского университета, Серия 1, Математика, Механика, No 4, 2007, С. 17–21.

10. Сапоженко А. А., О числе антицепей в ранжированных частично упорядоченных множествах // Дискретная математика — М.: Наука — 1989 — т. 1, вып. 1, — С. 74–93.

11. Сапоженко А. А., О числе антицепей в многослойных ранжированных частично упорядоченных множествах// Дискретная математика — М.: Наука — 1989 — т. 1, вып. 2, — С. 110–128.

12. Сапоженко А. А., Проблема Дедекинда и метод граничных функционалов // в кн. Математические вопросы кибернетики, Вып. 6, М. Наука, 2000г. С.161–220.

13. Сапоженко А. А., О числе множеств, свободных от сумм в абелевых группах // Вестник московского университета, Серия 1, Математика, Механика, №4, 2002, с. 14–18.

14. Сапоженко А. А., Доказательство гипотезы Камерона-Эрдеша // В кн. Математические вопросы кибернетики вып., М., ФМЛ, вып. 12, 2003 г. С. 5–14.

15. Фрейман Г. А., Сложение конечных множеств // Изв. Высш. Учебн. Завед., Математика, **6 (13)**, (1959), 202–213.

16. Alon N., Independent sets in regular graphs and Sum-Free Subsets of Finite Groups // Israel Journal of Math., 73 (1991), No 2, 247–256.

17. Cameron P., Erdős P., On the number of integers with various properties// in R. A. Mollin (ed). Number Theory: Proc. First Conf. Can. Number Th. Ass., Banff, 1988, — de Gruyter. 1990 — P. 61–79.

18. Green B., The Cameron-Erdos conjecture, Bull. Lond. Math. Soc. 36(2004), no. 6, 769-778.

19. Knezer M., Ein Satz über abelischen Gruppen mit Anwendungen auf die Geometry der Zahlen, Math.Zeit. 61 (1955), 429–434.

20. Lev V. F., Luczak T., Schoen T., Sum-free sets in Abelian groups // Israel Journ.Math. 125 (2001) 347, 347–367.

21. Olson, J. E., On the sum of two sets in a group, Journal of Number Theory, **18** (1984), 110–120.

22. Sapozhenko A. A., On the Number of Independent Sets in Bipartite Graphs with Large Minimum Degree // DIMACS Technical Report 2000-25, August 2000, 24–31.

ИНФОРМАЦИОННЫЕ СВОЙСТВА НЕДООПРЕДЕЛЕННЫХ ДАННЫХ

Л. А. Шоломов

Институт системного анализа РАН,
117312 Москва, просп. 60-летия Октября, 9
e-mail: sholomov@isa.ru

1. Введение

Прежде чем дать строгие определения понятий, связанных с недоопределенными данными, рассмотрим примеры.

Пусть имеется последовательность цифр, написанная нечетко. В ней может, например, встретиться символ, похожий на 3 и на 5, либо символ, похожий на 1, 4 и 7. Такие символы будем называть недоопределенными и обозначать $a_{3,5}$ и $a_{1,4,7}$, а сами цифры $0, 1, \ldots, 9$ будем называть основными символами. В общем случае для недоопределенного символа будем использовать обозначение a_T, где T — множество основных символов, одним из которых он может быть замещен (доопределен). Результатом решения задачи распознавания нечетко написанной последовательности является некоторая последовательность основных символов, доопределяющая исходную.

Подобная ситуация имеет место и при реализации частично определенных управляющих систем. Если, например, поведение системы задается двумя частичными булевыми функциями и при рассматриваемых значениях переменных первая функция не определена, а вторая равна 1, то возникает недоопределенная пара $(*1)$, доопределимая до (01) и (11). Пары (00), (01), (10), (11) образуют основной алфавит, и если для их обозначения использовать символы a_0, a_1, a_2, a_3, то недоопределенной паре $(*1)$ будет соответствовать символ $a_{1,3}$. После того как система реализована (схемно или программно), она становится всюду определенной, что соответствует замене недоопределенных символов в описании систем основными.

С недоопределенными данными имеют дело в задачах распознавания, синтеза управляющих систем, управления, принятия решений, генетики. Поэтому целесообразно изучить недоопределенные данные в качестве самостоятельного объекта подобно тому, как это делается в теории информации для полностью определенных данных. Как будет видно из последующего изложения, некоторые результаты и методы теории информации переносятся на недоопределенные данные, некоторые модифицируются соответствующим образом, а в некоторых случаях возникают новые эффекты. Установленные здесь факты оказываются полезными при решении конкретных задач [8, 10, 12].

2. Энтропия недоопределенных данных

Пусть $M = \{0, 1, \ldots, m-1\}$ — некоторое множество и каждому непустому подмножеству $T \subseteq M$ сопоставлен символ a_T. Алфавит всех символов a_T обозначим через A, а его подалфавит $\{a_0, a_1, \ldots, a_{m-1}\}$, символы которого соответствуют элементам множества M, — через A_0. Символы из A_0 будем называть *основными*, из A — *недоопределенными*. *Доопределением* символа $a_T \in A$ назовем всякий основной символ a_i, $i \in T$, а доопределением последовательности в алфавите A — любую последовательность в алфавите A_0, полученную из исходной заменой всех ее символов некоторыми доопределениями. Символ a_M, доопределимый любым основным символом, играет особую роль. Его будем называть *неопределенным* и обозначать $*$.

Излагаемые дальше результаты допускают статистическую и детерминированную формулировки. Статистическое изложение несколько проще и, следуя традициям теории информации, будем придерживаться его. Для статистических результатов часто будем приводить детерминированную интерпретацию.

Пусть имеется источник S, порождающий символы $a_T \in A$ независимо с вероятностями $p_T \geq 0$, $\sum_T p_T = 1$. Набор вероятностей $(p_T,\ T \subseteq M)$ обозначим через P и для источника S будем использовать обозначение (A, P). Такой источник будем называть *недоопределенным*, а при выполнении условия $p_T = 0$ для $a_T \notin A_0$ — *полностью определенным*. В случае $p_T = 0$ для $a_T \notin A_0 \cup \{*\}$ источник называется *частично определенным*. Подчеркнем, что мы различаем термины "недоопределенный" и "частично определенный". Иногда вместо обозначения (A, P) источника S будем использовать (A', P'), где алфавит A' получен из A удалением всех или некоторых символов a_T с $p_T = 0$, а P' образован из P удалением соответствующих нулевых компонент. В этих обозначениях полностью опреде-

ленный источник может быть записан как $(A_0, (p_0, \ldots, p_{m-1}))$, а частично определенный — как $(A_0 \cup \{*\}, (p_0, \ldots, p_{m-1}, p_*))$.

Зададимся некоторым набором вероятностей $Q = (q_i,\ i \in M)$ символов $a_i \in A_0$ ($q_i \geq 0$, $q_0 + \ldots + q_{m-1} = 1$) и введем функцию

$$\mathcal{H}(P, Q) = -\sum_{T \subseteq M} p_T \log \sum_{i \in T} q_i \tag{34}$$

(здесь и дальше логарифмы двоичные). *Энтропией источника* S назовем величину

$$\mathcal{H}(P) = \min_{Q} \mathcal{H}(P, Q). \tag{35}$$

Наряду с $\mathcal{H}(P)$ будем использовать обозначение $\mathcal{H}(S)$. Указанная формула энтропии была предложена М. М. Бонгардом [1, с. 92] из некоторых эвристических соображений в качестве меры неопределенности задач с несколькими ответами.

Для полностью определенного источника $(A_0, (p_0, \ldots, p_{m-1}))$, в силу известного соотношения

$$\min_{(q_0, \ldots, q_{m-1})} (-\sum_i p_i \log q_i) = -\sum_i p_i \log p_i, \tag{36}$$

величина $\mathcal{H}(P)$ совпадает с энтропией Шеннона $H(P) = -\sum_i p_i \log p_i$. Иногда энтропию Шеннона определяют и исследуют в терминах энтропии разбиений [6] (это соответствует случаю, когда множества T, для которых $p_T > 0$, не пересекаются). Энтропию $\mathcal{H}(P)$ можно рассматривать как обобщение этого понятия на случай произвольной системы множеств T. Вместо источников можно говорить об энтропии случайных опытов с недоопределенными исходами.

3. Вычисление энтропии

Энтропия недоопределенного источника задана неявно, как минимум по Q функции (34). Следующий критерий оказывается полезным для нахождения точек минимума.

Теорема 1. *Набор вероятностей Q минимизирует функцию $\mathcal{H}(P, Q)$ тогда и только тогда, когда при каждом i, $i \in M$, выполнено*

$$\sum_{T:\, i \in T} \frac{p_T}{\sum\limits_{j \in T} q_j} \leq 1, \tag{37}$$

где строгое неравенство может иметь место лишь при тех i, для которых $q_i = 0$.

Доказательство. Вогнутая по Q функция $-\mathcal{H}(P,Q)$ удовлетворяет условиям теоремы 4.4.1 из [3]. По этой теореме необходимым и достаточным условием ее максимума в точке Q является существование такого λ, что $-\partial\mathcal{H}(P,Q)/\partial q_i \le \lambda$, $i \in M$, где строгие неравенства могут соответствовать лишь нулевым значениям q_i. В нашем случае эти соотношения приобретают вид

$$\log e \sum_{T:\, i\in T} \frac{p_T}{\sum\limits_{j\in T} q_j} \le \lambda, \quad i \in M. \tag{38}$$

Поскольку равенства могут нарушаться лишь при нулевых q_i, домножив на q_i, получаем равенства

$$\sum_{T:\, i\in T} \frac{p_T q_i}{\sum\limits_{j\in T} q_j} = \frac{\lambda}{\log e} q_i, \quad i \in M. \tag{39}$$

Просуммировав их по $i \in M$ с учетом $\sum_T p_T = \sum_i q_i = 1$ и того, что

$$\sum_{i\in M} \sum_{T:\, i\in T} \frac{p_T q_i}{\sum\limits_{j\in T} q_j} = \sum_T p_T \frac{\sum\limits_{i\in T} q_i}{\sum\limits_{j\in T} q_j} = \sum_T p_T,$$

находим, что $\lambda = \log e$. Подставив это значение в (38), получаем требуемое утверждение. Теорема доказана.

На базе этой теоремы может быть указан численный метод нахождения $\mathcal{H}(P)$ [8]. Введем оператор $Q' = \mathcal{U}(Q)$, сопоставляющий набору $Q = (q_0, q_1, \ldots, q_{m-1})$ набор $Q' = (q'_0, q'_1, \ldots, q'_{m-1})$, где

$$q'_i = \sum_{T:\, i\in T} \frac{p_T q_i}{\sum\limits_{j\in T} q_j}, \quad i = 0, 1, \ldots, m-1.$$

Нетрудно проверить непосредственно, что оператор $\mathcal{U}$ переводит наборы вероятностей в наборы вероятностей. Домножив обе части (37) на q_i и учитывая, что строгое неравенство в (37) может иметь место лишь при

$q_i = 0$, получаем равенства,

$$\sum_{T:\, i\in T} \frac{p_T q_i}{\sum\limits_{j\in T} q_j} = q_i, \quad i \in M. \tag{40}$$

означающие, что минимизирующий набор Q является неподвижной точкой оператора $\mathcal{U}$. Приведем без доказательства утверждение из [8], дающее алгоритм вычисления $\mathcal{H}(P)$.

Теорема 2. *Если $Q^{(0)} = (q_0^{(0)}, \ldots, q_{m-1}^{(0)})$ — набор вероятностей с положительными компонентами и $Q^{(\nu)} = \mathcal{U}(Q^{(\nu-1)})$, $\nu = 1, 2, \ldots$, то при $\nu \to \infty$ последовательность $\mathcal{H}(P, Q^{(\nu)})$ сходится к $\mathcal{H}(P)$.*

Таким образом, нахождение численного значения энтропии трудностей не вызывает. В некоторых содержательно важных случаях может быть получено явное выражение энтропии. Это относится, например, к частично определенным источникам, энтропия которых может быть найдена на основе следующего утверждения.

Теорема 3. *Для любого r, $0 \le r \le p_*$ $(p_* = p_M)$, справедливо равенство*

$$\mathcal{H}(P) = (1 - r)\mathcal{H}(P'),$$

где $P' = (p'_T,\, T \subseteq M)$, $p'_T = \dfrac{p_T}{1-r}$ для $T \ne M$, $p'_* = \dfrac{p_* - r}{1-r}$.

Доказательство. Для любого набора вероятностей $Q = (q_i,\, i \in M)$ выполнено $r \log \sum_{i\in M} q_i = 0$, поэтому

$$-\sum_{T\subseteq M} p_T \log \sum_{i\in T} q_i = -(1-r)\Big\{\sum_{T\subset M} \frac{p_T}{1-r} \log \sum_{i\in T} q_i + \frac{p_* - r}{1-r} \log \sum_{i\in M} q_i\Big\}.$$

Взяв минимум по Q, получаем нужное утверждение. Теорема доказана.

Следствие 1. *Имеет место равенство*

$$\mathcal{H}(P) = (1 - p_*)\mathcal{H}(P^0),$$

где $P^0 = (p^0_T,\, T \subset M)$ (включение строгое) — набор, полученный из $P = (p_T,\, T \subseteq M)$ отбрасыванием компоненты $p_M = p_*$ и пересчетом вероятностей $p^0_T = p_T/(1 - p_*)$.

Этот факт получается из теоремы при $r = p_*$.

Следствие 2. *Энтропия частично определенного источника* $S = (A_0 \cup \{*\}, (p_0, \ldots, p_{m-1}, p_*))$ *задается выражением*

$$\mathcal{H}(S) = (1 - p_*) H\Big(\frac{p_0}{1 - p_*}, \ldots, \frac{p_{m-1}}{1 - p_*}\Big) = \\ = (1 - p_*) \log(1 - p_*) - \sum_{0 \le i \le m-1} p_i \log p_i.$$

Это вытекает из следствия 1, поскольку в рассматриваемом случае набор вероятностей P^0 соответствует всюду определенному источнику с вероятностями символов $p_i/(1 - p_*)$.

Теорема 3 и следствие 1 имеют важную интерпретацию, о которой будет сказано дальше.

4. Связь с принципом Шеннона

Задача сжатия недоопределенных данных состоит в том, чтобы каждой недоопределенной последовательности, сопоставить двоичный код по возможности малой длины, позволяющий восстановить какое-либо ее доопределение (но не саму последовательность). Точная постановка задачи и результаты будут приведены в последующих разделах.

Эта задача может быть описана в терминах более общей задачи кодирования источников при заданной точности воспроизведения. Приведем соответствующие понятия [3, 14] применительно к дискретным (конечным) источникам. Пусть некоторый источник S порождает символы b конечного алфавита B независимо с вероятностями $p(b)$ и они должны быть представлены у адресата символами c конечного алфавита C. Условия на точность воспроизведения задаются указанием множества W допустимых совместных распределений $(p(b, c),\ b \in B,\ c \in C)$. Теоретико-информационной мерой неопределенности источника S при точности воспроизведения W считают W-*энтропию* [2]

$$H_W(S) = \min_{(p(b,c)) \in W} \sum_{b,c} p(b, c) \log \frac{p(b, c)}{p(b) \sum\limits_{b'} p(b', c)}.$$

Она обобщает понятие ε-энтропии (скорости создания сообщений в терминологии К. Шеннона [3, 14]). Согласно принципу Шеннона W-энтропия характеризует степень сжимаемости сообщений источника S с точностью W. Это содержательный принцип, который для многих типов источников

и мер точности доказан. Формулировку, обоснование и обсуждение этого принципа (применительно к задачам квантования сообщений) можно найти в [4].

Для недоопределенных данных алфавитам B и C соответствуют A и A_0, а множество W состоит из всех совместных распределений ($p_{Ti} = p(a_T, a_i)$, $T \subseteq M$, $i \in M$), удовлетворяющих условию $p_{Ti} = 0$ для $i \notin T$. В этом случае W-энтропия, для которой будем использовать обозначение $H_W(P)$, приобретает вид

$$H_W(P) = \min_{(p_{Ti}) \in W} \sum_{T,i} p_{Ti} \log \frac{p_{Ti}}{p_T \sum\limits_U p_{Ui}}.$$

Следующая теорема [8] показывает, что $\mathcal{H}(P)$ и $H_W(P)$ являются различными представлениями одной и той же функции.

Теорема 4. *Имеет место равенство*

$$H_W(P) = \mathcal{H}(P).$$

Доказательство. Пусть значение $H_W(P)$ достигает на совместном распределении $(p^0_{Ti}) \in W$. Положим $q_i = \sum_T p^0_{Ti}$.

При заданном T, воспользовавшись для выпуклой функции $f(x) = x \log x$ при $x_i = p^0_{Ti} / (p_T q_i)$, $\alpha_i = q_i \Big/ \sum_{j \in T} q_j$ неравенством Иенсена

$$\sum_i \alpha_i f(x_i) \geq f(\sum_i \alpha_i x_i),$$

получаем с учетом $p^0_{Ti} = 0$, $i \notin T$,

$$\sum_i p^0_{Ti} \log \frac{p^0_{Ti}}{p_T q_i} = p_T \left(\sum_{j \in T} q_j \right) \sum_i \frac{q_i}{\sum\limits_{j \in T} q_j} \frac{p^0_{Ti}}{p_T q_i} \log \frac{p^0_{Ti}}{p_T q_i} \geq p_T \log \frac{1}{\sum\limits_{j \in T} q_j}.$$

Просуммировав эти неравенства по T, заключаем, что $H_W(P) \geq \mathcal{H}(P)$.

Обратно, пусть величина $\mathcal{H}(P)$ в (35) достигается на наборе $Q^0 = (q^0_0, q^0_1, \ldots, q^0_{m-1})$. Положим $p_{Ti} = p_T q^0_i \Big/ \sum_{j \in T} q^0_j$ при $i \in T$ и $p_{Ti} = 0$ при

$i \notin T$. Учитывая равенства $\sum_T p_{Ti} = q_i^0$, вытекающие из (39), получаем

$$H_W(P) \le \sum_{T,i} p_{Ti} \log \frac{p_{Ti}}{\sum\limits_j p_{Tj} \sum\limits_U p_{Ui}} = \sum_{T,i} p_{Ti} \log \frac{p_{Ti}}{p_T q_i^0} =$$

$$= \sum_{T,i} p_{Ti} \log \frac{1}{\sum\limits_{j \in T} q_j^0} = \mathcal{H}(P).$$

Теорема доказана.

Из нее и теоремы 5 следующего раздела следует справедливость принципа Шеннона для недоопределенных данных.

5. Комбинаторная энтропия

Будем рассматривать последовательности длины n в алфавите A. Для набора натуральных чисел $\mathbf{n} = (n_T,\ T \subseteq M)$, такого что $\sum_T n_T = n$, обозначим через $\mathcal{K}_n(\mathbf{n})$ множество всех последовательностей, в которых символ a_T, $T \subseteq M$, встречается n_T раз. Скажем, что некоторое множество последовательностей в алфавите A_0 *доопределяет* класс $\mathcal{K}_n(\mathbf{n})$, если в нем найдется доопределение для каждой последовательности из $\mathcal{K}_n(\mathbf{n})$. Обозначим через $N_n(\mathbf{n})$ минимальную мощность множества, доопределяющего $\mathcal{K}_n(\mathbf{n})$. Величину $\log N_n(\mathbf{n})$ назовем *комбинаторной энтропией* класса $\mathcal{K}_n(\mathbf{n})$. Она указывает наименьшее число двоичных символов, достаточное для кодирования последовательностей класса $\mathcal{K}_n(\mathbf{n})$, позволяющего восстанавливать некоторое их доопределение.

Теорема 5. *Существуют константы $c_1 = c_1(m)$ и $c_2 = c_2(m)$ такие, что комбинаторная энтропия класса $\mathcal{K}_n(\mathbf{n})$ заключена в пределах*

$$n\mathcal{H}(\mathbf{n}/n) - c_1 \log n \le \log N_n(\mathbf{n}) \le n\mathcal{H}(\mathbf{n}/n) + c_2 \log n.$$

Доказательство. Верхняя оценка. Воспользуемся широко применяемым в теории информации методом случайного кодирования. Для этого зададимся некоторым набором вероятностей $Q = (q_0, \dots, q_{m-1})$ и возьмем $N \ge 1$ случайных последовательностей длины n в алфавите A_0, компоненты которых независимо с вероятностями q_i принимают значения a_i $(i = 0, \dots, m-1)$. Вероятность того, что случайная последовательность доопределяет фиксированную последовательность из $\mathcal{K}_n(\mathbf{n})$, составляет $\prod_T \left(\sum_{i \in T} q_i\right)^{n_T}$, а вероятность того, что ни одна из N случайных последовательностей не является ее доопределением, равна

$\left(1-\prod_T \left(\sum_{i\in T} q_i\right)^{n_T}\right)^N$. Вероятность $p(N)$ отсутствия доопределения хотя бы у одной последовательности из $\mathcal{K}_n(\mathbf{n})$ не превосходит

$$|\mathcal{K}_n(\mathbf{n})|\left(1-\prod_T\left(\sum_{i\in T} q_i\right)^{n_T}\right)^N < 2^{mn}\left(1-\prod_T\left(\sum_{i\in T} q_i\right)^{n_T}\right)^N$$

($|\cdot|$ означает мощность множества). Непосредственный подсчет с учетом соотношения $\ln(1-x)\le -x$ показывает, что при всяком N, удовлетворяющем условию

$$\log N \ge \log n + \log\ln m - \sum_T n_T \log\sum_{i\in T} q_i = \log n + \log\ln m + n\mathcal{H}(\mathbf{n}/n, Q),$$

выполнено $p(N) < 1$ и потому существует доопределяющее множество мощности N. Это справедливо для любого Q и, в частности, для которого $\mathcal{H}(\mathbf{n}/n, Q) = \mathcal{H}(\mathbf{n}/n)$.

Н и ж н я я о ц е н к а. Обозначим через $t(\mathbf{n})$ максимальное число последовательностей из $\mathcal{K}_n(\mathbf{n})$, которое может быть доопределено одной последовательностью. Рассмотрим последовательность $\mathbf{y}$, являющуюся доопределением некоторой последовательности $\mathbf{x}\in\mathcal{K}_n(\mathbf{n})$. Пусть $\mathbf{y}$ имеет параметры $s_0, s_1, \ldots, s_{m-1}$ ($s_0+\ldots+s_{m-1}=n$). Обозначим через v_{Ti} число символов a_T последовательности $\mathbf{x}$, доопределенных в $\mathbf{y}$ символом a_i. Числа v_{Ti} удовлетворяют условиям:

$$\sum_T v_{Ti} = s_i\ (i\in M),\quad \sum_i v_{Ti} = n_T\ (T\subseteq M),\quad v_{Ti}=0\ \ i\notin T. \tag{41}$$

При фиксированных v_{Ti} последовательность $\mathbf{y}$ доопределяет

$$\frac{s_0!}{\prod_T v_{T0}!}\cdots\frac{s_{m-1}!}{\prod_T v_{T,m-1}!} = \frac{\prod_i s_i!}{\prod_{T,i} v_{Ti}!}$$

последовательностей из $\mathcal{K}_n(\mathbf{n})$, а всего она доопределяет

$$t_{s_0\ldots s_{m-1}}(\mathbf{n}) = \sum_{v_{Ti},(41)} \frac{\prod_i s_i!}{\prod_{T,i} v_{Ti}!}$$

последовательностей из этого класса, где сумма берется по всем наборам неотрицательных чисел v_{Ti}, удовлетворяющих условиям (41). В силу $0 \le v_{Ti} \le n$ и того, что количества индексов i и множеств T ограничены константами (зависящими от m), имеем

$$t_{s_0 \dots s_{m-1}}(\mathbf{n}) \le n^{c_1} \max_{v_{Ti},(41)} \frac{\prod\limits_i s_i!}{\prod\limits_{T,i} v_{Ti}!},$$

где $c_1 = c_1(m)$ — константа. Откуда

$$t(\mathbf{n}) = \max_{s_0,\dots,s_{m-1}} \max_{v_{Ti},(41)} t_{s_0 \dots s_{m-1}}(\mathbf{n}) \le$$

$$\le n^{c_1} \max_{s_0,\dots,s_{m-1}} \max_{v_{Ti},(41)} \frac{\prod\limits_i s_i!}{\prod\limits_{T,i} v_{Ti}!} \le n^{c_1} \max_{v_{Ti}} \frac{\prod\limits_i \left(\sum\limits_T v_{Ti}\right)!}{\prod\limits_{T,i} v_{Ti}!}, \qquad (42)$$

где для целых неотрицательных v_{Ti} выполнены условия

$$\sum_i v_{Ti} = n_T \ (T \subseteq M), \quad v_{Ti} = 0 \ \ i \notin T. \qquad (43)$$

Класс $\mathcal{K}_n(\mathbf{n})$ содержит $n! \,/\prod_T n_T!$ последовательностей. Отсюда и из (42) заключаем, что минимальная мощность $N_n(\mathbf{n})$ доопределяющего множества для класса $\mathcal{K}_n(\mathbf{n})$ удовлетворяет оценке

$$N_n(\mathbf{n}) \ge \frac{|\mathcal{K}_n(\mathbf{n})|}{t(\mathbf{n})} \ge n^{-c_1} \min_{(v_{Ti}),(43)} \frac{n! \prod\limits_{T,i} v_{Ti}!}{\prod\limits_T \left(\sum\limits_i v_{Ti}\right)! \prod\limits_i \left(\sum\limits_T v_{Ti}\right)!}.$$

Из формулы Стирлинга следует, что для любых целых $z, z_1, \dots, z_k$, где $z \ge 2$ и $z_1 + \dots + z_k = z$, выполнено

$$\log \frac{z!}{\prod\limits_j z_j!} = z \log z - \sum_j z_j \log z_j + \theta \log z,$$

где $-c_2 \le \theta \le c_2$, $c_2 = c_2(k)$ — константа. С учетом этого получаем

$$\log N_n(\mathbf{n}) \ge \min_{(v_{Ti}),(43)} \left(n\log n - \sum_T \left(\sum_i v_{Ti}\right) \log\left(\sum_i v_{Ti}\right) - \right.$$
$$\left. - \sum_i \left(\sum_T v_{Ti}\right) \log\left(\sum_T v_{Ti}\right) + \sum_{T,i} v_{Ti}\log v_{Ti} \right) - c\log n.$$

Минимизируемое выражение может быть преобразовано к виду

$$n\left(-\sum_T \left(\sum_i \frac{v_{Ti}}{n}\right) \log\left(\sum_i \frac{v_{Ti}}{n}\right) - \sum_i \left(\sum_T \frac{v_{Ti}}{n}\right) \log\left(\sum_T \frac{v_{Ti}}{n}\right) + \right.$$
$$\left. + \sum_{T,i} \frac{v_{Ti}}{n} \log \sum_{T,i} \frac{v_{Ti}}{n} \right) = n \sum_{T,i} \frac{v_{Ti}}{n} \log \frac{\dfrac{v_{Ti}}{n}}{\displaystyle\sum_U \frac{v_{Ui}}{n} \sum_j \frac{v_{Tj}}{n}}.$$

В силу (43) выполнено $v_{Ti}/n = 0$, $i \notin T$, и $\sum_i (v_{Ti}/n) = n_T/n$, поэтому

$$\log N_n(\mathbf{n}) \ge nH_W(\mathbf{n}/n) - c\log n.$$

Остается воспользоваться теоремой 4, согласно которой $H_W(\mathbf{n}/n) = \mathcal{H}(\mathbf{n}/n)$.

Теорема доказана.

Для указания параметров класса $\mathcal{K}_n(\mathbf{n})$ (и даже просто для задания n) требуется порядка $\log n$ единиц информации. Будем рассматривать случай когда энтропия $\log N_n(\mathbf{n})$ класса $\mathcal{K}_n(\mathbf{n})$ существенно больше $\log n$; тогда $\log N_n(\mathbf{n}) \sim n\mathcal{H}(\mathbf{n}/n)$. Этот результат, записанный с учетом теоремы 4 в виде $\log N_n(\mathbf{n}) \sim nH_W(\mathbf{n}/n)$, означает, что для частично определенных данных справедлив принцип Шеннона.

Теорема 5 позволяет указать важную содержательную интерпретацию теоремы 3 и следствия 1. Рассмотрим произвольный класс $\mathcal{K}_n(\mathbf{n})$. Класс, полученный из $\mathcal{K}_n(\mathbf{n})$ путем удаления из его последовательностей всех символов $*$, обозначим $\mathcal{K}_{n^0}(\mathbf{n}^0)$, где $n^0 = n - n_*$, $\mathbf{n}^0$ — результат отбрасывания в $\mathbf{n}$ компоненты n_*. Применив следствие 1 при $p_T = n_T/n$, $p_* = n_*/n$, $p_T^0 = n_T/n^0$ и домножив обе части на n, приходим к равенству $n\mathcal{H}(\mathbf{n}/n) = n^0\mathcal{H}(\mathbf{n}^0/n^0)$, которое с учетом теоремы 5 показывает, что эффект, обнаруженный Э. И. Нечипоруком [7] и состоящий в том, что последовательности в алфавите $\{0,1,*\}$ и последовательности, полученные из них удалением символов $*$, могут быть представлены кодами одинаковой с точностью до $O(\log n)$ длины, имеет место и в самой общей ситуации

недоопределенных последовательностей. Теорема 3 описывает более общую ситуацию и показывает, что этот эффект сохраняется при отбрасывании части неопределенных символов.

С классом $\mathcal{K}_n(\mathbf{n})$ свяжем функционал $h_n(\mathbf{n}) = n\mathcal{H}(\mathbf{n}/n)$. Для последовательности $\mathbf{a} \in A^n$ обозначим через $\mathcal{K}_n(\mathbf{a})$ содержащий ее класс $\mathcal{K}_n(\mathbf{n})$ и положим $h_n(\mathbf{a}) = h_n(\mathbf{n})$, $N_n(\mathbf{a}) = N_n(\mathbf{n})$. В этих обозначениях оценки теоремы 5 могут быть переписаны в виде

$$h_n(\mathbf{a}) - c_1 \log n \le \log N_n(\mathbf{a}) \le h_n(\mathbf{a}) + c_2 \log n. \tag{44}$$

6. Свойства энтропии

Рассмотрим некоторые свойства энтропии $\mathcal{H}(P)$ и сравним их со свойствами энтропии Шеннона $H(P)$.

Теорема 6. *Энтропия $\mathcal{H}(P)$ неотрицательна, причем $\mathcal{H}(P) = 0$ тогда и только тогда, когда пересечение всех T, для которых $p_T > 0$, непусто.*

Доказательство. Неотрицательность энтропии очевидна. Пусть минимум $\mathcal{H}(P, Q)$ в (35) достигается на наборе Q^0. Положим $T^0 = \{i \in M \mid q_i^0 > 0\}$. Если $\mathcal{H}(P) = \mathcal{H}(P, Q^0) = 0$, то для любого T с $p_T > 0$ выполнено $\sum_{t \in T} q_t^0 = 1$, а потому T содержит T^0 и пересечение всех таких T непусто. Обратно, если пересечение непусто, то назначив $q_i = 1$ для некоторого i из этого пересечения и $q_j = 0$ для всех $j \ne i$, получим набор Q, для которого $\mathcal{H}(P, Q) = 0$. Теорема доказана.

Таким образом, энтропия $\mathcal{H}(S)$ недоопределенного источника S равна 0 лишь если порождаемые им последовательности могут быть доопределены до последовательностей из одинаковых символов. Этот факт обобщает известный результат для полностью определенного источника, энтропия которого равна 0 лишь если он порождает последовательности, образованные одинаковыми символами.

Укажем верхнюю границу энтропии источника $S = (A, P)$ в функции от распределения $(p(t),\ 1 \le t \le m)$ числа t доопределений символов источника,

$$p(t) = \sum_{T:\, |T| = t} p_T.$$

Теорема 7. *Справедлива оценка*

$$\mathcal{H}(P) \le \log m - \sum_{1 \le t \le m} p(t) \log t,$$

достижимая для любого распределения $(p(t),\ 1 \le t \le m)$.

Доказательство. Эту оценку получим, вычислив $\mathcal{H}(P,Q)$ на наборе $Q=(1/m,\ldots,1/m)$,

$$\mathcal{H}(P)\le-\sum_T p_T\log\frac{|T|}{m}=\log m-\sum_t\sum_{|T|=t}p_T\log t=\log m-\sum_t p(t)\log t.$$

Оценка достигается на наборе P, в котором всем t-элементным множествам T соответствуют равные вероятности $p=p(t)\,/\binom{n}{t}$, $1\le t\le m$. Из соображений симметрии [3] следует, что в этом случае $\mathcal{H}(P,Q)$ минимизируется набором вероятностей $q_0=\ldots=q_{m-1}=1/m$. Теорема доказана.

Если источник полностью определен, то $p(1)=1$, $p(t)=0$ для $t\ge 2$, и оценка утверждения 4 превращается в известную оценку энтропии $H(P)\le\log m$.

Теорема 8. *Функция $\mathcal{H}(P)$ вогнута, т. е. для любых P, P' и числа α, $0\le\alpha\le 1$, выполнено*

$$\mathcal{H}(\alpha P+(1-\alpha)P')\ge\alpha\mathcal{H}(P)+(1-\alpha)\mathcal{H}(P').$$

Доказательство. Пусть минимум функции $\mathcal{H}(\alpha P+(1-\alpha)P',Q)$ достигается на наборе Q. Тогда

$$\begin{aligned}\mathcal{H}(\alpha P+(1-\alpha)P')&=-\alpha\sum_T p_T\log\sum_{i\in T}q_i-(1-\alpha)\sum_T p'_T\log\sum_{i\in T}q_i\ge\\&\ge\alpha\mathcal{H}(P)+(1-\alpha)\mathcal{H}(P').\end{aligned}$$

Теорема доказана.

В отличие от обычной энтропии $H(P)$, функция $\mathcal{H}(P)$ не является строго вогнутой. Из доказательства видно, что при $\alpha\ne 0,1$ равенство в утверждении теоремы имеет место лишь тогда, когда существует Q, минимизирующее $\mathcal{H}(P,Q)$ и $\mathcal{H}(P',Q)$ одновременно. Поскольку минимум в (36) достигается только при $Q=P$, для полностью определенных источников условием равенства является $P=P'$ и функция $H(P)$ строго вогнута.

Произведение $S\hat{S}$ недоопределенных источников $S=(A,P)$ и $\hat{S}=(\hat{A},\hat{P})$ представляет собой источник, порождающий пары $(a_T,\hat{a}_{\hat{T}})$ с некоторыми вероятностями $p_{T\hat{T}}$. При этом выполнены условия согласования

$$\sum_{\hat{T}}p_{T\hat{T}}=p_T,\quad\sum_T p_{T\hat{T}}=p_{\hat{T}}.\tag{45}$$

Доопределением символа $(a_T, \hat{a}_{\hat{T}})$ источника $S\hat{S}$ считается всякий символ $(a_i, \hat{a}_{\hat{\imath}})$, $i \in T$, $\hat{\imath} \in \hat{T}$. Источники *статистически независимы*, если $p_{T\hat{T}} = p_T \hat{p}_{\hat{T}}$. Энтропия $\mathcal{H}(S\hat{S})$ определяется аналогично (34)–(35)

$$\mathcal{H}(S\hat{S}) = \min_{Q} \Big\{ - \sum_{T,\hat{T}} p_{T\hat{T}} \log \sum_{i \in T, \hat{\imath} \in \hat{T}} q_{i\hat{\imath}} \Big\},$$

где $Q = (q_{i\hat{\imath}}, i \in M, \hat{\imath} \in \hat{M})$, $q_{i\hat{\imath}} \geq 0$, $\sum_{i,\hat{\imath}} q_{i\hat{\imath}} = 0$.

Теорема 9. *Для любых недоопределенных источников S и $\hat{S}$*

$$\mathcal{H}(S\hat{S}) \leq \mathcal{H}(S) + \mathcal{H}(\hat{S}),$$

а если S и $\hat{S}$ статистически независимы, то

$$\mathcal{H}(S\hat{S}) = \mathcal{H}(S) + \mathcal{H}(\hat{S}).$$

Доказательство. Пусть величины $\mathcal{H}(P)$ и $\mathcal{H}(\hat{P})$ в (35) достигаются на наборах $Q = (q_i, i \in M)$ и $\hat{Q} = (\hat{q}_{\hat{\imath}}, \hat{\imath} \in \hat{M})$. Для всех i и $\hat{\imath}$ положим $q_{i\hat{\imath}} = q_i \hat{q}_{\hat{\imath}}$. С учетом (45) получим

$$\begin{aligned} \mathcal{H}(S\hat{S}) &\leq - \sum_{T,\hat{T}} p_{T\hat{T}} \log \left(\sum_{i \in T, \hat{\imath} \in \hat{T}} q_i \hat{q}_{\hat{\imath}} \right) = - \sum_{T,\hat{T}} p_{T\hat{T}} \log \left(\sum_{i \in T} q_i \sum_{\hat{\imath} \in \hat{T}} \hat{q}_{\hat{\imath}} \right) = \\ &= - \sum_{T} p_T \log \sum_{i \in T} q_i - \sum_{\hat{T}} \hat{p}_{\hat{T}} \log \sum_{\hat{\imath} \in \hat{T}} \hat{q}_{\hat{\imath}} = \mathcal{H}(S) + \mathcal{H}(\hat{S}). \end{aligned} \tag{46}$$

Если S и $\hat{S}$ статистически независимы, то $p_{T\hat{T}} = p_T \hat{p}_{\hat{T}}$, и в силу теоремы 1, примененной к S и $\hat{S}$, для всех пар $(i, \hat{\imath})$ выполнено

$$\sum_{T \ni i, \hat{T} \ni \hat{\imath}} \frac{p_{T\hat{T}}}{\sum\limits_{j \in T, \hat{\jmath} \in \hat{T}} q_{j\hat{\jmath}}} = \sum_{T \ni i, \hat{T} \ni \hat{\imath}} \frac{p_T \hat{p}_{\hat{T}}}{\sum\limits_{j \in T, \hat{\jmath} \in \hat{T}} q_{j\hat{\jmath}}} = \sum_{T \ni i} \frac{p_T}{\sum\limits_{j \in T} q_j} \sum_{\hat{T} \ni \hat{\imath}} \frac{\hat{p}_{\hat{T}}}{\sum\limits_{\hat{\jmath} \in \hat{T}} \hat{q}_{\hat{\jmath}}} \leq 1.$$

Строгое неравенство имеет место, лишь когда оно справедливо для хотя бы одного сомножителя. В этом случае $q_i = 0$ либо $\hat{q}_{\hat{\imath}} = 0$, а потому $q_{i\hat{\imath}} = 0$. По теореме 1, примененной к источнику $S\hat{S}$, заключаем, что

$$\mathcal{H}(S\hat{S}) = - \sum_{T,\hat{T}} p_{T\hat{T}} \log \sum_{i \in T, \hat{\imath} \in \hat{T}} q_{i\hat{\imath}},$$

и неравенство (46) обращается в равенство. Теорема доказана.

В отличие от обычной энтропии H, независимость источников не является необходимой для равенства $\mathcal{H}(S\hat{S}) = \mathcal{H}(S) + \mathcal{H}(\hat{S})$. В частности, оно справедливо, когда области определения источников S и $\hat{S}$ не пересекаются, т. е. когда $p_{T\hat{T}} > 0$ лишь если $a_T = *$ или $\hat{a}_{\hat{T}} = *$ [10].

7. Теорема кодирования

Одними из центральных в теории информации являются результаты о сжатии данных, формулируемые в терминах кодирования последовательностей, порождаемых источниками. Согласно им, нельзя добиться, чтобы средняя длина кода, приходящаяся на символ источника, была меньше энтропии, и возможно закодировать так, чтобы средняя длина на символ превосходила энтропию сколь угодно мало [3]. Аналогичный факт оказывается верным и для недоопределенных источников рассматриваемого нами вида. Напомним, что в случае недоопределенных последовательностей требуется по коду восстановить не саму последовательность, а какое-либо ее доопределение.

Последовательности длины n будем называть *n-блоками* (либо просто блоками) и обозначать символом B. Будем рассматривать следующий способ кодирования нечеткого источника $S = (A, P)$. Возьмем некоторое множество $\mathcal{D} \subseteq (A_0)^n$, содержащее доопределения всех блоков из A^n, и каждому блоку $B \in A^n$ сопоставим некоторое его доопределение $D = D_B$ из $\mathcal{D}$ (разным B могут соответствовать одинаковые D_B). Закодируем блоки множества $\mathcal{D}$ двоичными наборами с соблюдением условия однозначного декодирования (разделимости кода) [13]. Кодом блока B будем считать код его доопределения D_B. Последовательность, порождаемая источником S, кодируется путем разбиения ее на n-блоки и приписывания друг к другу кодов полученных n-блоков. Кодирование источника $S = (A, P)$ называется *универсальным*, если оно не зависит от набора вероятностей P.

Обозначим через l_B длину слова, кодирующего блок B. Качество кодирования будем характеризовать средним числом кодовых символов на символ источника

$$\bar{l}_n = \frac{1}{n} \sum_B p(B) l_B,$$

где $p(B) = p_{T_1} p_{T_2} \dots p_{T_n}$ — вероятность блока $B = a_{T_1} a_{T_2} \dots a_{T_n}$, а сумма берется по всем n-блокам. Ставится задача оценки минимального значения $\bar{l}_n$ по всем кодированиям источника S.

Следующий результат обобщает на недоопределенные источники теорему кодирования полностью определенных источников.

Теорема 10. 1. *При любом n и любом способе кодирования*

$$\bar{l}_n \geq \mathcal{H}(S).$$

2. *Существует универсальное кодирование, для которого*

$$\bar{l}_n \leq \mathcal{H}(S) + O\left(\frac{\log n}{n}\right).$$

Из результата Р. Е. Кричевского [6] следует, что при универсальном кодировании понизить порядок остаточного члена в пункте 2 нельзя.

Доказательство. Нижняя оценка. Блоки (n-блоки) $a_{i_1} \dots a_{i_n}$ в алфавите A_0 будем обозначать через $D_{\mathbf{i}}$, $\mathbf{i} = (i_1, \dots, i_n)$, а блоки $a_{T_1} \dots a_{T_n}$ в алфавите A — через $D_{\mathbf{T}}$, $\mathbf{T} = T_1 \times \dots \times T_n$. Для вероятности блока $D_{\mathbf{T}}$ будем использовать обозначение $p_{\mathbf{T}}$.

Каждому блоку $D_{\mathbf{T}}$ соответствует единственный блок $D_{\mathbf{i}(\mathbf{T})}$, взятый в качестве его доопределения. Положим для $\mathbf{i} \in M^n$ и $\mathbf{T} \subseteq M^n$

$$\delta_{\mathbf{T}\mathbf{i}} = \begin{cases} 1, & \mathbf{i} = \mathbf{i}(\mathbf{T}), \\ 0, & \mathbf{i} \neq \mathbf{i}(\mathbf{T}) \end{cases} \tag{47}$$

и введем величины

$$q_{\mathbf{i}} = \sum_{\mathbf{T}} p_{\mathbf{T}} \delta_{\mathbf{T}\mathbf{i}}, \qquad \mathbf{i} \in M^n. \tag{48}$$

Для них выполнено

$$\sum_{\mathbf{i}} q_{\mathbf{i}} = \sum_{\mathbf{T}} p_{\mathbf{T}} \sum_{\mathbf{i}} \delta_{\mathbf{T}\mathbf{i}} = \sum_{\mathbf{T}} p_{\mathbf{T}} = 1. \tag{49}$$

Пусть $\mathbf{I} = \{\mathbf{i} \mid q_{\mathbf{i}} > 0\}$, тогда

$$\sum_{\mathbf{i} \in \mathbf{I}} q_{\mathbf{i}} = 1. \tag{50}$$

Через $l_{\mathbf{i}}$ обозначим длину кода для $D_{\mathbf{i}}$. Согласно определению и (47)

$$l_{D_{\mathbf{T}}} = l_{\mathbf{i}(\mathbf{T})} = \sum_{\mathbf{i} \in \mathbf{T}} \delta_{\mathbf{T}\mathbf{i}} l_{\mathbf{i}}.$$

Отсюда и из (48) с учетом введенных обозначений получаем

$$n\bar{l}_n = \sum_{\mathbf{T}} p_{\mathbf{T}} l_{D_{\mathbf{T}}} = \sum_{\mathbf{T}} p_{\mathbf{T}} \sum_{\mathbf{i} \in \mathbf{I}} \delta_{\mathbf{T}\mathbf{i}} l_{\mathbf{i}} = \sum_{\mathbf{i} \in \mathbf{i}} q_{\mathbf{i}} l_{\mathbf{i}}. \tag{51}$$

Набор длин $l_{\mathbf{i}}$, $\mathbf{i} \in \mathbf{I}$, удовлетворяет неравенству Макмиллана-Крафта [13] $\sum_{\mathbf{i}} 2^{-l_{\mathbf{i}}} \le 1$, и по свойству (36) энтропийной функции выполнено

$$\sum_{\mathbf{i}\in\mathbf{I}} q_{\mathbf{i}} l_{\mathbf{i}} = -\sum_{\mathbf{i}\in\mathbf{I}} q_{\mathbf{i}} \log 2^{-l_{\mathbf{i}}} \ge -\sum_{\mathbf{i}\in\mathbf{I}} q_{\mathbf{i}} \log q_{\mathbf{i}} = -\sum_{\mathbf{T},\mathbf{i}\in\mathbf{I}} p_{\mathbf{T}} \delta_{\mathbf{T}\mathbf{i}} \log q_{\mathbf{i}} =$$
$$= -\sum_{\mathbf{T}} p_{\mathbf{T}} \log q_{\mathbf{i}(\mathbf{T})} \ge -\sum_{\mathbf{T}} p_{\mathbf{T}} \log \sum_{\mathbf{i}\in\mathbf{T}} q_{\mathbf{i}}. \qquad (52)$$

Используя определение энтропии применительно к источнику S^n и утверждение 2, заключаем, что правая часть в (52) не превосходит $\mathcal{H}(S^n) = n\mathcal{H}(S)$. Отсюда, из (51) и (52) получаем утверждение пункта 1 теоремы.

В е р х н я я о ц е н к а. Для каждого класса $\mathcal{K}_n(\mathbf{n})$ возьмем доопределяющее множество мощности $N_n(\mathbf{n})$ и в соответствии с теоремой 5 занумеруем входящие в него блоки двоичными наборами $\tilde{\alpha}$ длины $n\mathcal{H}(\mathbf{n}/n) + O(\log n)$. Сами классы $\mathcal{K}_n$ также занумеруем двоичными наборами $\tilde{\beta}$ одинаковой длины, которую можно взять равной $O(\log n)$. Блоку B припишем кодовое слово $\tilde{\beta}\tilde{\alpha}$, где $\tilde{\beta}$ соответствует классу $\mathcal{K}_n(\mathbf{n})$, содержащему B, а $\tilde{\alpha}$ — номер некоторого доопределения блока B. Тогда

$$l_B = l(\tilde{\beta}\tilde{\alpha}) \le n\mathcal{H}(\mathbf{n}/n) + O(\log n). \qquad (53)$$

Нетрудно видеть, что построенное кодовое множество является префиксным и, следовательно, разделимым [13].

Обозначим через $p(P, \mathbf{n}, n)$ суммарную вероятность блоков класса $\mathcal{K}_n(\mathbf{n})$, равную $(n!/\prod_T n_T!)\prod_T p_T^{n_T}$. Поскольку $\sum_{\mathbf{n}} p(P, \mathbf{n}, n) = 1$, из (53) получаем

$$\bar{l}_n = \frac{1}{n}\sum_B p(B) l_B \le \sum_{\mathbf{n}} p(P, \mathbf{n}, n)\mathcal{H}\left(\frac{\mathbf{n}}{n}\right) + O\left(\frac{\log n}{n}\right). \qquad (54)$$

По свойству полиномиального распределения [2] выполнено

$$\sum_{\mathbf{n}} p(P, \mathbf{n}, n)\frac{n_T}{n} = p_T, \qquad (T \subseteq M) \qquad (55)$$

где n_T и p_T — компоненты наборов $\mathbf{n}$ и P. Применив к вогнутой функции $\mathcal{H}$ (теореме 8) неравенство Иенсена, получим с учетом (55)

$$\sum_{\mathbf{n}} p(P, \mathbf{n}, n)\mathcal{H}\left(\frac{\mathbf{n}}{n}\right) \le \mathcal{H}\left(\sum_{\mathbf{n}} p(P, \mathbf{n}, n)\frac{\mathbf{n}}{n}\right) = \mathcal{H}(P).$$

Отсюда и из (54) следует утверждение пункта 2. Теорема доказана.

8. Мера информации и условная энтропия недоопределенных данных

Пусть заданы недоопределенные источники $S = (A, P)$, $A = \{a_T, T \subseteq M\}$, и $\hat{S} = (B, \hat{P})$, $B = \{b_U, U \subseteq L\}$, и пусть $\{p_{TU}, T \subseteq M, U \subseteq L\}$ — их совместное распределение. Если, как обычно, считать меру информации $\mathcal{I}(S, \hat{S})$ в S о $\hat{S}$ связанной с условной энтропией $\mathcal{H}(\hat{S}|S)$ соотношением

$$\mathcal{I}(S, \hat{S}) = \mathcal{H}(\hat{S}) - \mathcal{H}(\hat{S}|S),$$

то для введения меры информации достаточно дать определение условной энтропии. Здесь возникают принципиальные трудности, связанные с тем, что модификации для общего случая недоопределенных данных статистического подхода К. Э. Шеннона и алгоритмического подхода А. Н. Колмогорова к определению условной энтропии приводят к разным результатам [11]. Оставляя в стороне общую ситуацию, рассмотрим содержательно важный случай, в котором оба подхода оказываются согласованными [11].

Будем говорить, что *символ* $a_T \in A$ *конкретней символа* $b_U \in B$, если $a_T \in A_0$ либо $b_U = *$, и что *источник* S *конкретней источника* $\hat{S}$, если из $p_{TU} > 0$ следует, что a_T конкретней b_U. Вводимые ниже определения относятся к случаю, когда S конкретней $\hat{S}$.

Для условных вероятностей $p(b_U|a_i)$, $U \subseteq L$, $i \in M$, будем использовать обозначение $p_{U|i}$. Если $p_i > 0$, то $p_{U|i} = p_{iU}/p_i$. В случае $p_i > 0$ по аналогии с (34)–(35) определим *энтропию источника* $\hat{S}$ *при условии* a_i, положив

$$\mathcal{H}(\hat{S}|i) = \min_{Q^{(i)}} \left\{ -\sum_{U \subseteq L} p_{U|i} \log \sum_{j \in U} q_j^{(i)} \right\}, \tag{56}$$

где $Q^{(i)} = (q_j^{(i)},\ j \in L)$, $q_j^{(i)} \geq 0$, $\sum_j q_j^{(i)} = 1$. Введем обозначение $\mathcal{U}_i = \{U \subseteq L | p_{U|i} > 0\}$. Применительно к набору $Q^{(i)}$, минимизирующему (56), условия (37) теоремы 1 приобретают вид

$$\sum_{U \in \mathcal{U}_i : j \in U} \frac{p_{U|i}}{\sum\limits_{u \in U} q_u^{(i)}} \leq 1, \quad j \in L. \tag{57}$$

Условную энтропию $\mathcal{H}(\hat{S}|S)$ источника $\hat{S}$ относительно S определим равенством

$$\mathcal{H}(\hat{S}|S) = \sum_{i \in M} p_i \mathcal{H}(\hat{S}|i). \tag{58}$$

Таким образом, если S конкретней $\hat{S}$, то условная энтропия $\mathcal{H}(\hat{S}|S)$ вводится по той же схеме, что и в теории информации для всюду определенных источников. Отличие состоит в том, что выражение (58), вообще говоря, не является математическим ожиданием величин $\mathcal{H}(\hat{S}|i)$, поскольку в типичном случае $\sum_i p_i$ меньше 1. Дальше будет показано, что так введенная условная энтропия обладает свойствами, аналогичными свойствам условной энтропии в классической теории информации.

9. Правило сложения энтропий

В теории информации важную роль играет правило сложения энтропий

$$H(S\hat{S}) = H(S) + H(\hat{S}|S).$$

Оно (в более слабом варианте) включено К. Шенноном в число свойств, аксиоматически определяющих вид энтропийной функции. В данном разделе показывается, что в случае, когда S конкретней $\hat{S}$, правило сложения энтропий справедливо и для недоопределенных источников.

Ниже используются обозначения, введенные в предыдущем разделе. Пусть значения $\mathcal{H}(S)$ и $\mathcal{H}(\hat{S}|i)$ в (35) и (56) достигаются на наборах $Q = (q_i,\, i \in M)$ и $Q^{(i)} = (q_j^{(i)},\, j \in L)$, $i \in M$. Образуем набор $\tilde{Q} = (q_{ij},\, i \in M, j \in L)$, положив $q_{ij} = q_i q_j^{(i)}$.

Лемма 1. *Значение $\mathcal{H}(S\hat{S})$ достигается на наборе $\tilde{Q}$.*

Доказательство. Учитывая, что S конкретней $\hat{S}$ и $p_{T*} = p_T$ для $|T| \geq 2$, запишем для произведения $S\hat{S}$ левую часть аналога неравенства (37), относящегося к q_{ij}, в виде

$$\sum_{U \in \mathcal{U}_i:\, j \in U} \frac{p_{iU}}{\sum_{u \in U} q_{iu}} + \sum_{T:\, |T| \geq 2, i \in T} \frac{p_T}{\sum_{l \in T} \sum_{u \in L} q_{lu}} = \Sigma_1 + \Sigma_2.$$

Преобразуя Σ_1 при $p_{iU} = p_i p_{U|i}$, получаем

$$\Sigma_1 = \frac{p_i}{q_i} \sum_{U \in \mathcal{U}_i:\, j \in U} \frac{p_{U|i}}{\sum_{u \in U} q_u^{(i)}}.$$

Отсюда в силу (57) имеем $\Sigma_1 \leq p_i/q_i$.

Для Σ_2, принимая во внимание $q_{lu} = q_l q_u^{(l)}$ и $\sum_u q_u^{(l)} = 1$, находим

$$\Sigma_2 = \sum_{T:\, |T| \geq 2, i \in T} \frac{p_T}{\sum_{l \in T} q_l \sum_{u \in L} q_u^{(l)}} = \sum_{T:\, |T| \geq 2, i \in T} \frac{p_T}{\sum_{l \in T} q_l}.$$

В результате

$$\Sigma_1 + \Sigma_2 \le \frac{p_i}{q_i} + \sum_{T:\,|T|\ge 2,\, i\in T} \frac{p_T}{\sum_{l\in T} q_l} = \sum_{T:\, i\in T} \frac{p_T}{\sum_{l\in T} q_l}.$$

С учетом (37) это дает $\Sigma_1 + \Sigma_2 \le 1$. Строгое неравенство здесь имеет место, лишь если оно возникало при использовании (57), либо (37). Но тогда $q_j^{(i)} = 0$ либо $q_i = 0$, а потому $q_{ij} = q_i q_j^{(i)} = 0$. Требуемое утверждение получается применением к $\tilde{Q}$ теоремы 1. Лемма доказана.

Теорема 11. *Если источник S конкретней $\hat{S}$, то*

$$\mathcal{H}(S\hat{S}) = \mathcal{H}(S) + \mathcal{H}(\hat{S}|S) \tag{59}$$

Доказательство. Воспользовавшись леммой 1 и тем, что S конкретней $\hat{S}$, запишем выражение для $\mathcal{H}(S\hat{S})$ в виде

$$\begin{aligned}\mathcal{H}(S\hat{S}) &= -\sum_{i\in M}\sum_{U\in\mathcal{U}_i} p_i p_{U|i} \log \sum_{j\in U} q_i q_j^{(i)} - \sum_{T:\,|T|\ge 2} p_T \log \sum_{i\in T}\sum_{j\in L} q_i q_j^{(i)} = \\ &= -\Sigma_1 - \Sigma_2. \end{aligned} \tag{60}$$

Преобразуем первую сумму, учитывая, что минимум в (56) достигается на наборе $Q^{(i)}$,

$$\begin{aligned}-\Sigma_1 &= -\sum_{i\in M} p_i \log q_i \sum_{U\in\mathcal{U}_i} p_{U|i} - \sum_{i\in M} p_i \sum_{U\in\mathcal{U}_i} p_{U|i} \log \sum_{j\in U} q_j^{(i)} = \\ &= -\sum_{i\in M} p_i \log q_i - \sum_{i\in M} p_i \mathcal{H}(\hat{S}|i).\end{aligned}$$

Для второй суммы имеем

$$-\Sigma_2 = -\sum_{T:\,|T|\ge 2} p_T \log\left(\sum_{i\in T} q_i \sum_{j\in L} q_j^{(i)}\right) = -\sum_{T:\,|T|\ge 2} p_T \log \sum_{i\in T} q_i.$$

Подставляя эти выражения в (60) и принимая во внимание

$$-\sum_{i\in M} p_i \log q_i - \sum_{T:\,|T|\ge 2} p_T \log \sum_{i\in T} q_i = \mathcal{H}(S),$$

с учетом (58) получаем утверждение теоремы.

В качестве примера применения теоремы найдем энтропию $\mathcal{H}(S\hat{S})$ произведения частично определенных источников $S = (A_0 \cup \{*\}, P)$, $A_0 = \{a_i,\, i \in M\}$, и $\hat{S} = (B_0 \cup \{*\}, \hat{P})$, $B_0 = \{b_j,\, j \in L\}$, обладающих тем свойством, что область определения одного из них (пусть источника $\hat{S}$) не выходит за пределы области определения другого. Это означает, что в их совместном распределении ненулевыми могут быть лишь вероятности $p_{ij} = p(a_i, b_j)$, $p_{i*} = p(a_i, *)$ и $p_{**} = p(*, *)$. Положим $p'_i = \sum_{j\in L} p_{ij}$, $p_i = p'_i + p_{i*}$. Из следствия теоремы 3 находим

$$\mathcal{H}(S) = \Big(\sum_{i\in M} p_i\Big) \log\Big(\sum_{i\in M} p_i\Big) - \sum_{i\in M} p_i \log p_i. \tag{61}$$

Несложные выкладки с использованием теоремы 3 и ее следствия дают

$$\mathcal{H}(\hat{S}|S) = \sum_{i\in M} p'_i \log p'_i - \sum_{i\in M,\, j\in L} p_{ij} \log p_{ij}. \tag{62}$$

По теореме 11 энтропия $\mathcal{H}(S\hat{S})$ равна сумме значений (61) и (62). Прямое вычисление этой величины, не опирающееся на теорему, требует гораздо более громоздких выкладок [9].

10. Алгоритмический подход

Алгоритмический подход к измерению количества информации, предложенный А. Н. Колмогоровым [5], имеет дело с последовательностями символов, и условная энтропия интерпретируется как сложность одной последовательности относительно другой. Относительная сложность измеряется минимальной длиной двоичнго набора (двоичной программы), позволяющго по одной последовательности построить другую (подробнее в [5]). Ниже излагается некоторая модификация алгоритмического подхода к недоопределенным последовательностям [11]. Она относится к случаю, когда одна последовательность конкретней другой (определение см. ниже)

Пусть $\mathbf{a} = a_{T_1} \ldots a_{T_n}$ и $\mathbf{b} = b_{U_1} \ldots b_{U_n}$ — последовательности одинаковой длины n в алфавитах $A = \{a_T,\, T \subseteq M\}$ и $B = \{b_U,\, U \subseteq L\}$ соответственно. Обозначим через n_{TU} число появлений в них среди пар (a_{T_i}, b_{U_i}) $(1 \le i \le n)$ пары (a_T, b_U) и положим $\mathbf{n}(\mathbf{a}, \mathbf{b}) = (n_{TU},\, T \subseteq M,\, U \subseteq L)$. Введем класс всех пар $(\mathbf{x}, \mathbf{y})$ с теми же параметрами, что у $(\mathbf{a}, \mathbf{b})$:

$$\mathcal{K}_n(\mathbf{a}, \mathbf{b}) = \{(\mathbf{x}, \mathbf{y}) \mid (\mathbf{x}, \mathbf{y}) \in A^n \times B^n,\, \mathbf{n}(\mathbf{x}, \mathbf{y}) = \mathbf{n}(\mathbf{a}, \mathbf{b})\}.$$

Будем считать, что *последовательность* $\mathbf{a}$ *конкретней* $\mathbf{b}$, т. е. что при каждом i $(1 \le i \le n)$ символ a_{T_i} конкретней b_{U_i}. В этом случае $n_{TU} > 0$

лишь если a_T конкретней b_U, поэтому во всех парах $(\mathbf{x}, \mathbf{y}) \in \mathcal{K}_n(\mathbf{a}, \mathbf{b})$ последовательность $\mathbf{x}$ конкретней $\mathbf{y}$. Рассмотрим источники $S_\mathbf{a}$ и $S_\mathbf{b}$ с алфавитами A и B и совместным распределением $\left\{p_{TU} = \dfrac{n_{TU}}{n}\right\}$. Источник $S_\mathbf{a}$ конкретней $S_\mathbf{b}$, поэтому определена условная энтропия $\mathcal{H}(S_\mathbf{b}|S_\mathbf{a})$. Введем функционал $h_n(\mathbf{b}|\mathbf{a}) = n\mathcal{H}(S_\mathbf{b}|S_\mathbf{a})$.

При каждом i, $i \in M$, образуем по $\mathbf{a}$ и $\mathbf{b}$ последовательность $\mathbf{b}_i$ длины n, которая совпадает в j-м разряде $(1 \le j \le n)$ с b_{U_j}, если $a_{T_j} = a_i$, и содержит в j-м разряде $*$ в противном случае. Величину $h_n(\mathbf{b_i})$ определим подобно $h_n(\mathbf{a})$ (перед неравенствами (44)).

Лемма 2. *Если* $\mathbf{a}$ *конкретней* $\mathbf{b}$, *то*

$$h_n(\mathbf{b}|\mathbf{a}) = \sum_{i \in M} h_n(\mathbf{b}_i).$$

Доказательство. Обозначим через $\mathbf{b}^{(i)}$ последовательность длины $n_i = \sum_U n_{iU}$, полученную из $\mathbf{b}$ удалением символов b_{U_j} для всех таких j, что $a_{T_j} \ne a_i$. Имеют место равенства

$$n\frac{n_i}{n}\mathcal{H}(S_\mathbf{b}|i) = n\frac{n_i}{n}\min_{Q^{(i)}}\left\{-\sum_U \frac{n_{iU}}{n_i}\log\sum_{j\in U} q_j^{(i)}\right\} =$$
$$= n_i\mathcal{H}\left(\frac{\mathbf{n}(\mathbf{b}^{(i)})}{n_i}\right) = h_{n_i}(\mathbf{b}^{(i)}).$$

Их суммирование по i с учетом того, что значению p_i в (58) соответствует n_i/n, дает

$$n\mathcal{H}(S_\mathbf{b}|S_\mathbf{a}) = \sum_{i\in M} h_{n_i}(\mathbf{b}^{(i)}). \tag{63}$$

Последовательность $\mathbf{b}^{(i)}$ образована из $\mathbf{b}_i$ удалением $n-n_i$ неопределенных символов. Применив к $\mathbf{b}_i$ теорему 3 при $r = 1 - n_i/n$, получаем $h_n(\mathbf{b}_i) = h_{n_i}(\mathbf{b}^{(i)})$. Заменив в (63) величины $h_{n_i}(\mathbf{b}^{(i)})$ на $h_n(\mathbf{b}_i)$ и учитывая, что левая часть (63) совпадает с $h_n(\mathbf{b}|\mathbf{a})$, приходим к утверждению леммы.

Комбинаторный смысл условной энтропии проясняет следующее утверждение.

Теорема 12. *Если* $\mathbf{a}$ *конкретней* $\mathbf{b}$, *то найдется константа* c *(зависящая от мощности алфавитов) такая, что для любых* $(\mathbf{x}, \mathbf{y}) \in \mathcal{K}_n(\mathbf{a}, \mathbf{b})$ *существует двоичный набор длины не больше* $h_n(\mathbf{b}|\mathbf{a}) + c\log n$, *позволяющий по*

любому доопределению последовательности $\mathbf{x}$ *найти некоторое доопределение для* $\mathbf{y}$.

Доказательство. Пусть $(\mathbf{x}, \mathbf{y})$, $\mathbf{x} = x_1 \dots x_n$, $\mathbf{y} = \mathbf{y_1} \dots \mathbf{y_n}$ — произвольная пара из $\mathcal{K}_n(\mathbf{a}, \mathbf{b})$. При каждом i $(0 \le i \le m-1)$ образуем по $\mathbf{x}$ и $\mathbf{y}$ последовательность $\mathbf{y}_i$ длины n, которая совпадает в j-м разряде $(1 \le j \le n)$ с $\mathbf{y}$, если $x_j = a_i$, и содержит в j-м разряде $*$ в противном случае. Поскольку $\mathbf{x}$ конкретней $\mathbf{y}$, то каждый значащий (т. е. отличный от $*$) разряд из $\mathbf{y}$ попадает ровно в одну последовательность $\mathbf{y}_i$.

Число различных классов $\mathcal{K}_n(\mathbf{n})$ для последовательностей длины n в алфавите B не превосходит $(n+1)^{|B|}$, где $|\cdot|$ — мощность множества (каждому символу $b_U \in B$ отвечает параметр n_U, не больший n). Занумеруем классы двоичными последовательностями одинаковой длины $\lceil |B| \log(n+1) \rceil$, где $\lceil \cdot \rceil$ означает ближайшее сверху целое. Класс $\mathcal{K}_n(\mathbf{n})$, содержащий $\mathbf{y}_i$, будем обозначать $\mathcal{K}_n(\mathbf{y}_i)$.

Каждой последовательности $\mathbf{y}_i$ сопоставим двоичное кодовое слово $\tilde{\sigma}_i = \tilde{\lambda}_i \tilde{\alpha}_i \tilde{\beta}_i$. Здесь $\tilde{\alpha}_i$ — двоичная запись номера класса $\mathcal{K}_n(\mathbf{y}_i)$, $\tilde{\beta}_i$ — двоичная запись номера доопределения последовательности $\mathbf{y}_i$ в множестве, доопределяющем класс $\mathcal{K}_n(\mathbf{y}_i)$ и удовлетворяющем верхней оценке теоремы 5. Согласно (44) длина слова $\tilde{\alpha}_i \tilde{\beta}_i$ ограничена величиной $h(\mathbf{y}_i) + c \log n$. Далее, пусть двоичная запись длины слова $\tilde{\alpha}_i \tilde{\beta}_i$ имеет вид $\lambda_1 \lambda_2 \dots \lambda_s$. Тогда положим $\tilde{\lambda}_i = \lambda_1 \lambda_1 \lambda_2 \lambda_2 \dots \lambda_s \lambda_s 01$; длина этого набора не превосходит $c' \log n$. Суммируя оценки, заключаем, что длина слова $\tilde{\sigma}_i$ оценивается сверху величиной $h(\mathbf{y}_i) + O(\log n)$.

Образуем слово $\tilde{\sigma} = \tilde{\sigma}_0 \tilde{\sigma}_1 \dots \tilde{\sigma}_{m-1}$. С учетом сделанных оценок и леммы 2 его длина удовлетворяет неравенству

$$l(\tilde{\sigma}) \le \sum_{0 \le i \le m-1} h(\mathbf{y}_i) + O(\log n) = h(\mathbf{y}|\mathbf{x}) + O(\log n).$$

Пара $(\mathbf{x}, \mathbf{y}) \in \mathcal{K}_n(\mathbf{a}, \mathbf{b})$ имеет те же параметры, что и $(\mathbf{x}, \mathbf{y})$, поэтому $h_n(\mathbf{y}|\mathbf{x}) = h_n(\mathbf{b}|\mathbf{a})$ и оценка переписывается в виде $l(\tilde{\sigma}) \le h(\mathbf{b}|\mathbf{a}) + O(\log n)$.

По слову $\tilde{\sigma}$ могут быть найдены доопределения всех $\mathbf{y}_i$. Действительно, слово $\tilde{\lambda}_0$ является минимальным началом слова $\tilde{\sigma}$, состоящим из парных букв, заканчивающихся 01. Оно задает длину слова $\tilde{\alpha}_0 \tilde{\beta}_0$ и позволяет его найти. По нему однозначно восстанавливаются слова $\tilde{\alpha}_0$ (оно имеет фиксированную длину) и $\tilde{\beta}_0$. Слово $\tilde{\alpha}_0$ определяет класс $\mathcal{K}_n(\mathbf{y}_0)$, а $\tilde{\beta}_0$ задает доопределение для $\mathbf{b}_0$. Далее от слова $\tilde{\sigma}$ отсекается $\tilde{\sigma}_0 = \tilde{\lambda}_0 \tilde{\alpha}_0 \tilde{\beta}_0$, и те же рассуждения применяются к следующему слову $\tilde{\sigma}_1$.

По произвольному доопределению $\mathbf{x}'$ последовательности $\mathbf{x}$ и слову $\tilde{\sigma}$ можно найти некоторое доопределение $\mathbf{y}'$ для $\mathbf{y}$. В качестве y'_j $(1 \le j \le n)$

следует взять j-й разряд доопределения $\mathbf{y}'_i$, где i определяется значением $a_i = x'_j$. Действительно, если y_j является значащим символом, то в силу того, что $\mathbf{x}$ конкретней $\mathbf{y}$, символ x_j является основным и совпадает с x'_j. Поэтому символ y_j принадлежит указанной последовательности $\mathbf{y}_i$ и ее доопределение $\mathbf{y}'_i$ содержит доопределение для y_j.

Теорема доказана.

Замечание. Из теоремы 5 и мощностных соображений нетрудно заключить, что существует константа c' такая, что для почти всех (при $n \to \infty$) пар последовательностей $(\mathbf{x}, \mathbf{y}) \in \mathcal{K}_n(\mathbf{a}, \mathbf{b})$ оценка теоремы 12 не может быть понижена до $h_n(\mathbf{b}|\mathbf{a}) - c' \log n$.

Таким образом, если источник S конкретней $\hat{S}$, то условная энтропия $\mathcal{H}(\hat{S}|S)$, может быть введена подобно тому, как это делается в теории информации. Она удовлетворяет правилу сложения энтропий (теорема 11) и допускает алгоритмическую интерпретацию в терминах относительной сложности (теорема 12). В общем случае недоопределенных данных правило сложения энтропий и сложностная интерпретация оказываются несовместными. Проявления этого встречаются в [12].

Если S конкретней $\hat{S}$, мера *информации в S о $\hat{S}$*

$$\mathcal{I}(S, \hat{S}) = \mathcal{H}(\hat{S}) - \mathcal{H}(\hat{S}|S),$$

приводится с помощью правила сложения энтропий к известному виду [3]

$$\mathcal{I}(S, \hat{S}) = \mathcal{H}(S) + \mathcal{H}(\hat{S}) - \mathcal{H}(S\hat{S}).$$

Симметричная форма здесь не означает симметрию информации, ибо требования к S и $\hat{S}$ различны. Поскольку все полностью определенные источники равно конкретны, для них информация оказывается симметричной [3]. В общем случае недоопределенных данных симметрия отсутствует.

Работа выполнена при поддержке Отделения информационных технологий и вычислительных систем РАН (программа "Фундаментальные проблемы информационных технологий и систем") и Российского фонда фундаментальных исследований (проекты 06-01-00577 и 06-07-89293).

Литература

1. Бонгард М. М., О понятии "полезная информация" // Проблемы кибернетики. Вып. 9. — М.: Физматгиз, 1963. — С. 71–102.

2. Вероятность и математическая статистика. Энциклопедия. — М: Большая Российская энциклопедия, 1999.

3. Галлагер Р., Теория информации и надежная связь. — М.: Советское радио, 1974.

4. Добрушин Р. Л., Единые способы оптимального квантования сообщений // Проблемы кибернетики. Вып. 22. — М.: Наука, 1970. — С. 107–156.

5. Колмогоров А. Н., Алгоритм, информация, сложность. — М.: Знание, 1991.

6. Кричевский Р. Е., Сжатие и поиск информации. — М.: Радио и связь, 1989.

7. Нечипорук Э. И., О сложности вентильных схем, реализующих булевские матрицы с неопределенными элементами // ДАН СССР. — 1965. — Т. 163, №1. — С. 40–42.

8. Шоломов Л. А., Информационные свойства функционалов сложности для систем недоопределенных булевых функций // Проблемы кибернетики. Вып. 34. — М: Наука, 1978. — С. 133–150.

9. Шоломов Л. А., Энтропия системы частично определенных последовательностей с вложенными областями определения // Нелинейная динамика и управление. Вып. 3. — М.: Физматлит, 2003. — С. 305–320.

10. Шоломов Л. А., Сжатие частично определенной информации // Нелинейная динамика и управление. Вып. 4. М.: Физматлит, 2004. С. 385–399.

11. Шоломов Л. А., О мере информации нечетких и частично-определенных данных // Доклады Академии наук. 2006. Т. 410. №1. С. 321–325.

12. Шоломов Л. А., О сложности последовательной реализации частичных булевых функций схемами // Дискретный анализ и исследование операций. Сер. 1, 2007. Т. 12, №. 3. С. 110–139.

13. Яблонский С. В., Введение в дискретную математику. — М.: Высшая школа, 2006.

14. Berger T., Rate distortion theory. A mathtmatical basis for data compression. New Jersey: Prentice-Hall, 1971.

ТРИАНГУЛЯЦИИ ПОЛИТОПОВ, f-ВЕКТОРЫ И БУЛЕВЫ ФУНКЦИИ

В. Н. Шевченко

Нижегородский государственный университет им. Н.И. Лобачевского, факультет вычислительной математики и кибернетики, 603950, г. Нижний Новгород, пр. Гагарина, 23

e-mail: shev@uic.nnov.ru

Цель доклада — "ввести читателя в чарующий мир выпуклых многогранников". Этой цитатой из прекрасной книги [2] я обычно начинаю свой курс "Комбинаторная теория многоранников"для магистров факультета ВМК ННГУ. Здесь предлагается краткая версия этого курса.

В п. 1 рассматриваются конечные системы линейных неравенств, которым соответствуют многогранные конусы (далее — просто конусы) в однородном случае и выпуклые многогранники (называемые далее полиэдрами или политопами, если они ограничены) — в неоднородном. Для обоих случаев приводится алгоритм, позволяющий описать параметрически множество решений системы, а также решить обратную задачу. Ставится вопрос об эффективности этого алгоритма и о сложности поставленной задачи.

В п. 2 вводится множество $\Gamma_i(P)$ i-мерных граней политопа P и его f-вектор $f(P)$, компонентами которого являются числа $f_i(P) = |\Gamma_i(P)|$. Рассматривается еще один алгоритм, решающий поставленные в п. 1 задачи более эффективно.

Триангуляция $T(P)$ политопа P, множество $\Gamma_i(T(P))$ ее i-мерных граней и ее f-вектор $f(T(P))$ рассматриваются в п. 3. Там же вводятся булевы функции, равные нулю на характеристических векторах граней $T(P)$.

1. Пусть R — поле вещественных чисел,

$$R_+ = \{\alpha \in R / \alpha \geq 0\},$$

R^d — d-мерное (евклидово) пространство над R и $R^{m \times n}$ — множество вещественных матриц с m строками и n столбцами. Если $M \in R^d$ и $\alpha \in R_+$, то $\alpha M = \{\alpha x, x \in M\}$, если $M_1 \in R^d$ и $M_2 \in R^d$, то $M_1 + M_2 = \{x^{(1)} + x^{(2)}, x^{(1)} \in M_1, x^{(2)} \in M_2\}$.

Множество M из R^d называется *выпуклым*, если для любого $\alpha \in R$ такого, что $0 \leq \alpha \leq 1$

$$x \in M, y \in M \Rightarrow (1-\alpha)x + \alpha y \in M.$$

Говорят, что M *r-мерно* (и пишут $\dim M = r$), если r — максимальное число аффинно независимых векторов (точек) из M. По определению пустое множество выпукло и $\dim \oslash = -1$.

Ясно, что пересечение выпуклых множеств выпукло и R^d — пример выпуклого множества. Поэтому пересечение всех выпуклых множеств $M^{\backprime}$ из R^d, содержащих M, является наименьшим (по включению) выпуклым множеством, содержащим M. Оно называется *выпуклой оболочкой* множества M и обозначается $\operatorname{conv} M$. Если M — конечное множество, то $\operatorname{conv} M$ называется *политопом*. Множество P решений конечной системы линейных неравенств

$$\sum_{k=1}^{d} a_{ik} x_k \leq a_{i0}, i = 1, \ldots, m \tag{1}$$

называется *полиэдром*.

Непустое множество K из R^d назовем *конусом*, если $K + K = K$ и $\forall \lambda \in R_+ \ \lambda K \subseteq K$. Наименьший (по включению) из конусов, содержащий множество M, называется *конической оболочкой* M (обозначение $\operatorname{conv} M$). Если M — конечное множество, то $\operatorname{cone} M$ называется *конечно-порожденным конусом*. Конус K называется *конечно-определенным*, если его можно задать как множество решений некоторой однородной системы линейных неравенств.

Например, с матрицей A из $R^{m \times n}$ можно связать четыре конуса соответственно в m-мерном пространстве столбцов, в n-мерном пространстве строк, в n-мерном пространстве столбцов, в m-мерном пространстве строк:

$$A^{\angle} = \{Ax, x \geq 0\},$$

$$A_{\angle} = \{yA, y \geq 0\},$$

$$A^* = \{x \ / \ Ax \geq 0\},$$

$$A_* = \{y \ / \ yA \geq 0\},$$

из которых первые два — конечно-порожденные, а последние два — конечно-определенные.

Следующий фундаментальный факт состоит в том, что понятия конечно-порожденного и конечно-определенного конусов совпадают (имеется еще один синоним этих понятий — *полиэдральный*) [4, 5, 10, 11, 23].

Пример 1. Пусть L — подпространство в R^d, базис которого составляют столбцы $b_1, b_2, \ldots, b_r$ матрицы B, а базис ортогонального дополнения $L^\perp$ — строки $a_1, a_2, \ldots, a_{d-r}$ матрицы A. Тогда L можно представить двумя способами: $L = (B, -B)^\angle$ и $L = \{x \in R^d \;/\; Ax \geq 0, -Ax \geq 0\} = \binom{A}{-A}^*$. Аналогично, $L^\perp = \binom{A}{-A}_\angle = (B, -B)_*$. Заметим, что эти представления *неприводимы*, т. е. ни один из элементов, порождающих конус, выбросить нельзя. Положив $B_0 = -\sum_{j=1}^{r} b_j$, получим минимальное (с наименьшим числом порождающих векторов) представление $L = (b_0, B)^\angle$.

Теорема 1. (Г. Минковский, Ю. Фаркаш, Г. Вейль)

1.1. $\forall\ A \in R^{m \times d}\ \exists\ B \in R^{d \times n}\ /\ A^* = B^\angle$,

1.2. $\forall\ B \in R^{d \times n}\ \exists\ A \in R^{m \times d}\ /\ B^\angle = A^*$,

1.3. $A^* = B^\angle \ \Leftrightarrow\ A_\angle = B_*$.

Для доказательства первого утверждения Г. Минковский предложил следующий алгоритм (обозначим его Ω_1). На первом этапе найдем ранг r матрицы A и при $r < d$ минимальное представление (см. пример 1) подпространства $L = \{x \in R^d \;/\; Ax = 0\}$ — получим первые $d - r + 1$ столбцов матрицы B. При $r = d$ этот этап опускается. На втором этапе найдем список $I_1, I_2, \ldots, I_s$ "максимальных предранговых" подмножеств I таких, что $I \subseteq \{1, \ldots, m\}$, ранг подматрицы $A(I)$ равен $r - 1$, но для любого I', содержащего I, ранг $A(I')$ равен r. Для каждого $k = 1, \ldots, s$ находим $x^{(k)}$ такой, что $A(I_k)x^{(k)} = 0$, но $x^{(k)} \notin L$. Если найдутся такие i' и i'', при которых $a_{i'}x^{(k)} > 0$, и $a_{i''}x^{(k)} < 0$, то переходим к следующему значению k. Иначе заносим $x^{(k)}$, если $Ax^{(k)} \geq 0$, или $(-x^{(k)})$, если $Ax^{(k)} \leq 0$, очередным столбцом в матрицу B. Столбцы построенной таким образом матрицы B составляют минимальное из порождающих A^* множество, называемое *остовом* конуса A^*. Нетрудно видеть, что при $r = d$ векторы остова определяются однозначно с точностью до умножения на положительное число и $n \leq \binom{m}{d-1}$. Можно ли уменьшить этот перебор, будет ясно из дальнейшего.

Пример 2. Если $m = d = r$, то в качестве B можно взять A^{-1}.

Итак, алгоритм Ω_1 дает решение двух задач:

Pr_1 : по матрице A найти матрицу B такую, что $A^* = B^\angle$,

Pr_2 : по матрице B найти матрицу A такую, что $B^\angle = A^*$, т. е. дает переход от одного способа задания полиэдрального конуса к другому и, в частности, дает возможность избавиться от лишних неравенств (следствий системы $Ax \geq 0$), построив остов конуса $A_\angle = B_*$. Он позволяет также находить сумму и пересечение полиэдральных конусов.

Вернемся к полиэдру P, заданному системой (1), и поставим ей в соответствие однородную систему линейных неравенств

$$x_0 \geq 0,\ a_{i0}x_0 - \sum_{k=1}^{d} a_{ik}x_k \geq 0, \quad i = 1, \ldots, m, \tag{2}$$

задающую конечно-определенный конус $K_P = A^*$ в R^{d+1}. Пусть матрица $B = (b_{kj}), k = 0, 1, \ldots, d, j = 1, \ldots, n$, задает остов конуса A^*. Если $b_{0j} = 0$ при всех $j = 1, \ldots, n$, то $P = \oslash$. Иначе, не уменьшая общности, можно считать, что $b_{0j} > 0$ при $j = 1, \ldots, s$ и $b_{0j} = 0$ при $j = s+1, \ldots, n$. Для $k = 1, \ldots, d$ положим $v_{kj} = b_{kj} \,/\, b_{0j}$ при $j \leq s$, $v_{kj} = q_{kj}$ при $j = s+1, \ldots, n$ и обозначим через v_j столбец, k-я компонента которого равна v_{kj}.

Теорема 1 позволяет перейти от задания полиэдра P в виде системы (1) к его параметрическому представлению

$$P = \{\sum_{j=1}^{n} \alpha_j v_j, \sum_{j=1}^{s} \alpha_j = 1, \alpha_j \in R_+ (j = 1, \ldots, n)\}, \tag{3}$$

и обратно. Говорят, что P *порождается точками* $v_1, \ldots, v_s$ *и направлениями* $v_{s+1}, \ldots, v_n$, называемыми *рецессивными*. Ясно, что

$$(v_{s+1}, \ldots, v_n)^\angle = \{x \in R^d \,/\, \sum_{k=1}^{d} a_{ik}x_k \leq 0, i = 1, \ldots, m\}$$

и условие $s = n$ необходимо и достаточно для ограниченности P.

Теорема 2. [4, 5, 10, 11, 18, 22, 23]

2.1. (Т. Моцкин) P — *полиэдр в* $R^d \Leftrightarrow P = P_1 + K$, *где* P_1 — *политоп, а* K — *конечно-порожденный конус.*

2.2. (Г. Минковский, Э. Штейниц, Г. Вейль) P — *политоп* $\Leftrightarrow P$ — *ограниченный полиэдр.*

Наличие двух способов описания полиэдра P : в виде (1) и в виде (3) — позволяет решать разнообразные задачи о полиэдрах, например, находить их сумму и пересечение, сводя их к соответствующим задачам о конусах. Итак, алгоритм Ω_1 дает решение двух (сводящихся одна к другой) задач:

Pr_3 : по системе (1) найти представление (3),

Pr_4 : по представлению (3) найти систему (1).

Имеются ли более эффективные, чем Ω_1, алгоритмы? Какова сложность задачи Pr_1? Известно [9], что при $d = 2, 3$ эта величина растет с ростом n пропорционально величине $n \log n$.

2. Пусть $K = A^* = B^{\angle},\ C = AB$,

$$A = (a_i, i = 1, \ldots, m) = (a_{ik}) \in R^{m \times d}$$

$$B = (b_j, j = 1, \ldots, n) = (b_{kj}) \in R^{d \times n}$$

$$c_{ij} = a_i b_j = \sum_{k=1}^{d} a_{ik} b_{kj} \geq 0, C = (c_{ij}) \in R^{m \times n},$$

при $I \subseteq \{1, \ldots, m\}$ $A(I) = (a_i, i \in I)$, при $J \subseteq \{1, \ldots, n\},\ B(J) = (b_j, j \in J)$, $J_i = \{j / c_{ij} = 0\},\ I_j = \{i / c_{ij} = 0\}$.

Для каждого $a \in A_{\angle}$ определим грань K_a конуса K равенством $K_a = \{x \in K \,/\, ax = 0\} = \{x \in K \,/\, ax \leq 0\}$ и рассмотрим множество $\Gamma(K) = \{K_a,\ a \in A_{\angle}\}$ всех граней конуса K, частично упорядоченное отношением включения. Ясно, что K_a — конечно-порожденный конус и при $J(a) = \{j \,/\, ab_j = 0\}$ $K_a = B^{\angle}(J_a)$. Отсюда следует конечность множества $\Gamma(K)$ и равенство

$$K_{a'} \bigcap K_{a''} = K_{a'+a''}. \tag{4}$$

В частности, взяв $a = 0$, получим $K_0 = K$ — максимальный элемент $\Gamma(K)$. Минимальный элемент $\Gamma(K)$ — подпространство $L = \{x \in R^d \,/\, Ax = 0\}$ — можно получить, положив $a = \sum_{i=1}^{m} a_i$.

На языке частично упорядоченных множеств (терминологию и необходимые сведения можно получить в [1–3, 21, 23]) равенство (4) означает, что $\Gamma(K)$ замкнуто относительно операции min.

Аналогичное утверждение относительно max можно доказать, обозначив через I, I', I'' соответственно множества положительных коэффициентов в разложениях строк a, a', a'' из $A_{\angle}$:

$$a = \sum_{i \in I} \mu_i a_i, \quad a' = \sum_{i \in I'} \mu'_i a_i, \quad a'' = \sum_{i \in I''} \mu''_i a_i.$$

Тогда $K_a = \bigcap_{i \in I} B^{\angle}(J_i)$ и, следовательно, при $I = I' \bigcup I''$ $K_{a'} + K_{a''} = K_a$. Отсюда следует, что множество $\Gamma(K)$ является *решеткой* (относительно включения).

Полиэдральный конус K' называется *комбинаторно эквивалентным* конусу K ($K \sim K'$), если существует взаимно однозначное отображение множества $\Gamma(K)$ на множество $\Gamma(K')$, сохраняющее отношение включения (при этом решетки называют *изоморфными* и пишут $\Gamma(K) \approx \Gamma(K')$).

Нетрудно теперь перенести понятие грани и полученные результаты на полиэдр P, перейдя, как в п.1, от (3) к конусу $K = K_P = B^{\angle}$. Сделаем это, считая далее для простоты, что P — политоп, т. е. $s = n$. Тогда можно положить при $j = 1, \ldots, n$ $\quad b_{0j} = 1$ и $b_{kj} = v_{kj}$ $\quad (k = 1, \ldots, d)$. Теперь если $0 \neq (x_0, x) \in K$, то $x_0 \neq 0$ и точка $\pi(x_0, x) = x \,/\, x_0 \in P$. Ясно, что $\pi(b_j) = v_j$ и $\pi(B^{\angle}(J)) = \mathrm{conv}(v_j, j \in J)$. Если $F' \in \Gamma(K)$ и $F' \neq \{0\}$, то назовем $F = \pi(F')$ гранью политопа P. Положив $\pi(\{0\}) = \varnothing$, обозначим через $\Gamma(P) = \pi(\Gamma(K))$ множество граней политопа P.

Теорема 3. (см., например, [23]) *Для любого политопа* P

3.1. $\Gamma(P)$ — *градуированная решетка длины* $\dim P = \dim K_P - 1$ *с ранговой функцией* $r(F) = \dim F + 1 \quad \forall F \in \Gamma(P)$.

3.2. *Если* $G \subseteq F$, *то интервал* $[G, F]$, *т. е. множество граней* H *из* $\Gamma(P)$ *таких, что* $G \subseteq H \subseteq F$, *изоморфен* $\Gamma(P')$ *для некоторого политопа* P' *размерности* $r(F) - r(G) - 1$.

3.3. *Каждый интервал* $[G, F]$ *длины* 2 (*т. е.* $\dim F - \dim G = 2$) *имеет ровно* 4 *элемента.*

Пример 3. Если $d = 2$ и $P = \mathrm{conv}\,(v_1, v_2, v_3, v_4)$ — квадрат, то следующая таблица задает грани P и их ранги.

Таблица 1

1	0	1 0 0 0	1 0 1 0	1
2	0	0 1 0 0	1 1 0 0	1
3	0	0 0 1 0	0 1 0 1	1
4	0	0 0 0 1	0 0 1 1	1
$r(F)$	0	1 1 1 1	2 2 2 2	3

Обозначим через $\Gamma_k(P)$ множество k-мерных граней r-мерного политопа P и положим $\partial P = \Gamma(P) \setminus \{P\}$, $f(P) = (f_{-1}(P), f_0(P), \ldots, f_r(P))$, где $f_k(P) = |\Gamma_k(P)|$ — число k-мерных граней политопа ,

$$f(\lambda, P) = \sum_{k=-1}^{r} f_k(P) \lambda^{k+1}$$

и $f(\lambda, \partial P) = f(\lambda, P) - \lambda^{r+1}$.

Множество ∂P называется *граничным комплексом* политопа P.

Пример 4. Политоп P называется r-*симплексом*, если $\dim P = f_0(P) - 1 = r$. Нетрудно видеть, что для него $\Gamma_k(P)$ составляет $(k+1)$-мерный слой $(r+1)$-мерного булева куба, $f_k(P) = \binom{r+1}{k+1}$, $f(\lambda, P) = (1+\lambda)^{r+1}$.

Пример 5. Рассмотрим d-мерный куб $C_d = \{x = (x_1, \ldots, x_d) / -1 \leq x_i \leq 1, i = 1, \ldots, d\}$. Соответствующая ему система (2) имеет вид

$$x_0 - x_i \geq 0,\ x_0 + x_i \geq 0\ (i = 1, \ldots, d) \tag{5}$$

— неравенство $x_0 \geq 0$ является ее следствием и его можно отбросить. Ясно, что для получения любой грани F куба достаточно выбрать два непересекающихся подмножества I_1 и I_2 множества $\{1, \ldots, d\}$ и рассмотреть пересечение C_d с множеством решений системы

$$x_i = 1\ (i \in I_1),\ x_i = -1\ (i \in I_2).$$

Нетрудно видеть, что $\dim F = d - |I_1| - |I_2|$ и $f(\lambda, C_d) = 1 + \lambda(2+\lambda)^d$.

Пример 6. Политоп $P_d = \mathrm{conv}(-E_d, E_d)$, где E_d — единичная матрица порядка d, называемый кросс-политопом (при $d = 3$ это октаэдр). Для него $f(\lambda, P_d) = (1+2\lambda)^d + \lambda^{d+1}$.

Естественно возникает задача

Pr_5 : найти критерий реализуемости f-вектора, т. е. необходимые и достаточные условия, которым должны удовлетворять компоненты целочисленного неотрицательного вектора $f = (f_{-1}, f_0, \ldots, f_d)$ для того, чтобы он совпадал с $f(P)$ для некоторого d-мерного политопа P. Ясно, что $f_{-1} = f_d = 1$ и что при $d = 2$ такой критерий дает условие $f_0 = f_1$.

Приведем решение задачи Pr_5 при $d = 3$, полученное в 1922 г. Э. Штейницем [22].

Теорема 4. *Целочисленный вектор* $f = (1, f_0, f_1, f_2, 1)$ *является* f*-вектором некоторого 3-мерного политопа тогда и только тогда, когда*

$$f_1 - f_0 + f_2 = 2,\quad 4 \leq f_2 \leq 2f_0 - 4,\quad 4 \leq f_0 \leq 2f_2 - 4.$$

При $d \geq 4$ задача Pr_5 не решена. В общем случае известно необходимое условие Эйлера-Пуанкаре

$$\sum_{k=-1}^{d} f_k (-1)^k = 0. \tag{6}$$

Весьма полезна также доказанная в [16] (в более общем варианте) Г. Бругессером и П. Мани.

Теорема 5. *Граничный комплекс ∂P имеет линейную развертку.*

Конструкция, лежащая в основе доказательства этого утверждения, имеет прозрачный геометрический смысл. Рассматривается прямая $v_\lambda = \lambda q$, где $0 \in int\, P$ ($int\, P = P \setminus \partial P$ — множество внутренних точек политопа P), λ — вещественное число, $a_i q \neq 0$ при $i = 1, \ldots, m$ и $\lambda_i = a_{i0}/a_i q$ — значение параметра λ, при котором прямая пересекает аффинную оболочку грани F_i. Ясно, что существует такое q, что $\lambda_i \neq \lambda_k$ при $i \neq k$. Тогда, переупорядочив $\Gamma_{d-1}(P) = \{F_1, \ldots, F_m\}$ неравенствами

$$\frac{1}{\lambda_1} > \cdots > \frac{1}{\lambda_m}, \tag{7}$$

получим развертку граневого комплекса ∂P, называемую *линейной*. Определение общего понятия *развертки* политопиального (и, в частности, симплициального) комплекса можно найти в [2, 3, 21, 23].

Концепция разворачиваемости привела к появлению ряда алгоритмов (Черниковой, Моцкина–Бургера, Фурье–Моцкина, метод двойного описания) — будем называть их ΦM-алгоритмами, — позволяющих решать задачи $Pr_i (i = 1, 2, 3, 4)$ в режиме *on-line* [6, 9, 12, 15, 17, 23].

Под *ΦM-алгоритмом* Ω_2 будем понимать следующую процедуру, позволяющую по заданной последовательности точек $v_1, \ldots, v_n$ построить неприводимую систему (1), описывающую политоп $P = P_n = \mathrm{conv}\,(v_1, \ldots, v_n)$. Здесь не требуется, чтобы $v_j \in \Gamma_0(P)$; потребуем лишь, чтобы политоп P_{d+1} был d-симплексом и $0 \in int\, P_{d+1}$ (ограничения технического характера). Предположив, что задача решена для политопа P (при $n = d + 1$ это просто) и появилась новая точка v_{n+1}, покажем ее решение для политопа $P' = P_{n+1}$. Положим $\mu_i = a_{i0} - a_i v_{n+1}$ и разобьем множество $1, \ldots, m$ на три подмножества

$$I_- = \{i/\mu_i < 0\},\ I_+ = \{i/\mu_i > 0\},\ I_0 = \{i/\mu_i = 0\}.$$

Далее для каждого $i \in I_-$ сформируем множество

$$M_i = \{i' \in I_+ / F_i \cap F_{i'} \in \Gamma_{d-2}(P)\}$$

и для каждого $i' \in M_i$ образуем неравенство

$$(\mu_{i'} a_i - \mu_i a_{i'})x \leq \mu_{i'} a_{i0} - \mu_i a_{i'0}.$$

Если к системе (1) добавить все полученные таким способом неравенства и затем выбросить неравенства с номерами из I_-, то новая система будет описывать политоп P'. О неприводимости новой системы следует позаботиться дополнительно.

Существуют реализации ФМ-алгоритмов, трудоемкость которых пропорциональна величине $n^{\lfloor d/2\rfloor+1}$. С другой стороны, известно [2, 3, 5, 20, 23], что существуют такие d-политопы P с n вершинами, для которых

$$f_{d-1}(P)=\binom{n-\lfloor (d+1)/2\rfloor}{\lfloor d/2\rfloor}+\binom{n-\lfloor d/2\rfloor-1}{\lfloor (d-1)/2\rfloor}, \tag{8}$$

что дает неулучшаемую нижнюю оценку трудоемкости задач P_{r_i} $(i=1,2,3,4)$

3. Множество $V=\{v_1,...,v_n\}$, где $v_j\in\mathbf{R}^d$, назовем (d,n)-*точечной конфигурацией*, если $P=\operatorname{conv}V$ есть d-политоп. *Триангуляцией политопа P с узлами из множества V* назовем множество $T(V)=\{S_1,\dots,S_t\}$ таких d-симплексов S_τ, для которых выполнены три следующие условия [3, 21, 23]:

1)$\Gamma_0(S_\tau)\subseteq V$,

2) $\bigcup_{\tau=1}^t S_\tau=P$ и

3) пересечение любых двух d-симплексов является гранью каждого из них.

Тогда множество $\triangle(T(V))=\bigcup_{\tau=1}^t\Gamma(S_\tau)$ дает пример симплициального комплекса (с. к.) — одного из основных понятий топологии и комбинаторной геометрии [2, 3, 5, 18, 21, 23]. Коллекция (непустая) C подмножеств F (называемых *гранями* с. к. C) конечного множества $V=\{v_1,\dots,v_n\}$ называется *симлициальным комплексом на множестве* V, если $F\in C$ и $G\subseteq F$ влечет $G\in C$. Ясно, что для задания с. к. C достаточно указать список $S_1,S_2,\dots,S_t$ максимальных (по включению) граней с. к. C или список $N_1,\dots,N_l$ минимальных (по включению) подмножеств множества V, не принадлежащих C. Аналогично прежнему можно определить $\Gamma_k(C)=\bigcup_{\tau=1}^t(\Gamma_k(S_\tau)$, $f_k(C)=|\Gamma_k(C)|$, вектор $f(C)$ и многочлен $f(\lambda,C)$. Для любой грани F с. к. C определим характеристический вектор $\xi_F=(\xi_1,\dots,\xi_n)$, где $\xi_i=1$ при $v_i\in F$, и $\xi_i=0$ при $v_i\notin F$, и булеву функцию $\varphi_C(v_1,\dots,v_n)$, положив $\varphi_C=0$ тогда и только тогда, когда $F\in C$. Нетрудно видеть, что справедлива

Теорема 6. *Для любого с. к. C булева функция $\varphi_C(v_1,\dots,v_n)$ монотонна и*

$$\varphi_C(v_1,\dots,v_n)=\&_{\tau=1}^t(\vee_{j\notin S_\tau}v_j)=\vee_{\lambda=1}^l\&_{j\in N_\lambda}v_j. \tag{9}$$

Будем писать $\triangle$ вместо $\triangle(T(V))$, если это не приводит к недоразумениям. Заметим, что при любой фиксированной размерности d сложность функции $\varphi_{\triangle}(v_1, \ldots, v_n)$ ограничена сверху некоторым полиномом от n. Для решения вопросов, связанных с ее эффективным представлением (в различных базисах) отошлем к обзору [8]. Весьма полезным может оказаться применение пороговых булевых функций, так как, используя результаты В. К. Коробкова [7], множество нулей функции $\varphi_{\partial P}$ можно описать системой линейных неравенств

$$\sum_{j \in N_k} v_j \leq |N_k| - 1, k = 1, \ldots, l \tag{10}$$

Пример 7. Зададим (2, 5)-точечную конфигурацию таблицей 2 и характеристические векторы симплексов S_1, S_2, S_3, S_4 — таблицей 3

Таблица 2

j	1	2	3	4	5
v_{1j}	1	0	0	2	-1
v_{2j}	0	1	-1	0	0

Таблица 3

	S_1	S_2	S_3	S_4
1	1	1	1	0
2	1	1	0	1
3	1	0	1	1
4	0	1	1	0
5	0	0	0	1

Тогда $T(V) = \{S_1, S_2, S_3, S_4\}$ — триангуляция выпуклого четырехугольника $P = \mathrm{conv}\,(v_2, v_3, v_4, v_5)$,

$$\begin{aligned} f(\lambda, \triangle) &= 1 + 5\lambda + 8\lambda^2 + 4\lambda^3, \\ \varphi_{\triangle}(v_1, v_2, v_3, v_4, v_5) &= (v_4 \vee v_5)(v_3 \vee v_5)(v_2 \vee v_5)(v_1 \vee v_4) \\ &= (v_2 v_3 v_4 \vee v_5)(v_1 \vee v_4) = v_1 v_5 \vee v_4 v_5 \vee v_1 v_2 v_3 v_4. \end{aligned}$$

Пример 8. Из списка максимальных граней с. к. $\triangle$ из предыдущего примера выбросим S_1 и рассмотрим получающийся при этом подкомплекс C_1. Очевидно, множество $\{S_2, S_3, S_4\}$ триангуляцией не является. Для него

$f(\lambda, C_1) = 1 + 5\lambda + 8\lambda^2 + 3\lambda^3$,

$\varphi_{C_1}(v_1, v_2, v_3, v_4, v_5) = (v_2 v_3 \vee v_5)(v_1 \vee v_4) = v_1 v_5 \vee v_1 v_4 \vee v_1 v_2 v_3$.

Один из алгоритмов построения триангуляции основан на следующем результате.

Теорема 7. [13] *Пусть A — матрица со строками $a_1, \ldots, a_m$, B — матрица со столбцами $b_1, \ldots, b_n$ такие, что $A^* = B^{\angle}$, $J_i = \{j / a_i b =$*

$0\}$, $B(J_i) = \{b_j, j \in J_i\}$ *и* $b \notin B^{\angle}$. *Тогда*

$$(b, B)^{\angle} = \bigcup_{i \in I} (b, B(J_i))^{\angle} \bigcup B^{\angle}, \tag{10}$$

где $I = \{i / a_i b < 0\}$, *и пересечение любых конусов, объединяемых в правой части равенства* (10), *является общей гранью каждого из них.*

Опишем модификацию ΦM-алгоритма (обозначим его через Ω_3), решающего следующую задачу

Pr_6 : по данной точечной конфигурации V найти ее триангуляцию и систему (1), описывающую политоп $P = \operatorname{conv} V$.

Будем предполагать для облегчения изложения, что первые $(d+1)$ столбцов матрицы V афинно независимы, и перейдем, как и в п. 1, на язык конусов, поставив в соответствие $d \times n$-матрице V матрицу $B_n \in R^{(d+1)\times n}$, у которой $b_{0j} = 1 \quad (j = 1, 2, \ldots, n)$.

При $n = d + 1$ задача Pr_6 решается просто: триангуляция состоит из единственного симплекса, а коэффициенты системы (1) получаются из матрицы B_{d+1}^{-1}. Считая, что задача Pr_6 решена для предыдущего значения n, рассмотрим следующий вектор b_{n+1}. Если $b_{n+1} \in B^{\angle}$, то решение задачи не меняем. Иначе найдем множество $I = \{i / a_i b_{n+1} < 0\}$ и дополним полученную на предыдущем шаге триангуляцию согласно теореме 7.

Этот и другие способы триангуляции описываются в огромном количестве работ, из которых укажем лишь несколько [3, 14, 15, 19, 21, 23].

Заметим, что триангуляцию из примера 7 можно получить алгоритмом Ω_3, подавая точки v_j в указанном порядке. Изменив порядок, можно получить другую триангуляцию и даже с другим f-вектором. Уже это обстоятельство порождает много вопросов (например, аналог задачи Pr_5 для триангуляций), для точной постановки которых введенная здесь булева функция $\varphi_{\triangle}$ может оказаться весьма полезной.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект 05-01-00552а)

Литература

1. Биркгоф Г., Теория решеток. М.: Наука, 1984.

2. Бренстед А., Введение в теорию выпуклых многогранников. М. : Мир, 1988.

3. Бухштабер В. М., Панов Т. Е. Торические действия в топологии и комбинаторике М.: МЦНМО, 2004.

4. Вейль Г., Элементарная теория выпуклых многогранников. В кн.: Матричные игры. М., 1961.

5. Емеличев В. А., Ковалев М. М., Кравцов М. К., Многогранники, графы, оптимизация. М.: Наука, 1981.

6. Золотых Н. Ю., Программная реалимзация алгоритма Моцкина Бюргера построения остова многогранного конуса и ее применение // Труды Второй международной конференции "Математические алгоритмы". Н. Новгород: Изд-во Нижегородского университета, 1997, 72–74.

7. Коробков В. К., О некоторых целочисленных задачах линейного программирования // Проблемы Кибернетики. 1965 г. Вып. 14. М.: Наука, 297–299.

8. Коршунов А. Д., Монотонные булевы функции // Успехи математических наук. 2003. т. 58, вып. 5(353). 89–162.

9. Препарата Ф., Шеймос М. Вычислительная геометрия: Введение. М.: Мир, 1989.

10. Схрейвер А. Теория линейного и целочисленного программирования: В 2-х т.: Пер. с англ. — М.: Мир, 1991.

11. Черников С. Н., Линейные неравенства. М.: Наука 1968.

12. Черникова Н. В., Алгоритм для нахождения общей формулы формулы неотрицательных решений системы линейных неравенств // Ж. вычисл. матем. и матем. физ., 1965, т.5, — 2, 334–337.

13. Шевченко В. Н., Качественные вопросы целочисленного программирования. М.: Наука, 1995.

14. Шевченко В. Н., О разбиении выпуклого политопа на симплексы без новых вершин // Известия ВУЗ. Математика. 1997, №12 (427), 89–99.

15. Шевченко В. Н., Груздев Д. В., Модификация алгоритма Фурье-Моцкина для построения триангуляции и ее звездной развертки. // Дискретный анализ и исследование операций. Сер. 2. Новосибирск: Изд-во ин-та математики, 2006. т.13, №1. 1–101.

16. Bruggesser H., Mani P., Shellable decompositions of cells and spheres. // Math Scand, 1971, 29, 197, 205.

17. Burger E., Uber homogene lineare Ungleichungssysteme // Z. angew. Math. und Mech, 1956, v/36, Nr/3/4. — 135–139.

18. Grunbaum B., Convex polytopes. N-Y: Wiley and Sons, 1967.

19. Lee C., Regular triangulations of convex politopes // Applied Geometry and Discrete Mathematics — The Victor Klee Festschrift. DIMACS Series in Discrete Mathematics and Theoretical Computer Science. 1991.— V.4 — AMS — P. 443–456.

20. McMullen P., The maximum numbers of faces of a convex polytope // Mathematika. 1970, V.17, 179–184.

21. Stanley R. P., Combinatorics and Commutative Algebra. Progress in mathematics V. 41: Birkhauser, Boston, 1983.

22. Steinitz E., Polyeder und Raumeinteilungen / Encyclopadie der mathematischen Wissenschaften, 1922, Band 3 (Geometrie), 1–139.

23. Ziegler G., Lectures on polytopes. Berlin: Springer-Verlag, 1995.

О СЛОЖНОСТИ РЕАЛИЗАЦИИ СИСТЕМ БУЛЕВЫХ ФУНКЦИЙ СХЕМАМИ ИЗ ФУНКЦИОНАЛЬНЫХ ЭЛЕМЕНТОВ В БАЗИСАХ, СОДЕРЖАЩИХ ЛИНЕЙНУЮ ФУНКЦИЮ

К. А. Зыков

Московский государственный университет им. М. В. Ломоносова, механико-математический факультет, 119992 Москва, Ленинские горы
e-mail: zykost@mail.ru

Трудности, связанные с получением нелинейных нижних оценок сложности схем из функциональных элементов в базисах, содержащих линейную функцию, широко известны. Получать такие оценки для конкретных систем функций удается только при весьма существенных ограничениях.

Высокие нижние оценки получены в работах Д. Ю. Черухина [8, 9]. В частности, в [8] рассматривалась реализация оператора циклической свертки

$$z_j = \bigoplus_{i+k\equiv n \pmod n} x_i y_k, \quad j = 1, \ldots, n,$$

схемами в базисе из всех булевых функций. При этом весом элемента считалось число его входов. В случае, когда глубина схемы не превосходит двух, получена нижняя оценка сложности вида $\Omega(n^{3/2})$.

В статье [10] ее авторы также рассматривают схемы из элементов с произвольным количеством входов, но в более узком базисе. А именно, в базисе из элементов двоичного сложения. Будем говорить, что булева матрица с элементами a_{ji} задает систему линейных булевых функций

$$f_i = \bigoplus_{j=1}^{m} a_{ji} x_j, \quad j = 1, \ldots, m, \quad i = 1, \ldots, n.$$

В [10] получены нелинейные нижние оценки сложности реализации систем функций, задаваемых матрицами Адамара (т. е. $n \times n$ матрицами, в которых любые две строки отличаются ровно в $n/2$ позициях). При этом роль существенного ограничения играет, как и в рассмотренных ранее работах, ограничение на глубину схем. В частности, для схем глубины 2, получена нижняя оценка сложности вида $\Omega(n \log_2 n)$. Доказательство этого результата опирается на следующие алгебраические свойства матрицы Адамара. Пусть S — произвольное подмножество входов схемы, а T — произвольное подмножество выходов схемы. Через $H_{S,T}$ обозначим подматрицу матрицы Адамара H, в которой оставлены только строки, соответствующие множеству S, и столбцы, соответствующие множеству T.

Утверждение 1. *Пусть* $|S| = |T| = n^{1/2+2\epsilon}$. *Тогда для матрицы Адамара* H

$$\operatorname{rank}(H_{S,T}) \geq \epsilon \log_2 n.$$

Утверждение 2. (Валиант [15]). *Число внутренних вершин рассматриваемой схемы, связанных как с вершинами из* S, *так и из* T, *не менее* $\operatorname{rank}(H_{S,T})$.

Утверждение 3. (Алон, Маасс [11]). *Если для любых* S *и* T *таких, что* $|S| = |T| = n^{1/2+2\epsilon}$ *число внутренних вершин рассматриваемой схемы, связанных как с вершинами из* S, *так и из* T, *не менее* $\epsilon \log_2 n$, *то граф имеет по крайней мере* $\Omega(n \log_2 n)$ *ребер.*

Отметим, что в [10] приведено также доказательство несколько более слабой оценки $\Omega(n \log_2 n / \log_2 \log_2 n)$, использующее только комбинаторные рассуждения (теорему Эрдеша-Радо о подсолнечнике).

В то же время, как показал Баблитс [13], нижняя оценка $\Omega(n^2 / \log_2 n)$ функции Шеннона для схем без ограничения на глубину (доказывается мощностным методом) достигается уже на схемах глубины 2. Для этого достаточно разбить переменные на группы по $[\log_2 n/2]$ и реализовать все суммы в каждой группе схемами глубины 1.

Хорошо иллюстрирует трудности получения нижних оценок в базисах содержащих линейную функцию сравнение следующих результатов.

Э. И. Нечипорук показал [6], что для матриц без прямоугольников (т. е. без подматриц размера 2×2 из единиц) сложность реализации вентильными схемами и схемами из функциональных элементов в базисе $\{\&, \vee\}$ равна сложности тривиальной схемы (т. е. такой, что каждая функция реализуется отдельной схемой). С другой стороны, С. Б. Гашков показал, что для схем в базисе из двувходовых элементов двоичного сложения это не так.

Простейшим примером является схема из шести элементов, реализующая систему функций $f_1, \ldots, f_5$:

$$f_1 = x_1 \oplus x_2, \quad f_2 = x_2 \oplus x_4, \quad f_3 = x_4 \oplus x_5,$$
$$g_1 = x_3 \oplus f_2, \quad f_4 = f_1 \oplus g_1, \quad f_5 = g_1 \oplus f_3.$$

Сложность тривиальной реализации соответствующей матрицы

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix},$$

очевидно, равна числу единиц матрицы минус число ненулевых столбцов, что в данном случае составляет 7.

В связи с этим возникают следующие вопросы.

1. Какой экономии можно достичь для матриц без прямоугольников?

2. Возможна ли "экономия" сложности при реализации двух независимых систем функций? А именно, выполняется ли равенство $L(M) = L(A) + L(B)$ для матриц M, A, B, где матрица M имеет вид

$$M = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}.$$

Последний вопрос остается открытым. На первый же частичный ответ дают следующие результаты.

Автору [3], с использованием матрицы Нечипорука [6], удалось построить схему, в которой часть элементов простейшей схемы, приведенной выше, используется для растущего числа подсхем. В результате получена последовательность матриц F_n, для которой

$$\frac{L(F_n)}{L_T(F_n)} \lesssim 1/2,$$

где $L_T(F)$ — сложность тривиальной реализации матрицы F. Вопрос о возможности большей экономии остается открытым.

В рассмотренной выше конструкции глубина схем стремится к бесконечности. Легко показать, что для схем глубины 2 сложность совпадает со сложностью тривиальной реализации: $L^2(A) = L_T(A)$. Для схем глубины

3 конструкция работы [3] позволяет получить последовательность матриц G_n, для которой

$$\frac{L^3(G_n)}{L_T(G_n)} \lesssim 3/4.$$

Ранее, С. Б. Гашков построил последовательность матриц M_n, для которой

$$\frac{L^3(M_n)}{L_T(M_n)} \lesssim 2/3.$$

Применение метода из работы [3] к конструкции Гашкова позволило улучшить оценку до $7/12$, что вместе с легко получаемой нижней оценкой приводит к следующему результату [4]:

$$1/2 \lesssim \frac{L^3(Q_n)}{L_T(Q_n)} \lesssim 7/12.$$

Еще одной открытой проблемой является поведение функции Шеннона в рассматриваемом базисе при реализации систем линейных функций, задаваемых "узкими" матрицами. Для систем из q_n функций зависящих от p_n переменных при выполнении условий $\frac{p_n}{\log_2 q_n} \to \infty$ и $\frac{q_n}{\log_2 p_n} \to \infty$ известна асимптотика $L(p_n, q_n) \sim p_n q_n / \log_2(p_n q_n)$. Здесь нижняя оценка — мощностная, а верхняя следует из оценки Пиппенджера [14] сложности вентильных схем содержащих не более одного пути от произвольного входа к выходу.

Для "узких" матриц ($\frac{p_n}{\log_2 q_n} \to \alpha$, $1 \le \alpha < \infty$) удается получить неравенства $(\alpha - 1)q_n \lesssim L(p_n, q_n) \lesssim [\alpha]q_n$. Нижняя оценка здесь мощностная, а верхняя получается аналогично оценке В. А. Орлова для вентильных схем [7].

Причем известно [16], что нижняя оценка здесь достигается при $\alpha \in N$, $\alpha \neq 1$ и $p_n = \alpha \log_2 q_n - r_n$, где $r_n \to \infty$, а верхняя — при $1 \le \alpha < 2$.

С точки зрения сравнения с тривиальной реализацией систем функций можно рассматривать и результат, полученный Н. Блюмом и М. Сейсеном [12]. Они изучали совместную реализацию функций $AND_n = x_1 \& \ldots \& x_n$ и $NOR_n = \bar{x}_1 \& \ldots \& \bar{x}_n$ схемами из функциональных элементов в базисе из всех двувходовых элементов и показали, что $L(AND_n, NOR_n) = 2(n-1)$. Более того, в [12] доказано, что все минимальные схемы для вычисления системы из этих двух функций представляют собой независимые схему для вычисления AND_n и схему для вычисления NOR_n.

В работах автора [1, 2] тоже рассматривались схемы в базисе из всех двувходовых элементов. При этом на систему реализуемых функций накладывалось только ограничение на взаимное расположение существенных переменных. Оно задавалось циклической матрицей $M_{r,n}$ с элементами

$$m_{1,1} = \cdots = m_{1,r} = 1, \quad m_{1,r+1} = \cdots = m_{1,n} = 0,$$

$$m_{i,1} = m_{i-1,n}, \quad m_{i,j} = m_{i-1,j-1} \text{ при } i, j = 2, \ldots, n.$$

Приведем пример матрицы $M_{r,n}$ при $n = 5$, $r = 3$:

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Пусть $\tilde{x} = (x_1, \ldots, x_n)$. Будем говорить, что структура существенных переменных системы функций $f_1(\tilde{x}), \ldots, f_d(\tilde{x})$ задается матрицей $A = (a_{i,j})$, если функция f_j $(j = 1, \ldots, d)$ зависит от переменной x_i $(i = 1, \ldots, n)$ тогда и только тогда, когда $a_{i,j} = 1$. В следующей теореме номера функций и переменных будем брать по модулю n, т. е. будем считать, что $f_n = f_0$ и т. д.

Теорема 1. *Пусть $f_1(\tilde{x}), \ldots, f_n(\tilde{x})$ — произвольная система булевых функций, структура существенных переменных которой задается матрицей $M_{r,n}$. Пусть, кроме того, $f_{i_1}(\tilde{x}), \ldots, f_{i_y}(\tilde{x})$ — подсистема системы $f_1(\tilde{x}), \ldots, f_n(\tilde{x})$, а $M_{r,n}^{i_1,\ldots,i_d}$ — матрица размера $n \times d$, которая получается из матрицы $M_{r,n}$ вычеркиванием всех столбцов, кроме столбцов с номерами $i_1, \ldots, i_d$ (т. е. структура существенных переменных системы $f_{i_1}(\tilde{x}), \ldots, f_{i_d}(\tilde{x})$ задается матрицей $M_{r,n}^{i_1,\ldots,i_d}$.) Тогда при $n - 1 > r > 2$, $r \geq n/2$, и $d > 1$ имеет место соотношение $L(f_{i_1}, \ldots, f_{i_d}) \geq z + d - c$, где z — число ненулевых строк матрицы $M_{r,n}^{i_1,\ldots,i_d}$ (т. е. число переменных, от которых зависит хотя бы одна функция системы), а константа c равна 4 если $r = n/2$ и выполняется условие (I) и 3 — в противном случае.*

(I) Для некоторого j система функций $f_{i_1}, \ldots, f_{i_d}$ содержит функции f_j и f_{j+r} и не содержит ни одной из функций $f_{j+1}, \ldots, f_{j+r-1}$, либо $d = 4$ и система $f_{i_1}, \ldots, f_{i_d}$ совпадает с системой f_j, f_{j+a}, f_{j+r}, f_{j+a+r} для некоторых j и a.

В частности, при $d = n$ получаем оценку $L(f_1, \ldots, f_n) \geq 2n - 3$.

Полученная в теореме нижняя оценка является в некотором смысле неулучшаемой, так как при $r = n/2$ и $r = 2n/3$ для любых удовлетворяющих условию n и d найдется система линейных функций, для которой данная оценка является точной.

Отметим, что эта теорема допускает обобщение на k-значную логику для схем в базисе из всех двуместных функций $q(x, y)$, реализующих перестановки при любой фиксации любой из переменных. (То есть при всех x_0, y_0 каждая из функций $q(x_0, y)$, $q(x, y_0)$ принимает все k значений [5].)

Работа выполнена при финансовой поддержке РФФИ (проект 05–01–00994) и программы поддержки ведущих научных школ РФ (проект НШ–5400.2006.1).

Литература

1. Зыков К. А. Реализация некоторых систем булевых функций схемами из двувходовых элементов // *Дискретная математика.* — Т. 5, вып. 3. — 1993. — С. 125–149.

2. Зыков К. А. О линейной нижней оценке сложности реализации некоторых систем булевых функций схемами из двувходовых элементов // *Теоретические и прикладные аспекты математических исследований (сборник трудов конференции молодых ученых механико-математического факультета МГУ).* —М.: Из-во МГУ. — 1994. — С. 18–22.

3. Зыков К. А. О сравнении сложности двух способов реализации некоторых линейных булевых преобразований // *Дискретная математика.* — Т. 8, вып. 2. — 1996. — С. 151–159.

4. Зыков К. А. О сложности реализации линейных булевых преобразований схемами глубины 3 // *Вестник Московского университета.* Сер. 1. Математика. Механика. — 1998. № 2. — С. 68–70.

5. Зыков К. А. О сложности реализации систем функций k-значной логики соответствующих некоторым циклическим матрицам // *Сборник трудов девятого международного семинара "Дискретная математика и ее приложения"* — 2007.

6. Нечипорук Э. И. Об одной булевской матрице // *Проблемы кибернетики.* — Вып. 21. — 1969. — С. 237–240.

7. Орлов В. А. Реализация "узких" матриц вентильными схемами // *Проблемы кибернетики.* — Вып. 22. — 1970. — С. 45–52.

8. Черухин Д. Ю. Нижняя оценка сложности в классе схем глубины 2 без ограничения на базис // *Вестник Московского университета.* Сер. 1. Математика. Механика. — 2005. № 4. — С. 54–56.

9. Черухин Д. Ю. О схемах из функциональных элементов с ограниченной глубиной ветвления // *Докл. РАН.* — 2005. — Т. 405, № 4. — С. 1–4.

10. Alon N., Karchmer M., Wigderson A. Linear circuits over GF(2) // *SIAM J. Comput.,* — 1990. — V. 19, № 6. — P. 1064–1067.

11. Alon N., Maass W. Meanders, Ramsey theory and lower bounds for branching programs // *Proc. 27t Annual IEEE Symposium on Foundations of Computer Science,* IEEE Computer Society, Washington, DC. — 1986. — P. 410–417.

12. Blum N., Seysen M. Characterization of all optimal networks for a simultaneous computation of AND and NOR // *Acta Informatica.* — 1984. — № 21. — P. 171–181. [Имеется перевод: Блюм Н., Сейсен М. Характеристика всех оптимальных схем из функциональных элементов для одновременного вычисления AND и NOR // *Кибернетический сборник.* — М.: Мир. — 1990. — Вып. 27. — С. 104–117.

13. Bublitz S. Decomposition of graphs and monotone formula size of homogeneous functions // *Acta Inform.,* — 1986. — V. 23. — P. 689–696.

14. Pippenger N. The minimum number of edges in graphs with prescribed paths // *Math. Systems Theory.* — 1979. — V. 12. — 333.

15. Valiant L. Graph-theoretic arguments in low-level complexity // *Lecture Notes in Computer Science.* — 1977. — V. 53. — Springer-Verlag, Berlin, New York. — P. 162@–176.

16. Zykov K. A., Kasim-Zadeh O. M., Tarannikov Yu. V. Complexity and combinatorial aspects of informatics systems // *Proceedings of the Conference on Applied Mathematics and Computer Science*, 28–29 October 1996, Moscow, Издательство Франко-русского центра им. А. М. Ляпунова. — 1997 — с. 82–90.

СОДЕРЖАНИЕ