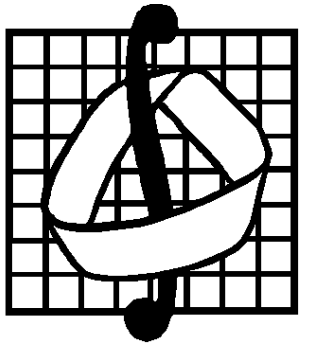
МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В. ЛОМОНОСОВА
МЕХАНИКО-МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ



А. В. Чашкин

ЛЕКЦИИ ПО ДИСКРЕТНОЙ МАТЕМАТИКЕ

Учебное пособие

Москва · 2007

А. В. Чашкин. Лекции по дискретной математике.

Учебное пособие содержит материалы лекций и семинарских занятий, составивших обязательный полугодовой курс дискретной математики, прочитанный автором студентам четвертого курса механико-математического факультета Московского государственного университета им. М. В. Ломоносова. Для студентов и аспирантов. **Предварительная версия.**

Оглавление

Лекция 1

Комбинаторные числа и тождества

1.1 Перестановки, размещения, сочетания

Пусть $A = \{a_1, \ldots, a_n\}$ — конечное множество. Совокупность из k элементов множества A (не обязательно различных) называется *k-выборкой* множества A. Выборка называется *упорядоченной*, если каждому ее элементу поставлен в соответствие номер — натуральное число, не превосходящее k так, что разным элементам соответствуют разные числа. Упорядоченные выборки будем называть также *наборами*. Элементы упорядоченных выборок будем заключать в круглые скобки, а элементы неупорядоченных выборок — в фигурные скобки. Например, (a_1, a_2, a_2) и (a_2, a_1, a_2) — две различных упорядоченных выборки, а $\{a_1, a_2, a_2\}$ и $\{a_2, a_1, a_2\}$ — одна и та же неупорядоченная выборка.

Перестановкой n-элементного множества $A = \{a_1, \ldots, a_n\}$ называется любой набор $(a_{i_1}, \ldots, a_{i_n})$, состоящий из элементов A, в котором каждый элемент из A встречается ровно один раз. Например, у трехэлементного множества $\{a_1, a_2, a_3\}$ существует ровно шесть различных перестановок:

$$(a_1, a_2, a_3), \quad (a_1, a_3, a_2), \quad (a_2, a_1, a_3),$$
$$(a_2, a_3, a_1), \quad (a_3, a_1, a_2), \quad (a_3, a_2, a_1).$$

Найдем число P_n различных перестановок n-элементного множества. Для этого из n-элементного множества будем последовательно выбирать элементы и формировать из них упорядоченную выборку: первый выбранный элемент станет первым элементом упорядоченной выборки, второй — вторым и т. д. Нетрудно видеть, что первый элемент можно выбрать n способами. Второй элемент будет выбираться из $(n-1)$ оставшихся элементов, поэтому его можно выбрать $(n-1)$ способом. Продолжая выбор, заметим, что после выбора первых k элементов останется $(n-k)$ невыбранных элементов. Следовательно, $(k+1)$-й элемент можно выбрать $(n-k)$ способами. Перемножив числа способов, которыми можно выбрать первый, второй, …, $(n-1)$-й и n-й элементы, получим величину, равную числу способов, кото-

рыми можно упорядочить n-элементное множество. Таким образом,

$$P_n = n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1 = n! \tag{1.1}$$

Размещением из n элементов по k называется произвольная перестановка k-элементного подмножества n-элементного множества. Для обозначения числа размещений из n элементов по k используется символ A_n^k. Рассуждениями аналогичными приведенным выше при определении величины P_n, нетрудно показать, что

$$A_n^k = n(n-1) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}. \tag{1.2}$$

Сочетанием из n элементов по k называется произвольное k-элементное подмножество n-элементного множества. Число сочетаний из n элементов по k обозначается через $\binom{n}{k}$ (иногда также используется символ C_n^k). Так как у одного k-элементного подмножества существует ровно $k!$ различных перестановок, то из (1.2) легко следует, что

$$\binom{n}{k} = \frac{n!}{(n-k)!k!} = \frac{n(n-1) \cdot \ldots \cdot (n-k+1)}{k(k-1) \cdot \ldots \cdot 2 \cdot 1}. \tag{1.3}$$

Из равенства (1.3) легко вытекают следующие часто используемые свойства сочетаний:

$$\binom{n}{k} = \binom{n}{n-k}, \qquad \binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}. \tag{1.4}$$

Второе из этих равенств докажем также при помощи комбинаторных рассуждений. Пусть A — множество всех k-элементных подмножеств множества $\{1, 2, \ldots, n\}$. Это множество разобъем на два класса A_1 и A_2 так, что в первый класс отнесем все подмножества, содержащие n, а во второй класс — подмножества без этого элемента. Нетрудно видеть, что A_1 состоит из $\binom{n-1}{k-1}$ подмножеств, а A_2 — из $\binom{n-1}{k}$. Так как каждое k-элементное подмножество попадает либо в класс A_1, либо в класс A_2, то $|A| = |A_1| + |A_2|$, и, следовательно, $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$.

Сочетанием с повторениями из n элементов по k называется неупорядоченная k-выборка n-элементного множества. Например, из трех элементов a_1, a_2 и a_3 можно составить шесть сочетаний с повторениями по два элемента:

$$a_1a_1, \quad a_1a_2, \quad a_1a_3, \quad a_2a_2, \quad a_2a_3, \quad a_3a_3.$$

Каждое сочетание с повторениями из n элементов по k однозначно определяется тем, сколько раз каждый элемент множества входит в рассматриваемое сочетание. Пусть в некоторое такое сочетание элемент a_i входит m_i раз, где $i = 1, 2, \ldots, n$. Этому сочетанию поставим в соответствие набор

$$\underbrace{1\ldots1}_{m_1}0\underbrace{1\ldots1}_{m_2}0\ldots\ldots0\underbrace{1\ldots1}_{m_n} \tag{1.5}$$

из k единиц, сгруппированных в n блоков, и $n-1$ нулей, разделяющих эти блоки. В этом наборе первый блок из m_1 единиц соответствует элементу $_1$, второй блок из m_2 единиц — элементу a_2, и т. д. Приведенным выше двухэлементным сочетаниям соответствуют следующие шесть наборов:

$$1100, \quad 1010, \quad 1001, \quad 0110, \quad 0101, \quad 0011.$$

Очевидно, что набор вида (1.5) однозначно определяет соответствующее ему сочетание с повторениями. Поэтому число H_n^k сочетаний с повторениями из n элементов по k равно числу наборов из k единиц и $n-1$ нулей. Каждый такой набор можно рассматривать как набор значений характеристической функции k-элементного подмножества $(n+k-1)$-элементного множества. Следовательно,

$$H_n^k = \binom{n+k-1}{k} = \frac{(n+k-1)!}{(n-1)!k!}.$$

1.2 Бином Ньютона

Числа сочетаний и сочетаний с повторениями появляются в известной формуле бинома Ньютона

$$(1+x)^y = 1 + \sum_{k=1}^{\infty} \frac{y(y-1)\ldots(y-k+1)}{k!}\, x^k \tag{1.6}$$

в случаях, когда y принимает целые положительные и целые отрицательные значения, соответственно. Поэтому эти числа называются также биномиальными коэффициентами. Покажем, что их появление в (1.6) не случайно. Для этого установим справедливость формулы бинома Ньютона при целых значениях y при помощи комбинаторных рассуждений.

Если показатель степени бинома — целое неотрицательное число, то равенство (1.6) записывается в виде

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k. \tag{1.7}$$

Справедливость равенства (1.7) следует из того, что после раскрытия скобок в выражении $(1+x)^n$ коэффициент при k-й степени переменной x будет равен числу способов, которыми можно выбрать k раз переменную x из n двучленов $(1+x)$.

Если показатель степени бинома — целое отрицательное число, то после изменения знака перед переменной x равенство (1.6) превращается в равенство

$$(1-x)^{-n} = \sum_{k=0}^{\infty} \binom{n+k-1}{k} x^k, \tag{1.8}$$

которое справедливо при $|x|<1$. Для доказательства (1.8) воспользуемся хорошо известным частным случаем формулы (1.8) — формулой суммы убывающей геометрической прогрессии $(1-x)^{-1}=\sum_{k=0}^{\infty}x^k$. Тогда

$$(1-x)^{-n}=\underbrace{(1+x+x^2+\dots)(1+x+x^2+\dots)\cdots(1+x+x^2+\dots)}_{n\ \text{раз}}. \quad (1.9)$$

Нетрудно видеть, что после раскрытия скобок в правой части (1.9) коэффициент при k-й степени переменной x будет равен числу способов, которыми можно выбрать n степеней переменной x из n рядов $1+x+x^2+\dots$ так, чтобы сумма этих степеней была равна k. Рассматривая выбор x^p из q-го ряда $1+x+x^2+\dots$ в (1.9) как выбор p элементов q-го вида из n возможных видов, заключаем, что коэффициент при x^k равен числу сочетаний с повторениями их n элементов по k, т. е. $\binom{n+k-1}{k}$.

Рассмотрим несколько примеров использования формулы бинома Ньютона. Подставляя в (1.7) вместо x единицу и минус единицу получим тождества

$$\sum_{k=0}^{n}\binom{n}{k}=2^n,\qquad \sum_{k=0}^{n}(-1)^k\binom{n}{k}=0.$$

Вычисляя сумму и разность этих тождеств и деля результаты пополам, получаем, что

$$\sum_{k=0}^{\lfloor n/2\rfloor}\binom{n}{2k}=2^{n-1},\qquad \sum_{k=0}^{\lfloor n/2\rfloor}\binom{n}{2k+1}=2^{n-1}.$$

Дифференцирование (1.7) с подстановкой единицы вместо x и интегрирование (1.7) по x от нуля до единицы дают следующие тождества

$$\sum_{k=0}^{n}k\binom{n}{k}=n2^{n-1},\qquad \sum_{k=0}^{n}\frac{1}{k+1}\binom{n}{k}=\frac{2^n-1}{n+1}.$$

Раскроем скобки в равенстве $(1+x)^n(1+x)^m=(1+x)^{n+m}$. Так как коэффициенты при x^p в правой и левой частях равны, то из (1.7) и правила умножения многочленов следует равенство

$$\sum_{k=0}^{p}\binom{n}{k}\binom{m}{p-k}=\binom{n+m}{p}, \quad (1.10)$$

называемое сверткой Вандермонда. Следующее тождество

$$\sum_{k=0}^{m-1}\binom{n+k}{n}=\binom{n+m}{n+1} \quad (1.11)$$

нетрудно доказать индукцией по m. Действительно, при $m=1$ рассматриваемое тождество тривиально: $\binom{n}{n}=\binom{n+1}{n+1}$. Предположим, что оно верно

при $m-1$. Тогда

$$\sum_{k=0}^{m-1}\binom{n+k}{n}=\sum_{k=0}^{m-2}\binom{n+k}{n}+\binom{n+m-1}{n}=$$
$$=\binom{n+m-1}{n+1}+\binom{n+m-1}{n}=\binom{n+m}{n+1}.$$

Тождество (1.11) доказано.

Далее будем использовать обозначения $n!!$ для произведения всех тех натуральных чисел, которые не превосходят n и имеют такую же четность как и n, т. е.

$$2n!!=2n(2n-2)\cdot\ldots\cdot 2,\quad (2n+1)!!=(2n+1)(2n-1)\cdot\ldots\cdot 1.$$

Рассмотрим пример использования формулы (1.6) с нецелым показателем степени. Разложим в ряд функцию $(\sqrt{1-4x})^{-1}$. Нетрудно видеть, что

$$\frac{1}{\sqrt{1-4x}}=1+\sum_{k=1}^{\infty}\frac{-\frac12(-\frac12-1)\cdots(-\frac12-k+1)}{k!}(-4)^kx^k=$$
$$=1+\sum_{k=1}^{\infty}\frac{(-\frac12)(-\frac32)\cdots(-\frac{2k-1}{2})}{k!}(-4)^kx^k=1+\sum_{k=1}^{\infty}\frac{2^k(2k-1)!!}{k!}x^k=$$
$$=1+\sum_{k=1}^{\infty}\frac{2^kk!(2k-1)!!}{k!k!}x^k=1+\sum_{k=1}^{\infty}\frac{(2k)!!(2k-1)!!}{k!k!}x^k=$$
$$=1+\sum_{k=1}^{\infty}\frac{(2k)!}{k!k!}x^k=\sum_{k=0}^{\infty}\binom{2k}{k}x^k.$$

Поэтому

$$\frac{1}{1-4x}=\left(\sum_{k=0}^{\infty}\binom{2k}{k}x^k\right)^2=\sum_{k=0}^{\infty}\left(\sum_{m=0}^{k}\binom{2k-2m}{k-m}\binom{2m}{m}\right)x^k.$$

Таким образом, учитывая разложение $(1-4x)^{-1}=\sum_{m=0}^{\infty}4^kx^k$, приходим к тождеству

$$\sum_{m=0}^{k}\binom{2k-2m}{k-m}\binom{2m}{m}=4^k.$$

Использование равенства (1.6) является сильным средством получения различных соотношений с биномиальными коэффициентами. Однако в ряде случаев более действенными оказываются методы, использующие комбинаторную природу биномиальных коэффициентов. В качестве примера таких методов рассмотрим комбинаторные доказательства тождеств (1.10) и (1.11). Сначала докажем равенство (1.10):

$$\sum_{k=0}^{p}\binom{n}{k}\binom{m}{p-k}=\binom{n+m}{p}.$$

Для этого множество всех p-элементных подмножеств $(n+m)$-элементного множества $A=\{1,2,\ldots,n+m\}$ разобъем на p классов $A_1,\ldots,A_p$ так, что класс A_k будет состоять из всех тех подмножеств множества A, в которые входит ровно k чисел, каждое из которых не превосходит n, и $p-k$ чисел, каждое из которых больше n. Первые k чисел можно выбрать $\binom{n}{k}$ способами, а оставшиеся $p-k$ чисел — $\binom{m}{p-k}$ способами (если $k>n$ или $p-k>m$, то такого подмножества не существует и, соответственно, $\binom{n}{k}=0$ или $\binom{m}{p-k}=0$). Следовательно, каждый класс A_k состоит из $\binom{n}{k}\binom{m}{p-k}$ подмножеств. Наконец заметим, что каждое p-элементное подмножество множества A принадлежит одному из классов A_k, т. е. $|A|=\sum_{k=0}^{p}|A_k|$. Тождество (1.10) доказано.

Теперь приведем комбинаторное доказательство тождества (1.11):

$$\sum_{k=0}^{m-1}\binom{n+k}{n}=\binom{n+m}{n+1}.$$

Множество всех $(n+1)$-элементных подмножеств $(n+m)$-элементного множества $\{1,2,\ldots,n+m\}$ разобъем на m классов $A_1,\ldots,A_m$ так, что в j-й класс A_j попадут все те подмножества, у которых минимальный элемент равен j. Нетрудно видеть, что A_j состоит из $\binom{n+m-j}{n}$ подмножеств. Поэтому, полагая $k=m-j$, имеем

$$\binom{n+m}{n+1}=\sum_{j=1}^{m}|A_j|=\sum_{j=1}^{m}\binom{n+m-j}{n}=\sum_{j=m}^{1}\binom{n+m-j}{n}=\sum_{k=0}^{m-1}\binom{n+k}{n}.$$

1.3 Формула включений и исключений

Полезным средством при решении комбинаторных задач является формула включений и исключений, позволяющая находить мощность объединения различных множеств, если известны мощности их пересечений.

Теорема 1.1. *Для любых конечных множеств $A_1,\ldots,A_n$ справедливо равенство*

$$\begin{aligned}|A_1\cup\cdots\cup A_n|=&\sum_{1\le i\le n}|A_i|-\sum_{1\le i_1<i_2\le n}|A_{i_1}\cap A_{i_2}|+\ldots\\&\cdots+(-1)^{k+1}\sum_{1\le i_1<\cdots<i_k\le n}|A_{i_1}\cap\cdots\cap A_{i_k}|+\ldots\\&\cdots+(-1)^{n+1}|A_1\cap\ldots\cap A_n|.\end{aligned}\tag{1.12}$$

Доказательство. Пусть целое m не меньше нуля и не больше n. Допустим, что некоторый элемент a принадлежит ровно m множествам. Тогда a принадлежит $\binom{m}{2}$ попарным пересечениям множеств $A_1,\ldots,A_n$, $\binom{m}{3}$ тройным пересечениям этих множеств, и, в общем случае, $\binom{m}{k}$ пересечениям по

k множеств. Следовательно, в сумме, стоящей в правой части (1.12), этот элемент будет учтен ровно

$$\binom{m}{1}-\binom{m}{2}+\cdots+(-1)^{k+1}\binom{m}{k}+\cdots+(-1)^{m+1}\binom{m}{m}.\tag{1.13}$$

раз. Из (1.7) следует, что

$$(1-1)^m=1-\binom{m}{1}+\binom{m}{2}-\cdots+(-1)^k\binom{m}{k}+\cdots+(-1)^m\binom{m}{m}.$$

Поэтому сумма (1.13) равна единице. Следовательно, в правой части (1.12) каждый элемент, принадлежащий объединению множеств A_i, учитывается ровно один раз и, поэтому, вся сумма равна мощности объединения этих множеств. Теорема доказана.

Функция $\varphi(m)$, равная количеству натуральных чисел, не превосходящих и взаимнопростых с m, называется *функцией Эйлера*. В следующей теореме при помощи формулы включений и исключений для функции Эйлера устанавливается явная формула.

Теорема 1.2. *Если $m=p_1^{k_1}p_2^{k_2}\ldots p_n^{k_n}$, то*

$$\varphi(m)=m\Big(1-\frac{1}{p_1}\Big)\Big(1-\frac{1}{p_2}\Big)\ldots\Big(1-\frac{1}{p_n}\Big).$$

Доказательство. Пусть множество A_i состоит из всех натуральных чисел, каждое из которых не превосходит m и делится на p_i. Тогда множество $A_1\cup\cdots\cup A_n$ состоит из всех натуральных чисел, каждое из которых не превосходит m и имеет с m хотя бы один общий делитель больший единицы. Следовательно,

$$\varphi(m)=m-|A_1\cup\cdots\cup A_n|.$$

Нетрудно видеть, что $|A_{i_1}\cap\cdots\cap A_{i_s}|=\frac{m}{p_{i_1}\cdots p_{i_s}}$ для любых $i_1,\ldots,i_s$. Поэтому для вычисления мощности объединения множеств A_i можно воспользоваться формулой включений–исключений:

$$\begin{aligned}|A_1\cup\cdots\cup A_n|&=\sum_{1\le i\le n}|A_i|+\cdots+(-1)^{s+1}\sum_{1\le i_1<\cdots<i_s\le n}|A_{i_1}\cap\cdots\cap A_{i_s}|+\ldots\\&=\sum_{1\le i\le n}\frac{m}{p_i}+\cdots+(-1)^{s+1}\sum_{1\le i_1<\cdots<i_s\le n}\frac{m}{p_{i_1}\cdots p_{i_s}}+\ldots\end{aligned}$$

Теперь осталось заметить, что после раскрытия скобок в формуле

$$m-m\Big(1-\frac{1}{p_1}\Big)\Big(1-\frac{1}{p_2}\Big)\ldots\Big(1-\frac{1}{p_n}\Big)$$

получится такое же выражение, как и в правой части последнего равенства. Теорема доказана.

1.4 Факториальные степени

Произведение $x^{\underline{k}} = x(x-1)\cdot\ldots\cdot(x-k+1)$ называется k-й *нижней факториальной степенью* числа x, а произведение $x^{\overline{k}} = x(x+1)\cdot\ldots\cdot(x+k-1)$ — k-й *верхней факториальной степенью* этого числа. Факториальные степени естественным образом возникают при решении комбинаторных задач — например $n^{\underline{k}}$ равно числу упорядоченных k-элементных подмножеств n-элементного множества.

Коэффициенты $\begin{bmatrix} n \\ k \end{bmatrix}$ и $\begin{Bmatrix} n \\ k \end{Bmatrix}$ в равенствах

$$x^{\overline{n}} = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} x^k, \qquad x^n = \sum_{k=0}^{n} \begin{Bmatrix} n \\ k \end{Bmatrix} x^{\underline{k}}$$

называются *числами Стирлинга первого* и *второго рода*, соответственно.

Так как

$$(-x)^{\overline{n}} = (-x)(-x+1)(-x+2)\ldots(-x+(n-1)) = (-1)^n x^{\underline{n}},$$

то, учитывая равенство $(-1)^m = (-1)^{-m}$, получаем формулу

$$x^{\underline{n}} = (-1)^n (-x)^{\overline{n}} = (-1)^n \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} (-1)^k x^k = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} (-1)^{n-k} x^k$$

для выражения нижней факториальной степени через обычные степени. Аналогичным образом получается равенство

$$x^n = \sum_{k=0}^{n} \begin{Bmatrix} n \\ k \end{Bmatrix} (-1)^{n-k} x^{\overline{k}}$$

для выражения обычной степени числа x через его верхние факториальные степени.

Из определения чисел Стирлинга легко следует, что равенства

$$\begin{bmatrix} n \\ n \end{bmatrix} = 1, \quad \begin{bmatrix} n \\ 0 \end{bmatrix} = 0, \quad \begin{Bmatrix} n \\ n \end{Bmatrix} = 1, \quad \begin{Bmatrix} n \\ 0 \end{Bmatrix} = 0 \tag{1.14}$$

справедливы при всех $n \geq 1$. Для вычисления чисел Стирлинга с нижним индексом, не равным верхнему индексу или нулю, можно воспользоваться простыми рекуррентными равенствами

$$\begin{bmatrix} n \\ k \end{bmatrix} = \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} + (n-1) \begin{bmatrix} n-1 \\ k \end{bmatrix}, \quad \begin{Bmatrix} n \\ k \end{Bmatrix} = \begin{Bmatrix} n-1 \\ k-1 \end{Bmatrix} + k \begin{Bmatrix} n-1 \\ k \end{Bmatrix}, \tag{1.15}$$

которые, как нетрудно заметить, похожи на соответствующее равенство для биномиальных коэффициентов из (1.4). Покажем справедливость первого из равенств (1.15). Сделаем это индукцией по n. В основание индукции положим равенства для чисел Стирлинга первого рода из (1.14) при $n = 1$. Тогда

$$\begin{aligned} \sum_{k=1}^{n} \begin{bmatrix} n \\ k \end{bmatrix} x^k = x^{\overline{n}} = x^{\overline{n-1}}(x-n+1) = x^{\overline{n-1}} \cdot x + x^{\overline{n-1}} \cdot (n-1) = \\ = x \sum_{k=0}^{n-1} \begin{bmatrix} n-1 \\ k \end{bmatrix} x^k + (n-1) \sum_{k=0}^{n-1} \begin{bmatrix} n-1 \\ k \end{bmatrix} x^k = \\ = \sum_{k=1}^{n-1} \left(\begin{bmatrix} n-1 \\ k-1 \end{bmatrix} + (n-1) \begin{bmatrix} n-1 \\ k \end{bmatrix} \right) x^k + \begin{bmatrix} n-1 \\ n-1 \end{bmatrix} x^n. \end{aligned}$$

Первое равенство в (1.15) доказано. Второе равенство доказывается аналогично.

Равенства (1.15) вместе с равенствами (1.14) позволяют придать числам Стирлинга простой комбинаторный смысл. Покажем, что число $\begin{Bmatrix} n \\ k \end{Bmatrix}$ равно количеству способов, которыми n-элементное множество можно разбить ровно на k непустых подмножеств. Прежде всего отметим, что n-элементное множество можно единственным образом разбить на n непустых подмножеств и не существует ни одного разбиения такого множества на нулевое число подмножеств. Следовательно, число разбиений одноэлементного множества на k непустых подмножеств при $k = 0$ и 1 равно $\begin{Bmatrix} 1 \\ 0 \end{Bmatrix}$ и $\begin{Bmatrix} 1 \\ 1 \end{Bmatrix}$, соответственно. Далее предположим, что число разбиений m-элементного множества ровно на k непустых подмножеств равно $\begin{Bmatrix} m \\ k \end{Bmatrix}$ при всех $m \leq n-1$. Затем рассмотрим разбиения множества $\{1, 2, \ldots, n\}$ на k непустых подмножеств. Все разбиения разобьем на два класса. К первому отнесем те разбиения, в которых элемент n образует самостоятельное подмножество. Очевидно, что в силу сделанного предположения первый класс состоит из $\begin{Bmatrix} n-1 \\ k-1 \end{Bmatrix}$ подмножеств. Ко второму классу отнесем все остальные разбиения. Каждое разбиение из второго класса можно получить из некоторого разбиения множества $\{1, 2, \ldots, n-1\}$ на k непустых подмножеств, если к одному из этих k подмножеств добавить элемент n. Следовательно, опять в силу сделанного предположения второй класс состоит из $k\begin{Bmatrix} n-1 \\ k \end{Bmatrix}$ подмножеств. Теперь доказываемое утверждение следует из второго равенства в (1.15). Похожим образом можно показать, что число Стирлинга первого рода $\begin{bmatrix} n \\ k \end{bmatrix}$ равно числу элементов симметрической группы S_n, каждый из которых представляется в виде произведения ровно k непересекающихся циклов. Отсюда немедленно следует, что

$$\sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} = n!$$

Аналогичная сумма чисел Стирлинга второго рода

$$B_n = \sum_{k=0}^{n} \left\{ {n \atop k} \right\}$$

называется n-м *числом Белла* и очевидно равна числу всех разбиений n-элементного множества. Для чисел Белла справедлива следующая рекуррентная формула:

$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k, \tag{1.16}$$

где $B_0 = 1$. Для доказательства равенства (1.16) достаточно заметить, что существует ровно $\binom{n}{k} B_k$ разбиений множества из $n+1$ элемента, в которых $(n+1)$-й элемент попадает в $(n-k+1)$-элементное подмножество.

Далее, используя комбинаторный смысл чисел Стирлинга второго рода, покажем, что

$$\left\{ {n \atop k} \right\} = \sum_{m=0}^{k} (-1)^m \frac{(k-m)^n}{m!(k-m)!}. \tag{1.17}$$

Сначала решим вспомогательную задачу — найдем число способов $R(n,k)$, которыми можно разложить n разных предметов по k нумерованным ящикам так, чтобы не было пустых ящиков. В общем случае, когда допускаются размещения с пустыми ящиками, разложить n предметов по k ящикам можно k^n способами. Из всех таких размещений составим k подмножеств $A_1, \ldots, A_k$, так, что A_i будет состоять из всех тех размещений, при которых i-й ящик остается пустым. Тогда $R(n,k) = k^n - |A_1 \cup \cdots \cup A_k|$, где мощность объединения находится при помощи формулы включений–исключений. Следовательно,

$$R(n,k) = k^n - |A_1 \cup \cdots \cup A_k| = k^n - \sum_{1 \le i \le k} |A_i| + \sum_{1 \le i_1 < i_2 \le k} |A_{i_1} \cap A_{i_2}| - \ldots$$
$$\cdots + (-1)^m \sum_{1 \le i_1 < \cdots < i_m \le k} |A_{i_1} \cap \cdots \cap A_{i_m}| + \cdots + (-1)^k |A_1 \cap \cdots \cap A_k|.$$

Размещения из множества $A_{i_1} \cap \cdots \cap A_{i_m}$ можно получить размещая предметы по ящикам, номера которых не принадлежат множеству $\{i_1, \ldots, i_m\}$. Поэтому

$$|A_{i_1} \cap \cdots \cap A_{i_m}| = (k-m)^n.$$

Подставляя это равенство в предыдущую формулу и учитывая, что число различных множеств $A_{i_1} \cap \cdots \cap A_{i_m}$ равно $\binom{k}{m}$, получаем

$$R(n,k) = \sum_{m=0}^{k} (-1)^m (k-m)^n \binom{k}{m} = \sum_{m=0}^{k} (-1)^m \frac{(k-m)^n k!}{m!(k-m)!}. \tag{1.18}$$

Нетрудно видеть, что число $R(n,k)$ ровно в $k!$ раз больше числа таких размещений n различных предметов по k ненумерованным ящикам, при которых нет ни одного пустого ящика. Рассматривая содержимое каждого ненумерованного ящика как подмножество множества размещаемых предметов, получаем, что $\left\{ {n \atop k} \right\} \cdot k! = R(n,k)$. Теперь равенство (1.17) легко следует из (1.18).

1.5 Метод траекторий

В задачах, связанных с перечислением элементов различных множеств, часто можно существенно упростить исходную задачу, если отобразить рассматриваемое множество в новое множество, элементы которого обладают какими-либо свойствами, облегчающими их подсчет. Именно таким образом выше была найдена формула для числа сочетаний с повторениями. Теперь используем этот прием в более сложной задаче — задаче о числе последовательностей длины $2n$ из n нулей и n единиц, в которых в любом начальном отрезке единиц не меньше чем нулей. Множество таких последовательностей обозначим через N_{2n}. Описываемый ниже метод называется методом траекторий и представляет независимый от рассматриваемой задачи интерес.

Каждой последовательности из нулей и единиц поставим в соответствие выходящую из начала координат плоскости (x,y) ломаную, составленную из звеньев $\nearrow$ и $\searrow$ — векторов с координатами $(1,1)$ и $(1,-1)$, в которой звено $\nearrow$ соответствует единице, а звено $\searrow$ — нулю. Такие ломаные будем называть траекториями. На рис. 1.1 изображены траектории, соответствующие последовательностям 100110 и 101100. Так как в каждой последовательно-

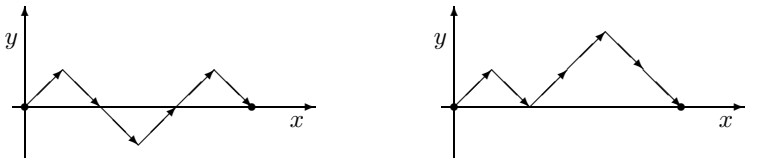


Рис. 1.1

сти из N_{2n} число нулей равно числу единиц, то все траектории, соответствующие последовательностям из N_{2n}, будут соединять начало координат с точкой $(2n,0)$, а так как в любом начальном отрезке каждой такой последовательности единиц не меньше чем нулей, то все траектории будут проходить выше оси x так, как правая траектория на рис. 1.1. Посчитать такие траектории можно при помощи принципа отражения, суть которого заключена в следующей лемме.

Лемма 1.1. *Число траекторий, начинающихся в начале координат, заканчивающихся в точке* $(2n,0)$ *и пересекающих ось* x*, равно числу траекторий, начинающихся в точке* $(0,-2)$ *и заканчивающихся в точке* $(2n,0)$.

Доказательство. Каждая траектория α, которая начинается в начале координат $A=(0,0)$, заканчивается в точке $B=(2n,0)$ и опускается ниже оси x, обязательно имеет хотя бы одну общую точку с прямой $y=-1$. Пусть C — первая точка на прямой $y=-1$, в которую приходит траектория α. Траектории α поставим в соответствие траекторию α', которая начинается в точке $D=(0,-2)$, заканчивается в точке $B=(2n,0)$, совпадает с α между точками C и B, а между точками A и C является отражением α относительно прямой $y=-1$ (см. рис. 1.2). Очевидно, что такое соответствие является взаимнооднозначным. Лемма доказана.

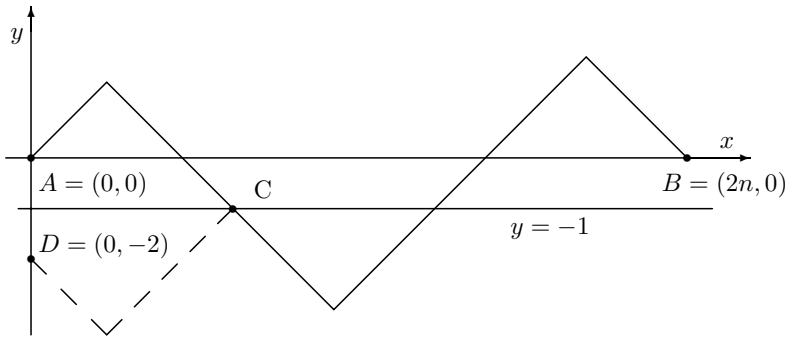


Рис. 1.2

Каждая траектория, начинающаяся в точке D и заканчивающаяся в точке B, состоит из $n+1$ звеньев вида $\nearrow$ и $n-1$ звеньев вида $\searrow$. Поэтому общее число таких траекторий равно $\binom{2n}{n+1}$. Следовательно, число искомых траекторий равно разности числа всех траекторий из начала координат в точку $(2n,0)$ и $\binom{2n}{n+1}$, т. е.

$$|N_{2n}| = \binom{2n}{n} - \binom{2n}{n+1} = \binom{2n}{n}\Big(1-\frac{n}{n+1}\Big) = \frac{1}{n+1}\binom{2n}{n}.$$

Число $\frac{1}{n+1}\binom{2n}{n}$ называется n-м числом Каталана.

1.6 Задачи

1.1. На плоскости проведено n прямых так, что никакие две не параллельны и никакие три не пересекаются в одной точке. На сколько областей делят плоскость эти прямые?

1.2. Найти число (m,n)-матриц из ±1, у которых произведения всех элементов каждой строки и каждого столбца положительны.

1.3. За круглым столом сидят n человек. Сколькими способами из них можно выбрать k человек так, чтобы среди выбранных не было соседей.

1.4. Найти число натуральных решений уравнения $x_1+\dots+x_k=n$.

1.5. Сколько существует n-разрядных натуральных чисел, у которых цифры находятся в невозрастающем порядке?

1.6. Показать, что

$$(x_1+\dots+x_k)^n = \sum_{m_1+\dots m_k=n} \frac{n!}{m_1!\cdots m_k!}\, x_1^{m_1}\cdots x_k^{m_k}.$$

1.7. Найти

$$1-3\binom{n}{2}+9\binom{n}{4}-27\binom{n}{6}+\dots$$

1.8. Найти

$$\binom{4n}{0}-\binom{4n}{2}+\binom{4n}{4}-\dots$$

1.9. Найти

$$\binom{n}{1}+\binom{n}{2}+\binom{n}{5}+\binom{n}{6}+\binom{n}{9}+\binom{n}{10}+\dots.$$

1.10. Найти

$$\binom{2n}{0}-\binom{2n-1}{1}+\binom{2n-2}{2}-\dots+(-1)^n\binom{n}{n}.$$

1.11. Найти

$$\binom{n}{0}\binom{n-1}{0}-\binom{n}{1}\binom{n}{1}+\binom{n}{2}\binom{n+1}{2}-\dots+(-1)^n\binom{n}{n}\binom{2n-1}{n}.$$

1.12. Доказать равенство

$$\binom{n-2}{k-2}+2\binom{n-3}{k-2}+\dots+(n-k+1)\binom{k-2}{k-2}=\binom{n}{k}.$$

1.13. Доказать равенство

$$\binom{n-3}{k-3}\binom{2}{2}+\binom{n-4}{k-3}\binom{3}{2}+\dots+\binom{k-3}{k-3}\binom{n-k+2}{2}=\binom{n}{k}.$$

1.14. Доказать равенство

$$\sum_k \binom{n}{r+km} = \frac{2^n}{m}\sum_{k=1}^{m}\cos^n\frac{\pi k}{m}\cos\frac{(n-2r)\pi k}{m}.$$

1.15. Доказать равенство

$$\sum_{k=0}^{m}(n+k)^{\underline{n}} = \frac{(n+m+1)^{\underline{n+1}}}{n+1}.$$

1.16. Показать, что

$$(a+b)^{\underline{n}} = \sum_{k=0}^{n}\binom{n}{k}a^{\underline{k}}b^{\underline{n-k}}, \qquad (a+b)^{\overline{n}} = \sum_{k=0}^{n}\binom{n}{k}a^{\overline{k}}b^{\overline{n-k}}.$$

1.17. Доказать равенства

$$\left[{n+1 \atop k+1}\right] = \sum_{i=k}^{n} n^{\underline{n-i}}\left[{i \atop k}\right], \qquad \left\{{n+1 \atop k+1}\right\} = \sum_{i=0}^{n-k}\binom{n}{i}\left\{{n-i \atop k}\right\}.$$

1.18. Доказать равенство

$$\sum_{i=k}^{n}(-1)^{i+k}\left\{{n \atop i}\right\}\left[{i \atop k}\right] = \delta_{k,n}.$$

1.19. Сколькими различными способами можно разместить k предметов по n ящикам если: a) предметы и ящики различимы; b) предметы и ящики неразличимы; c) предметы различимы, ящики неразличимы; d) предметы неразличимы, ящики различимы.

1.20. В квадрате с длиной стороны 6 находятся три фигуры, площади которых равны 17, 18 и 19. Доказать, что среди этих фигур найдутся две, площадь пересечения которых не меньше 6.

1.21. Пусть A — множество из 60-ти элементов, A_1, A_2, A_3, A_4 — подмножества множества A, каждое из которых состоит из 30 элементов. Показать, что найдутся два подмножества, пересечение которых состоит не менее чем из 10 элементов.

1.22. Найти число перестановок π из S_n таких, что $\pi(k) \neq k$ для $k = 1, \ldots, n$.

1.23. Найти число последовательностей из n единиц и n нулей, в каждой из которых среди первых i знаков, $i = 1, 2, \ldots, 2n$, число единиц может превышать число нулей не более чем на 5.

1.24. Найти число траекторий, начинающихся в начале координат, заканчивающихся в точке (n, m) и лежащих выше оси x.

1.25. Найти число траекторий, начинающихся в начале координат, заканчивающихся в точке (n, m), лежащих выше оси x и ниже прямой $y = m$.

Лекция 2

Оценки комбинаторных функций

В различных задачах дискретной математики часто приходится находить мощности тех или иных множеств. При этом возможны ситуации, когда точные формулы, описывающие мощности множеств, существенным образом затрудняют решение задач из-за своей чрезмерной сложности и громоздкости, в то время как менее точные, но более просто устроенные формулы таких трудностей не создают. Поэтому методы получения простых оценок значений различных комбинаторных функций составляют важную часть методов дискретной математики.

2.1 Оценки $n!$

Функция $n!$ является составной частью многих комбинаторных формул. Ниже для этой функции доказываются три пары двухсторонних неравенств, причем точность неравенств каждой следующей пары выше точности неравенств предыдущей.

Теорема 2.1. *При* $n \geq 6$

$$\left(\frac{n}{3}\right)^n \leq n! \leq \left(\frac{n}{2}\right)^n.$$

Доказательство. Теорему докажем индукцией по n. При $n = 6$ справедливость доказываемых неравенств проверяется непосредственно. Докажем верхнюю оценку. Допустим, что $n! \leq \left(\frac{n}{2}\right)^n$. Тогда

$$\begin{aligned}\left(\frac{n+1}{2}\right)^{n+1} &= \left(\frac{n+1}{2}\right)^n \frac{n+1}{2} = \left(\frac{n}{2}\right)^n \left(1+\frac{1}{n}\right)^n \frac{n+1}{2} \geq \\ &\geq n!\left(1+1+\binom{n}{2}\frac{1}{n^2}+\cdots+\binom{n}{n}\frac{1}{n^n}\right)\frac{n+1}{2} > \\ &> n!\cdot 2\cdot\frac{n+1}{2} = (n+1)!.\end{aligned}$$

Аналогичным образом докажем нижнюю оценку. Допустим, что $n! \geq \left(\frac{n}{3}\right)^n$.

Тогда, учитывая, что $n! \geq 2^{n-1}$ при $n \geq 2$, имеем

$$\left(\frac{n+1}{3}\right)^{n+1} = \left(\frac{n+1}{3}\right)^n \frac{n+1}{3} = \left(\frac{n}{3}\right)^n \left(1+\frac{1}{n}\right)^n \frac{n+1}{3} \leq$$
$$\leq n! \left(1+1+\binom{n}{2}\frac{1}{n^2}+\binom{n}{3}\frac{1}{n^3}+\cdots+\frac{1}{n^n}\right)\frac{n+1}{3} <$$
$$< n! \left(1+1+\frac{n^2}{2!}\frac{1}{n^2}+\frac{n^3}{3!}\frac{1}{n^3}+\cdots+\frac{n^n}{n!}\frac{1}{n^n}\right)\frac{n+1}{3} <$$
$$< n! \left(1+1+\frac{1}{2^1}+\frac{1}{2^2}+\cdots+\frac{1}{2^{n-1}}\right)\frac{n+1}{3} < (n+1)!$$

Теорема доказана.

Теорема 2.2. *При $n \geq 1$*

$$\mathbf{e}\cdot\left(\frac{n}{\mathbf{e}}\right)^n \leq n! \leq n\mathbf{e}\cdot\left(\frac{n}{\mathbf{e}}\right)^n.$$

Доказательство. При $n = 1, 2$ справедливость неравенств проверяется подстановкой этих значений n. Далее легко видеть, что при $k \geq 2$ для $\ln k$ справедливы неравенства

$$\int_{k-1}^{k} \ln x dx < \ln k < \int_{k}^{k+1} \ln x dx. \tag{2.1}$$

Поэтому

$$\int_{1}^{n} \ln x dx < \ln n! < \int_{2}^{n+1} \ln x dx. \tag{2.2}$$

Преобразуя правое неравенство (2.2) при условии $n \geq 3$, получим, что

$$\ln n! < \int_2^{n+1} \ln x dx = (x\ln x - x)\Big|_2^{n+1} =$$
$$= (n+1)\ln(n+1) - (n+1) - 2\ln 2 + 2 =$$
$$= (n+1)\ln\frac{n+1}{\mathbf{e}} - 2\ln 2 + 2 =$$
$$= (n+1)\ln\frac{n}{\mathbf{e}} + (n+1)\ln(1+\frac{1}{n}) - 2\ln 2 + 2 < (n+1)\ln\frac{n}{\mathbf{e}} + 2.$$

Следовательно,

$$n! < n\mathbf{e}\cdot\left(\frac{n}{\mathbf{e}}\right)^n.$$

Аналогичным образом преобразуем левое неравенство (2.2):

$$\ln n! > \int_1^n \ln x dx = (x\ln x - x)\Big|_1^n = n\ln n - n + 1 = n\ln\frac{n}{\mathbf{e}} + 1.$$

Следовательно,

$$n! > \mathbf{e}\cdot\left(\frac{n}{\mathbf{e}}\right)^n.$$

Теорема доказана.

Далее установим порядок функции $n!$. Сделаем это при помощи новых неравенств для $\ln k$, которые значительно точнее чем, использованные в доказательстве теоремы 2.2 неравенства (2.1).

Лемма 2.1. *При $k \geq 2$*

$$\int_{k-1}^{k} \ln x dx + \Big(\ln 2k - \ln(2k-1)\Big) \leq \ln k \leq \int_{k-1}^{k} \ln x dx + \frac{1}{2}\Big(\ln k - \ln(k-1)\Big).$$

Доказательство. На рис 2.1 изображена часть графика функции $\ln x$ находящаяся между точками $x = k-1$ и $x = k$. На этом рисунке в точке

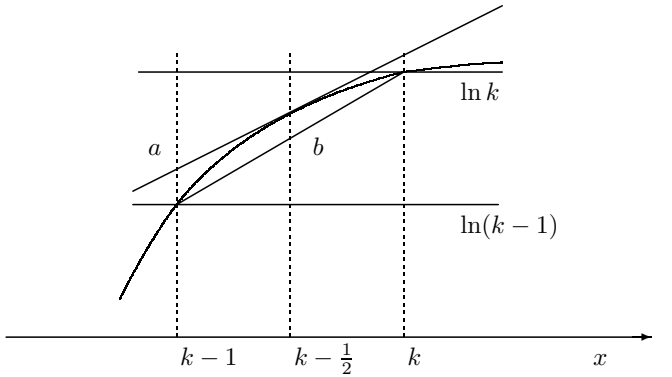


Рис. 2.1

$x = k - \frac{1}{2}$ к кривой $\ln x$ проведена касательная a, отрезок b соединяет точки с координатами $(k-1, \ln(k-1))$ и $(k, \ln k)$. Нетрудно видеть, что $\int_{k-1}^{k} \ln x dx$ превосходит разность величин $\ln k$ и площади треугольника, ограниченного отрезком b и прямыми $x = k-1$ и $y = \ln k$. Поэтому

$$\int_{k-1}^{k} \ln x dx \geq \ln k - \frac{1}{2}\Big(\ln k - \ln(k-1)\Big).$$

Также легко видеть, что $\int_{k-1}^{k} \ln x dx$ не превосходит площади трапеции, ограниченной прямой a и прямыми $x = k-1$, $x = k$ и $y = 0$. Так как площадь трапеции равна произведению длины средней линии, равной в данном случае $\ln(k - \frac{1}{2})$, и высоты, равной единице, то

$$\int_{k-1}^{k} \ln x dx \leq \ln\Big(k - \frac{1}{2}\Big) = \ln k + \ln\Big(1 - \frac{1}{2k}\Big) = \ln k - \ln(2k) + \ln(2k-1).$$

Оценивая при помощи полученных неравенств величину $\ln k$, получим для нее следующие оценки:

$$\int_{k-1}^{k} \ln x dx + \ln(2k) - \ln(2k-1) \leq \ln k \leq \int_{k-1}^{k} \ln x dx + \frac{1}{2}\Big(\ln k - \ln(k-1)\Big).$$

Лемма доказана.

Теорема 2.3.

$$0,8 \cdot \mathbf{e}\sqrt{n}\left(\frac{n}{\mathbf{e}}\right)^n \le n! \le \mathbf{e}\sqrt{n}\left(\frac{n}{\mathbf{e}}\right)^n.$$

Доказательство. Для доказательства теоремы достаточно просуммировать неравенства леммы 2.1 по всем k от 2 до n. Суммируя правые неравенства, видим, что

$$\sum_{k=2}^{n} \ln k \le \int_1^n \ln x dx + \frac{1}{2}\Big(\ln n - \ln 1\Big) = n\ln n - n + 1 + \frac{1}{2}\ln n.$$

Следовательно,

$$n! \le \mathbf{e}\sqrt{n}\left(\frac{n}{\mathbf{e}}\right)^n.$$

Для того, чтобы оценить сумму левых неравенств положим

$$a_1 = \sum_{k=2}^{n}(\ln(2k) - \ln(2k-1)), \quad a_2 = \sum_{k=2}^{n}(\ln(2k+1) - \ln(2k)).$$

Легко видеть, что $a_1 + a_2 = \ln(2n+1) - \ln 3$. А так как $a_1 > a_2$, то

$$a_1 > \frac{1}{2}\ln(2n+1) - \frac{1}{2}\ln 3 > \frac{1}{2}\ln n - \frac{1}{2}\ln\frac{3}{2}.$$

Таким образом,

$$\sum_{k=2}^{n} \ln k > n\ln n - n + 1 + \frac{1}{2}\ln n - \frac{1}{2}\ln\frac{3}{2}.$$

Теперь заметим, что $\sqrt{2/3} > 0,8$, и, следовательно,

$$n! > 0,8 \cdot \mathbf{e} \cdot \sqrt{n}\left(\frac{n}{\mathbf{e}}\right)^n.$$

Теорема доказана.

2.2 Формула Стирлинга

Неравенства теоремы 2.3 значительно точнее неравенств первых двух теорем. Тем не менее верхняя и нижняя оценки теоремы 2.3 все еще значительно отличаются. Нетрудно видеть, что это происходит из-за того, что использованные в доказательстве теоремы неравенства леммы 2.1, достаточно точные при больших k, становятся грубыми при малых значениях k. Если использовать неравенства леммы 2.1 только при больших значениях k, то можно надеяться на увеличение точности неравенств для $n!$. Именно так сделано в доказательстве следующей теоремы.

Теорема 2.4.

$$\sqrt{2\pi n} \cdot \left(\frac{n}{\mathbf{e}}\right)^n \mathbf{e}^{-1/4n} \le n! \le \sqrt{2\pi n} \cdot \left(\frac{n}{\mathbf{e}}\right)^n \mathbf{e}^{1/4n}.$$

Доказательство теоремы 2.4 основано на неравенствах леммы 2.1 и оценках биномиального коэффициента $\binom{2n}{n}$, которые будут установлены ниже в лемме 2.3. Для доказательства этой леммы потребуется следующее вспомогательное утверждение.

Лемма 2.2.

$$\int_0^{\pi/2} \sin^{2n} dx = \frac{(2n-1)(2n-3)\cdot\ldots\cdot 3\cdot 1}{2n(2n-2)\cdot\ldots\cdot 4\cdot 2}\cdot\frac{\pi}{2},$$
$$\int_0^{\pi/2} \sin^{2n+1} dx = \frac{2n(2n-2)\cdot\ldots\cdot 4\cdot 2}{(2n+1)(2n-1)\cdot\ldots\cdot 3\cdot 1}.$$

Доказательство. Обозначим интеграл $\int_0^{\pi/2}\sin^n x dx$ через I_n. Тогда

$$\begin{aligned} I_n &= -\int_0^{\pi/2} \sin^{n-1} d\cos x = \\ &= -\sin^{n-1} x\cos x\Big|_0^{\pi/2} + \int_0^{\pi/2} \cos x d\sin^{n-1} x = \\ &= (n-1)\int_0^{\pi/2} \cos^2 x \sin^{n-2} x dx = \\ &= (n-1)\int_0^{\pi/2} (1-\sin^2)\sin^{n-2} x dx = (n-1)(I_{n-2} - I_n). \end{aligned}$$

Следовательно, имеет место рекуррентная формула

$$I_n = \frac{n-1}{n} I_{n-2},$$

последовательное применение которой к интегралам I_{2n} и I_{2n+1} дает следующие равенства:

$$\begin{aligned} I_{2n} &= \frac{(2n-1)(2n-3)\cdot\ldots\cdot 3\cdot 1}{2n(2n-2)\cdot\ldots\cdot 4\cdot 2}\cdot I_0, \\ I_{2n+1} &= \frac{2n(2n-2)\cdot\ldots\cdot 4\cdot 2}{(2n+1)(2n-1)\cdot\ldots\cdot 3\cdot 1}\cdot I_1. \end{aligned} \tag{2.3}$$

Так как $I_0 = \pi/2$ и $I_1 = 1$, то подставив эти значения в (2.3) получим требуемые равенства. Лемма доказана.

Напомним, что через $n!!$ обозначается произведения всех натуральных чисел, не превосходящих n и имеющих такую же четность как и n, т. е.

$$2n!! = 2n(2n-2)\cdot\ldots\cdot 2, \quad (2n+1)!! = (2n+1)(2n-1)\cdot\ldots\cdot 1.$$

Используя эти обозначения, равенства леммы 2.2 можно записать так:

$$\int_0^{\pi/2} \sin^{2n} dx = \frac{(2n-1)!!}{2n!!} \cdot \frac{\pi}{2}, \quad \int_0^{\pi/2} \sin^{2n+1} dx = \frac{2n!!}{(2n+1)!!}.$$

Лемма 2.3.

$$\frac{2^{2n}}{\sqrt{\pi n}} \mathbf{e}^{-1/4n} \le \binom{2n}{n} \le \frac{2^{2n}}{\sqrt{\pi n}}.$$

Доказательство. Так как $\sin x$ между 0 и $\pi/2$ изменяется от 0 до 1, то $\sin^{2n+1} x \le \sin^{2n} x \le \sin^{2n-1} x$ при $x \in [0, \pi/2]$. Следовательно,

$$\int_0^{\pi/2} \sin^{2n+1} x dx \le \int_0^{\pi/2} \sin^{2n} x dx \le \int_0^{\pi/2} \sin^{2n-1} x dx.$$

Применяя лемму 2.2, получим следующие неравенства

$$\frac{2n!!}{(2n+1)!!} \le \frac{(2n-1)!!}{2n!!} \cdot \frac{\pi}{2} \le \frac{(2n-2)!!}{(2n-1)!!},$$

которые, как легко видеть, преобразуются к виду

$$\frac{2n!! \cdot 2n!!}{(2n+1)!! \cdot (2n-1)!!} \le \frac{\pi}{2} \le \frac{2n!! \cdot (2n-2)!!}{(2n-1)!! \cdot (2n-1)!!}.$$

Извлекая квадратные корни из новых неравенств, получим, что

$$\frac{1}{\sqrt{2n+1}} \cdot \frac{2n!!}{(2n-1)!!} \le \sqrt{\frac{\pi}{2}} \le \frac{1}{\sqrt{2n}} \cdot \frac{2n!!}{(2n-1)!!}.$$

Далее разделим все члены получившихся неравенств на $2n!!$ и $\sqrt{\pi/2}$ и умножим на $(2n-1)!!$ и 2^{2n}. Тогда

$$\frac{1}{\sqrt{1+1/2n}} \cdot \frac{2^{2n}}{\sqrt{\pi n}} \le 2^{2n} \cdot \frac{(2n-1)!!}{2n!!} \le \frac{2^{2n}}{\sqrt{\pi n}}. \tag{2.4}$$

Наконец, заметим, что

$$\frac{(2n-1)!!}{2n!!} = \frac{(2n-1)!! \cdot 2n!!}{2n!! \cdot 2n!!} = \frac{(2n)!}{2^{2n} \cdot n! \cdot n!} = \binom{2n}{n} \cdot 2^{-2n}.$$

Теперь, учитывая, что $\mathbf{e}^{-x} \le 1/(1+x)$ при $0 < x < 1$, подставим последнее равенство в (2.4) и получим требуемые оценки для $\binom{2n}{n}$. Лемма доказана.

Доказательство Теоремы 2.4. Так как

$$\binom{2n}{n} = \frac{2n!}{n! \cdot n!} = \frac{2n(2n-1) \cdot \ldots \cdot (n+1)}{n!},$$

то легко видеть, что

$$n! = 2n(2n-1) \cdot \ldots \cdot (n+1) \Big/ \binom{2n}{n}. \tag{2.5}$$

Оценим логарифм произведения $2n(2n-1) \cdot \ldots \cdot (n+1)$. Для этого воспользуемся неравенствами

$$\ln k \ge \int_{k-1}^{k} \ln x dx + \ln(2k) - \ln(2k-1), \tag{2.6}$$

$$\ln k \le \int_{k-1}^{k} \ln x dx + \frac{1}{2}\Big(\ln k - \ln(k-1)\Big), \tag{2.7}$$

которые были доказаны в лемме 2.1. Суммируя неравенства (2.7) по всем k от $n+1$ до $2n$, видим, что

$$\sum_{k=n+1}^{2n} \ln k \le \int_n^{2n} \ln x dx + \frac{1}{2}\Big(\ln(2n) - \ln n\Big) =$$

$$= \int_n^{2n} \ln x dx + \frac{1}{2}\ln 2 = n \ln n + 2n \ln 2 - n + \frac{1}{2}\ln 2.$$

Для того, чтобы оценить аналогичную сумму неравенств (2.6), как и ранее при доказательстве теоремы 2.3 положим

$$a_1 = \sum_{k=n+1}^{2n} (\ln(2k) - \ln(2k-1)), \quad a_2 = \sum_{k=n+1}^{2n} (\ln(2k+1) - \ln(2k)).$$

Легко видеть, что $a_1 + a_2 = \ln(4n+1) - \ln(2n+1)$. А так как $a_1 > a_2$, то

$$a_1 > \frac{1}{2}\ln(4n+1) - \frac{1}{2}\ln(2n+1) = \frac{1}{2}\ln\Big(2 \cdot \frac{2n+1/2}{2n+1}\Big) =$$

$$= \frac{1}{2}\ln 2 + \frac{1}{2}\ln\Big(1 - \frac{1}{4n+2}\Big) \ge \frac{1}{2}\ln 2 - \frac{1}{4n}.$$

Таким образом,

$$\sum_{k=n+1}^{2n} \ln k > n \ln n + 2n \ln 2 - n + \frac{1}{2}\ln 2 - \frac{1}{4n}.$$

Следовательно,

$$\sqrt{2}\Big(\frac{n}{\mathbf{e}}\Big)^n 2^{2n} \mathbf{e}^{-1/4n} \le 2n(2n-1) \cdot \ldots \cdot (n+1) \le \sqrt{2}\Big(\frac{n}{\mathbf{e}}\Big)^n 2^{2n}.$$

Из леммы 2.3 следует, что

$$\frac{\sqrt{\pi n}}{2^{2n}} \le 1 \Big/ \binom{2n}{n} \le \frac{\sqrt{\pi n}}{2^{2n}} \mathbf{e}^{1/4n}.$$

Теперь умножим почленно последние неравенства и получим

$$\sqrt{2\pi n} \cdot \Big(\frac{n}{\mathbf{e}}\Big)^n \mathbf{e}^{-1/4n} \le n! \le \sqrt{2\pi n} \cdot \Big(\frac{n}{\mathbf{e}}\Big)^n \mathbf{e}^{1/4n}.$$

Теорема доказана.

Отношение верхнего и нижнего неравенств теоремы 2.4 не превосходит $\mathbf{e}^{1/2n}$ и при $n \to \infty$ стремится к единице. Поэтому из теоремы 2.4 легко следует известная формула Стирлинга

$$n! = \sqrt{2\pi n} \cdot \left(\frac{n}{\mathbf{e}}\right)^n (1 + o(1)).$$

Неравенства для $n!$, установленные в теореме 2.4, можно усилить показав (например, см. [32]), что

$$\sqrt{2\pi n} \cdot \left(\frac{n}{\mathbf{e}}\right)^n \mathbf{e}^{1/(12n+1)} \le n! \le \sqrt{2\pi n} \cdot \left(\frac{n}{\mathbf{e}}\right)^n \mathbf{e}^{1/12n}.$$

Более точные оценки $n!$ можно найти в [8].

2.3 Биномиальные коэффициенты

Используя неравенства теоремы 2.1, для биномиальных коэффициентов нетрудно установить следующие часто используемые оценки

$$\left(\frac{2(n-k)}{k}\right)^k \le \binom{n}{k} \le \left(\frac{3n}{k}\right)^k,$$

справедливые при $k \ge 6$. Более точные формулы для биномиальных коэффициентов можно получить применяя более точные оценки $n!$.

Далее при помощи формулы Стирлинга установим три асимптотически точные формулы для $\binom{n}{k}$, справедливые при некоторых ограничениях на параметр k. В первой из этих формул биномиальный коэффициент выражается через *функцию энтропии* $H(x)$. Функция энтропии для каждого x из интервала $(0, 1)$ определяется равенством

$$H(x) = -x \log_2 x - (1-x) \log_2 (1-x).$$

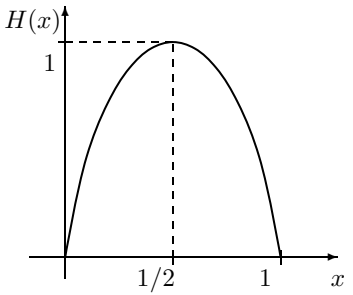


Рис. 2.2

При использовании функции $H(x)$ удобно считать, что точки нуль и единица принадлежат ее области определения. Поэтому доопределим $H(x)$ по непрерывности в нуле и единице положив $H(0) = H(1) = 0$. Нетрудно видеть, что $H(x)$ неотрицательна, выпукла вверх, симметрична относительно прямой $x = 1$ и достигает своего максимального значения равного единице при $x = 1/2$. График функции $H(x)$ изображен на рис. 2.2.

Теорема 2.5. *Если* $\min(k, n-k) \to \infty$, *то*

$$\binom{n}{k} = \frac{2^{nH(\frac{k}{n})}}{\sqrt{2\pi k(n-k)/n}} (1 + o(1)).$$

Доказательство. Применяя формулу Стирлинга, видим, что[1)]

$$\begin{aligned}\binom{n}{k} &= \frac{n!}{k!(n-k)!} \sim \frac{\sqrt{2\pi n}}{\sqrt{2\pi k \cdot 2\pi(n-k)}} \cdot \frac{(n/\mathbf{e})^n}{(k/\mathbf{e})^k((n-k)/\mathbf{e})^{n-k}} = \\ &= \frac{1}{\sqrt{2\pi k(n-k)/n}} \cdot \frac{n^n}{k^k (n-k)^{n-k}} = \\ &= \frac{1}{\sqrt{2\pi k(n-k)/n}} \cdot \left(\frac{k}{n}\right)^{-k} \cdot \left(1 - \frac{k}{n}\right)^{-(n-k)} = \\ &= \frac{1}{\sqrt{2\pi k(n-k)/n}} \cdot 2^{nH(\frac{k}{n})}.\end{aligned}$$

Теорема доказана.

Если k мало или близко к $n/2$, то для $\binom{n}{k}$ существуют более простые формулы, которые получим в двух следующих теоремах. Для доказательства этих теорем потребуется оценка функции $\ln(1+x)$, которую можно легко получить из разложения этой функций в ряд Тейлора в нуле. Так как

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots + (-1)^{k+1}\frac{x^k}{k} + \dots,$$

то при $-1 < x < 1$

$$\ln(1+x) = x - \frac{x^2}{2} + \mathcal{O}(x^3). \tag{2.8}$$

Теорема 2.6. *Если* $n \to \infty$ *и* $t = o(n^{2/3})$, *то*

$$\binom{n}{n/2-t} = \frac{2^n \mathbf{e}^{-2t^2/n}}{\sqrt{\pi n/2}} (1 + o(1)).$$

Доказательство. Воспользуемся предыдущей теоремой и преобразуем показатель экспоненты, стоящей в правой части ее равенства:

$$\begin{aligned}H\Big(\frac{n/2-t}{n}\Big) &= H\Big(\frac{1}{2}\Big(1 - \frac{2t}{n}\Big)\Big) = \\ &= -\frac{1}{2}\Big(1 - \frac{2t}{n}\Big)\log_2 \frac{1}{2}\Big(1 - \frac{2t}{n}\Big) - \frac{1}{2}\Big(1 + \frac{2t}{n}\Big)\log_2 \frac{1}{2}\Big(1 + \frac{2t}{n}\Big) = \\ &= 1 - \frac{1}{2}\Big(\Big(1 - \frac{2t}{n}\Big)\log_2\Big(1 - \frac{2t}{n}\Big) + \Big(1 + \frac{2t}{n}\Big)\log_2\Big(1 + \frac{2t}{n}\Big)\Big).\end{aligned} \tag{2.9}$$

[1)] Формула $a(n) \sim b(n)$ означает, что $\lim_{n\to\infty} a(n)/b(n) = 1$.

Используя равенство (2.8) нетрудно показать, что

$$-\Big(1-\frac{2t}{n}\Big)\log_2\Big(1-\frac{2t}{n}\Big)-\Big(1+\frac{2t}{n}\Big)\log_2\Big(1+\frac{2t}{n}\Big)=-\Big(\frac{4t^2}{n^2}+\mathcal{O}\Big(\frac{t^3}{n^3}\Big)\Big)\log_2\mathbf{e}.$$

Поэтому

$$nH\Big(\frac{n/2-t}{n}\Big)=n\Big(1-\frac{1}{2}\Big(\frac{4t^2}{n^2}+\mathcal{O}\Big(\frac{t^3}{n^3}\Big)\Big)\log_2\mathbf{e}\Big)=n-\Big(\frac{2t^2}{n}+\mathcal{O}\Big(\frac{t^3}{n^2}\Big)\Big)\log_2\mathbf{e}.$$

Теперь подставляя полученное равенство в равенство теоремы 2.5 и учитывая условие $t=o(n^{2/3})$, получаем, что

$$\binom{n}{n/2-t}=\frac{2^n\mathbf{e}^{-2t^2/n+\mathcal{O}(t^3/n^2)}}{\sqrt{2\pi(n/2-t)(n/2+t)/n}}\,(1+o(1))=\frac{2^n\mathbf{e}^{-2t^2/n}}{\sqrt{\pi n/2}}\,(1+o(1)).$$

Теорема доказана.

Нетрудно видеть, что теорему 2.6 можно переформулировать следующим образом: если $n\to\infty$ и $t=o(n^{2/3})$, то

$$\binom{n}{n/2-t}=\binom{n}{\lfloor n/2\rfloor}\mathbf{e}^{-2t^2/n}\,(1+o(1)).$$

Теорема 2.7. *Если $n\to\infty$ и $k=o(n^{2/3})$, то*

$$\binom{n}{k}=\frac{n^k\mathbf{e}^{-k^2/2n}}{k!}\,(1+o(1)).$$

Доказательство. Применяя формулу Стирлинга, равенство (1.7) и условие $k=o(n^{2/3})$ видим, что

$$\begin{aligned}\binom{n}{k}&=\frac{n!}{k!(n-k)!}\sim\frac{\sqrt{2\pi n}}{\sqrt{2\pi(n-k)}}\cdot\frac{(n/\mathbf{e})^n}{k!((n-k)/\mathbf{e})^{n-k}}=\\&=\frac{n^k}{k!}\cdot\frac{\mathbf{e}^{-k}}{\sqrt{1-k/n}}\cdot\Big(1-\frac{k}{n}\Big)^{k-n}\sim\frac{n^k\mathbf{e}^{-k}}{k!}\cdot\mathbf{e}^{(k-n)\ln(1-k/n)}=\\&=\frac{n^k\mathbf{e}^{-k}}{k!}\cdot\mathbf{e}^{(k-n)\left(-k/n-k^2/(2n^2)-\mathcal{O}\left(k^3/n^3\right)\right)}=\\&=\frac{n^k\mathbf{e}^{-k}}{k!}\cdot\mathbf{e}^{k-k^2/(2n)-\mathcal{O}\left(k^3/n^2\right)}=\frac{n^k\mathbf{e}^{-k^2/2n}}{k!}\,(1+o(1)).\end{aligned}$$

Теорема доказана.

2.4 Суммы биномиальных коэффициентов

Установим несколько оценок для сумм биномиальных коэффициентов. Первая оценка доказывается так же, как и известное в теории вероятностей неравенство Чебышева.

Теорема 2.8. *Пусть $1\le\varphi(n)\le\sqrt{n}/2$. Тогда*[2)]

$$\sum_{k=0}^{n/2-\sqrt{n\varphi(n)}}\binom{n}{k}\le\frac{2^{n-3}}{\varphi(n)}.$$

Доказательство. Оценим сумму биномиальных коэффициентов, нижний индекс которых отличается от $\frac{n}{2}$ более чем на t единиц:

$$\begin{aligned}\sum_{k:|n/2-k|>t}\binom{n}{k}&=\sum_{k:|n/2-k|>t}\frac{(n/2-k)^2}{(n/2-k)^2}\binom{n}{k}\le\\&\le\frac{1}{t^2}\sum_{k:|n/2-k|>t}\Big(\frac{n}{2}-k\Big)^2\binom{n}{k}\le\frac{1}{t^2}\sum_{k=0}^{n}\Big(\frac{n}{2}-k\Big)^2\binom{n}{k}.\end{aligned}\tag{2.10}$$

Найдем сумму, стоящую в правой части неравенства (2.10). Легко видеть, что

$$\begin{aligned}\sum_{k=0}^{n}\binom{n}{k}\Big(\frac{n}{2}-k\Big)^2&=\sum_{k=0}^{n}\binom{n}{k}\Big(\frac{n^2}{4}-nk+k^2\Big)=\\&=\frac{n^2}{4}\sum_{k=0}^{n}\binom{n}{k}-\sum_{k=0}^{n}\binom{n}{k}(n-k)k.\end{aligned}\tag{2.11}$$

Первая сумма в правой части (2.11) равна $n^2 2^{n-2}$. Найдем вторую сумму:

$$\begin{aligned}\sum_{k=0}^{n}(n-k)k\binom{n}{k}&=\sum_{k=1}^{n-1}(n-k)k\binom{n}{k}=\\&=\sum_{k=1}^{n-1}\frac{(n-k)kn!}{(n-k)!k!}=\sum_{k=1}^{n-1}\frac{n(n-1)(n-2)!}{(n-k-1)!(k-1)!}=\\&=n(n-1)\sum_{k=0}^{n-2}\binom{n-2}{k}=n(n-1)2^{n-2}.\end{aligned}$$

Из двух предыдущих неравенств следует, что

$$\sum_{k=0}^{n}\binom{n}{k}\Big(\frac{n}{2}-k\Big)^2=n^2 2^{n-2}-n(n-1)2^{n-2}=n2^{n-2}.$$

Подставляя полученное равенство в правую часть (2.10) и полагая t равным $\sqrt{n\varphi(n)}$, находим, что

$$\sum_{k:|n/2-k|>\sqrt{n\varphi(n)}}\binom{n}{k}\le\frac{n2^{n-2}}{n\varphi(n)}=\frac{2^{n-2}}{\varphi(n)}.$$

[2)] Знак $\sum_a^b$ с нецелыми индексами означает суммирование по всем целым, лежащим между a и b.

Теорема доказана.

Неравенство теоремы 2.8 достаточно грубое (далее оно будет существенно усилено в теореме 2.10). Тем не менее метод, которым оно получено, представляет значительный интерес и может быть успешно использован в различных задачах.

Теорема 2.9. *При $1 \le t \le \frac{n}{2}$ справедливо неравенство*

$$\sum_{k=0}^{t} \binom{n}{k} \le 2^{nH(\frac{t}{n})}.$$

Доказательство. Полагая, что $0 < x < 1$, имеем

$$\sum_{k=0}^{t} \binom{n}{k} \le \sum_{k=0}^{t} x^{k-t} \binom{n}{k} = \frac{1}{x^t} \sum_{k=0}^{t} x^k \binom{n}{k} \le \frac{1}{x^t} \sum_{k=0}^{n} x^k \binom{n}{k} = \frac{(1+x)^n}{x^t}. \tag{2.12}$$

Дифференцируя функцию $f(x) = \frac{(1+x)^n}{x^t}$ по x, видим, что ее производная

$$\Big(\frac{(1+x)^n}{x^t}\Big)' = \frac{n(1+x)^{n-1}}{x^t} - \frac{t(1+x)^n}{x^{t+1}} = \\ = \frac{(1+x)^{n-1}}{x^{t+1}}\Big(nx - t(1+x)\Big)$$

между нулем и единицей имеет единственный корень $x_0 = \frac{t}{n-t}$. Так как при этом $f(x)$ неограниченно возрастает при x стремящемся к нулю справа, и $f(1) = 2^n$, то на интервале $(0,1)$ функция $f(x)$ достигает своего минимального значения в точке x_0. Поэтому

$$\sum_{k=0}^{t} \binom{n}{k} \le \Big(\frac{t}{n-t}\Big)^{-t}\Big(1 + \frac{t}{n-t}\Big)^n = \Big(\frac{t}{n-t}\Big)^{-t}\Big(\frac{n}{n-t}\Big)^n = \\ = \Big(\frac{t}{n}\Big)^{-t}\Big(\frac{n}{n-t}\Big)^{n-t} = \Big(\frac{t}{n}\Big)^{-t}\Big(1 - \frac{t}{n}\Big)^{-(n-t)} = \\ = \Big(\Big(\frac{t}{n}\Big)^{-t/n}\Big(1 - \frac{t}{n}\Big)^{-(1-t/n)}\Big)^n = 2^{nH(\frac{t}{n})}.$$

Теорема доказана.

Тривиальным следствием доказанной теоремы является неравенство

$$\binom{n}{k} \le 2^{nH(\frac{k}{n})}, \tag{2.13}$$

справедливое при $0 \le k \le n$.

Неравенство теоремы 2.9 называется энтропийным. Неравенство, доказываемое в следующей теореме, называется неравенством Чернова, по имени американского математика, получившего это неравенство в более общей форме в 1952 году.

Теорема 2.10. *При $0 \le t \le \frac{n}{2}$ справедливо неравенство*

$$\sum_{k=0}^{n/2-t} \binom{n}{k} \le 2^n \mathbf{e}^{-2t^2/n}.$$

Доказательство. Из теоремы 2.9 и доказанного на стр. 29 равенства (2.9) следует, что

$$\sum_{k=0}^{n/2-t} \binom{n}{k} \le 2^{nH\left(\frac{n/2-t}{n}\right)} = 2^{nH\left(\frac{1}{2}\left(1-\frac{2t}{n}\right)\right)} \le \\ \le 2^{n\left(1-\frac{1}{2}\left(\left(1-\frac{2t}{n}\right)\log_2\left(1-\frac{2t}{n}\right)+\left(1+\frac{2t}{n}\right)\log_2\left(1+\frac{2t}{n}\right)\right)\right)}. \tag{2.14}$$

Для того, чтобы оценить показатель экспоненты в правой части неравенства (2.14) покажем, что

$$f(x) = (1-x)\ln(1-x) + (1+x)\ln(1+x) - x^2 \ge 0$$

при $x \in (-1,1)$. Прежде всего заметим, что $f(x)$ — четная функция. Следовательно, справедливость доказываемого неравенства достаточно установить только для полуинтервала $[0,1)$, а так как $f(0) = 0$, то достаточно показать, что на этом полуинтервале производная функции $f(x)$ неотрицательна. Дифференцируя $f(x)$, находим

$$f'(x) = -\frac{1-x}{1-x} - \ln(1-x) + \frac{1+x}{1+x} + \ln(1+x) - 2x = \\ = \ln(1+x) - \ln(1-x) - 2x.$$

Нетрудно видеть, что $f'(0) = 0$ и вторая производная

$$f''(x) = \frac{1}{1+x} + \frac{1}{1-x} - 2 = \frac{2}{1-x^2} - 2$$

функции $f(x)$ на $[0,1)$ неотрицательна. Таким образом, $f'(x) \ge 0$ на $[0,1)$, и поэтому,

$$(1-x)\ln(1-x) + (1+x)\ln(1+x) \ge x^2$$

при всех x из интервала $(-1,1)$. Следовательно,

$$-\Big(1-\frac{2t}{n}\Big)\log_2\Big(1-\frac{2t}{n}\Big) - \Big(1+\frac{2t}{n}\Big)\log_2\Big(1+\frac{2t}{n}\Big) \le -\frac{4t^2}{n^2}\log_2 \mathbf{e}. \tag{2.15}$$

Подставляя неравенство (2.15) в неравенство (2.14), нетрудно видеть, что

$$\sum_{k=0}^{n/2-t} \binom{n}{k} \le 2^{n\left(1-\frac{1}{2}\cdot\frac{4t^2}{n^2}\log_2 \mathbf{e}\right)} = 2^n \mathbf{e}^{-2t^2/n}.$$

Теорема доказана.

2.5 Задачи

2.1. Показать, что при $0 \le x < 1$

$$1 + x \le \mathbf{e}^x \le \frac{1}{1-x}, \qquad 1 - x \le \mathbf{e}^{-x} \le \frac{1}{1+x}.$$

2.2. Показать, что при $-1 < x < 1$

$$\mathbf{e}^{x/(1+x)} \le 1 + x \le \mathbf{e}^x.$$

2.3. Показать, что

$$\left(\frac{n}{k}\right)^k \le \binom{n}{k} \le (n-k+1)^k.$$

2.4. Показать, что при фиксированном k и $n \to \infty$

$$\binom{n}{k} = \frac{n^k}{k!}\,(1 + o(1)).$$

2.5. Показать, что если $n \to \infty$ и $k = o(n^{3/4})$, то

$$\binom{n}{k} = \frac{n^k \mathbf{e}^{-k^2/2n - k^3/6n^2}}{k!}\,(1 + o(1)).$$

2.6. Показать, что

$$\binom{n}{k} \le \frac{2^{nH(\frac{k}{n})}}{\sqrt{2\pi k(n-k)/n}}.$$

2.7. Пусть $x_1 + \cdots + x_m = 1$, $x_i > 0$ и $H(x_1, \ldots, x_m) = -\sum_{i=1}^m x_i \log_2 x_i$. Показать, что если $k_1 + \cdots + k_m = n$ и $k_i > 0$, то

$$P(k_1, \ldots, k_m) = \frac{n!}{k_1! \cdots k_m!} \le 2^{nH(\frac{k_1}{n}, \ldots, \frac{k_m}{n})}.$$

2.8. Найти и оценить $\max(P(k_1, \ldots, k_m))$ по всем наборам $k_1, \ldots, k_m$ таким, что $k_1 + \cdots + k_m = n$ и $k_i > 0$.

2.9. Пусть $1 \le \varphi(n) \le \sqrt{n}/2$ и $0 < p < 1$. Показать, что

$$\sum_{k=0}^{np - \sqrt{n\varphi(n)}} \binom{n}{k} p^k (1-p)^{n-k} \le \frac{p(1-p)}{2\varphi(n)}.$$

2.10. Показать, что при $k \to \infty$ и $k = o(n)$

$$\sum_{m=0}^{k} \binom{n}{m} = \frac{2^{nH(\frac{k}{n})}}{\sqrt{2\pi k(n-k)/n}}\,(1 + o(1)).$$

2.11. Показать, что при $k \to \infty$ и $k = o(n^{2/3})$

$$\sum_{m=0}^{k} \binom{n}{m} = \frac{n^k \mathbf{e}^{-k^2/2n}}{k!}\,(1 + o(1)).$$

2.12. Пусть $n_1 + n_2 + \cdots + n_m = n$ и n делится на m. Показать, что

$$\sum_{i=1}^{m} \binom{n_i}{k} \ge m\binom{n/m}{k}.$$

Лекция 3

Производящие функции

Метод производящих функций является мощным средством для работы с различными множествами дискретной природы. Основная идея этого метода заключается в отображении исследуемых множеств в множество степенных рядов и последующей работе с рядами при помощи развитого аппарата теории функций. Ниже рассматривается применение метода производящих функций для решения линейных рекуррентных соотношений с постоянными коэффициентами и нахождения числа неприводимых многочленов над конечным полем. После этого вводится понятие производящей функции множества и доказывается ряд теорем, позволяющих устанавливать соответствие между действиями над множествами и действиями над функциями.

3.1 Линейные рекуррентные последовательности

Производящей функцией последовательности $\{F_n\}_{n=0}^{\infty}$, где $F_n \in \mathbb{R}$, называется ряд

$$F(x) = \sum_{n=0}^{\infty} x^n F_n. \tag{3.1}$$

Если ряд (3.1) сходится абсолютно в какой-либо окрестности нуля, то его можно умножить на другой ряд, возвести его в степень, продифференцировать и т. д. Используя такие действия, можно попытаться найти явный вид функции $F(x)$ в случае, когда последовательность $\{F_n\}_{n=0}^{\infty}$ неизвестна и описывается только какими-нибудь своими свойствами. Если явную формулу для $F(x)$ удается найти, то далее можно попытаться найти явную формулу и для F_n — общего члена последовательности. Сделать это можно, вычислив n-ю производную функции $F(x)$ или разложив эту функцию в ряд каким-либо иным способом.

Подобным образом доказывается следующая теорема о решении линейных рекуррентных соотношений с постоянными коэффициентами.

Теорема 3.1. *Пусть последовательность $F_0, F_1, F_2, \ldots, F_n, \ldots$ удовлетворяет линейному рекуррентному соотношению*

$$F_n = a_1 F_{n-1} + a_2 F_{n-2} + \cdots + a_k F_{n-k}. \tag{3.2}$$

с постоянными коэффициентами a_i при $n \geq k$. Тогда при $n \geq 0$

$$F_n = \sum_{i=1}^{m} \alpha_i^n P_i(n), \tag{3.3}$$

где α_i — корень многочлена $f(x) = x^k - a_1 x^{k-1} - a_2 x^{k-2} - \cdots - a_k$ кратности p_i, P_i — многочлен степени $p_i - 1$, коэффициенты которого определяются так, чтобы равенство (3.3) *было справедливо для первых k членов рассматриваемой последовательности.*

Доказательство. Последовательности $\{F_n\}$ поставим в соответствие производящую функцию

$$F(x) = \sum_{n=0}^{\infty} x^n F_n.$$

Прежде всего заметим, что из соотношения (3.2) следует не более чем степенной рост модуля n-го члена рассматриваемой последовательности. Поэтому существует окрестность нуля, в которой ряд $F(x)$ сходится абсолютно.

Правую и левую части рекуррентного соотношения (3.2) умножим на x^n и просуммируем по всем целым n от k до ∞. В результате получим равенство

$$\sum_{n=k}^{\infty} x^n F_n = \sum_{n=k}^{\infty} x^n (a_1 F_{n-1} + a_2 F_{n-2} + \cdots + a_k F_{n-k}).$$

Разбивая сумму, стоящую в правой части последнего равенства, на k частей и вынося из под знака i-й суммы множитель $a_i x^i$, получим новое равенство

$$\begin{aligned} \sum_{n=k}^{\infty} x^n F_n = a_1 x \sum_{n=k}^{\infty} x^{n-1} F_{n-1} + a_2 x^2 \sum_{n=k}^{\infty} x^{n-2} F_{n-2} + \ldots \\ \cdots + a_k x^k \sum_{n=k}^{\infty} x^{n-k} F_{n-k}. \end{aligned} \tag{3.4}$$

Так как $F(x) = \sum_{n=0}^{\infty} x^n F_n$, то легко видеть, что для левой части (3.4) справедливо равенство

$$\sum_{n=k}^{\infty} x^n F_n = F(x) - (F_0 + xF_1 + \cdots + x^{k-1} F_{k-1}),$$

а для i-й суммы правой части (3.4) — равенство

$$a_i x^i \sum_{n=k}^{\infty} x^{n-i} F_{n-i} = a_i x^i \sum_{m=k-i}^{\infty} x^m F_m = \\ = a_i x^i (F(x) - (F_0 + xF_1 + \cdots + x^{k-i-1} F_{k-i-1})).$$

Поэтому равенство (3.4) можно записать в виде

$$F(x) - (F_0 + xF_1 + \ldots + x^{k-1}F_{k-1}) = \\ = a_1 x^1 (F(x) - (F_0 + xF_1 + \ldots + x^{k-2}F_{k-2})) + \\ + a_2 x^2 (F(x) - (F_0 + xF_1 + \ldots + x^{k-3}F_{k-3})) + \ldots + a_k x^k F(x).$$

Перенося в последнем неравенстве в левую часть все слагаемые содержащие множитель $F(x)$, а в правую — все остальные, получим новое равенство

$$F(x)(1 - a_1 x - a_2 x^2 - \ldots - a_k x^k) = \\ = F_0(1 - a_1 x - a_2 x^2 - \ldots - a_{k-1} x^{k-1}) + \\ + F_1(1 - a_1 x - a_2 x^2 - \ldots - a_{k-2} x^{k-2}) + \ldots + x^{k-1} F_{k-1},$$

в правой части которого стоит многочлен степени не выше $k-1$. Обозначим этот многочлен через $H_{k-1}(x)$. Тогда легко видеть, что

$$F(x) = \frac{H_{k-1}(x)}{1 - a_1 x - a_2 x^2 - \ldots - a_k x^k}. \tag{3.5}$$

Предположим, что многочлен, стоящий в знаменателе правой части, имеет m различных корней $\frac{1}{\alpha_1}, \frac{1}{\alpha_2}, \ldots, \frac{1}{\alpha_m}$, кратности которых соответственно равны $p_1, p_2, \ldots, p_m$. Раскладывая знаменатель на множители, получим, что

$$F(x) = \frac{H_{k-1}(x)}{(1-\alpha_1 x)^{p_1}(1-\alpha_2 x)^{p_2} \ldots (1-\alpha_k x)^{p_m}}, \tag{3.6}$$

где очевидно $p_1 + p_2 + \cdots + p_m = k$. Представим правую часть последнего равенства в виде суммы простейших дробей. Тогда

$$F(x) = \frac{\beta^1_{p_1}}{(1-\alpha_1 x)^{p_1}} + \cdots + \frac{\beta^1_{p_1-j}}{(1-\alpha_1 x)^{p_j-1}} + \cdots + \frac{\beta^1_1}{(1-\alpha_1 x)} + \\ + \frac{\beta^2_{p_2}}{(1-\alpha_2 x)^{p_2}} + \cdots + \frac{\beta^2_{p_2-j}}{(1-\alpha_2 x)^{p_2-j}} + \cdots + \frac{\beta^2_1}{(1-\alpha_2 x)} + \ldots \\ + \frac{\beta^m_{p_m}}{(1-\alpha_m x)^{p_m}} + \cdots + \frac{\beta^m_{p_m-j}}{(1-\alpha_2 x)^{p_m-j}} + \cdots + \frac{\beta^m_1}{(1-\alpha_m x)},$$

где β^j_i — константы, возможно комплексные. Так как

$$\frac{1}{(1-\alpha x)^k} = \sum_{n=0}^{\infty} (\alpha x)^n \binom{n+k-1}{n},$$

то, раскладывая в ряд сумму

$$\frac{\beta_p}{(1-\alpha x)^p} + \cdots + \frac{\beta_{p-j}}{(1-\alpha x)^{p-j}} + \cdots + \frac{\beta_1}{(1-\alpha x)},$$

находим, что в этом ряду коэффициент при x^n равен

$$\alpha^n \sum_{j=0}^{p} \beta_{p-j} \binom{n+p-1-j}{n}.$$

Поэтому для F_n справедливо равенство

$$F_n = \sum_{i=1}^{m} \alpha_i^n \sum_{j=0}^{p} \beta^i_{p_i-j} \binom{n+p_i-1-j}{n}. \tag{3.7}$$

Теперь заметим, что для любых постоянных $\beta_{p-1}, \ldots, \beta_0$ найдутся такие постоянные $\gamma_{p-1}, \ldots, \gamma_0$, что

$$\sum_{j=0}^{p} \beta_{p-j} \binom{n+p-1-j}{n} = \sum_{j=0}^{p} \gamma_{p-j} n^{n+p-1-j}.$$

Поэтому (3.7) можно переписать в виде равенства

$$F_n = \sum_{i=1}^{m} \alpha_i^n \sum_{j=0}^{p} \gamma^i_{p_i-j} n^{n+p_i-1-j}, \tag{3.8}$$

в котором постоянные γ^j_i определяются при помощи подстановки в (3.8) вместо n величин $0, 1, \ldots, k-1$ и последующего решения получившейся системы линейных уравнений.

Рекуррентному равенству (3.2) сопоставим его характеристический многочлен

$$f(x) = x^k - a_1 x^{k-1} - a_2 x^{k-2} - \cdots - a_k. \tag{3.9}$$

Заменяя в этом многочлене переменную x на $\frac{1}{y}$ и умножая затем результат на y^k, получим новый многочлен

$$f^*(y) = 1 - a_1 y - a_2 y^2 - \cdots - a_k y^k,$$

который совпадает с многочленом, стоящим в знаменателе правой части (3.5). Так как $f^*(y) = y^k f\left(\frac{1}{y}\right)$, то легко видеть, что если α является корнем уравнения $f^*(y) = 0$, то $\frac{1}{\alpha}$ будет корнем уравнения $f(y) = 0$. Таким образом, из (3.8) следует, что n-й член рекуррентной последовательности (3.2) представляется в виде

$$F_n = \sum_{i=1}^{m} (\alpha_i)^n P_i(n),$$

где α_i — корень характеристического многочлена (3.9) кратности p_i, P_i — многочлен степени $p_i - 1$, коэффициенты которого определяются при помощи первых k членов рассматриваемой последовательности. Теорема доказана.

В рекуррентном соотношении (3.2) перенесем все ненулевые слагаемые в левую часть. В результате получим равенство, в левой части которого находится линейная комбинация $k+1$ элемента последовательности, а в правой — нуль. Соотношение такого вида называется однородным и является частным случаем общего рекуррентного соотношения

$$F_n - a_1F_{n-1} - a_2F_{n-2} - \ldots - a_kF_{n-k} = f(n). \tag{3.10}$$

Если в (3.10) функция $f(n)$ не равна нулю, то соотношение называется неоднородным. Для решения неоднородных рекуррентных соотношений также можно использовать метод производящих функций. Как это можно сделать, покажем на следующем примере.

Пусть в последовательности $\{F_n\}_0^\infty$ первые два члены равны единице, а остальные удовлетворяют неоднородному рекуррентному соотношению

$$F_n = 3F_{n-1} - 2F_{n-2} + n. \tag{3.11}$$

Пусть как и ранее $F(x) = \sum_{n=0}^{\infty} F_nx^n$. Тогда

$$\sum_{n=2}^{\infty} F_nx^n = 3\sum_{i=2}^{\infty} F_{n-1}x^n - 2\sum_{n=2}^{\infty} F_{n-2}x^n + \sum_{n=2}^{\infty} n\cdot x^n,$$

или

$$F(x) - F_0 - xF_1 = 3xF(x) - 3xF_0 - 2x^2F(x) + \frac{x}{(1-x)^2} - x.$$

Подставив в последнее равенство единицы вместо F_0 и F_1 и выполнив несложные преобразования, получим, что

$$F(x) = \frac{1-3x}{1-3x+2x^2} + \frac{x}{(1-x)^2(1-3x+2x^2)}.$$

Так как $1-3x+2x^2 = (1-x)(1-2x)$, то

$$F(x) = \frac{H_3(x)}{(1-x)^3(1-2x)},$$

где $H_3(x)$ — многочлен третьей степени. Следовательно,

$$F_n = an^2 + bn + c + d2^n. \tag{3.12}$$

Из (3.11) при $n = 2$ и 3 легко находим $F_2 = 3$ и $F_3 = 10$. Подставив в (3.12) вместо n числа $0, 1, 2$ и 3, получим систему линейных уравнений для определения значений a, b, c и d:

$$\begin{cases} 1 = c + d, \\ 1 = a + b + c + 2d, \\ 3 = 4a + 2b + c + 4d, \\ 10 = 9a + 3b + c + 8d. \end{cases}$$

Решая эту систему, находим $a = -\frac{1}{2}$, $b = -\frac{5}{2}$, $c = -2$, $d = 3$. Следовательно,

$$F_n = -\frac{1}{2}n^2 - \frac{5}{2}n - 2 + 3\cdot 2^n.$$

Из рассмотренного примера видно, что возможность решения соотношения (3.10) существенно зависит от вида функции $f(n)$. В частности, если $f(n)$ является квазимногочленом, т. е. $f(n) = \alpha^n P(n)$, где α — константа, а $P(n)$ — многочлен, соотношение (3.10) решается практически также, как и однородное.

3.2 Число неприводимых многочленов

Применим метод производящих функций для нахождения числа неприводимых многочленов над полем $\mathbb{Z}_p$. Число неприводимых многочленов степени n, у которых коэффициент при старшей степени равен единице, обозначим через $P(n)$.

Лемма 3.1. *Для последовательности $P(n)$ справедливо рекуррентное равенство*

$$p^n = \sum_{m\,|\,n} mP(m). \tag{3.13}$$

Доказательство. Пусть $p_{1m}, p_{2m}, \ldots, p_{P(m)m}$ — все неприводимые многочлены степени m. Нетрудно видеть, что, раскрывая скобки в произведении

$$\prod_{m=1}^{\infty}\prod_{k=1}^{P(m)} \left(1 + p_{km} + (p_{km})^2 + \cdots + (p_{km})^l + \ldots\right), \tag{3.14}$$

получим сумму $\sum$ всевозможных произведений неприводимых многочленов, причем каждое произведение встретится в этой сумме ровно один раз. Так как каждый многочлен единственным образом раскладывается в произведение неприводимых многочленов, то в $\sum$ будет содержаться ровно p^n произведений степени n. Каждому неприводимому многочлену степени m поставим в соответствие одночлен x^m, а произведению (3.14) — произведение

$$\prod_{m=1}^{\infty}\prod_{k=1}^{P(m)} \left(1 + x^m + (x^m)^2 + \cdots + (x^m)^l + \ldots\right) = \prod_{m=1}^{\infty}\left(\frac{1}{1-x^m}\right)^{P(m)}. \tag{3.15}$$

Так как существует ровно p^n многочленов степени n, у которых коэффициент при x^n равен единице, то легко видеть, что в ряду, получившемся после раскрытия скобок в (3.15), коэффициент при x^n будет равен p^n. Следовательно,

$$\frac{1}{1-px} = \prod_{m=1}^{\infty}\left(\frac{1}{1-x^m}\right)^{P(m)}. \tag{3.16}$$

Логарифмируя правую и левую части (3.16), получим новое равенство

$$\ln \frac{1}{1-px} = \sum_{m=1}^{\infty} P(m) \ln \frac{1}{1-x^m}.$$

Теперь, применяя формулу $\ln \frac{1}{1-x} = \sum_{n=1}^{\infty} \frac{1}{n} x^n$, разложим в ряд правую и левую части последнего равенства:

$$\sum_{n=1}^{\infty} \frac{1}{n} p^n x^n = \sum_{m=1}^{\infty} \sum_{k=1}^{\infty} \frac{1}{k} P(m) x^{km} = \sum_{n=1}^{\infty} \left(\sum_{km=n} \frac{1}{k} P(m) \right) x^n.$$

Приравнивая в получившемся равенстве коэффициенты при n-й степени x, находим

$$\frac{1}{n} p^n = \sum_{km=n} \frac{1}{k} P(m) = \sum_{m \mid n} \frac{m}{n} P(m).$$

Лемма доказана.

Для того, чтобы из равенства (3.13) в явном виде выразить функцию $P(n)$ воспользуемся формулой обращения Мебиуса, которую докажем далее в лемме 3.3. Сначала определим функцию Мебиуса

$$\mu(n) = \begin{cases} 1, & \text{если } n = 1; \\ (-1)^k, & \text{если } n \text{ — произведение } k \text{ простых чисел;} \\ 0, & \text{если } n \text{ делится на квадрат простого числа,} \end{cases}$$

и покажем, что имеет место следующее утверждение.

Лемма 3.2. *Справедливо равенство*

$$\sum_{m \mid n} \mu(m) = \begin{cases} 1, & \text{если } n = 1; \\ 0, & \text{если } n > 0. \end{cases} \tag{3.17}$$

Доказательство. Если $n = 1$, то единица является единственным делителем, и, следовательно, $\mu(1) = 1$. При $n > 1$ представим n в виде произведения простых чисел: $n = p_1^{q_1} \cdots p_r^{q_r}$. Легко видеть, что в сумме (3.17) нужно учитывать только делители без кратных множителей. Поэтому

$$\sum_{m \mid n} \mu(m) = \sum_{k=0}^{r} \sum_{1 \le i_1 < \cdots < i_k \le r} \mu(p_{i_1} \cdots p_{i_k}) = \sum_{k=0}^{r} \binom{r}{k} (-1)^k = 0.$$

Лемма доказана.

Лемма 3.3. *Функции $f(n)$ и $h(n)$, определенные на множестве целых положительных чисел, удовлетворяют равенству*

$$f(n) = \sum_{m \mid n} h(m) \quad \textit{при всех } n \in \mathbb{N} \tag{3.18}$$

тогда и только тогда, когда

$$h(n) = \sum_{m \mid n} \mu\Big(\frac{n}{m}\Big) f(m) \quad \textit{при всех } n \in \mathbb{N}. \tag{3.19}$$

Доказательство. Покажем, что из (3.18) следует (3.19). Для этого прежде всего заметим, что

$$\sum_{m \mid n} \mu\Big(\frac{n}{m}\Big) f(m) = \sum_{m \mid n} \mu(m) f\Big(\frac{n}{m}\Big),$$

так как суммы, стоящие в обеих частях равенства, отличаются только порядком следования слагаемых. Затем в правую часть последнего равенства вместо $f(m)$ подставим правую часть равенства (3.18). Меняя в получившейся двойной сумме порядок суммирования и применяя лемму 3.2, получим следующую цепочку равенств:

$$\begin{aligned} \sum_{m \mid n} \mu(m) f\Big(\frac{n}{m}\Big) &= \sum_{m \mid n} \mu(m) \sum_{k \mid \frac{n}{m}} h(k) = \\ &= \sum_{m \mid n} \sum_{k \mid \frac{n}{m}} \mu(m) h(k) = \sum_{km \mid n} \mu(m) h(k) = \\ &= \sum_{k \mid n} \sum_{m \mid \frac{n}{k}} \mu(m) h(k) = \sum_{k \mid n} h(k) \sum_{m \mid \frac{n}{k}} \mu(m) = h(n). \end{aligned}$$

Таким образом, справедливость равенства (3.19) установлена. Обратное утверждение доказывается аналогично. Лемма доказана.

Теорема 3.2. *Для числа $P(n)$ неприводимых многочленов степени n справедливо равенство*

$$P(n) = \frac{1}{n} \sum_{m \mid n} \mu\Big(\frac{n}{m}\Big) p^m.$$

Доказательство. Из леммы 3.1 следует, что равенство (3.18) леммы 3.3 справедливо при $f(n) = p^n$ и $h(n) = nP(n)$ для всех натуральных n. Поэтому утверждение теоремы следует непосредственно из леммы 3.3. Теорема доказана.

3.3 Производящие функции множеств

При доказательстве леммы 3.1 каждому неприводимому многочлену степени m была поставлена в соответствие функция x^m. Затем при помощи этого соответствия и формулы (3.14), порождающей множество всех многочленов над $\mathbb{Z}_p$, было получено уравнение (3.16), которое включало искомые величины $P(m)$ и из которого в конце концов была получена явная формула для $P(m)$. Аналогичным способом могут быть решены многие задачи. В основе этого способа лежит понятие производящей функции множества.

Рассмотрим множество A и функцию $w : A \to \{0, 1, \dots\}$, принимающую на элементах этого множества целые неотрицательные значения. Функция w называется весовой функцией (весом) на множестве A. Производящей функцией множества A относительно весовой функции w называется сумма

$$F_A^w(x) = \sum_{\alpha \in A} x^{w(a)}.$$

Пусть, например, множество A состоит из всех двоичных последовательностей конечной длины, а значение $w(\alpha)$ весовой функции на последовательности α равно длине этой последовательности. Тогда

$$F_A^w(x) = \sum_{n=1}^{\infty} 2^n x^n = \frac{2x}{1-2x}.$$

При доказательстве леммы 3.1 в качестве веса многочлена использовалась его степень, а производящей функцией множества многочленов была функция $\frac{1}{1-px}$.

В приводимых далее теоремах устанавливаются связи между операциями над множествами и операциями над производящими функциями этих множеств.

Теорема 3.3. *Пусть w — весовая функция на множестве N, A и B — непересекающиеся подмножества множества N. Тогда*

$$F_{A\cup B}^w(x) = F_A^w(x) + F_B^w(x).$$

Доказательство. Непосредственно из определения производящей функции следует, что

$$F_{A\cup B}^w(x) = \sum_{\sigma \in A\cup B} x^{w(\sigma)} = \sum_{\sigma \in A} x^{w(\sigma)} + \sum_{\sigma \in \mathcal{B}} x^{w(\sigma)} = F_A^w(x) + F_B^w(x).$$

Теорема доказана.

Теорема 3.4. *Пусть w_a, w_b и w — весовые функции на множествах A, B и прямом произведении $A \times B$. Если для всех (α, β) из $A \times B$ справедливо равенство*

$$w((\alpha, \beta)) = w_a(\alpha) + w_b(\beta),$$

то

$$F_{A\times B}^w(x) = F_A^w(x) \cdot F_B^w(x).$$

Доказательство. Непосредственно из определения производящей функции следует, что

$$F_{A\times B}^w(x) = \sum_{(\alpha,\beta)\in A\times B} x^{w((\alpha,\beta))} = \sum_{\alpha \in A} x^{w_a(\alpha)} \cdot \sum_{\beta \in B} x^{w_b(\beta)} = F_A^w(x) \cdot F_B^w(x).$$

Теорема доказана.

Далее, как правило, будем опускать знак прямого произведения. При этом следует отметить, что формальное удаление знака прямого произведения в некоторых случаях может приводить к появлению объединения пересекающихся множеств, и, как следствие, к некорректному использованию теоремы 3.4. Например, в множестве последовательностей из единиц с весом, равным длине последовательности, рассмотрим подмножество, состоящее из последовательностей длины два, три и четыре. Операцию прямого произведения последовательностей естественным образом можно трактовать как их конкатенацию, т. е. $11 \times 11 = 1111$. В этом случае нетрудно видеть, что $F_{\{11\cup 111\cup 1111\}}(x) = x^2+x^3+x^4$. С другой стороны для рассматриваемого множества справедливо равенство $11\cup 111\cup 1111 = (1\cup 11)(1\cup 11)$, применяя к правой части которого теорему 3.4, получаем очевидно неправильную производящую функцию $F_{(1\cup 11)(1\cup 11)}(x) = (x + x^2)^2 = x^2 + 2x^3 + x^4$. В этой функции лишнее слагаемое x^3 возникает из-за того, что после раскрытия скобок и удаление знаков прямого произведения последовательность 111 получается двумя разными способами из двух прямых произведений 1×11 и 11×1.

Итерацией множества A называется множество $A^* = \bigcup_{n=0}^{\infty} A^n$, где под A^0 будем понимать пустое множество. Далее пустое множество будем обозначать символом λ и полагать, что его вес равен нулю. Справедлива следующая теорема о производящей функции итерации.

Теорема 3.5. *Пусть w — весовая функция на множестве A. Тогда*

$$F_{A^*}^w(x) = \frac{1}{1 - F_A^w(x)},$$

при условии, что функция $\frac{1}{1-F_A^w(x)}$ разлагается в ряд в нуле.

Доказательство. Так как $A^* = \bigcup_{n=0}^{\infty} A^n$, то из двух предыдущих теорем следует, что

$$F_{A^*}^w(x) = \sum_{n=0}^{\infty} F_{A^n}^w(x) = \sum_{n=0}^{\infty} (F_A^w(x))^n = \frac{1}{1 - F_A^w(x)}.$$

Теорема доказана.

Вместе с операциями объединения, прямого произведения и итерации будем также иногда использовать операцию разности множеств. Нетрудно показать, что если на множествах A и B определена одна и та же весовая функция w и $A \subseteq B$, то $F_{B\setminus A}^w(x) = F_B^w(x) - F_A^w(x)$.

Рассмотрим два примера использования приведенных выше теорем. Сначала найдем производящую функцию для множества $A = \cup_{n=1}^{\infty}\{0,1\}^n$ всех $(0,1)$-последовательностей конечной длины с весовой функцией, равной длине последовательности. Множество A можно задать различными формулами, например такими, как $A = (0\cup1)^* - \lambda$ и $A = (0\cup1)^*(0\cup1)$. Вычисляя производящие функции в соответствии с этими формулами, видим, что

$$F^w_{(0\cup1)^*-\lambda}(x) = \frac{1}{1-2x} - 1 = \frac{2x}{1-2x}, \quad F^w_{(0\cup1)^*(0\cup1)}(x) = \frac{1}{1-2x}\cdot 2x = \frac{2x}{1-2x},$$

т. е. разные формулы, задающие одно и то же множество, приводят к одинаковым производящим функциям.

Теперь найдем производящую функцию для множества C, состоящего из всех $(0,1)$-последовательностей конечной длины, не имеющих двух соседних нулей, и где как и ранее весовая функция последовательности равна ее длине. Каждую последовательность в C разобъем на три части. Первая часть — начало, состоит только из единиц, а если последовательность начинается с нуля, то будем считать, что длина первой части равна нулю. Третья часть — конец, состоит из нуля, если последовательность оканчивается нулем, а если последовательность оканчивается единицей, то считаем, что длина третьей части равна нулю. Вторая часть находится между первой и третьей и может быть единственным образом поделена на блоки, каждый из которых начинается нулем, после которого следует ненулевое число единиц. Так, например, в последовательности 1110110101 первая часть равна 111, вторая часть — 0110101, а длина третьей части равна нулю. Последовательность 10 состоит только из начала 1 и конца 0, длина второй части равна нулю. Нетрудно видеть, что множество C можно задать формулой $1^*(011^*)^*(0\cup\lambda)$, причем каждый элемент этого множества будет порождаться формулой единственным образом. Тогда

$$F^w_C(x) = F^w_{1^*}(x)F^w_{(011^*)^*}(x)F^w_{0\cup\lambda}(x) = \\ = \frac{1}{1-x}\left(1 - \frac{x^2}{1-x}\right)^{-1}(1+x) = \frac{1+x}{1-x-x^2}.$$

Теперь можно получить явную формулу для числа последовательностей из C длины n. Это можно сделать, разложив функцию $F^w_C(x)$ в ряд так, как это было сделано с производящими функциями рекуррентных последовательностей в доказательстве теоремы 3.1.

Пусть B — некоторое множество. Будем полагать, что каждый элемент этого множества можно каким-либо образом разбить на фрагменты. Среди фрагментов, на которые разбиваются элементы B, выделим подмножество и его элементы назовем s-фрагментами. Например, если B состоит из всех $(0,1)$-последовательностей конечной длины, то в качестве фрагментов можно рассматривать нули и единицы, а к s-фрагментам отнести нули.

Рассмотрим три множества A, B, C. Будем говорить, что C является композицией множеств A и B (обозначается $C = B \circ A$), если существует такое множество s-фрагментов элементов из B, что каждый элемент

множества C может быть единственным образом получен из некоторого элемента множества B заменой в этом элементе всех его s-фрагментов элементами множества A. Например, пусть $A = 0(0^*)$, B состоит из $(0,1)$-последовательностей конечной длины, не имеющих двух соседних нулей, множество s-фрагментов элементов из B состоит из единственного фрагмента "0". Так как любая $(0,1)$-последовательность единственным образом может быть получена из подходящей последовательности множества B, в которой каждый нуль заменен какой-либо последовательностью из нулей, то множество C всех конечных $(0,1)$-последовательностей является композицией множеств A и B, т. е. $C = B \circ A$.

Пусть $C = B \circ A$. На множестве B введем весовую функцию w_s так, что для элемента β из B значение $w_s(\beta)$ равно числу s-фрагментов в элементе β. Для рассмотренного выше примера $w_s(\beta)$ равна числу нулей в наборе β.

Теорема 3.6. *Пусть w и w' — весовые функции на множествах A и $B\circ A$. Если для любого элемента $\sigma \in B\circ A$, полученного из элемента $\beta \in B$ веса $w_s(\beta) = m$ и элементов $\alpha_1, \ldots, \alpha_m \in A$, справедливо равенство*

$$w'(\sigma) = w(\alpha_1) + \ldots + w(\alpha_m),$$

то

$$F^{w'}_{B\circ A}(x) = F^{w_s}_B\left(F^w_A(x)\right),$$

при условии, что указанная подстановка допустима.

Доказательство. Пусть $\sigma \in B\circ A$, $\alpha^m = (\alpha_1, \ldots, \alpha_m) \in A^m$ и $\beta \in B$, где $w_s(\beta) = m$. Используем равенство $\sigma = (\beta, \alpha^m)$ для того, чтобы показать, что элемент σ получен подстановкой элементов $\alpha_1, \ldots, \alpha_m$ в элемент β. Тогда

$$F^{w'}_{B\circ A}(x) = \sum_{\sigma\in B\circ A} x^{w'(\sigma)} = \sum_{m\ge0}\sum_{\substack{\sigma=(\beta,\alpha^m)\\ w_s(\beta)=m}} x^{w'(\sigma)} = \\ = \sum_{m\ge0}\sum_{\substack{\sigma=(\beta,\alpha^m)\\ w_s(\beta)=m}} x^{w(\alpha_1)+\cdots+w(\alpha_m)} = \\ = \sum_{m\ge0}\sum_{w_s(\beta)=m}\Big(\sum_{\alpha\in A} x^{w(\alpha)}\Big)^m = \\ = \sum_{m\ge0}\sum_{w_s(\beta)=m}\Big(\sum_{\alpha\in A} x^{w(\alpha)}\Big)^{w_s(\beta)} = \\ = \sum_{\beta\in B}\Big(\sum_{\alpha\in A} x^{w(\alpha)}\Big)^{w_s(\beta)} = F^{w_s}_B\Big(F^w_A(x)\Big).$$

Теорема доказана.

Воспользуемся теоремой 3.6 для решения известной задачи о числе формул. Последовательность из левых и правых скобок называется формулой

тогда и только тогда, когда она удовлетворяет следующему индуктивному определению:

(i) последовательность ( ) является формулой;

(ii) если f — формула, то (f) — формула;

(iii) если f_1 и f_2 — формулы, то $f_1 f_2$ — формула.

Нетрудно видеть, что каждая формула содержит одинаковое число левых и правых скобок. Длиной формулы называется число входящих в эту формулу левых (или правых) скобок.

Найдем число формул длины $k \geq 1$. На множестве A, состоящем из всех формул, в том числе и пустой, не содержащей ни одной скобки, введем весовую функцию w, равную длине формулы. Пусть $F_A^w(x)$ — производящая функция множества A. Рассмотрим множество B, состоящее из всех формул, каждая из которых представима в виде (f), где f — формула. Так например, формула (( ) ( )) принадлежит B, а формула ( )(( ) ( )) не принадлежит B. Нетрудно видеть, что функция $xF_A^w(x)$ будет производящей функцией множества B. Далее рассмотрим множество C, которое состоит из пустой формулы и всех формул вида ( ) ... ( ), в которых после каждой левой скобки стоит правая скобка. Очевидно, что C является итерацией формулы ( ), и поэтому $F_C^w(x) = \frac{1}{1-x}$. Заметим, что любую формулу из A можно единственным образом получить из некоторой формулы из множества C, если в этой формуле каждую пару скобок ( ) заменить на подходящую формулу из B. Например формула ( )(( ) ( )) $\in A$ получается из формулы ( )( ) $\in C$, если в этой формуле первую пару скобок заменить на формулу ( ) $\in B$, а вторую — на (( ) ( )) $\in B$. Следовательно, $A = C \circ B$, где s-фрагментами в формулах из C являются пары скобок ( ). Легко видеть, что в данном случае производящие функции $F_C^w(x)$ и $F_C^{w_s}(x)$ совпадают, а равенство

$$w(f) = w(f_1) + \ldots + w(f_m)$$

справедливо для любой формулы $f \in A$, которая получается из какой-либо формулы из C заменой в этой формуле пар скобок ( ) на формулы $f_1, \ldots, f_m \in B$. Таким образом, можно воспользоваться теоремой 3.6. Применяя эту теорему, получим уравнение

$$F_A^w(x) = F_C^{w_s}(F_B^w(x)) = \frac{1}{1 - xF_A^w(x)},$$

которое легко преобразуется в квадратное относительно производящей функции $F_A^w(x)$ уравнение

$$x(F_A^w(x))^2 - F_A^w(x) + 1 = 0.$$

Так как функция $F_A^w(x)$ должна раскладываться в ряд в окрестности нуля, то из двух корней квадратного уравнения оставим корень, удовлетворяющий этому условию. Следовательно,

$$F_A^w(x) = \frac{1 - \sqrt{1-4x}}{2x}. \tag{3.20}$$

Функцию $\sqrt{1-4x}$ разложим в ряд при помощи формулы бинома Ньютона:

$$\begin{aligned}
\sqrt{1-4x} &= 1 + \sum_{k=1}^{\infty} \frac{\frac{1}{2}(\frac{1}{2}-1)\cdots(\frac{1}{2}-k+1)}{k!}(-4)^k x^k = \\
&= 1 + \sum_{k=1}^{\infty} \frac{\frac{1}{2}(-\frac{1}{2})(-\frac{3}{2})\cdots(-\frac{2k-3}{2})}{k!}(-4)^k x^k = 1 - \sum_{k=1}^{\infty} \frac{2^k(2k-3)!!}{k!}x^k = \\
&= 1 - \sum_{k=1}^{\infty} \frac{2^k(2k-3)!!(k-1)!}{k!(k-1)!}x^k = 1 - \sum_{k=1}^{\infty} \frac{2(2k-3)!!(2k-2)!!}{k!(k-1)!}x^k = \\
&= 1 - \sum_{k=1}^{\infty} \frac{2(2k-2)!}{k(k-1)!(k-1)!}x^k = 1 - \sum_{k=1}^{\infty} \frac{2}{k}\binom{2k-2}{k-1}x^k.
\end{aligned}$$

Подставляя полученный ряд в формулу (3.20) после простых преобразований получаем разложение в ряд функции $F_A^w(x)$:

$$F_A^w(x) = \sum_{k=0}^{\infty} \frac{1}{k+1}\binom{2k}{k}x^k.$$

Таким образом, существует ровно $\frac{1}{k+1}\binom{2k}{k}$ формул длины k[1)].

Теоремы 3.3–3.6 сформулированы и доказаны для производящих функций одной переменной, однако нетрудно видеть, что каждая из этих теорем справедлива и для производящих функций большего числа переменных, если ее условия выполняются относительно какой-нибудь из переменных.

В качестве примера использования производящих функций многих переменных рассмотрим новое решение задачи о числе $(0,1)$-последовательностей, в которых нет двух соседних нулей. Пусть $A = 1(1^*)$, B — множество $(0,1)$-последовательностей, в которых нет двух стоящих рядом одинаковых символов, C — множество $(0,1)$-последовательностей, в которых нет двух стоящих рядом нулей. Пусть далее множество s-фрагментов B состоит из единственного фрагмента "1". В этом случае нетрудно видеть, что $C = B \circ A$. На множестве A в качестве весовой функции w_a будем использовать длину последовательности, а на множествах B и C определим весовую функцию w так, что $w(\beta) = (i,j)$, где i равно числу нулей, j — числу единиц в β. Функции w и w_a таковы, что вторая компонента w_2 функции w и функция w_a удовлетворяют условию теоремы 3.6, т. е. для любого σ, полученного заменой s-фрагментов в последовательности из B на последовательности $\alpha_1, \ldots, \alpha_m$ из A, справедливо равенство

$$w_2(\sigma) = w_a(\alpha_1) + \ldots + w_a(\alpha_m).$$

Поэтому нетрудно видеть, что

$$F_{B\circ A}^w(x,y) = F_B^w\left(x, F_{\mathcal{A}}^{w_a}(y)\right).$$

[1)] Ранее эта задача была решена методом траекторий.

Так как $B = (\lambda \cup 1)(01)^*(\lambda \cup 0)$, то $F_B^w(x,y) = \frac{(1+x)(1+y)}{1-xy}$. Подставляя в эту формулу вместо переменной y производящую функцию множества A, найдем производящую функцию множества $B \circ A$:

$$F_{B\circ A}^w(x,y) = \frac{(1+x)(1+\frac{y}{1-y})}{1-\frac{xy}{1-y}} = \frac{1+x}{1-y-xy}. \qquad (3.21)$$

Теперь, так как нас интересует только количество последовательностей определенной длины и не важно сколько в этих последовательностях нулей и единиц, нам для получения производящей функции множества C достаточно в формуле (3.21) отождествить переменные. В результате получим формулу $F_C^w(x,x) = \frac{1+x}{1-x-x^2}$, которая, как нетрудно видеть, совпадает с полученной ранее другим способом.

3.4 Задачи

3.1. Найти a_n, если $a_n = 2a_{n-1} + \sin\frac{\pi n}{2}$ и $a_1 = 1$.

3.2. Найти a_n, если $a_n = a_{n-1} - a_{n-2}$ и $a_1 = 1$, $a_2 = 3$.

3.3. Найти a_n, если $a_n = 3a_{n-1} - 2a_{n-2} + 2^n$ и $a_1 = a_2 = 1$.

3.4. Найти a_n, если $a_n = 4a_{n-1} - 4a_{n-2} + n$ и $a_1 = a_2 = 1$.

3.5. Найти a_n, если $a_n = -2a_{n-1} + 8a_{n-2} + (-2)^n$ и $a_1 = 0$, $a_2 = 1$.

3.6. Найти a_n, если $a_n = -(a_{n-1} + a_{n-2} + \cdots + a_{n-16})$ и $a_1 = \cdots = a_{16} = 1$.

3.7. Найти x_n, y_n, если $x_n = x_{n-1} + y_{n-1} + 1$, $y_n = 3x_{n-1} + y_{n-1} - 1$ и $x_1 = y_1 = 1$.

3.8. Найти асимптотику числа целых неотрицательных решений уравнения $\sum_{i=1}^m i \cdot x_i = n$ при $n \to \infty$.

3.9. Пусть $A(x), B(x)$ — многочлены с целыми коэффициентами, $\deg B(x) = k$, $f(x) = \sum_{n=0}^\infty f_n x^n = A(x)/B(x)$. Показать, что начиная с некоторого n_0 коэффициенты f_n удовлетворяют рекуррентному соотношению $f_n = \sum_{i=1}^k c_k f_{n-k}$ с постоянными c_k.

3.10. Найти число последовательностей из нулей и единиц длины n, в которых каждый блок из единиц имеет четную длину.

3.11. Найти число последовательностей из 0 и 1 длины n, в которых длина каждого блока из 0 кратна трем.

3.12. Найти число последовательностей из 0, 1, 2 и 3 длины n, в которых каждый блок из 1 имеет четную длину, а длины блоков из 2 и 3 кратны трем.

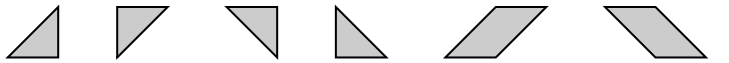
3.13. Сколькими способами можно замостить прямоугольник высоты 1 и длины n, используя плитки высоты 1 следующих видов:

В этой и следующих задачах плитки вращать нельзя.

3.14. Сколькими способами можно замостить прямоугольник высоты 1 и длины n, используя плитки высоты 1 следующих видов:

3.15. Сколькими способами можно замостить прямоугольник высоты 1 и длины n, используя плитки высоты 1 следующих видов:

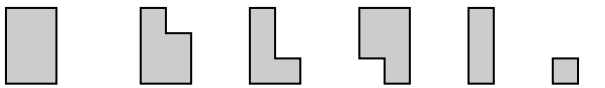
3.16. Сколькими способами можно замостить прямоугольник высоты 1 и длины n, используя плитки высоты 1 следующих видов:



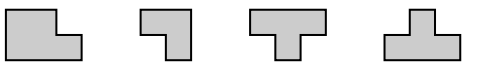
3.17. Сколькими способами можно замостить прямоугольник высоты 2 и длины n, используя плитки следующих видов:

3.18. Сколькими способами можно замостить прямоугольник высоты 1 и длины n, используя плитки высоты 1 следующих видов:

3.19. Сколькими способами можно замостить прямоугольник высоты 3 и длины n, используя плитки следующих видов:



3.20. Сколькими способами можно замостить прямоугольник высоты 2 и длины n, используя плитки следующих видов:



3.21. Найти количество n-разрядных десятичных чисел, в которых нет двух стоящих рядом четных цифр.

3.22. Найти количество n-разрядных десятичных чисел, в которых после цифры 2 не стоит цифра 5.

3.23. Найти число последовательностей длины 23 из 0, 1 и 2, в которых все максимальные подпоследовательности из единиц имеют нечетную длину.

3.24. Функция $E(x) = \sum_{n=1}^{\infty} a_n \frac{x^n}{n!}$ называется *экспоненциальной производящей функцией* последовательности $\{a_n\}$. Найти экспоненциальную производящую функцию $E_k^2(x)$ для чисел Стирлинга второго рода $\left\{ {m \atop k} \right\}$ при данном k.

3.25. Найти экспоненциальную производящую функцию $E_k^1(x)$ для чисел Стирлинга первого рода $\left[{m \atop k} \right]$ при данном k.

Лекция 4

Теорема Пойа

Рассмотрим класс задач следующего типа. Пусть дано множество D с определенным на его элементах отношением эквивалентности. Требуется найти число классов эквивалентности этого множества, которые удовлетворяют некоторым дополнительным условиям. Классическим примером задачи такого типа является задачи о числе способов, которыми можно раскрасить грани трехмерного кубика. В этой задаче два раскрашенных кубика считаются эквивалентными, если один кубик можно повернуть так, что после поворота одинаково ориентированные грани кубиков будут окрашены одним и тем же цветом. Дополнительным условием в этой задаче может быть условие на число граней, покрашенных определенным цветом. Например, задача может состоять в том, чтобы найти число различных кубиков, у каждого из которых три белых и три черных грани. Техника решения подобных задач была разработана венгерским математиком Д. Пойа в конце тридцатых годов двадцатого века для решения ряда перечислительных задач теории графов. Похожие результаты были опубликованы Дж. Редфилдом (см. [26]) в 1927 г. в менее удобной для использования форме. Видимо по этой причине работа Редфилда была забыта и не оказала влияния на дальнейшие исследования в области перечислительной комбинаторики.

4.1 Действие группы на множестве

Будем говорить, что группа G действует на множестве D, если каждому элементу g группы G поставлено в соответствие взаимнооднозначное отображение $\varphi(g)$ множества D в себя так, что $\varphi(g_1 g_2) = \varphi(g_1)\varphi(g_2)$ для любых g_1 и g_2 из G. Иначе говоря группа G действует на множестве D, если определен гомоморфизм φ группы G в множество взаимнооднозначных отображений множества D в себя. Рассматривая далее действие группы G на множестве D, будем опускать символ гомоморфизма и будем рассматривать элементы группы G непосредственно как преобразования множества D: результат действия элемента g группы G на элементе d множества D будем обозначать через $g(d)$ или gd.

Пусть группа G действует на множестве D. Стабилизатором элемента d_0

из D называется множество $\mathrm{St}(d_0) = \{g \in G \,|\, g(d_0) = d_0\}$. Орбитой элемента d_0 из D называется множество $\mathrm{Or}(d_0) = \{d \in D \,|\, d = g(d_0), \text{где } g \in G\}$, число элементов орбиты называется ее длиной. Известно, что стабилизатор любого элемента d является подгруппой в группе G.

Лемма 4.1. *Пусть конечная группа G действует на конечном множестве D. Тогда для любого d из D*

$$|\mathrm{Or}(d)| \cdot |\mathrm{St}(d)| = |G|.$$

Доказательство. Покажем, что длина орбиты произвольного элемента d из D равна числу смежных классов группы G по подгруппе $\mathrm{St}(d)$. Для этого достаточно показать, что элементы из одного смежного класса переводят d в один и тот же элемент множества D, а элементы из разных смежных классов — в разные элементы множества D.

Если g_1 и g_2 лежат в одном и том же смежном классе группы G по подгруппе $\mathrm{St}(d)$, то $g_2 = g_1 s$, где $s \in \mathrm{St}(d)$. Поэтому

$$g_2(d) = g_1 s(d) = g_1(s(d)) = g_1(d),$$

т. е. элементы из одного и того же смежного класса группы G по подгруппе $\mathrm{St}(d)$ отображают d в один и тот же элемент множества D. Теперь покажем, что элементы из разных смежных классов группы G по подгруппе $\mathrm{St}(d)$ отображают d в разные элементы множества D. Допустим, что $g_1(d) = g_2(d)$, тогда

$$d = g_1^{-1} g_1(d) = g_1^{-1}(g_1(d)) = g_1^{-1}(g_2(d)) = g_1^{-1} g_2(d),$$

и, следовательно, $g_1^{-1} g_2 \in \mathrm{St}(d)$. Но тогда в $\mathrm{St}(d)$ найдется такой элемент s, что $g_2 = g_1 s$, и поэтому g_1 и g_2 лежат в одном и том же смежном классе группы G по подгруппе $\mathrm{St}(d)$.

Так как число смежных классов группы G по подгруппе $\mathrm{St}(d)$ равно $|G|/|\mathrm{St}(d)|$, а длина орбиты элемента d равна числу смежных классов группы G по подгруппе $\mathrm{St}(d)$, то $|\mathrm{Or}(d)| \cdot |\mathrm{St}(d)| = |G|$. Лемма доказана.

4.2 Лемма Бернсайда

Пусть группа G действует на конечном множестве D. Элементы d_1 и d_2 из D назовем эквивалентными, если $d_1 = g d_2$ для некоторого g из G. Нетрудно видеть, что множество D под действием группы G распадается на классы эквивалентности, состоящие из попарно эквивалентных элементов. Число классов эквивалентности можно найти при помощи следующей леммы Бернсайда.

Лемма 4.2. *Пусть группа G действует на конечном множестве D. Тогда для N — числа классов эквивалентности, порождаемых на множестве D действием группы G, справедливо равенство*

$$N = \frac{1}{|G|} \sum_{g \in G} \psi(g),$$

где $\psi(g)$ — число элементов d множества D таких, что $gd = d$.

Доказательство. Введем функцию $\psi(d, g)$ так, что

$$\psi(d, g) = \begin{cases} 1, & \text{если } gd = d; \\ 0, & \text{если } gd \neq d. \end{cases}$$

Тогда, учитывая лемму 4.1,

$$\sum_{g \in G} \psi(g) = \sum_{g \in G} \sum_{d \in D} \psi(d, g) = \sum_{d \in D} \sum_{g \in G} \psi(d, g) =$$
$$= \sum_{Or_i} \sum_{d \in Or_i} \sum_{g \in G} \psi(d, g) = \sum_{Or_i} \sum_{d \in Or_i} |\mathrm{St}(d)| = \sum_{Or_i} |G| = N|G|.$$

Разделив левую и правую части получившегося равенства на $|G|$, получаем требуемую формулу для N. Лемма доказана.

4.3 Цикловой индекс

Если элемент g группы G, действуя на множестве D, разбивает это множество на k_i орбит длины i, где $i = 1, \ldots, s$, то *цикловым индексом I_g элемента g* называется одночлен $z_1^{k_1} z_2^{k_2} \cdots z_s^{k_s}$. *Цикловым индексом группы G* называется многочлен

$$P_G(z_1, z_2, \ldots, z_k, \ldots) = \frac{1}{|G|} \sum_{g \in G} I_g(z_1, z_2, \ldots, z_k, \ldots).$$

Рассмотрим введенные определения на примере упомянутой выше задачи о раскраске граней трехмерного кубика. В этой задаче группа G вращений кубика действует на множестве его граней. Прежде всего заметим, что произвольное ребро с упорядоченными вершинами при помощи вращений можно преобразовать в любое другое ребро и при этом образ ребра может быть ориентирован двумя разными способами. Так как образ произвольного ребра однозначно определяет выполненное вращение и всего в кубике содержится двенадцать ребер, то очевидно, что группа G состоит из 24 элементов. Нетрудно видеть (см. рис. 4.1), что вращать кубик можно вокруг

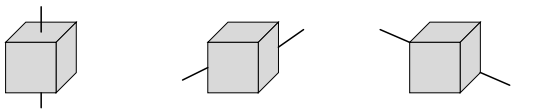


Рис. 4.1

осей, проходящих через центры противоположных граней на углы 90°, 180° и 270° — всего $3 \cdot 3 = 9$ различных вращений, вокруг осей, проходящих через центры противоположных ребер на углы 180° — всего 6 различных

вращений, и вокруг осей, проходящих через противоположные вершины на углы 120° и 240° — всего $4 \cdot 2 = 8$ различных вращений. Общее число перечисленных вращений с учетом тождественного вращения, оставляющего все грани на своих местах, равно 24 и, следовательно, вращений, отличных от указанных выше, в группе G нет. Перенумеруем грани кубика так, как это показано на рис. 4.2, где номер закрашенной грани указан рядом с

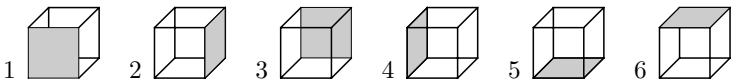


Рис. 4.2

соответствующим экземпляром кубика. Перечислим элементы группы G и выпишем их индексы.

1. Нейтральный элемент e. Этот элемент оставляет все грани кубика на месте, и, поэтому, $I_e = z_1^6$.

2. Вращения на 90° и 270° вокруг осей, проходящих через центры противоположных граней. Рассмотрим вращение вокруг оси, проходящей через центры пятой и шестой граней на 90° по часовой стрелке. Нетрудно видеть, что это вращение оставляет пятую и шестую грани на месте и последовательно переводит первые четыре грани друг в друга, т. е. на множестве граней возникают два цикла длины 1 — (5) и (6), и один цикл длины 4 — (1234). Остальные вращения действуют аналогичным образом, и, поэтому, для каждого из 6 подобных вращений цикловой индекс равен $z_1^2 z_4$.

3. Вращения на 180° вокруг осей, проходящих через центры противоположных граней. Вновь нетрудно видеть, что такое вращение вокруг оси, проходящей через центры пятой и шестой граней, оставляет эти грани на месте и меняет местами первую и третью, и вторую и четвертую грани, т. е. возникают два цикла длины 1 — (5) и (6), и два цикла длины 2 — (13) и (24). Следовательно, для каждого из 3 подобных вращений цикловой индекс равен $z_1^2 z_2^2$.

4. Вращения на 180° вокруг осей, проходящих через центры противоположных ребер. Вращения вокруг оси, изображенной в средней части рис. 4.1, меняют местами первую и четвертую, вторую и третью, и пятую и шестую грани, т. е. возникают три цикла длины 2 — (14), (23) и (56). Таким образом для каждого из 6 подобных вращений цикловой индекс равен z_2^3.

5. Вращения на 120° и 240° вокруг осей, проходящих через противоположные вершины. Нетрудно видеть, что вращение на 120° вокруг оси, изображенной в правой части рис. 4.1, переводят друг в друга первую, четвертую и шестую грани, а также вторую, третью, и пятую грани, т. е. возникают два цикла длины 3 — (146) и (253). Таким образом для каждого из 8 подобных вращений цикловой индекс равен z_3^2.

Объединяя полученные результаты, находим цикловой индекс

$$P_G = \frac{1}{24}\Big(z_1^6 + 6z_1^2 z_4 + 3z_1^2 z_2^2 + 6z_2^3 + 8z_3^2\Big), \tag{4.1}$$

группы вращений трехмерного кубика при действии этой группы на множестве граней кубика.

4.4 Функции и их классы эквивалентности

Пусть D и R — конечные множества, K — коммутативное кольцо, $w : R \to K$ — весовая функция на множестве R. Для каждой функции f из множества $\mathcal{F} = \{f : D \to R\}$ определим ее вес w положив $w(f) = \prod_{d \in D} w(f(d))$. Функции f_1 и f_2 назовем эквивалентными, если найдется такой элемент g группы G, что $f_1(d) = f_2(gd)$ для каждого $d \in D$. Очевидно, что множество $\mathcal{F}$ распадается на классы эквивалентности $F_1, \ldots, F_k$, и так как веса эквивалентных функций из одного класса совпадают, то можно говорить о весе класса эквивалентности функций из $\mathcal{F}$. Вес класса F обозначим через $W(F)$.

Введенные определения снова рассмотрим на примере задачи о раскраске граней кубика. Грани будем раскрашивать в два цвета — черный и белый. В этом случае множество D состоит из шести граней кубика, а множество R — из черного и белого цветов. Кубик, грани которого покрашены в черный и белый цвета, будем рассматривать как функцию из D в R, которая ставит в соответствие каждой грани ее цвет. Группой G, действующей на множестве D, будет рассмотренная выше группа вращений кубика, а две функции будут эквивалентными, если соответствующие им раскрашенные кубики можно преобразовать друг в друга при помощи вращений из группы G. Например, нетрудно видеть, что все кубики с одной черной и пятью белыми гранями эквивалентны друг другу. В качестве кольца K возьмем кольцо многочленов от переменных x и y с целыми коэффициентами, при этом белому цвету припишем вес x, а черному — y. Таким образом, весом раскрашенного кубика, и весом соответствующей ему функции, будет одночлен шестой степени от переменных x и y. Если нас интересует число различных кубиков с тремя черными и тремя белыми гранями, нам надо найти число классов эквивалентности, вес которых равен x^3y^3. Сделать это можно при помощи теоремы Пойа.

4.5 Основная теорема

Сформулируем и докажем теорему Пойа о сумме весов классов эквивалентности F функций из D в R, полагая, что на множестве D действует группа G, а на множестве R определена весовая функция w со значениями в коммутативном кольце K.

Теорема 4.1. *Сумма весов классов эквивалентности равна*

$$\sum_F W(F) = P_G\Big(\sum_{r \in R} w(r), \sum_{r \in R} (w(r))^2, \ldots, \sum_{r \in R} (w(r))^k, \ldots\Big),$$

где P_G — цикловой индекс группы.

Доказательство. Рассмотрим элемент g группы G, под действием которого множество D распадается на k_1 циклов длины единица, k_2 циклов длины два, и т. д. вплоть до k_s циклов длины s. Без ограничения общности будем полагать, что циклы длины единица формируются первыми k_1 элементами множества D, циклы длины два формируются следующими $2k_2$ элементами так, что каждый цикл имеет вид $(d_i d_{i+1})$, и т. д. Последние sk_s элементов множества D образуют k_s циклов вида $(d_j d_{j+1} \dots d_{j+s-1})$. Нетрудно видеть, что вектор значений $v(f)$ любой функции f, которая определена на D, принимает значения в R, и которая под действием элемента g переходит в себя, выглядит следующим образом. На первых k_1 местах произвольным образом располагаются любые элементы множества R. Следующие $2k_2$ мест заполнены k_2 парами одинаковых элементов из R. Это необходимо и достаточно для выполнения равенства $f(d) = f(g(d))$ при $d \in \{d_{k_1+1}, \dots, d_{k_1+2k_2}\}$. Следующие $3k_3$ мест заполнены k_3 тройками одинаковых элементов из R и т. д. Наконец последние sk_s разрядов вектора $v(f)$ представляют собой последовательность из k_s блоков длины s, каждый из которых состоит из одинаковых элементов. Нетрудно видеть, что все такие векторы можно получить, раскрыв скобки в произведении

$$\Big(\sum_{r\in R} r\Big)^{k_1}\Big(\sum_{r\in R} rr\Big)^{k_2}\dots\Big(\sum_{r\in R} \underbrace{r\dots r}_{s\text{ раз}}\Big)^{k_s}, \tag{4.2}$$

полагая при этом, что умножение в (4.2) некоммутативно. Таким образом,

$$\sum_{f=g(f)} v(f) = \Big(\sum_{r\in R} r\Big)^{k_1}\Big(\sum_{r\in R} rr\Big)^{k_2}\dots\Big(\sum_{r\in R} \underbrace{r\dots r}_{s\text{ раз}}\Big)^{k_s}. \tag{4.3}$$

Например, если $s = 2$, $k_1 = 1$, $k_2 = 2$ и $R = \{x, y\}$, то

$$\begin{aligned}(x+y)(xx+yy)(xx+yy) =& x\,xx\,xx + x\,xx\,yy + x\,yy\,xx + x\,yy\,yy + \\ &+ y\,xx\,xx + y\,xx\,yy + y\,yy\,xx + y\,yy\,yy.\end{aligned} \tag{4.4}$$

Теперь вычислим сумму весов всех функций, которые под действием элемента g переходят в себя. Для этого в (4.3) заменим каждый элемент r его весом $w(r)$. Тогда, в силу мультипликативности функции w,

$$\sum_{f=g(f)} w(v(f)) = \Big(\sum_{r\in R} w(r)\Big)^{k_1}\Big(\sum_{r\in R} (w(r))^2\Big)^{k_2}\dots\Big(\sum_{r\in R} (w(r))^s\Big)^{k_s}. \tag{4.5}$$

Допустим, что веса функций, оставляемых элементом g на месте, принимают значения $w_1, \dots, w_m$. Тогда сумму весов рассматриваемых функций можно представить в виде

$$\sum_{w_i} w_i \psi_i(g), \tag{4.6}$$

где $\psi_i(g)$ — число функций веса w_i. Сумма именно такого вида получится после открытия скобок в правой части равенства (4.5) и последующего

приведения подобных слагаемых. Продолжая рассмотренный выше пример, положим $w(x) = t$, $w(y) = s$ и вычислим сумму весов всех функций, векторы значений которых перечислены в правой части равенства (4.4). Нетрудно видеть, что

$$\begin{aligned}&(w(x)+w(y))(w(xx)+w(yy))(w(xx)+w(yy)) = \\ &\qquad = (t+s)(t^2+s^2)(t^2+s^2) = t^5 + t^4 s + 2t^3 s^2 + 2t^2 s^3 + t s^4 + s^5,\end{aligned}$$

где коэффициент при одночлене $t^i s^j$ равен количеству тех функций, вес которых равен $t^i s^j$.

Возвращаясь к равенству (4.5), заметим, что произведение в его правой части есть ничто иное, как индекс элемента g, в который вместо переменных z_k подставлены суммы $\sum_{r\in R}(w(r))^k$. Следовательно, сумма весов всех функций, которые под действием элемента g переходят в себя, равна

$$I_g\Big(\sum_{r\in R} w(r), \sum_{r\in R}(w(r))^2, \dots, \sum_{r\in R}(w(r))^s\Big). \tag{4.7}$$

Вычислив сумму величин (4.7) по всем элементам группы G и разделив результат на порядок группы G, видим, что в силу (4.6)

$$\begin{aligned}&P_G\Big(\sum_{r\in R} w(r), \sum_{r\in R}(w(r))^2, \dots, \sum_{r\in R}(w(r))^k, \dots\Big) = \\ &\qquad = \frac{1}{|G|}\sum_{g\in G}\sum_{w_i} w_i\psi_i(g) = \sum_{w_i} w_i\left(\frac{1}{|G|}\sum_{g\in G}\psi_i(g)\right).\end{aligned}$$

Из леммы Бернсайда следует, что при фиксированном значении веса w сумма $\frac{1}{|G|}\sum_{g\in G}\psi_i(g)$ равна числу классов эквивалентности, возникающих на множестве функций веса w_i в результате действия группы G на множестве D. Следовательно, левая часть последнего равенства равна сумме весов всех классов эквивалентности. Теорема доказана.

Воспользуемся доказанной теоремой и найдем число различных двухцветных кубиков с тремя черными и тремя белыми гранями. Для этого в найденный выше (см. (4.1)) цикловой индекс группы вращений трехмерного кубика

$$P_G = \frac{1}{24}\Big(z_1^6 + 6z_1^2 z_4 + 3z_1^2 z_2^2 + 6z_2^3 + 8z_3^2\Big)$$

вместо каждой переменной z_i подставим сумму $x^i + y^i$. В результате получим многочлен

$$\begin{aligned}\frac{1}{24}\Big((x+y)^6 &+ 6(x+y)^2(x^4+y^4) + \\ &+ 3(x+y)^2(x^2+y^2)^2 + 6(x^2+y^2)^3 + 8(x^3+y^3)^2\Big).\end{aligned} \tag{4.8}$$

Теперь найдем коэффициент, который будет стоять при одночлене x^3y^3 после раскрытия скобок и приведения подобных слагаемых. В первое слагаемое $(x+y)^6$ одночлен x^3y^3 входит с коэффициентом 20, во втором и четвертом слагаемых такого одночлена нет, так как они содержат только четные степени переменных x и y, в третье слагаемое одночлен x^3y^3 входит с коэффициентом 12, в пятое — с коэффициентом 16. Следовательно, коэффициент при x^3y^3 в (4.8) равен

$$\frac{1}{24}\Big(20+0+12+0+16\Big)=2.$$

Таким образом, грани трехмерного кубика можно раскрасить двумя различными способами при условии, что три грани будут окрашены белым цветом, а три — черным.

В случае, когда веса функций не важны, и надо найти только число классов эквивалентности, можно воспользоваться следующим простым следствием теоремы Пойа.

Следствие 4.1. *Число классов эквивалентности равно*

$$P_G\Big(|R|,|R|,\dots,|R|,\dots\Big),$$

где P_G — цикловой индекс группы.

Найдем число различных способов, которыми можно покрасить грани кубика тремя цветами. Для этого в цикловой индекс

$$P_G=\frac{1}{24}\Big(z_1^6+6z_1^2z_4+3z_1^2z_2^2+6z_2^3+8z_3^2\Big)$$

группы вращений вместо каждой переменной z_i подставим тройку. В результате получим

$$P_G(3,3,3,3)=\frac{1}{24}\Big(3^6+6\cdot3^3+3\cdot3^4+6\cdot3^3+8\cdot3^2\Big)=57.$$

4.6 Задачи

4.1. Функции $f: Z_{17}^*\to Z_2$ и $g: Z_{17}^*\to Z_2$ называются эквивалентными, если существует такое целое k, что для любого $x\in Z_{17}^*$ и $y=2^kx \pmod{17}$ справедливо равенство $f(x)=g(y)$. Найти количество неэквивалентных функций из Z_{17}^* в Z_2, принимающих значение 1 ровно на 7 аргументах.

4.2. Функции $f: \{1,2,\dots,9\}\to Z_3$ и $g: \{1,2,\dots,9\}\to Z_3$ называются эквивалентными, если существует такой элемент $\pi\in\langle(1357),(2468)\rangle$, что для любого $x\in\{1,2,\dots,9\}$ и $y=\pi(x)$ справедливо равенство $f(x)=g(y)$. Найти количество неэквивалентных функций из $\{1,2,\dots,9\}$ в Z_3, любая из которых принимает каждое значение ровно на трех аргументах.

4.3. Сколько существует различных бус из 18 бусин, если шесть бусин окрашены красной краской, а двенадцать бусин — синей краской.

4.4. Сколькими различными способами можно окрасить вершины правильного 21-угольника тремя красками, если каждой краской окрасить семь вершин.

4.5. Сколькими различными способами можно окрасить вершины правильного 2^n-угольника двумя красками.

4.6. Сколькими различными способами можно раскрасить ребра трехмерного куба черной и белой красками, если пять ребер покрашены черной краской, а семь — белой.

4.7. Сколькими различными способами можно раскрасить ребра трехмерного куба тремя красками.

4.8. Сколькими различными способами можно раскрасить вершины трехмерного куба черной и белой красками, если пять вершин покрашены черной краской, а три — белой.

4.9. Сколько существует бус из 14 бусин, если шесть бусин окрашены красной краской, пять бусин — синей краской, три — зеленой.

4.10. Сколькими различными способами можно раскрасить вершины правильного тетраэдра двумя красками.

4.11. Сколькими различными способами можно раскрасить ребра правильного тетраэдра тремя красками.

4.12. Стоимость красного камня равна одной единице, зеленого — двум, желтого — трем. Сколько существует различных ожерелий из 15 камней, если стоимость каждого ожерелья равна 30 единицам.

4.13. Сколькими различными способами можно раскрасить вершины и грани трехмерного куба, если вершины покрашены двумя разными красками, грани — тремя.

4.14. Сколькими различными способами можно раскрасить вершины и ребра трехмерного куба, если вершины покрашены тремя разными красками, шесть ребер покрашены синей краской и шесть красной.

4.15. Сколькими различными способами можно пометить грани трехмерного куба натуральными числами, если сумма меток всех граней равна n.

4.16. Сколькими различными способами можно пометить вершины правильного 10-угольника натуральными числами, если сумма меток всех вершин равна n, причем сумма четных меток равна сумме нечетных меток.

4.17. В каждую клетку прямоугольной таблица из восьми строк и восьми столбцов вписано целое неотрицательное число. Две таблицы называются эквивалентными, если одна из таблиц перейдет в другую после вращения вокруг центра симметрии. Найти число неэквивалентных таблиц, если сумма чисел во всех клетках равна n, причем сумма четных чисел равна сумме нечетных.

Лекция 5

Графы

5.1 Основные понятия и определения

Неориентированным графом (или графом) называется пара (V, E), где $V = \{v_1, v_2, \dots\}$ — множество вершин, $E = \{e_1, e_2, \dots\}$ — множество ребер, в котором каждый элемент e_k является неупорядоченной парой $\{v_i, v_j\}$. Пара (V, E), в которой множество E состоит из упорядоченных пар (v_i, v_j), называется ориентированным графом, а элементы из E — *дугами*, или ориентированными ребрами. Вершины v_i и v_j, составляющие ребро или дугу, называются концевыми вершинами ребра или дуги, а про ребро и дугу говорят, что они соединяют свои концевые вершины.

Обычно граф изображают на плоскости в виде множества точек, соответствующих вершинам, и множества линий, которые соединяют вершины и соответствуют ребрам. При изображении ориентированных графов линии снабжаются стрелками, указывающими ориентацию дуги, т. е. порядок вершин в паре. Про дугу (v_i, v_j) говорят, что она выходит из вершины v_i

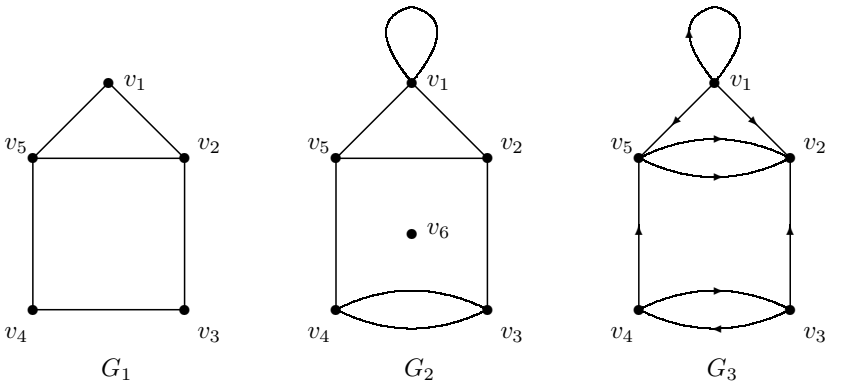


Рис. 5.1

и входит в вершину v_j. Вершины v_i и v_j называются смежными, если они соединены ребром или дугой e. При этом говорят, что e инцидентно вершинам v_i и v_j, а вершины v_i и v_j инцидентны e. Если два ребра (две дуги) инцидентны одной и той же вершине, то эти ребра (дуги) называются *смежными*. Если в E содержится $k \geq 2$ экземпляров пары $\{v_i, v_j\}$, то ребро $\{v_i, v_j\}$ называется кратным с кратностью k. Аналогичным образом в ориентированных графах определяются кратные дуги. В изображенном на рис. 5.1 графе G_2 присутствует ребро $\{v_3, v_3\}$ кратности два, а в изображенном там же графе G_3 — кратная дуга (v_5, v_2) и две противоположно ориентированных дуги (v_3, v_4) и (v_4, v_3). Ребро и дуга называются петлями, если их концевые вершины совпадают. Далее, если присутствие в графе петель и кратных ребер специально не оговаривается, то под графом будем понимать граф без петель и кратных ребер, такой как граф G_1 на рис. 5.1. Аналогичное замечание справедливо и для ориентированных графов.

Степенью $s(v)$ вершины v графа G называется число ребер инцидентных этой вершине. *Полустепенью захода* $s^+(v)$ вершины v ориентированного графа G называется число дуг, входящих в v, а *полустепенью исхода* $s^-(v)$ — число дуг, выходящих из этой вершины. Степенью $s(v)$ вершины v в ориентированном графе называется сумма $s^+(v) + s^-(v)$. Вершину нулевой степени будем называть изолированной. Такой вершиной является вершина v_6 графа G_2 на рис. 5.1. Граф называется *регулярным*, если степени всех его вершин равны.

Маршрутом[1] в графе G (ориентированном графе G) назовем чередующуюся последовательность вершин и ребер $\alpha = v_1 e_1 v_2 e_2 \dots v_k$, в которой при $i = 1, \dots, k-1$ каждая пара вершин v_i, v_{i+1} связана ребром e_i (дугой e_i), при этом будем говорить, что маршрут α связывает вершины v_1 и v_k и проходит через вершины $v_1, v_2, \dots, v_k$ и ребра (дуги) $e_1, e_2, \dots, e_{k-1}$. *Длиной маршрута* называется число ребер, через которые маршрут проходит. Маршрут, состоящий из несовпадающих ребер, будем называть *цепью*, а состоящий из несовпадающих вершин — *простой цепью*. Маршрут или цепь α в ориентированном графе называются *ориентированными*, если все дуги e_i в них направлены от v_i к v_{i+1}. Замкнутый маршрут, т. е. маршрут с совпадающими первой и последней вершинами, в неориентированном графе называется *циклом*. Замкнутый ориентированный маршрут в ориентированном графе называется *контуром*.

Подграфом графа $G = (V, E)$ называется такой граф $G_1 = (V_1, E_1)$, что $V_1 \subseteq V$ и $E_1 \subseteq E$. Так, например, граф G_1 на рис. 5.1 можно рассматривать в качестве подграфа графа G_2. Подграф G_1 графа G называется *остовным*, если он содержит все вершины графа G.

Граф называется *связным*, если для любых двух его вершин найдется цепь, связывающая эти вершины. Компонентой связности графа G называется такой его связный подграф $G_1 = (V_1, E_1)$, что никакой другой подграф $G_2 = (V_2, E_2)$, для которого $V_1 \subset V_2$, не является связным, и E_1 содержит все ребра, обе концевых вершины которых лежат в V_1. Графы G_1 и G_3 на рис. 5.1 связные, а граф G_2 нет, так как его вершина v_6 является изолированной. Ориентированный граф называется *сильно связным*, если две

[1] Часто также используется термин путь.

его любые вершины связаны ориентированной цепью. Граф G_3 не является сильно связным — в этом графе, например, нет ориентированных цепей, связывающих v_1 и v_3.

Графы $G_1 = (V_1, E_1)$ и $G_2 = (V_2, E_2)$ называются *изоморфными*, если существует взаимно однозначное отображение $f : V_1 \to V_2$, при котором $\{v_i, v_j\} \in E_1$ тогда и только тогда, когда $\{f(v_i), f(v_j)\} \in E_2$. Нетрудно видеть, что графы G_1 и G_2 на рис. 5.2 изоморфные. Отображение f, задающее их изоморфизм, можно определить следующим образом:

$$f(v_1) = u_1, \quad f(v_2) = u_3, \quad f(v_3) = u_5, \quad f(v_4) = u_2, \quad f(v_5) = u_4, \quad f(v_6) = u_6.$$

Также нетрудно видеть, что графы G_1 и G_3 изоморфными не являются. Это следует например из того, что степени всех вершин графа G_1 равны трем, а в графе G_3 есть вершины, степени которых равны двум и четырем. Изоморфизм ориентированных графов определяется также как и изомор-

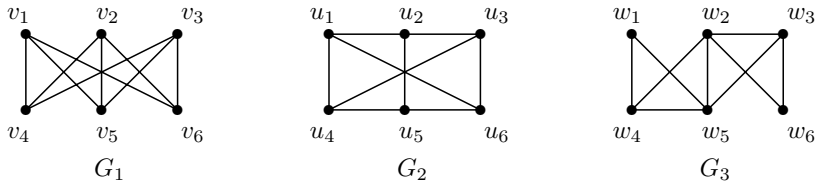


Рис. 5.2

физм неориентированных — в определении последнего достаточно заменить ребра $\{v_i, v_j\}$ на дуги (v_i, v_j).

Для задания графов часто используют матрицы смежности и инцидентности. *Матрицей смежности* графа $G = (V, E)$ называется матрица из $|V|$ строк и $|V|$ столбцов, в которой на пересечении i-й строки и j-го столбца стоит единица, если вершины v_i и v_j смежные, и нуль, если эти вершины несмежные. В матрице смежности ориентированного графа на пересечении i-й строки и j-го столбца стоит единица, если существует дуга, связывающая вершины v_i и v_j и направленная от v_i к v_j. В графах с кратными ребрами единицы в матрицах смежности заменяются на кратности ребер. Так, например, матрицы смежности графов G_1 и G_3, изображенных на рис. 5.1, имеют следующий вид:

$$M_{G_1} = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{pmatrix}, \qquad M_{G_3} = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 2 & 0 & 0 & 0 \end{pmatrix}.$$

Матрицей инцидентности графа $G = (V, E)$ называется матрица из $|V|$ строк и $|E|$ столбцов, в которой на пересечении i-й строки и j-го столбца стоит единица, если вершина v_i инцидентна ребру e_j.

Важное подмножество множества всех графов образуют деревья, которые естественным образом появляются в различных задачах теории графов. *Деревом* называется связный граф без циклов. Нетрудно видеть, что в любом дереве число вершин p на единицу больше числа ребер q, и что в дереве любые две вершины связаны ровно одной простой цепью. Дерево с выделенной вершиной называется *корневым*, а сама выделенная вершина — *корнем* этого дерева. Если в дереве степень вершины v равна единице, то v называется висячей вершиной, если при этом v не является корнем, то она также называется *листом*. Длина максимальной цепи от корня дере-

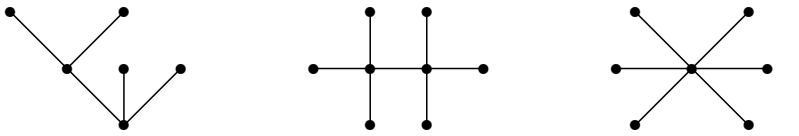


Рис. 5.3

ва до его листа называется *высотой дерева*. Примеры различных деревьев приведены на рис. 5.3.

Другой важный класс графов образуют двудольные графы. Граф $G = (V, E)$ называется *двудольным*, если множество вершин V является объединением таких непересекающихся подмножеств V_1 и V_2, что концевые вершины любого ребра из E лежат в разных подмножествах V_i. Подмножества V_1 и V_2 называются долями графа G. Графы G_1 и G_2 на рис. 5.2 двудольные, а граф G_3 двудольным не является. Отметим, что каждое дерево будет двудольным графом, поэтому все графы на рис. 5.3 двудольные. Двудольный граф G с долями V_1 и V_2 и множеством ребер E будем иногда обозначать тройкой (V_1, V_2, E).

Граф G называется *полным* графом на n вершинах, если любые две вершины этого графа смежные. Полный граф на n вершинах обозначается через K_n. Двудольный граф G с долями из n и m вершин называется *полным*, если любые две его вершины, принадлежащие разным, долям смежные. Полный двудольный граф G с долями из n и m вершин обозначается через $K_{n,m}$.

5.2 Теорема Холла

Произвольное подмножество попарно несмежных ребер графа G называется его *паросочетанием*. Паросочетание называется *совершенным*, если каждая вершина графа инцидентна какому-нибудь ребру паросочетания. Граф G_1 на рис. 5.2 обладает совершенным паросочетанием, например таким — $\{\{v_1, v_4\}, \{v_2, v_5\}, \{v_3, v_6\}\}$, в то время как ни в одном из графов на рис. 5.3 совершенного паросочетания нет.

Будем говорить, что паросочетание M двудольного графа $G = (V_1, V_2, E)$ покрывает долю V_i, если каждая вершина этой доли инцидентна какому-

нибудь ребру этого паросочетания. Пусть A — подмножество доли V_1 графа $G = (V_1, V_2, E)$. Тенью $S(A)$ множества A называется подмножество вершин доли V_2, каждая из которых смежна хотя бы с одной вершиной из A. В следующей теореме Ф. Холла устанавливается необходимое и достаточное условие существования паросочетания.

Теорема 5.1. *Конечный двудольный граф G с долями V_1 и V_2 обладает паросочетанием, покрывающим долю V_1, тогда и только тогда, когда $|S(X)| \geq |X|$ для любого непустого $X \subseteq V_1$.*

Доказательство. Необходимость условия теоремы очевидна, так как если тень некоторого подмножества X состоит из меньшего чем само X числа вершин, то между вершинами из X и вершинами из его тени нельзя установить взаимно однозначное соответствие.

Установим достаточность условия теоремы. Сделаем это индукцией по числу вершин n в V_1. В основание индукции положим очевидный случай $n = 1$. Допустим, что достаточность доказана при всех значениях $|V_1|$ не превосходящих $m - 1$. Пусть $G = (V_1, V_2, E)$ — произвольный двудольный граф, в котором доля V_1 состоит из m вершин. Рассмотрим два возможных случая.

(i) Для любого непустого подмножества $A \subset V_1$ справедливо строгое неравенство $|S(A)| > |A|$. В этом случае в V_1 и V_2 выберем по одной вершине, которые связаны ребром e, и удалим эти вершины со всеми инцидентными им ребрами из G. В новом графе $G' = (V_1', V_2', E')$ доля V_1' состоит из $m - 1$ вершин и для каждого непустого подмножества $X \subseteq V_1'$ справедливо неравенство $|S(X)| \geq |X|$, так как после удаления ребра e размер тени любого подмножества из V_1 уменьшился не более чем на единицу. По предположению индукции в G' существует паросочетание M', покрывающее V_1'. Тогда объединение M' с ребром e будет требуемым паросочетанием в G.

(ii) В V_1 найдется такое подмножество A, что $|S(A)| = |A|$. По предположению индукции в двудольном графе $G_A = (A, S(A), E_A)$, долями которого являются множества A и $S(A)$, а ребрами — все ребра графа G инцидентные вершинам этих множеств, существует совершенное паросочетание M_A. Рассмотрим двудольный граф $G' = (V_1', V_2', E')$, где $V_1' = V_1 \setminus A$, $V_2' = V_2 \setminus S(A)$, E' — множество ребер инцидентных одновременно V_1' и V_2'. Если для какого-нибудь $B \subset V_1$ выполняется неравенство $|S(B)| < |B|$, то в силу того, что A и B не пересекаются, справедливы неравенства

$$|S(A \cup B)| \leq |S(A)| + |S(B)| = |A| + |S(B)| < |A| + |B| = |A \cup B|,$$

которые очевидно противоречат условию теоремы. Следовательно, $|S(B)| \geq |B|$ для любого непустого $B \subseteq V_1'$. Поэтому в силу индуктивного предположения в графе G' существует паросочетание M', покрывающее долю V_1'. Легко видеть, что объединение M_A и M' будет паросочетанием в графе G, и это паросочетание будет покрывать V_1. Теорема доказана.

5.3 Теорема Менгера

Множество вершин $N \subseteq V$ графа $G = (V, E)$ называется *vw-рассекающим множеством*, если любая цепь, связывающая в G несмежные вершины v и w, проходит хотя бы через одну вершину из N и N не содержит v и w. Будем говорить, что две цепи, связывающие в графе G несмежные вершины v и w, не пересекаются, если они не имеют общих вершин кроме v и w. В следующей теореме, доказанной Менгером в 1927 г., устанавливается связь между числом вершинно непересекающихся цепей, соединяющих две различные вершины графа, и числом вершин в соответствующем рассекающем множестве.

Теорема 5.2. *Максимальное число вершинно непересекающихся цепей, соединяющих две различные несмежные вершины v и w связного графа, равно минимальному числу вершин в vw-рассекающем множестве.*

Доказательство. Так как каждая цепь, соединяющая вершины v и w, пересекает vw-рассекающее множество, то число вершинно непересекающихся цепей не больше числа вершин в любом vw-рассекающем множестве. Докажем обратное неравенство. Сделаем это индукцией по числу вершин графа. Допустим, что в любом связном графе из $m - 1$ или менее вершин максимальное число вершинно непересекающихся цепей, соединяющих две различные несмежные вершины v и w этого графа, не меньше минимального числа вершин в vw-рассекающем множестве. Пусть граф G содержит m вершин, среди которых есть несмежные вершины v и w, и его минимальное vw-рассекающее множество состоит из k вершин. Без ограничения общности будем считать, что каждая вершина принадлежит по крайней мере одному k-элементному vw-рассекающему множеству. В противном случае после удаления из G любой неудовлетворяющей этому условию вершины минимальное vw-рассекающее множество в новом графе по-прежнему будет состоять из k вершин, и в силу предположения индукции в нем найдется k вершинно непересекающихся цепей из v в w. Теперь заметим, что всякое vw-рассекающее множество A делит вершины графа (все кроме v, w и вершин из самого A) на два непересекающихся подмножества A_v и A_w. Первое подмножество состоит из всех тех вершин, которые связаны с v цепями, не проходящими через A, второе — из тех вершин, которые связаны с w цепями, также не проходящими через A. Рассмотрим два возможных случая.

1) Допустим, что в G существует минимальное vw-рассекающее множество A, которое состоит из k вершин $a_1, \ldots, a_k$ и такое, что оба множества A_v и A_w содержат хотя бы по одной вершине. В этом случае граф G преобразуем в два новых графа G_v и G_w следующим образом. Граф G_v получим, удалив из G все вершины, принадлежащие A_v, и соединив v ребрами со всеми вершинами из A. Аналогичным образом получим G_w — удалим все вершины, принадлежащие A_w, и соединим w ребрами со всеми вершинами из A. Пример таких построений приведен на рис. 5.4. Каждый из графов G_v и G_w состоит не более чем из $m - 1$ вершин, и в каждом из этих графов минимальное число вершин в vw-рассекающем множестве равно k.

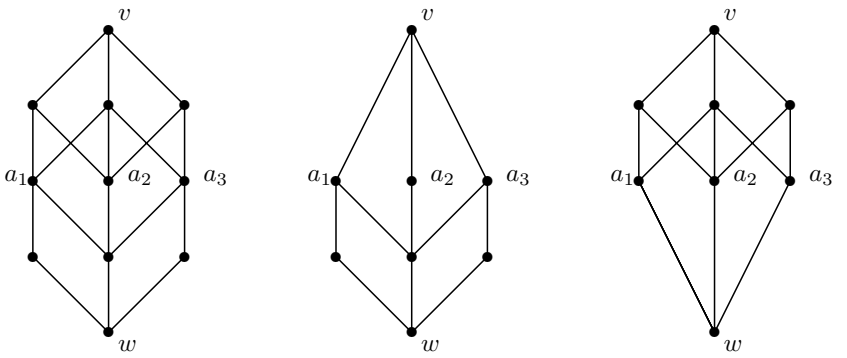


Рис. 5.4

Следовательно, в силу индуктивного предположения в каждом из этих графов максимальное число вершинно непересекающихся цепей, соединяющих вершины v и w, равно k. Пусть $\alpha_1, \ldots, \alpha_k$ и $\beta_1, \ldots, \beta_k$ — такие цепи в G_v и G_w, причем цепи α_i и β_i для всех $i = 1, \ldots, k$ проходят через вершину a_i множества A. Нетрудно видеть, что если для каждого $i = 1, \ldots, k$ соединить лежащие в G фрагменты цепей α_i и β_i в одну цепь γ_i, то получим k вершинно непересекающихся цепей, которые в G соединяют вершины v и w. Таким образом в G существует не менее k вершинно непересекающихся цепей, связывающих v и w .

2) Теперь рассмотрим случай, когда в графе G каждое минимальное vw-рассекающее множество A состоит из k вершин и при этом хотя бы одно из соответствующих множеств A_v или A_w пусто. Нетрудно видеть, что в

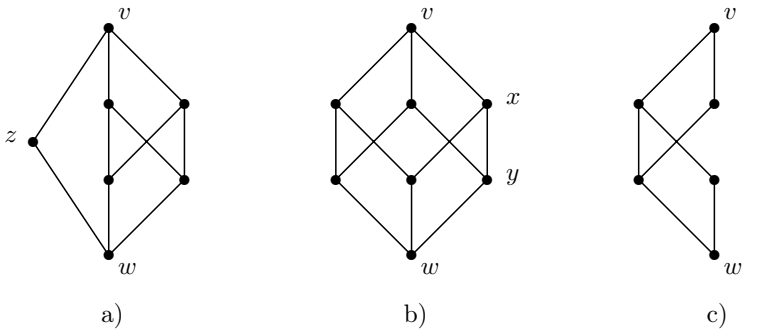


Рис. 5.5

этом случае каждая вершина графа, которая принадлежит рассекающему множеству A, смежна с v, если пусто A_v, или смежна с w, если пусто A_w. Примеры таких графов изображены на рис. 5.5. Если предположить, что существует вершина $z \in A$, где A_v пусто и z несмежна с v, то любая цепь,

связывающая z с v, должна проходить через какую-либо другую вершину из A, и, следовательно, A не будет минимальным vw-рассекающим множеством.

Предположим, что в G найдется вершина z смежная одновременно с v и w. Пример такого графа показан на рис. 5.5 a. Удалим вершину z вместе с инцидентными ей ребрами из G. В новом графе G' минимальное vw-рассекающее множество будет состоять из $k-1$ вершин и, в силу индуктивного предположения, в G' вершины v и w связывают $k-1$ вершинно непересекающихся цепей $\alpha_1, \ldots, \alpha_{k-1}$. Очевидно, что все эти цепи присутствуют и в графе G и не пересекаются с цепью $\alpha_k = (vzw)$. Следовательно, число вершинно непересекающихся цепей, связывающих в G вершины v и w, не меньше k.

Осталось рассмотреть графы, в которых каждая вершина смежна только с v или только с w. Так как каждая вершина в таком графе G принадлежит хотя бы одному рассекающему множеству, то в этом случае в графе обязательно найдется цепь длины три, которая связывает v и w. Выберем в G такую цепь длины три, полагая, что она проходит через вершины x и y (см. рис. 5.5 b). Удалим из G вершины x и y вместе со всеми инцидентными им ребрами. В получившемся графе G' любое минимальное vw-рассекающее множество A' будет состоять из $k-1$ вершин (см. рис. 5.5 c). Если найдется A', состоящее из $k-2$ вершин, то очевидно, что $A = A' \cup \{x, y\}$ будет минимальным vw-рассекающим множеством в G, причем таким, что ни A_v, ни A_w не будут пустыми. Это следует из того, что в графе G степени вершин v и w не меньше k, и, поэтому, по крайней мере одна вершина смежная с v и одна вершина смежная с w не попадут в A и будут разделены этим множеством, т. е. A_v и A_w должны содержать не менее чем по одной вершине. Таким образом, в силу индуктивного предположения в G' вершины v и w связывают $k-1$ вершинно непересекающихся цепей $\alpha_1, \ldots, \alpha_{k-1}$, все эти цепи присутствуют в графе G и не пересекаются с цепью $\alpha_k = (vxyw)$. Следовательно, число вершинно непересекающихся цепей, связывающих в G вершины v и w, не меньше k. Теорема доказана.

Множество ребер $M \subseteq E$ графа $G = (V, E)$ называется *vw-разделяющим множеством*, если любая цепь, связывающая в G вершины v и w, проходит хотя бы через одно ребро из M. Следующую теорему, реберный вариант теоремы 5.2, приведем без доказательства, которое почти дословно совпадает с доказательством теоремы 5.2.

Теорема 5.3. *Максимальное число реберно непересекающихся цепей, соединяющих две различные вершины v и w связного графа, равно минимальному числу ребер в vw-разделяющем множестве.*

Теорема 5.3 естественным образом легко переносится на ориентированные графы. Имеет место следующее утверждение.

Теорема 5.4. *Максимальное число реберно непересекающихся ориентированных цепей, соединяющих две различные вершины v и w связного графа, равно минимальному числу дуг в vw-разделяющем множестве.*

Теорема 5.4 называется реберной теоремой Менгера для ориентированных графов. С этой теоремой тесно связана теорема Форда–Фалкерсона о максимальном потоке и минимальном разрезе. Эта теорема, доказанная Л. Фордом и Д. Фалкерсоном в 1956 г., стала основой развитой и имеющей различные приложения теории потоков в сетях. Прежде чем сформулировать теорему Форда–Фалкерсона дадим необходимые определения.

Сетью N называется ориентированный граф G с двумя выделенными вершинами v и w, на дугах которого определена неотрицательная функция φ. Вершины v и w называются полюсами сети, функция φ — пропускной способностью, а ее значение $\varphi(e)$ на дуге e — пропускной способностью дуги e. На дугах графа G определим функцию ψ так, что $0 \le \psi(e) \le \varphi(e)$ для каждой дуги e. Полустепенью исхода $\rho^-(u)$ вершины u назовем сумму значений функции ψ, взятую по всем дугам, выходящим из u, а полустепенью захода $\rho^+(u)$ вершины u — сумму значений функции ψ, взятую по всем дугам, входящим в u. Функция ψ называется *потоком* из полюса v в полюс w через сеть N, если $\rho^-(v) - \rho^+(v) \ge 0$, $\rho^+(w) - \rho^-(w) \ge 0$ и $\rho^+(u) = \rho^-(u)$ для любой вершины u отличной от v и w. Вершина v называется источником сети, а вершина w — стоком сети. Разность $\rho^+(u) - \rho^-(u)$ называется *потоком через вершину* u. Величиной потока ψ называется разность $\rho^-(v) - \rho^+(v)$. Так как поток через любую вершину отличную от v и w равен нулю, то нетрудно видеть, что $\rho^-(v) - \rho^+(v) = \rho^+(w) - \rho^-(w)$. Если величина потока ψ из v в w через сеть N не меньше величины любого другого потока из v в w через эту сеть, то поток ψ называется *максимальным*. *Разрезом* C сети (G, v, w, φ) называется любое vw-разделяющее множество ориентированного графа G. Сумма пропускных способностей входящих в разрез дуг называется пропускной способностью разреза. Если пропускная способность разреза C не больше пропускной способности любого другого разреза, то разрез C называется *минимальным*.

Сеть с определенными на ее дугах пропускными способностями изображена в левой части рис. 5.6. В этой сети все дуги ориентированы от полюса v к полюсу w. Нетрудно видеть, что пропускная способность минимальной

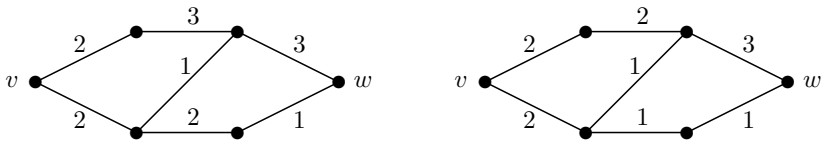


Рис. 5.6

ного разреза сети равна четырем. Эта же сеть изображена в правой части рис. 5.6, где рядом с каждой дугой указана величина определенного на этой сети потока. Величина этого потока через сеть равна четырем, т. е. равна величине минимального разреза. Так как величина потока не может быть больше пропускной способности разреза, то очевидно, что указанный поток будет максимальным.

Теперь сформулируем теорему Форда–Фалкерсона. Доказательство этой

теоремы опустим, так как его можно легко получить при помощи теоремы 5.4.

Теорема 5.5. *В любой сети величина любого максимального потока равна пропускной способности любого минимального разреза.*

5.4 Теорема Дилуорса

Сформулируем и докажем теорему Дилуорса, в которой устанавливается важное экстремальное свойство частично упорядоченных множеств. Частично упорядоченные множества удобно представлять в виде бесконтурных ориентированных графов, ставя в соответствие элементам множества вершины графов, а отношение частичного порядка между двумя элементами изображать посредством дуги, связывающей соответствующие вершины.

Теорема 5.6. *Минимальное число непересекающихся цепей, которыми можно покрыть конечное частично упорядоченное множество, равно максимальному числу попарно несравнимых элементов этого множества.*

Доказательство. Так как никакая цепь не может содержать двух несравнимых элементов, то очевидно, что m — минимальное число непересекающихся цепей, которыми можно покрыть частично упорядоченное множество, не меньше чем M — максимальное число попарно несравнимых элементов этого множества.

Покажем, что $m \le M$. Сделаем это индукцией по числу элементов частично упорядоченного множества. Допустим, что $m \le M$ для любого множества из не более чем n элементов. Пусть частично упорядоченное множество D состоит из $n+1$ элементов, среди которых найдется k попарно несравнимых. Рассмотрим множество $D \setminus d$, где d — произвольный элемент D. Если максимальное число попарно несравнимых элементов в $D \setminus d$ равно $k-1$, то по предположению индукции $D \setminus d$ будет объединением $k-1$ непересекающихся цепей. Добавив к этим цепям новую цепь из единственного элемента d, получим требуемое покрытие D, состоящее из k цепей.

Теперь рассмотрим случай, когда множество $D \setminus d$ содержит k попарно несравнимых элементов. В силу предположения индукции, множество $D \setminus d$ можно покрыть k цепями $D_1, \ldots, D_k$. Каждую цепь D_i разобьем на три части A_i, B_i и C_i так, что $a \succ d$ для всякого a из A_i, $c \prec d$ для всякого c из C_i, и все элементы B_i несравнимы с d. Если какое-либо B_i пусто, то элемент d можно добавить в цепь D_i, и в этом случае D будет покрыто k цепями. Поэтому далее полагаем, что все множества B_i не пусты.

Покажем, что среди множеств $F_i = ((\cup_{j=1}^{k} A_j) \cup (\cup_{j=1}^{k} B_j)) \setminus A_i$ найдется множество F_p, в котором число попарно несравнимых элементов не превосходит $k-1$. Допустим это не так и в каждом множестве F_i найдется k попарно несравнимых элементов $\{s_1^i, \ldots, s_k^i\}$, где $s_j^i \in D_j$. Очевидно, что $s_i^i \in B_i$, так как элементы цепи A_i не входят в F_i. Положим $s_i = \min(s_i^1, \ldots, s_i^k)$. Так как $s_i^i \in B_i$, то и $s_i \in B_i$. Покажем, что элементы s_i и s_j несравнимы для любых неравных i и j. Действительно, если $s_i = s_i^m$ и $s_i \prec s_j$, то $s_i^m \prec s_j \preceq s_j^m$,

что очевидно противоречит попарной несравнимости элементов $s_1^m, \ldots, s_k^m$. Таким образом элементы $s_1, \ldots, s_k$ попарно несравнимы, а так как каждый $s_i \in B_i$, то они несравнимы и с элементом d. Следовательно, предположив, что в каждом множестве F_i есть k попарно несравнимых элементов, мы пришли к заключению, что в D есть $k+1$ попарно несравнимый элемент, т. е. пришли к противоречию.

Совершенно аналогично можно показать, что среди множеств $H_i = ((\cup_{j=1}^k C_j) \cup (\cup_{j=1}^k B_j)) \setminus C_i$ найдется множество H_q, в котором число попарно несравнимых элементов не превосходит $k-1$.

Пусть T — максимальное подмножество попарно несравнимых элементов из $D \setminus (A_p \cup C_q \cup \{d\})$. Так как $a \succ d \succ c$ для любых a из A_i и c из C_j, то T целиком лежит либо в F_p, либо в H_q. Поэтому по предположению индукции множество $D \setminus (A_p \cup C_q \cup \{d\})$ есть объединение не более чем $k-1$ цепей. Так как $A_p \cup \{d\} \cup C_q$ является цепью, то множество D можно покрыть не более чем k непересекающимися цепями. Теорема доказана.

5.5 Раскраски вершин

Будем говорить, что вершины графа G правильным образом раскрашены k цветами, если каждой вершине приписан один из k цветов так, что всем смежным вершинам графа G приписаны разные цвета. *Хроматическим числом* $\chi(G)$ графа G называется минимально возможное число цветов k, которым можно правильным образом раскрасить его вершины.

В задачах, связанных с раскрасками графов, часто используется операция перекрашивания вершины. Перекраска в графе G вершины $\boldsymbol{v}$, покрашенной цветом c_i, в цвет c_j осуществляется следующим образом. В графе G выделяется подграф G_{ij}, состоящий из вершин, покрашенных цветами c_i и c_j, и всех инцидентных этим вершинам ребер. Затем в компоненте связности, содержащей вершину $\boldsymbol{v}$, цвета c_i и c_j меняются местами. На рис. 5.7

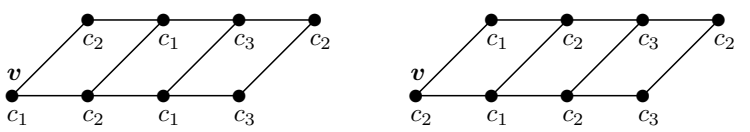


Рис. 5.7

изображен граф до и после перекраски вершины $\boldsymbol{v}$ в цвет c_2. В этом графе подграф G_{12} состоит из двух компонент связности. Одна компонента содержит пять левых вершин, в том числе и вершину $\boldsymbol{v}$, а вторая — одну изолированную вершину, покрашенную цветом c_2. После перекраски все вершины первой компоненты поменяли свои цвета, а единственная вершина второй компоненты сохранила свой цвет. Нетрудно видеть, что правильно раскрашенный граф после перекраски вершины остается правильно раскрашенным.

В следующей теореме, доказанной Р. Л. Бруксом в 1941 г., устанавливается верхняя оценка хроматического числа произвольного графа.

Теорема 5.7. *Если максимальная степень вершин связного графа G не превосходит n, $n \geq 3$ и G не является графом K_{n+1}, то $\chi(G) \leq n$.*

Доказательство. Воспользуемся индукцией по числу вершин графа. Допустим, что вершины любого графа, удовлетворяющего условиям теоремы и содержащего не более чем m вершин, можно раскрасить красками n цветов. Пусть граф G удовлетворяет условиям теоремы и содержит $m+1$ вершин, вершина $\boldsymbol{u}$ этого графа смежна с вершинами $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$, а граф G' получается из G после удаления $\boldsymbol{u}$ и всех инцидентных с $\boldsymbol{u}$ ребер. По предположению индукции вершины G' можно раскрасить n цветами. Будем полагать, что в графе G' каждая вершина $\boldsymbol{v}_i$ покрашена цветом c_i. Далее полагаем, что все (кроме $\boldsymbol{u}$) вершины графа G покрашены теми же цветами, что и аналогичные вершины G'. Если среди цветов $c_1, \ldots, c_n$ есть два одинаковых, то вершину $\boldsymbol{u}$ можно покрасить тем цветом, которого нет среди этих цветов (см. рис. 5.8). Поэтому далее считаем, что все c_j различны.

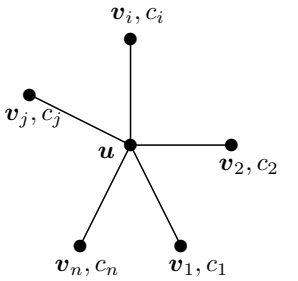


Рис. 5.8

Пусть G_{ij} — подграф графа G, образованный вершинами, покрашенными цветами c_i и c_j. Если вершины $\boldsymbol{v}_i$ и $\boldsymbol{v}_j$ лежат в разных компонентах связности графа G_{ij}, то вершину $\boldsymbol{v}_i$ можно перекрасить в цвет c_j, не изменяя цвета вершины $\boldsymbol{v}_j$. Очевидно, что после такой перекраски вершину $\boldsymbol{u}$ можно покрасить цветом c_i и в результате получить правильную раскраску графа G. Поэтому далее полагаем, что во всех подграфах G_{ij} вершины $\boldsymbol{v}_i$ и $\boldsymbol{v}_j$ лежат в одних и тех же компонентах связности. При этом возможны следующие четыре случая:

1) найдутся такие i и j, что граф G_{ij} не является простой цепью, связывающей $\boldsymbol{v}_i$ и $\boldsymbol{v}_j$;

2) при любых i и j граф G_{ij} является простой цепью, связывающей $\boldsymbol{v}_i$ и $\boldsymbol{v}_j$, и среди этих цепей найдутся две, имеющие общую неконцевую вершину;

3) при любых i и j граф G_{ij} является простой цепью, связывающей $\boldsymbol{v}_i$ и $\boldsymbol{v}_j$, никакие две цепи не имеют общих неконцевых вершин, и хотя бы одна из этих цепей не является ребром;

4) при всех i и j вершины $\boldsymbol{v}_i$ и $\boldsymbol{v}_j$ смежные.

Рассмотрим эти случаи.

1) Сначала будем полагать, что степень одной из вершин $\boldsymbol{v}_i$ или $\boldsymbol{v}_j$ больше единицы. Пусть такой вершиной будет $\boldsymbol{v}_j$. Тогда $\boldsymbol{v}_j$ смежна не менее чем с двумя вершинами, покрашенными цветом c_i (см. рис. 5.9). Так как $\boldsymbol{v}_j$ смежна не более чем с n вершинами, одна из которых не окрашена, а две окрашены одинаковым цветом, то, очевидно, что среди n цветов найдется цвет c, которым не окрашена ни одна из вершин смежных с $\boldsymbol{v}_j$. Поэтому можно перекрасить вершину $\boldsymbol{v}_j$ в цвет c, и после этого покрасить вершину

$\boldsymbol{u}$ цветом c_j. В результате получим правильную раскраску вершин графа G. Если степени вершин $\boldsymbol{v}_i$ и $\boldsymbol{v}_j$ равны единице, то в графе G_{ij} обязательно найдется вершина, степень которой не меньше трех. Пусть $\boldsymbol{v}$ — такая вер-

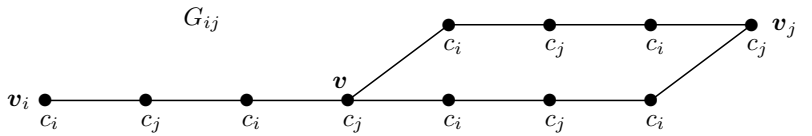


Рис. 5.9

шина, причем цепь, связывающая $\boldsymbol{v}$ и $\boldsymbol{v}_i$, проходит только через вершины степени два (см. рис. 5.9). Тогда $\boldsymbol{v}$ смежна не менее чем с тремя вершинами, покрашенными одинаковым цветом, например, c_i. Следовательно, как и в предыдущем случае, существует такой цвет c, что среди вершин смежных с $\boldsymbol{v}$, нет вершины, покрашенной цветом c. Поэтому в графе G можно перекрасить вершину $\boldsymbol{v}$ в цвет c. После этого в перекрашенном подграфе G_{ij} вершины $\boldsymbol{v}_i$ и $\boldsymbol{v}_j$ будут лежать в разных компонентах связности, и, следовательно, мы приходим к рассмотренной выше ситуации, в которой возможна правильная раскраска вершин G.

2) Допустим, что найдутся такие i, j и k, что графы G_{ij} и G_{jk} имеют кроме $\boldsymbol{v}_j$ еще хотя бы одну общую вершину $\boldsymbol{v}$. Подобная ситуация показана на рис. 5.10, где кривыми изображены цепи. Легко видеть, что вершина $\boldsymbol{v}$ смежна с двумя вершинами, покрашенными цветом c_i, и двумя вершинам, покрашенными цветом c_j. Следовательно, как и в предыдущем случае, существует такой цвет c, что среди вершин смежных с $\boldsymbol{v}$ нет вершины, покрашенной цветом c. Поэтому в графе G можно перекрасить вершину $\boldsymbol{v}$ в цвет c. После этого в перекрашенном подграфе G_{ij} вершины $\boldsymbol{v}_i$ и $\boldsymbol{v}_j$ будут лежать в разных компонентах связности, и, следовательно, мы снова приходим к рассмотренной выше ситуации, в которой возможна правильная раскраска вершин G.

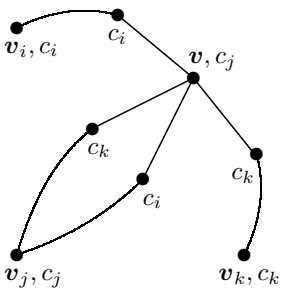


Рис. 5.10

3) Пусть теперь при любых i, j и k графы G_{ij} и G_{jk} не имеют кроме $\boldsymbol{v}_j$ ни одной общей вершины, а граф G_{st} кроме вершин $\boldsymbol{v}_s$ и $\boldsymbol{v}_t$ содержит вершину $\boldsymbol{v}$, которая смежна с вершиной $\boldsymbol{v}_s$ и покрашена цветом c_t. Пусть c_r – цвет отличный от c_s и c_t, а $\boldsymbol{v}_r$ — вершина, покрашенная этим цветом. Фрагмент такого графа изображен на рис. 5.11. В графе G_{sr} поменяем местами цвета c_s и c_r. Нетрудно видеть, что цвета вершин, не принадлежащих графу G_{sr}, не изменятся. Покажем, что после перекраски возникнет один из рассмотренных выше случаев 1) или 2). Если это не так, то в перекрашенном графе G' графы G'_{rt} и G'_{st} являются простыми цепями и не имеют кроме $\boldsymbol{v}_t$ ни одной общей вершины. Тогда цепь, связывающая в G' вершины $\boldsymbol{v}_s$ и $\boldsymbol{v}_t$ и состоящая из вершин цвета c_r и c_t, обязательно проходит через

вершину $\boldsymbol{v}$ (см. рис. 5.12), так как в противном случае граф G'_{st} не будет цепью, связывающей $\boldsymbol{v}_s$ и $\boldsymbol{v}_t$. Поэтому в G, как и в G', вершины $\boldsymbol{v}$ и $\boldsymbol{v}_t$

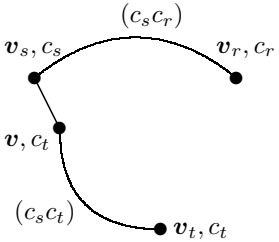


Рис. 5.11

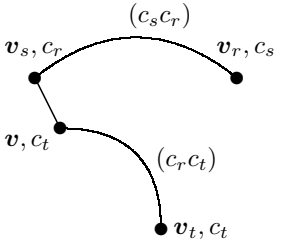


Рис. 5.12

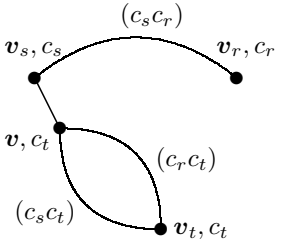


Рис. 5.13

связаны (c_rc_t)-цепью, состоящей из вершин цвета c_r и c_t (см. рис. 5.13). Но тогда (c_rc_t)-цепь между вершинами $\boldsymbol{v}$ и $\boldsymbol{v}_t$ графа G будет частью графа G_{rt}, и в этом случае или граф G_{rt} не является цепью, или цепи G_{st} и G_{rt} имеют общую вершину $\boldsymbol{v}$.

4) Если при всех i и j вершины $\boldsymbol{v}_i$ и $\boldsymbol{v}_j$ смежные, то нетрудно видеть, что в этом случае граф G является полным графом K_{n+1} и, следовательно, не удовлетворяет условиям теоремы.

Теорема доказана.

Нетрудно видеть, что для раскраски графа C_{2n+1}, являющегося циклом длины $2n+1$, требуется три различных цвета, и поэтому ограничение на число вершин в теореме Брукса убрать нельзя.

5.6 Раскраски ребер

Будем говорить, что ребра графа G правильным образом раскрашены k цветами, если каждому ребру графа G приписан один из k цветов так, что ребрам, инцидентным одной и той же вершине, приписаны разные цвета. *Хроматическим индексом* $\chi'(G)$ графа G называется минимально возможное число цветов k, которым можно правильным образом раскрасить его ребра.

При раскраске ребер, также как и при раскраске вершин, часто используется перекраска. Пусть в раскрашенном графе G ребро $\boldsymbol{e}$ покрашено цветом c_i. Перекраска этого ребра в цвет c_j осуществляется следующим образом. В графе G выделяется подграф G_{ij}, состоящий из ребер, покрашенных цветами c_i и c_j, и всех инцидентных этим ребрам вершин. Затем в компоненте связности, содержащей ребро $\boldsymbol{e}$, цвета c_i и c_j меняются местами. На рис. 5.14 изображен граф до и после перекраски ребра $\boldsymbol{e}$ в цвет c_2. В этом графе подграф G_{12} состоит из двух компонент связности. Первая компонента является циклом из шести ребер, среди которых находится и ребро $\boldsymbol{e}$, а вторая — из единственного ребра, покрашенного цветом c_1. После перекраски все ребра первой компоненты поменяли свои цвета, а ребро из второй

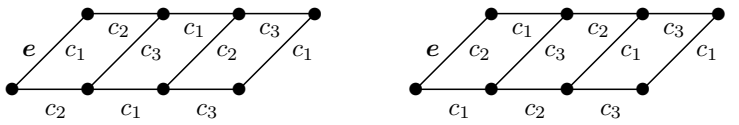


Рис. 5.14

компоненты сохранило свой цвет. Нетрудно видеть, что правильно раскрашенный граф после перекраски ребра остается правильно раскрашенным.

В следующей теореме, которая в несколько более общей формулировке была доказана В. Г. Визингом в 1964 г., устанавливаются верхняя и нижняя оценки хроматического индекса произвольного графа.

Теорема 5.8. *Если максимальная степень вершин графа G равна k, то $k \le \chi'(G) \le k+1$.*

Доказательство. Нижняя оценка теоремы очевидна. Для доказательства верхней оценки воспользуемся индукцией по числу ребер графа. В основание индукции положим дерево из k ребер, в котором степень одной из вершин равна k. Далее допустим, что любой граф, удовлетворяющий условиям теоремы и содержащий не более чем m ребер, можно раскрасить не более чем $k+1$ цветами. Пусть граф G удовлетворяет условиям теоремы и содержит $m+1$ ребро, вершина $\boldsymbol{u}$ этого графа смежна с вершинами $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$, а граф G' получается из G после удаления ребра $\{\boldsymbol{u}, \boldsymbol{v}_1\}$. По предположению индукции ребра G' можно раскрасить $k+1$ цветами. Далее полагаем, что все ребра графа G кроме ребра $\{\boldsymbol{u}, \boldsymbol{v}_1\}$ покрашены теми же цветами, что и аналогичные ребра G'.

Будем говорить, что цвет c_i отсутствует в вершине $\boldsymbol{v}$, если среди инцидентных этой вершине ребер нет ребра, покрашенного цветом c_i. Пусть цвет c_1 отсутствует в вершине $\boldsymbol{v}_1$. Если c_1 отсутствует и в $\boldsymbol{u}$, то ребро $\{\boldsymbol{u}, \boldsymbol{v}_1\}$ можно покрасить цветом c_1, и в этом случае граф G будет полностью покрашен. Поэтому далее полагаем, что инцидентные вершине $\boldsymbol{u}$ ребра покрашены цветами $c_1, \ldots, c_{k-1}$.

Из вершин $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ и цветов $c_1, \ldots, c_{k+1}$ составим упорядоченный набор вершин $\boldsymbol{v}_1, \boldsymbol{v}_{i_2}, \ldots, \boldsymbol{v}_{i_j}$ и упорядоченный набор цветов $c_1, c_{i_2}, \ldots, c_{i_j}$ в соответствии со следующей индуктивной процедурой.

Пусть выбраны вершины $\boldsymbol{v}_1, \boldsymbol{v}_{i_2}, \ldots, \boldsymbol{v}_{i_q}$ и цвета $c_1, c_{i_2}, \ldots, c_{i_q}$ так, что в вершине $\boldsymbol{v}_{i_t}$ отсутствует цвет c_{i_t}. Если среди еще невыбранных вершин найдется вершина $\boldsymbol{v}$ такая, что ребро $\{\boldsymbol{u}, \boldsymbol{v}\}$ покрашено цветом c_{i_q}, то полагаем $\boldsymbol{v}_{i_{q+1}} = \boldsymbol{v}$ и $c_{i_{q+1}} = c$, где c — цвет, отсутствующий в вершине $\boldsymbol{v}$. Если такой вершины $\boldsymbol{v}$ нет, то формирование наборов вершин и цветов закончено. Нетрудно видеть, что очередная вершина не может быть выбрана по одной из двух причин — 1) либо $c_{i_q} \in \{c_k, c_{k+1}\}$, 2) либо $c_{i_q} \in \{c_1, \ldots, c_{i_{q-1}}\}$. Рассмотрим два этих случая, полагая далее $\boldsymbol{v}_j = \boldsymbol{v}_{i_j}$ и $c_j = c_{i_j}$ для каждого $j = 2, \ldots, q$.

1) Соответствующий первому случаю фрагмент графа G изображен на рис. 5.15a, где рядом с вершинами указаны отсутствующие в этих вершинах

цвета (на рисунке не указаны отсутствующие в $\boldsymbol{u}$ и $\boldsymbol{v}_1$ вторые цвета). Цвет

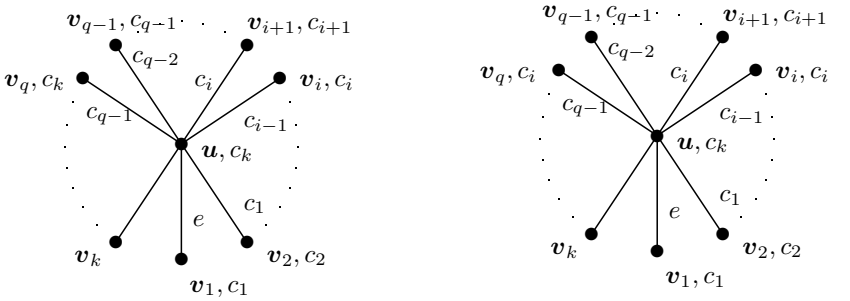


Рис. 5.15

c_k отсутствует в $\boldsymbol{u}$ и в $\boldsymbol{v}_q$, поэтому ребро $\{\boldsymbol{u}, \boldsymbol{v}_q\}$ можно перекрасить в цвет c_k, и при этом раскраска графа G останется правильной. В результате такой перекраски цвет c_{q-1} будет отсутствовать в вершинах $\boldsymbol{u}$ и в $\boldsymbol{v}_{q-1}$. Далее последовательно для каждого $j = q-1, \ldots, 2$ перекрасим ребро $\{\boldsymbol{u}, \boldsymbol{v}_j\}$ в цвет c_j. После перекраски цвет c_1 будет отсутствовать в $\boldsymbol{u}$ и в $\boldsymbol{v}_1$. Поэтому ребро $\{\boldsymbol{u}, \boldsymbol{v}_1\}$ можно покрасить в цвет c_1. Таким образом, существует правильная $(k+1)$-раскраска ребер графа G.

2) Рассмотрим второй случай. Соответствующий этому случаю фрагмент графа G изображен на рис. 5.15b. Пусть G_{ik} — остовный подграф графа G, содержащий все ребра, покрашенные цветами c_i и c_k. Так как степень каждой вершины в G_{ik} не превосходит двух, то нетрудно видеть, что каждая компонента связности графа G_{ik} является либо простым циклом, либо простой цепью, либо изолированной вершиной. В G_{ik} степени вершин $\boldsymbol{u}$, $\boldsymbol{v}_i$ и $\boldsymbol{v}_q$ равны единице. Поэтому эти три вершины не могут одновременно принадлежать одной и той же компоненте связности. Рассмотрим две возможности.

1. Вершина $\boldsymbol{u}$ не лежит в одной компоненте связности с $\boldsymbol{v}_i$. Тогда в компоненте, содержащей вершину $\boldsymbol{v}_i$, поменяем местами цвета c_k и c_i. В результате такой замены в вершине $\boldsymbol{v}_i$ будет отсутствовать цвет c_k. Так как после перекраски цвет c_k отсутствует в $\boldsymbol{u}$ и $\boldsymbol{v}_i$, то ребро $\{\boldsymbol{u}, \boldsymbol{v}_i\}$ можно перекрасить в цвет c_k. Если $i = 1$, то все ребра графа G покрашены. В противном случае последовательно для каждого $j = i-1, \ldots, 2$ ребро $\{\boldsymbol{u}, \boldsymbol{v}_j\}$ перекрасим в цвет c_j. После перекраски цвет c_1 будет отсутствовать в $\boldsymbol{u}$ и в $\boldsymbol{v}_1$. Поэтому ребро $\{\boldsymbol{u}, \boldsymbol{v}_1\}$ можно покрасить в цвет c_1.

2. Вершина $\boldsymbol{u}$ не лежит в одной компоненте связности с $\boldsymbol{v}_q$ и лежит в одной компоненте связности с $\boldsymbol{v}_i$. Тогда в компоненте, содержащей вершину $\boldsymbol{v}_q$, поменяем местами цвета c_k и c_i. После такой замены в вершине $\boldsymbol{v}_q$ будет отсутствовать цвет c_k, в $\boldsymbol{v}_i$ — по-прежнему c_i, а ребро $\{\boldsymbol{u}, \boldsymbol{v}_{i+1}\}$ останется окрашенным цветом c_i. Следовательно, настоящий случай переходит в рассмотренный выше случай 1).

Таким образом и во втором случае существует правильная $(k+1)$-раскраска ребер графа G. Теорема доказана.

5.7 Задачи

5.1. Показать, что в любом конечном графе без петель и кратных ребер найдутся две вершины одинаковой степени.

5.2. Показать, что число неизоморфных деревьев с n вершинами не превосходит 4^n.

5.3. Показать, что число неизоморфных связных графов с n вершинами и m ребрами не превосходит $a^{m+1}n^{m-n+1}$, где a — константа.

5.4. Показать, что в любом конечном регулярном двудольном графе существует совершенное паросочетание.

5.5. Доказать теоремы 5.3 и 5.4.

5.6. Доказать теорему 5.5.

5.7. Граф называется *эйлеровым*, если в нем существует цикл, проходящий через каждое ребро графа ровно один раз. Доказать, что связный граф будет эйлеровым тогда и только тогда, когда степень каждой его вершины четна.

5.8. Граф называется *планарным*, если его можно изобразить на плоскости так, что никакие его два ребра не пересекаются. Показать, что в любом планарном графе число вершин В, число ребер Р и число областей О, на которые граф разбивает плоскость, связаны соотношением $\text{В} + \text{О} - \text{Р} = 2$.

5.9. Показать, что в любом планарном графе без петель и кратных ребер найдется вершина степени не более чем пять.

5.10. Показать, что вершины любого планарного графа можно раскрасить в пять цветов так, что смежные вершины будут раскрашены в разные цвета.

5.11. Найти хроматические числа и хроматические индексы графов K_n и $K_{n,m}$.

5.12. Показать, что у произвольного двудольного (n,n)-графа G в каждой доле найдутся $m = \Theta(\log_2 n)$ вершин таких, что его подграф G', образованный этими вершинами и инцидентными им ребрами, будет либо пустым, либо графом $K_{m,m}$.

5.13. Цикл, проходящий через каждую вершину графа ровно один раз, называется *гамильтоновым*. Показать, что если в неориентированном n-вершинном графе G, не имеющем петель и кратных ребер, сумма степеней любых двух несмежных вершин не меньше n, то граф G имеет гамильтонов цикл.

5.14. Снегоуборочная машина должна проехать по всем улицам квадратного района размером $m \times n$ км и вернуться в гараж. Расстояние между любыми соседними перекрестками 100м. Какой наименьший путь она проедет?

5.15. При каких n можно обойти ходом коня по одному разу все клетки шахматной доски размером $n \times n$ и вернуться обратно?

5.16. В городе для любых двух из любых трех перекрестков найдется соединяющий их путь, не проходящий через третий перекресток. Доказать, что любые два перекрестка можно соединить не пересекающимися путями.

5.17. Двадцать городов соединены 172 авиалиниями. Доказать, что из любого города можно по воздуху добраться в любой другой.

Лекция 6

Булевы функции

6.1 Булев куб

Константы 0 и 1 называются *булевыми* константами. Упорядоченные наборы из нулей и единиц будем называть *двоичными* или *булевыми наборами*. Символы 0 и 1, входящие в двоичный набор, называются разрядами набора. Число разрядов набора называется его длиной. Разряды каждого набора длины n нумеруются целыми числами от 1 до n слева направо: крайний левый разряд получает номер 1, крайний правый — номер n. Множество всех булевых наборов длины n называется булевым кубом размерности n и обозначается символом $\mathbb{B}^n$. Часто булевы наборы называются также вершинами n-мерного единичного куба.

Номером и *весом* набора $\boldsymbol{u} = (u_1, \ldots, u_n)$ из $\mathbb{B}^n$ называются величины

$$|\boldsymbol{u}| = \sum_{i=1}^{n} u_i \cdot 2^{n-i}, \qquad \|\boldsymbol{u}\| = \sum_{i=1}^{n} u_i.$$

Номера наборов задают на множестве $\mathbb{B}^n$ отношение линейного порядка $\leq$. Будем говорить, что набор $\boldsymbol{u}$ не больше набора $\boldsymbol{v}$, если $|\boldsymbol{u}| \leq |\boldsymbol{v}|$. Порядок, определяемый отношением $\leq$, называется *лексикографическим*. Множество всех наборов длины n и веса k образует *k-й слой* куба $\mathbb{B}^n$, обозначаемый через $\mathbb{B}_k^n$.

Расстоянием Хемминга между вершинами $\boldsymbol{u}$ и $\boldsymbol{v}$ куба $\mathbb{B}^n$ называется число $d(\boldsymbol{u}, \boldsymbol{v}) = \sum_{i=1}^{n} |u_i - v_i|$, равное количеству несовпадающих разрядов $\boldsymbol{u}$ и $\boldsymbol{v}$. Нетрудно показать, что расстояние d является метрикой, т. е. d — положительная симметрическая функция двух аргументов, принимающая значение нуль тогда и только тогда, когда два ее аргумента совпадают, и для которой справедливо неравенство треугольника: $d(\boldsymbol{u}, \boldsymbol{v}) \leq d(\boldsymbol{u}, \boldsymbol{w}) + d(\boldsymbol{w}, \boldsymbol{v})$ для любых трех наборов $\boldsymbol{u}$, $\boldsymbol{v}$ и $\boldsymbol{w}$ из $\mathbb{B}^n$. Наборы $\boldsymbol{u}$ и $\boldsymbol{v}$ называются *соседними*, если $d(\boldsymbol{u}, \boldsymbol{v}) = 1$. Если $d(\boldsymbol{u}, \boldsymbol{v}) = n$, то наборы называются *противоположными*. Соседние наборы различаются между собой только в одном разряде, противоположные наборы — во всех разрядах.

Пары соседних вершин булева куба называются *ребрами*. Если $\boldsymbol{u}, \boldsymbol{v}$ — соседние вершины, различающиеся в i-м разряде, то говорят, что ребро $(\boldsymbol{u}, \boldsymbol{v})$

проходит в i-м направлении и соединяет эти вершины. Пусть $i_1, \ldots, i_{n-k}$ — попарно различные натуральные числа, не превосходящие n, $u_1, \ldots, u_{n-k}$ — булевы константы. Множество $\{\boldsymbol{v} \in \mathbb{B}^n \mid v_{i_j} = u_j, j = 1, 2, \ldots, n-k\}$ называется *k-мерной гранью* куба $\mathbb{B}^n$. Легко видеть, что k-мерная грань n-мерного булева куба содержит 2^k различных вершин.

Кроме рассмотренного выше линейного порядка $\leq$ на множестве наборов $\mathbb{B}^n$ существует естественный частичный порядок $\preceq$. Говорят, что набор $\boldsymbol{u}$ *не больше* набора $\boldsymbol{v}$ ($\boldsymbol{u} \preceq \boldsymbol{v}$), если $u_i \leq v_i$ при всех $i = 1, 2, \ldots, n$. Если $\boldsymbol{u} \preceq \boldsymbol{v}$ и $\boldsymbol{u} \neq \boldsymbol{v}$, то говорят, что что набор $\boldsymbol{u}$ *строго меньше* набора $\boldsymbol{v}$ ($\boldsymbol{u} \prec \boldsymbol{v}$). Наборы $\boldsymbol{u}$ и $\boldsymbol{v}$ называются *сравнимыми*, если либо $\boldsymbol{u} \preceq \boldsymbol{v}$, либо $\boldsymbol{v} \preceq \boldsymbol{u}$. Если ни одно из этих отношений не выполняется, то наборы называются *несравнимыми*. Последовательность вершин $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ называется *цепью*, если $d(\boldsymbol{u}_i, \boldsymbol{u}_{i+1}) = 1$ и $\boldsymbol{u}_i \preceq \boldsymbol{u}_{i+1}$ для всех $i = 1, 2, \ldots, k-1$. Вершина $\boldsymbol{u}_k$ называется наибольшей вершиной цепи $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$, а вершина $\boldsymbol{u}_1$ — наименьшей вершиной этой цепи. Число вершин в цепи называется ее длиной. Говорят, что цепь связывает вершины $\boldsymbol{u}$ и $\boldsymbol{v}$ и проходит через вершину $\boldsymbol{w}$, если $\boldsymbol{u}$ и $\boldsymbol{v}$ являются, соответственно, первой и последней вершинами цепи, а $\boldsymbol{w}$ принадлежит этой цепи. Цепь называется *максимальной*, если она не является частью цепи большей длины. Множество попарно несравнимых вершин называется *антицепью*. Антицепь называется *максимальной*, если она не является подмножеством другой антицепи, состоящей из большего количества вершин.

6.2 Булевы функции

1. Функция $f(x_1, \ldots, x_n)$, отображающая $\mathbb{B}^n$ в $\mathbb{B}^1$, называется *n-местной булевой* функцией. Множество всех булевых функций обозначается через P_2, а множество всех булевых функций, зависящих от n переменных, — через $P_2(n)$. Каждая булева функция имеет конечную область определения, что позволяет полностью задать функцию f из $P_2(n)$, перечислив все наборы из $\mathbb{B}^n$ и указав значения f на этих наборах. В частности, булева функция $f(x_1, \ldots, x_n)$ может быть задана таблицей, состоящей из 2^k строк, каждой из которых поставлен в соответствие булев набор длины k, и 2^{n-k} столбцов, каждому из которых поставлен в соответствие булев набор длины $n-k$. Параметр k принимает значения от 0 до n. В такой таблице (Таб. 6.1) значение функции f на наборе $(\sigma_1, \ldots, \sigma_k, \sigma_{k+1}, \ldots, \sigma_n)$ помещается на пересечении строки, соответствующей набору $(\sigma_1, \ldots, \sigma_k)$, и столбца, соответствующего набору $(\sigma_{k+1}, \ldots, \sigma_n)$. Если в таблице 6.1 параметр k равен нулю, то говорят, что функция задается вектором-строкой своих значений, а если $k = n$ — вектором-столбцом. Так как каждый элемент таблицы, задающей булеву функцию, равен либо нулю, либо единице, то легко видеть, что число различных таблиц и, соответственно, число различных булевых функций, зависящих от n переменных, равно 2^{2^n}. Число наборов из $\mathbb{B}^n$, на которых функция f принимает единичные значения, называется *весом* $\|f\|$ этой функции, т. е. $\|f\| = \sum_{\boldsymbol{u} \in \mathbb{B}^n} f(\boldsymbol{u})$.

Таблица 6.1

			0	0	…	σ_{k+1}	…	1	1	x_{k+1}
			0	0	…	σ_{k+2}	…	1	1	x_{k+2}
			…	…	…	…	…	…	…	…
			0	1	…	σ_n	…	0	1	x_n
x_1	…	x_k								
0	…	0				⋮				
0	…	1				⋮				
…	…	…				⋮				
σ_1	⋯	σ_k	⋯	⋯	⋯	$f(\boldsymbol{\sigma})$				
…	…	…								
1	…	0								
1	…	1								

2. Рассмотрим множества $P_2(1)$ и $P_2(2)$, состоящие из булевых функций, зависящих от одной и двух переменных. Первое множество состоит из четырех булевых функций, задаваемых векторами длины два: $\binom{0}{0}$, $\binom{0}{1}$, $\binom{1}{0}$, $\binom{1}{1}$. Первая и четвертая функции называются *тождественными константами*, нулем и единицей, и обозначаются, соответственно, символами 0 и 1. Вторая функция называется *тождественной* и обозначается так же, как и ее аргумент. Третья функция называется *отрицанием* или *инверсией*.

Теперь рассмотрим множество $P_2(2)$, состоящее из 16 функций, зависящих от переменных x и y. Среди этих функций две константы 0 и 1 и четыре функции x, $f_\neg(x)$, y, $f_\neg(y)$, каждая из которых зависит только от одной переменной. Векторы-столбцы семи из десяти оставшихся функций перечислены в таблице 6.2. Все эти функции имеют собственные названия. Первая функция $f_\&$ называется *конъюнкцией*. Эта функция часто также на-

Таблица 6.2

x y	$f_\&$	$f_\vee$	$f_\oplus$	$f_\sim$	$f_\mid$	$f_\downarrow$	$f_\to$
0 0	0	0	0	1	1	1	1
0 1	0	1	1	0	1	0	1
1 0	0	1	1	0	1	0	0
1 1	1	1	0	1	0	0	1

зывается умножением. Вторая функция называется *дизъюнкцией*. Нетрудно заметить, что $f_\&(x,y) = \min(x,y)$ и $f_\vee(x,y) = \max(x,y)$. Следующая функция $f_\oplus$ называется *суммой по модулю два*, иногда ее также называют исключающим или. Четвертая функция $f_\sim$ называется *эквивалентностью*, эта функция равна единице, если значения ее аргументов совпадают. Пятая функция называется *штрихом Шеффера*, шестая — *стрелкой Пирса*, седьмая — *импликацией*.

3. Переменная x_i функции $f(x_1,\dots,x_n)$ называется *существенной*, если найдутся такие булевы постоянные $u_1,\dots,u_{i-1},u_{i+1},\dots,u_n$, что

$$f(u_1,\dots,u_{i-1},0,u_{i+1},\dots,u_n) \neq f(u_1,\dots,u_{i-1},1,u_{i+1},\dots,u_n)$$

Несущественная переменная называется также *фиктивной*.

Определим несколько простейших преобразований булевых функций: подстановку констант, отождествление переменных, добавление и удаление фиктивных переменных.

Если для функций $f(x_1,\dots,x_n)$ и $g(x_{k+1},\dots,x_n)$ при всех возможных значениях переменных $x_{k+1},\dots,x_n$ справедливо равенство

$$f(\alpha_1,\dots,\alpha_k,x_{k+1},\dots,x_n) = g(x_{k+1},\dots,x_n),$$

то будем говорить, что функция g получена из функции f *подстановкой констант* $\alpha_1,\dots,\alpha_k$ вместо переменных $x_1,\dots,x_k$. Функцию g будем называть *подфункцией* функции f. Очевидным образом данное определение распространяется на случай подстановки констант $\alpha_1,\dots,\alpha_k$ вместо произвольных переменных $x_{i_1},\dots,x_{i_k}$.

Если для функций $f(x_1,\dots,x_n)$ и $g(x,x_{k+1},\dots,x_n)$ при всех возможных значениях переменных $x_{k+1},\dots,x_n$ справедливы равенства

$$f(0,\dots,0,x_{k+1},\dots,x_n) = g(0,x_{k+1},\dots,x_n),$$
$$f(1,\dots,1,x_{k+1},\dots,x_n) = g(1,x_{k+1},\dots,x_n),$$

то будем говорить, что функция g получена из функции f *отождествлением переменных* $x_1,\dots,x_k$. Как и ранее, данное определение очевидным образом распространяется на случай отождествления произвольных переменных $x_{i_1},\dots,x_{i_k}$.

Рассмотрим булеву функцию $f(x_1,\dots,x_n)$ с фиктивными переменными $x_1,\dots,x_k$ и булеву функцию $g(x_{k+1},\dots,x_n)$, получающуюся из f подстановкой констант $\alpha_1,\dots,\alpha_k$ вместо первых k переменных. Так как эти переменные у функции f фиктивные, то результат подстановки будет один и тот же для любых $\alpha_1,\dots,\alpha_k$. Поэтому можно полагать, что

$$g(x_{k+1},\dots,x_n) = f(0,\dots,0,x_{k+1},\dots,x_n).$$

Будем говорить, что функция g получена из функции f удалением фиктивных переменных $x_1,\dots,x_k$, а функция f получена из функции g добавлением фиктивных переменных $x_1,\dots,x_k$. Как и ранее, данное определение очевидным образом распространяется на случай удаления (добавления) произвольных фиктивных переменных $x_{i_1},\dots,x_{i_k}$.

Булевы функции f и g называются *равными*, если функция f получена из функции g удалением и добавлением фиктивных переменных.

4. Пусть $D \subseteq \mathbb{B}^n$. Функция $f(x_1,\dots,x_n)$, определенная на области D и принимающая значения 0 и 1, называется *частичной булевой функцией*. Если набор $\boldsymbol{x} \in \mathbb{B}^n$ и $\notin D$, то будем говорить, что функция f не определена на этом наборе, или, что f принимает на нем неопределенное значение "$*$".

Как и обычную булеву функцию, частичную булеву функцию можно задать таблицей из 2^n строк. Однако в таблице значений частичной функции в последнем столбце булевы величины будут стоять только в $|D|$ строках, соответствующих наборам из области определения частичной функции. В остальных местах будут стоять символы $*$. Очевидно, что на области D можно определить $2^{|D|}$ различных частичных функций.

Доопределением частичной булевой функции $f(x_1,\dots,x_n)$, определенной на области $D \subseteq \mathbb{B}^n$, называется такая булева функция $\hat{f}:\mathbb{B}^n \to \mathbb{B}$, что $\hat{f}(\boldsymbol{x}) = f(\boldsymbol{x})$ для любого $\boldsymbol{x}$ из D.

Доопределение частичной функции f можно получить заменив в таблице значений f каждый символ $*$ нулем или единицей. Поэтому легко видеть, что доопределение частичной функции не единственно. У каждой частичной функции, определенной на области D, есть $2^{2^n-|D|}$ различных доопределений.

Ниже приведены векторы значений двух частичных функций f_1 и f_2 и всех их доопределений. Первая функция определена на двух наборах из четырех возможных, и поэтому имеет четыре доопределения, вторая функция определена на трех наборах и у нее два доопределения.

x	y
0	0
0	1
1	0
1	1

f_1	y	$f_\vee$	$f_\to$	**1**
$*$	0	0	1	1
1	1	1	1	1
$*$	0	1	0	1
1	1	1	1	1

f_2	$f_\oplus$	$f_\mid$
$*$	0	1
1	1	1
1	1	1
0	0	0

6.3 Формулы

Выше булевы функции задавались перечислением своих значений на всей области определения. При таком задании все функции, зависящие от одного и того же числа переменных, оказываются одинаково сложными — для определения функции n переменных требуется таблица из 2^n строк. Далее рассматрим аналитический способ задания булевых функций посредством формул. Формульное представление булевых функций не только упрощает задание многих практически важных булевых функций, но и значительно облегчает различные действия с ними.

1. Пусть $X_n = \{x_1, x_2, \dots, x_n\}$ — множество булевых переменных, B — подмножество P_2. Выражение F, составленное из символов переменных из X_n и из символов функций из B называется *булевой формулой* в *базисе* B над множеством переменных X_n, если F удовлетворяет следующему индуктивному определению:

1. Переменная x_i является формулой ($i \in \{1,2,\dots,n\}$);

2. Если f — k-местная функция из B и $\mathsf{F}_1,\dots,\mathsf{F}_k$ — формулы, то выражение

$$f(\mathsf{F}_1,\dots,\mathsf{F}_k) \tag{6.1}$$

также является формулой. Формулы $\mathsf{F}_1,\dots,\mathsf{F}_k$ называются *подформулами* формулы (6.1), а функция f — внешней функцией этой формулы. Любая подформула каждой формулы F_i также называется подформулой формулы $\mathsf{F} = f(\mathsf{F}_1,\dots,\mathsf{F}_k)$.

Следуя индуктивному определению формулы, определим значение произвольной формулы $\mathsf{F}(x_1,x_2,\dots,x_n)$ на наборе $\alpha_1,\alpha_2,\dots,\alpha_n$ значений ее переменных:

1. Если $\mathsf{F} = x_i$, то $\mathsf{F}(\alpha_1,\dots,\alpha_n) = \alpha_i$;

2. Пусть $f \in B$, $\mathsf{F} = f(\mathsf{F}_1,\dots,\mathsf{F}_k)$ и значения формул $\mathsf{F}_1,\dots,\mathsf{F}_k$ определены для всех значений их переменных $x_1,\dots,x_n$. Тогда

$$\mathsf{F}(\alpha_1,\dots,\alpha_n) = f(\mathsf{F}_1(\alpha_1,\dots,\alpha_n),\dots,\mathsf{F}_k(\alpha_1,\dots,\alpha_n)).$$

Булева формула F над множеством переменных $x_1,\dots,x_n$ *реализует* булеву функцию $f(x_1,\dots,x_n)$, если

$$\mathsf{F}(\alpha_1,\dots,\alpha_n) = f(\alpha_1,\dots,\alpha_n)$$

при все наборах $(\alpha_1,\dots\alpha_n)$ из $\mathbb{B}^n$.

Пусть формула F, реализующая функцию $f(x_1,\dots,x_n)$, составлена из символов переменных $x_1,\dots,x_n$ и символов функций $f_1,\dots,f_m$. Тогда говорят, что формула F и функция f являются *суперпозициями* функций $f_1,\dots,f_m$. Далее формулы и реализуемые ими булевы функции будем обозначать одними и теми же символами в тех случаях, когда это не будет приводить к неоднозначному пониманию.

Сложностью $l(\mathsf{F})$ формулы F в базисе B называется число символов из B входящих в F.

Базис B называется *полным*, если любая функция из P_2 может быть реализована формулой в этом базисе.

2. Если базис B состоит только из двухместных и одноместных функций, то двухместные формулы $\mathsf{F}(x,y)$ будем записывать при помощи символов–связок в виде $(x \circ y)$, где $\circ$ — символ двухместной булевой функции, реализуемой формулой $\mathsf{F}(x,y)$. Наиболее часто встречающиеся символы двухместных булевых функций использованы в таблице 6.2 в качестве нижних индексов у символов соответствующих функций f. Так для обозначения конъюнкции чаще всего используется символ &, т. е. $f_\&(x,y) = (x\&y)$. Иногда конъюнкция обозначается также через $\wedge$ и $\cdot$, или функциональный символ опускается. Формулы для других двухместных булевых функций, перечисленных в таблице 6.2, записываются следующим образом:

$$f_\vee(x,y) = (x \vee y), \qquad f_\oplus(x,y) = (x \oplus y), \qquad f_\sim(x,y) = (x \sim y),$$
$$f_\mid(x,y) = (x \mid y), \qquad f_\downarrow(x,y) = (x \downarrow y), \qquad f_\to(x,y) = (x \to y).$$

Для эквивалентности вместо символа $\sim$ иногда используется символ $\equiv$. Одноместную формулу, реализующую функцию отрицания, будем записывать при помощи горизонтальной черты, покрывающей аргумент: $\mathsf{F}_\neg(x) = (\overline{x})$.

Далее для упрощения записи сложных формул иногда будем опускать скобки. Делать это будем в следующих случаях.

1. Во всех формулах будем опускать внешние скобки.

2. Полагая, что функция отрицания "сильнее" всех остальных функций, будем опускать скобки вокруг аргумента отрицания. Таким образом, если в формуле отсутствуют скобки, то сначала выполняется отрицание. Например, $\overline{(x_1 \to x_2)} = \overline{x_1 \to x_2}$.

3. Полагая, что функция & "сильнее" всех остальных двуместных функций, будем опускать скобки вокруг конъюнкции. Например, $(x_1 \& x_2) \oplus x_3 = x_1 \& x_2 \oplus x_3 = x_1 x_2 \oplus x_3$.

4. Легко видеть, что для дизъюнкции трех переменных справедливо равенство $(x_1 \vee x_2) \vee x_3 = x_1 \vee (x_2 \vee x_3)$. Аналогичные равенства имеют место также для конъюнкции и суммы по модулю два. Поэтому будем опускать скобки, если одна из функций $\&$, $\vee$ или $\oplus$ используется в формуле несколько раз подряд. Например, $(x_1 \vee x_2) \vee x_3 = x_1 \vee x_2 \vee x_3$.

3. Булевы формулы F_1 и F_2 называются *эквивалентными*, если они реализуют одну и ту же булеву функцию. Замена формулы F_1 на эквивалентную ей формулу F_2 называется *эквивалентным преобразованием* формулы F_1. Заметим, что любое эквивалентное преобразование формул устанавливает равенство реализуемых этими формулами функций.

Приведем ряд соотношений, определяющих простейшие эквивалентные преобразования булевых формул в двухместных базисах. Очевидно, что справедливы равенства $\overline{0} = 1$, $\overline{1} = 0$, и $\overline{\overline{x}} = x$, последнее из которых называется *правилом двойного отрицания*. Справедливость приводимых далее равенств для формул над множеством из одной переменной x для дизъюнкции, конъюнкции, суммы, эквивалентности, отрицания и констант легко следует из таблицы 6.2:

$$\begin{array}{llll} x \vee x = x, & x\&x = x, & x \oplus x = 0, & x \sim x = 1, \\ x \vee \overline{x} = 1, & x\&\overline{x} = 0, & x \oplus \overline{x} = 1, & x \sim \overline{x} = 0, \\ x \vee 0 = x, & x\&0 = 0, & x \oplus 0 = x, & x \sim 0 = \overline{x}, \\ x \vee 1 = 1, & x\&1 = x, & x \oplus 1 = \overline{x}, & x \sim 1 = x. \end{array} \tag{6.2}$$

Используя таблицу 6.2, нетрудно убедиться в эквивалентности различных формул над множеством из двух переменных. В частности справедливы соотношения

$$x\&y = \overline{\overline{x} \vee \overline{y}}, \quad x \vee y = \overline{\overline{x}\&\overline{y}}, \tag{6.3}$$

называемые *законами двойственности* или *законами де Моргана*. Из таблицы 6.2 также легко видеть, что

$$x \sim y = \overline{x \oplus y} = x \oplus y \oplus 1. \tag{6.4}$$

Подставляя в правые и левые части равенств (6.5) вместо переменных x, y и z различные булевы постоянные, видим, что конъюнкция связана с

дизъюнкцией и сложением по модулю два законами дистрибутивности:

$$\begin{array}{l} (x \vee y)\&z = (x\&z) \vee (y\&z), \\ (x\&y) \vee z = (x \vee z)\&(y \vee z), \\ (x \oplus y)\&z = (x\&z) \oplus (y\&z). \end{array} \tag{6.5}$$

Также легко подстановкой констант устанавливается справедливость следующих равенств:

$$x \oplus y = x\overline{y} \vee \overline{x}y, \qquad x \vee y = xy \oplus x \oplus y. \tag{6.6}$$

Используя (6.2)–(6.6), можно получить новые полезные равенства, позволяющие производить эквивалентные преобразования формул и устанавливать равенство реализуемых этими формулами функций. Например, для любых булевых функций f и g справедливо преобразование

$$f = f \cdot 1 = f \cdot (g \vee \overline{g}) = fg \vee f\overline{g},$$

называемое расщеплением. Обратное преобразование, переход от формулы $fg \vee f\overline{g}$ к формуле f, называется склеиванием. Следующая цепочка равенств задает преобразование, называемое поглощением:

$$f \vee fg = f \cdot 1 \vee fg = f \cdot (1 \vee g) = f \cdot 1 = f.$$

Наконец приведем еще одно часто используемое преобразование

$$f \vee \overline{f}g = fg \vee f\overline{g} \vee \overline{f}g = fg \vee f\overline{g} \vee \overline{f}g \vee fg = f \vee g,$$

которое, как нетрудно видеть, получается в результате последовательного применения преобразований расщепления и склеивания.

6.4 Нормальные формы

Ниже рассматриваются несколько стандартных способов представления булевых функций при помощи реализующих их формул, имеющих простую структуру.

1. Разложение функции по переменным. Для любого $x \in \mathbb{B}$ его *булевой степенью* называется функция[1)]

$$x^\sigma = \begin{cases} \overline{x}, & \text{если } \sigma = 0, \\ x, & \text{если } \sigma = 1. \end{cases}$$

Легко видеть, что $x^\sigma = x \oplus \sigma \oplus 1 = x\sigma \vee \overline{x}\,\overline{\sigma}$. Для любого $\boldsymbol{x} = (x_1, \ldots, x_n)$ и любого $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$ произведение

$$k_{\boldsymbol{\sigma}}(\boldsymbol{x}) = x_1^{\sigma_1} \cdot \ldots \cdot x_n^{\sigma_n},$$

[1)] В теории булевых функций булева степень встречается значительно чаще обычной степени. Поэтому за булевой степенью закрепилось обозначение, используемое вне теории булевых функций для обычной степени.

булевых степеней $\sigma_1, \ldots, \sigma_n$ переменных $x_1, \ldots, x_n$ называется *элементарной конъюнкцией*, ассоциированной с булевым набором $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$. Число переменных, входящих в конъюнкцию $k_{\boldsymbol{\sigma}}(\boldsymbol{x})$, называется *рангом* этой конъюнкции. Для функции $k_{\boldsymbol{\sigma}}(\boldsymbol{x})$ справедливо соотношение

$$k_{\boldsymbol{\sigma}}(\boldsymbol{a}) = \begin{cases} 1, & \text{если } \boldsymbol{a} = \boldsymbol{\sigma}, \\ 0, & \text{если } \boldsymbol{a} \neq \boldsymbol{\sigma}. \end{cases} \tag{6.7}$$

Отсюда немедленно следует, что $k_{\boldsymbol{a}}(\boldsymbol{x}) \cdot k_{\boldsymbol{b}}(\boldsymbol{x}) = 0$ при $\boldsymbol{a} \neq \boldsymbol{b}$.

Теорема 6.1. *Для каждой булевой функции $f(x_1, \ldots, x_n)$ при любом m, $1 \leq m \leq n$, справедливо представление*

$$f(x_1,...,x_m,x_{m+1},...,x_n) = \\ = \bigvee_{\sigma_1,\ldots,\sigma_m} x_1^{\sigma_1} \cdot \ldots \cdot x_m^{\sigma_m} \cdot f(\sigma_1,...,\sigma_m,x_{m+1},...,x_n). \tag{6.8}$$

Доказательство. Покажем, что для произвольного булева набора $(\alpha_1, \ldots, \alpha_n)$ значение функции, реализуемой формулой из правой части (6.8), равно $f(\alpha_1, \ldots, \alpha_n)$. Действительно, из (6.7) легко следует, что

$$\bigvee_{(\sigma_1,\ldots,\sigma_m)} \alpha_1^{\sigma_1} \cdot \ldots \cdot \alpha_m^{\sigma_m} \cdot f(\sigma_2,...,\sigma_m,\alpha_{m+1},...,\alpha_n) = \\ = \alpha_1^{\alpha_1} \cdot \ldots \cdot \alpha_m^{\alpha_m} \cdot f(\alpha_1,...,\alpha_m,\alpha_{m+1},...,\alpha_n) = f(\alpha_1,...,\alpha_n).$$

Теорема доказана.

Формула (6.8) называется *разложением* функции f по переменным $x_1, \ldots, x_m$. Очевидным образом (6.8) обобщается на случай разложения булевой функции по переменным $x_{i_1}, \ldots, x_{i_k}$. Важным частным случаем такого разложения является разложение по одной переменной. Разложение f по первой переменной выглядит следующим образом:

$$f(x_1, \ldots, x_k) = \overline{x}_1 f(0, x_2, \ldots, x_k) \vee x_1 f(1, x_2, \ldots, x_k).$$

2. Дизъюнктивные нормальные формы. *Совершенной дизъюнктивной нормальной формой* (СДНФ) функции $f(x_1, \ldots, x_n)$ называется ее разложение

$$f(x_1, \ldots, x_n) = \bigvee_{(\sigma_1,\ldots,\sigma_n)} x_1^{\sigma_1} \cdot \ldots \cdot x_n^{\sigma_n} \cdot f(\sigma_1,...,\sigma_n)$$

по всем ее переменным. Легко видеть, что СДНФ f есть дизъюнкция всех элементарных конъюнкций, ассоциированных с теми наборами $\boldsymbol{\sigma}$, на которых функция f равна единице. Так, например, импликация принимает единичные значения на наборах $(00), (01)$ и (11), поэтому в соответствии с определением СДНФ имеем:

$$x \to y = x^0 y^0 \vee x^0 y^1 \vee x^1 y^1 = \overline{x}\,\overline{y} \vee \overline{x} y \vee xy.$$

Конъюнкция $x\&y$ равна единице, только если $x = y = 1$, и поэтому сама является своей совершенной дизъюнктивной нормальной формой.

Совершенные дизъюнктивные нормальные формы устроены очень просто. В некотором смысле, СДНФ булевой функции это ее вектор значений, записанный на языке формул. С другой стороны, часто СДНФ даже простых функций состоят из очень большого числа элементарных конъюнкций. Так, например, СДНФ дизъюнкции $x_1 \vee \cdots \vee x_n$ содержит $2^n - 1$ элементарных конъюнкций. Более экономными (с точки зрения числа символов переменных, входящих в формулу) являются дизъюнктивные нормальные формы. *Дизъюнктивной нормальной формой* (ДНФ) булевой функции f называется реализующая f формула, являющаяся дизъюнкцией элементарных конъюнкций. В отличии от СДНФ дизъюнктивная нормальная форма функции f определяется неоднозначно, т. е. f может иметь несколько реализующих ее ДНФ. *Минимальной* дизъюнктивной нормальной формой функции f называется ДНФ, содержащая минимальное число символов переменных среди всех ДНФ функции f. Получить ДНФ можно из СДНФ при помощи эквивалентных преобразований.

Найдем минимальную ДНФ импликации. Сделаем это, преобразуя ее СДНФ при помощи приведенных в предыдущем параграфе равенств:

$$x \to y = \overline{x}\,\overline{y} \vee \overline{x} y \vee xy = \overline{x}\,\overline{y} \vee \overline{x} y \vee \overline{x} y \vee xy = \\ = \overline{x}(\overline{y} \vee y) \vee y(\overline{x} \vee x) = \overline{x} \vee y.$$

Так как импликация существенно зависит от двух переменных, то очевидно, что ее любая ДНФ содержит символы обеих переменных. Таким образом, формула $\overline{x} \vee y$ будет минимальной ДНФ импликации.

3. Конъюнктивные нормальные формы. Для любых булевых наборов $\boldsymbol{x} = (x_1, \ldots, x_n)$ и $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$ дизъюнкция

$$d_{\boldsymbol{\sigma}}(\boldsymbol{x}) = x_1^{\sigma_1} \vee \ldots \vee x_n^{\sigma_n},$$

булевых степеней $\sigma_1, \ldots, \sigma_n$ переменных $x_1, \ldots, x_n$ называется *элементарной дизъюнкцией*, ассоциированной с булевым набором $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$. Если все σ_i равны единице, то дизъюнкция $d_{\boldsymbol{\sigma}}(\boldsymbol{x})$ называется *монотонной*. Для функции $d_{\boldsymbol{\sigma}}(\boldsymbol{x})$ справедливо соотношение

$$d_{\boldsymbol{\sigma}}(\boldsymbol{a}) = \begin{cases} 0, & \text{если } \boldsymbol{a} \text{ и } \boldsymbol{\sigma} \text{ — противоположные наборы,} \\ 1, & \text{в остальных случаях.} \end{cases}$$

Пусть $f(x_1, \ldots, x_n)$ — произвольная булева функция. Представим отрицание f в виде совершенной дизъюнктивной нормальной формы:

$$\overline{f}(x_1, \ldots, x_n) = \bigvee_{\boldsymbol{\sigma}=(\sigma_1,\ldots,\sigma_n)} x_1^{\sigma_1} \cdot \ldots \cdot x_n^{\sigma_n} \cdot \overline{f}(\sigma_1,...,\sigma_n),$$

где дизъюнкция берется по всем таким наборам $\boldsymbol{\sigma}$, что $f(\boldsymbol{\sigma}) = 0$. Взяв отрицания от обоих частей равенства и применив законы двойственности,

видим, что для функции f справедливы равенства

$$f(x_1,\dots,x_n)=\overline{\bigvee x_1^{\sigma_1}\cdot\ldots\cdot x_n^{\sigma_n}}=\bigwedge\left(x_1^{\overline{\sigma}_1}\vee\ldots\vee x_n^{\overline{\sigma}_n}\right),\qquad(6.9)$$

в которых дизъюнкция и конъюнкция берутся по всем тем наборам $\boldsymbol{\sigma}=(\sigma_1,\dots,\sigma_n)$, для которых $f(\boldsymbol{\sigma})=0$. Формула, стоящая в правой части равенства (6.9), называется *совершенной конъюнктивной нормальной формой* (СКНФ) функции $f(x_1,\dots,x_n)$.

Найдем совершенные конъюнктивные нормальные формы стрелки Пирса и дизъюнкции переменных x и y. Так как стрелка Пирса принимает нулевые значения на наборах (01), (10) и (11), то в соответствии с формулой (6.9) имеем:

$$x\downarrow y=(x^{\overline{0}}\vee y^{\overline{1}})(x^{\overline{1}}\vee y^{\overline{0}})(x^{\overline{1}}\vee y^{\overline{1}})=(x\vee\overline{y})(\overline{x}\vee y)(\overline{x}\vee\overline{y}).$$

Дизъюнкция $x\vee y$ равна нулю, только если $x=y=0$, и поэтому сама является своей совершенной конъюнктивной нормальной формой.

Конъюнктивной нормальной формой (КНФ) булевой функции f называется реализующая f формула, являющаяся конъюнкцией элементарных дизъюнкций. Как и в случае дизъюнктивных нормальных форм КНФ называется *минимальной*, если она содержит минимальное число переменных среди всех КНФ функции f.

Найдем минимальную КНФ стрелки Пирса. Сделаем это, преобразуя ее СКНФ. Легко видеть, что

$$x\downarrow y=(x\vee\overline{y})(\overline{x}\vee y)(\overline{x}\vee\overline{y})=(\overline{x}\,\overline{y}\vee yx)(\overline{x}\vee\overline{y})=\overline{x}\,\overline{y}.$$

Так как стрелка Пирса существенно зависит от двух переменных, то очевидно, что ее любая КНФ содержит символы обеих переменных. Таким образом, формула $\overline{x}\,\overline{y}$ будет минимальной КНФ стрелки Пирса.

4. Алгебраическая нормальная форма. Для любого $x\in\mathbb{B}$ его *алгебраической степенью* (или просто степенью) называется функция[2)]

$$x^{(\sigma)}=\begin{cases}1, & \text{если } \sigma=0,\\ x, & \text{если } \sigma=1.\end{cases}$$

Для любого $\boldsymbol{x}=(x_1,\dots,x_n)$ и любого $\boldsymbol{\sigma}=(\sigma_1,\dots,\sigma_n)$ произведение

$$p_{\boldsymbol{\sigma}}(\boldsymbol{x})=x_1^{(\sigma_1)}\cdot\ldots\cdot x_n^{(\sigma_n)}\qquad(6.10)$$

степеней $\sigma_1,\dots,\sigma_n$ переменных $x_1,\dots,x_n$ называется булевым одночленом, ассоциированным с булевым набором $\boldsymbol{\sigma}=(\sigma_1,\dots,\sigma_n)$. Вес этого набора называется *степенью одночлена* (6.10) и обозначается $\deg p_{\boldsymbol{\sigma}}(\boldsymbol{x})$.

Булевы одночлены от переменных $x_1,\dots,x_n$ естественным образом нумеруются номерами ассоциированных булевых наборов — номером одночлена (6.10) является величина $|\boldsymbol{\sigma}|$. Далее для обозначения одночлена $p_{\boldsymbol{\sigma}}(\boldsymbol{x})$ будем также использовать обозначение $p_{|\boldsymbol{\sigma}|}(\boldsymbol{x})$ или просто $p_{|\boldsymbol{\sigma}|}$, если из контекста понятно от каких переменных зависит рассматриваемый одночлен.

[2)] См. примечание на стр. 87

Теорема 6.2. *Каждая булева функция $f(x_1,\dots,x_n)$ единственным образом представляется в виде*

$$f(x_1,\dots,x_n)=\bigoplus_{\boldsymbol{\sigma}=(\sigma_1,\dots,\sigma_n)}x_1^{(\sigma_1)}\cdot\ldots\cdot x_n^{(\sigma_n)}\cdot p_{\boldsymbol{\sigma}},\qquad(6.11)$$

где $p_{\boldsymbol{\sigma}}\in\{0,1\}$. Формула (6.11) *называется алгебраической нормальной формой функции f или ее многочленом Жегалкина.*

Доказательство. Прежде всего покажем, что каждая булева функция реализуется формулой в базисе $\{\&,\oplus,1\}$. Для одноместных функций это утверждение очевидно. Так как для любой n-местной функции f справедливо равенство

$$f(x_1,\dots,x_n)=x_n(f(x_1,\dots,x_{n-1},0)\oplus f(x_1,\dots,x_{n-1},1))\oplus f(x_1,\dots,x_{n-1},0),$$

то возможность реализации произвольной булевой функции формулой в базисе $\{\&,\oplus,1\}$ легко устанавливается индукцией по числу переменных. Пусть функция f реализуется формулой F. Раскрывая скобки и приводя подобные слагаемые, преобразуем F в сумму одночленов, т. е. в формулу вида (6.11). Теперь для окончательного доказательства теоремы осталось показать единственность такой формулы. Для этого найдем число различных многочленов от n переменных. Так как каждый одночлен, зависящий от n переменных (не обязательно существенно), однозначно определяется набором степеней переменных — булевым набором длины n, то очевидно, что число одночленов равно 2^n. Каждый одночлен либо входит, либо не входит в многочлен. Следовательно, число различных многочленов, вместе с нулевым, равно 2^{2^n}, т. е. многочленов столько же, сколько и булевых функций, зависящих от n переменных. Так как каждая булева функция реализуется хотя бы одним многочленом Жегалкина, то для каждой функции существует единственный реализующий ее многочлен. Теорема доказана.

Итак, каждая булева функция f единственным образом представляется в виде многочлена Жегалкина. *Степенью* функции f, обозначаемой через $\deg f$, называется максимальная степень одночленов, входящих в ее многочлен Жегалкина.

Методом неопределенных коэффициентов найдем многочлен Жегалкина дизъюнкции двух переменных. Для этого дизъюнкцию $x\vee y$ представим в виде многочлена

$$x\vee y=a\oplus bx\oplus cy\oplus dxy\qquad(6.12)$$

с неизвестными коэффициентами a,b,c и d. Подставляя в левую и правую части (6.12) нули вместо переменных x и y, получаем, что $a=0$. Полагая далее $x=1$, $y=0$, находим $a\oplus b=1$. Подстановки $x=0$, $y=1$ и $x=y=1$, дают, соответственно, $a\oplus c=1$ и $a\oplus b\oplus c\oplus d=1$. Таким образом, для определения четырех неизвестных коэффициентов получили систему из четырех уравнений. Решая эту систему, легко находим $a=0$, $b=c=d=1$. Следовательно, $x\vee y=x\oplus y\oplus xy$.

Преобразуем СДНФ произвольной булевой функции f, заменяя в СДНФ дизъюнкции формулами в базисе $\{\oplus, \&\}$. Из рассмотренного примера и равенства нулю произведения двух различных элементарных конъюнкций, зависящих от одних и тех же переменных, легко получаем формулу

$$f(x_1,\dots,x_n) = \bigoplus_{(\sigma_1,\dots,\sigma_n)} x_1^{\sigma_1}\cdot\ldots\cdot x_n^{\sigma_n}\cdot f(\sigma_1,\dots,\sigma_n), \tag{6.13}$$

которая в некоторых случаях оказывается более удобной, чем СДНФ или многочлен Жегалкина.

6.5 Задачи

6.1. Найти в $\mathbb{B}^n$: a) число различных k-мерных граней;
b) число различных k-мерных граней, проходящих через фиксированную вершину;
c) число всех граней.

6.2. Сколько вершин в среднем содержит одна грань n-мерного куба?

6.3. Найти число ребер, проходящих через вершины k-мерной грани n-мерного булева куба.

6.4. Найти число ребер, проходящих через вершины, лежащие в k-м слое n-мерного булева куба.

6.5. Найти: a) $\sum_{\boldsymbol{u}\in\mathbb{B}^n}\|\boldsymbol{u}\|$; b) $\sum_{\boldsymbol{u}\in\mathbb{B}^n}|\boldsymbol{u}|$; c) $\sum_{\boldsymbol{u}\in\mathbb{B}^n_k}|\boldsymbol{u}|$.

6.6. Для любых сравнимых наборов $\boldsymbol{a},\boldsymbol{b}\in\mathbb{B}^n$ интервалом с границами $\boldsymbol{a}$ и $\boldsymbol{b}$ называется множество $I(\boldsymbol{a},\boldsymbol{b}) = \{\boldsymbol{\gamma}\in\mathbb{B}^n \mid \boldsymbol{a}\preceq\boldsymbol{\gamma}\preceq\}$. Показать, что любой интервал является гранью, а грань — интервалом.

6.7. Найти в $\mathbb{B}^n$ число наборов несравнимых с данным набором $\boldsymbol{a}$.

6.8. Найти число пар попарно несравнимых вершин в $\mathbb{B}^n$.

6.9. Найти: a) число различных максимальных цепей в $\mathbb{B}^n$;
b) число различных максимальных цепей, проходящих через фиксированную вершину k-го слоя куба $\mathbb{B}^n$.

6.10. Пусть T — антицепь в $\mathbb{B}^n$, $T_k = T\cap\mathbb{B}^n_k$. Показать, что

$$\sum_{k=0}^{n}|T_k|\Big/\binom{n}{k}\le 1.$$

6.11. Пусть булева функция $f(x_1,\dots,x_n)$ задана вектором значений $(f_0,\dots,f_{2^n-1})$. Доказать, что если x_i является фиктивной переменной, то $f_j = f_{2^{n-i}+j}$ для всех целых j, принадлежащих множеству $\{k\,2^{n-i+1},\dots,(2k+1)\,2^{n-i+1}-1\}$, где $k=0,1,\dots 2^{i-1}-1$.

6.12. Пусть $f(x_1,\dots,x_n)$ существенно зависит ровно от k переменных. Показать, что $\|f\|$ делится на 2^{n-k}.

6.13. Найти в $P_2(n)$ число различных симметрических булевых функций.

6.14. Показать, что любая булева функция от трех переменных может быть получена из симметрической функции от семи переменных отождествлением переменных.

6.15. Показать, что любая булева функция может быть получена из симметрической функции отождествлением переменных.

6.16. Показать, что любая симметрическая функция отличная от константы существенно зависит от всех своих переменных.

6.17. Какое минимальное число неравных подфункций может быть у n-местной булевой функции, существенно зависящей от всех своих аргументов?

6.18. Доказать формулу обобщенного склеивания: $xz\vee y\overline{z}\vee xy = xz\vee y\overline{z}$.

6.19. При помощи эквивалентных преобразований упростить формулы:
a) $xyz\vee xy\overline{z}\vee x\overline{y}z\vee\overline{x}yz$; b) $xyz\vee\overline{x}yz\vee x\vee yz\vee(x\oplus y)$;
c) $x\overline{y}(x\to z)\sim\overline{z}$.

6.20. Показать, что любая формула в базисе $\{\vee,\&,\neg\}$ эквивалентна некоторой формуле в том же базисе, в которой отрицания появляются только над переменными.

6.21. Доказать, что:
a) $\overline{x}_1\&\overline{x}_2\&\dots\&\overline{x}_n = \overline{x_1\vee x_2\vee\dots\vee x_n}$;
b) $\overline{x}_1\vee\overline{x}_2\vee\dots\vee\overline{x}_n = \overline{x_1\&x_2\&\dots\&x_n}$.
c) $x_1\vee x_2\vee\dots\vee x_n = (x_1\oplus1)(x_2\oplus1)\cdot\ldots\cdot(x_n\oplus1)\oplus1$.

6.22. Найти число попарно различных булевых функций, получающихся из функции $\bigvee_{1\le i<j\le n}x_ix_j$ подстановкой констант вместо переменных $x_1,\dots,x_n$.

6.23. На скольких наборах из $\mathbb{B}^n$ равна единице функция f:
a) $f(x_1,\dots,x_n) = x_1\vee\overline{x}_1x_2\vee\overline{x}_1\overline{x}_2x_2\vee\dots\vee\overline{x}_1\dots\overline{x}_{n-1}x_n$;
b) $f(x_1,\dots,x_n) = \bigvee_{1\le i_1<\dots<i_k\le 2n}x_{i_1}\cdot\ldots\cdot x_{i_n}$.

6.24. На скольких наборах из $\mathbb{B}^n$ равна единице функция f:
a) $f(x_1,\dots,x_n) = \bigoplus_{k=1}^n\bigoplus_{1\le i_1<\dots<i_k\le n}x_{i_1}\cdot\ldots\cdot x_{i_k}$;
b) $f(x_1,\dots,x_n) = \bigoplus_{k=1}^n x_1\cdot\ldots\cdot x_k$;
c) $f(x_1,\dots,x_n) = \bigoplus_{1\le i_1<\dots<i_k\le n}x_{i_1}\cdot\ldots\cdot x_{i_k}$.

6.25. Показать, что каждая булева функция трех переменных имеет симметрическую подфункцию двух переменных.

6.26. Найти число различных дизъюнктивных нормальных форм над множеством переменных $x_1,\dots,x_n$.

6.27. Найти совершенные конъюнктивные нормальные формы всех булевых функций двух переменных.

6.28. Показать, что любая ДНФ функции $x_1\oplus\dots\oplus x_n$ состоит из:
a) 2^{n-1} элементарных конъюнкций; b) $n2^{n-1}$ символов переменных.

6.29. Показать, что любая КНФ функции $x_1\oplus\dots\oplus x_n$ состоит из:
a) 2^{n-1} элементарных дизъюнкций; b) $n2^{n-1}$ символов переменных.

Лекция 7

Полные системы булевых функций

7.1 Замкнутые классы булевых функций

Пусть R — произвольное множество булевых функций. *Замыканием* множества R называется множество всех функций, которые можно реализовать формулами в базисе R. Замыкание множества R будем обозначать через $[R]$. Множество булевых функций R называется (*функционально*) *замкнутым* множеством, если оно совпадает со своим замыканием, т. е. $R = [R]$.

Рассмотрим пять важнейших замкнутых множеств в P_2. Часто рассматриваемые ниже замкнутые множества называются также замкнутыми классами.

1. Будем говорить, что функция $f(x_1, \ldots, x_n)$ сохраняет нуль, если

$$f(0, \ldots, 0) = 0.$$

Множество, состоящее из всех булевых функций сохраняющих нуль, обозначается через T_0. Легко видеть, что функции 0, x, $x\&y$, $x \vee y$ и $x \oplus y$ принадлежат T_0, а функции 1, $\overline{x}$, $x \sim y$, $x \,|\, y$, $x \downarrow y$ и $x \to y$ не принадлежат T_0.

Так как тождественная функция сохраняет нуль, и для любых сохраняющих нуль функций $f_0, f_1, \ldots, f_k$ справедливо равенство

$$f(0, ..., 0) = f_0(f_1(0, ..., 0), \ldots, f_k(0, ..., 0)) = f_0(0, ..., 0) = 0,$$

т. е. реализуемая формулой $f_0(f_1, \ldots, f_k)$ функция f также сохраняет нуль, то легко видеть, что множество T_0 замкнуто.

Любой булев вектор длины 2^n с первой нулевой компонентой будет вектором значений функции из T_0. Поэтому, в T_0 содержится ровно 2^{2^n-1} функций из $P_2(n)$. Множество $T_0 \cap P_2(n)$ будем обозначать через $T_0(n)$.

2. Будем говорить, что функция $f(x_1, \ldots, x_n)$ сохраняет единицу, если

$$f(1, \ldots, 1) = 1.$$

Множество, состоящее из всех булевых функций сохраняющих единицу, обозначается через T_1.

Легко видеть, что функции 1, x, $x\&y$, $x \vee y$, $x \sim y$ и $x \to y$ принадлежат T_1, а функции 0, $\overline{x}$, $x \oplus y$, $x \,|\, y$ и $x \downarrow y$ не принадлежат T_1. Доказательство замкнутости множества T_1 аналогично доказательству замкнутости множества T_0. Также легко видеть, что в T_1 содержится ровно 2^{2^n-1} функций из $P_2(n)$. Множество $T_1 \cap P_2(n)$ будем обозначать через $T_1(n)$.

3. Будем говорить, что булева функция $f(x_1, \ldots, x_n)$ является *двойственной* к функции $g(x_1, \ldots, x_n)$, если

$$f(x_1, \ldots, x_n) = \overline{g}(\overline{x}_1, \ldots, \overline{x}_n).$$

Функцию двойственную к функции f будем обозначать через f^*. Легко видеть, что $(f^*)^* = f$ для любой булевой функции f. Из законов двойственности следует, что $(x\&y)^* = x \vee y$ и $(x \vee y)^* = x\&y$. Функция f называется *самодвойственной*, если $f = f^*$. Множество, состоящее из всех самодвойственных булевых функций, обозначается через S. Самодвойственными являются функции x, $\overline{x}$, $x_1 \oplus x_2 \oplus x_3$. Среди булевых функций, существенно зависящих ровно от двух переменных, нет ни одной самодвойственной функции.

Докажем замкнутость множества самодвойственных функций. Пусть $f_0, f_1, \ldots, f_k$ — произвольные самодвойственные функции. Рассмотрим новую функцию $f = f_0(f_1, \ldots, f_k)$. Так как добавление фиктивной переменной оставляет самодвойственную функцию самодвойственной, то без ограничения общности будем полагать, что все функции f_i зависят от одних и тех же переменных $x_1, \ldots, x_n$. Тогда

$$\begin{aligned} f(\overline{x}_1, \ldots, \overline{x}_n) &= f_0(f_1(\overline{x}_1, \ldots, \overline{x}_n), \ldots, f_k(\overline{x}_1, \ldots, \overline{x}_n)) = \\ &= f_0(\overline{f}_1(x_1, \ldots, x_n), \ldots, \overline{f}_k(x_1, \ldots, x_n)) = \\ &= \overline{f}_0(f_1(x_1, \ldots, x_n), \ldots, f_k(x_1, \ldots, x_n)) = \overline{f}(x_1, \ldots, x_n). \end{aligned}$$

Следовательно, функция f — самодвойственная. Таким образом, множество S замкнуто.

Так как каждая самодвойственная функция на противоположных наборах принимает противоположные значения, то для определения любой самодвойственной функции достаточно задать ее значения только на половине из 2^n наборов. Следовательно, в S содержится ровно $2^{2^{n-1}}$ функций из $P_2(n)$. Далее множество самодвойственных функций, зависящих от n переменных, будем обозначать через $S(n)$.

4. Функция $f(x_1, \ldots, x_n)$ называется *линейной*, если степень ее многочлена Жегалкина не превосходит единицу, т. е.

$$f(x_1, \ldots, x_n) = \alpha_1 x_1 \oplus \ldots \oplus \alpha_n x_n \oplus \alpha_0,$$

где α_i — булевы постоянные. Множество, состоящее из всех линейных булевых функций, обозначается через L. Очевидно, что среди функций из $P_2(2)$ линейными являются только 0, 1, x, $\overline{x}$, y, $\overline{y}$, $x \oplus y$ и $x \sim y$. Непосредственно из определения линейной функции следует замкнутость множества L.

Так как каждая булева функция однозначно определяется коэффициентами своего многочлена Жегалкина, а у каждой линейной функции все коэффициенты при одночленах степени два и выше равны нулю, то легко видеть, что в L содержится ровно 2^{n+1} функций из $P_2(n)$. Далее множество $L \cap P_2(n)$ будем обозначать через $L(n)$.

5. Функция $f(x_1, \ldots, x_n)$ называется *монотонной*, если

$$f(\alpha_1, \ldots, \alpha_n) \le f(\beta_1, \ldots, \beta_n)$$

для любых наборов $\boldsymbol{a} = (\alpha_1, \ldots, \alpha_n)$ и $\boldsymbol{b} = (\beta_1, \ldots, \beta_n)$ таких, что $\boldsymbol{a} \preceq \boldsymbol{b}$. Множество, состоящее из всех монотонных булевых функций, обозначается через M. В $P_2(2)$ монотонными являются функции 0, 1, x, y, $x\&y$ и $x \vee y$.

Докажем замкнутость множества монотонных функций. Пусть $f_0, \ldots, f_k$ — произвольные монотонные функции. Очевидно, что добавление фиктивной переменной оставляет монотонную функцию монотонной. Поэтому без ограничения общности будем полагать, что все функции f_i зависят от одних и тех же переменных $x_1, ..., x_n$. Пусть $\boldsymbol{a}$, $\boldsymbol{b}$ — такие наборы из $\mathbb{B}^n$, что $\boldsymbol{a} \preceq \boldsymbol{b}$. Рассмотрим новую функцию $f = f_0(f_1, \ldots, f_k)$. Так как $f_i(\boldsymbol{a}) \le f_i(\boldsymbol{b})$, то $(f_1(\boldsymbol{a}), \ldots, f_k(\boldsymbol{a})) \preceq (f_1(\boldsymbol{b}, \ldots, f_k(\boldsymbol{b}))$, и поэтому

$$f(\boldsymbol{a}) = f_0(f_1(\boldsymbol{a}), \ldots, f_k(\boldsymbol{a})) \le f_0(f_1(\boldsymbol{b}), \ldots, f_k(\boldsymbol{b})) = f(\boldsymbol{b}).$$

Следовательно, функция f — монотонная. Таким образом, множество M замкнуто.

Обозначим множество монотонных функций n переменных через $M(n)$. В отличие от множеств $T_0(n)$, $T_1(n)$, $S(n)$ и $L(n)$, мощности которых легко были найдены выше, точное число монотонных функций в $P_2(n)$ при больших n неизвестно. Лишь сравнительно недавно А. Д. Коршуновым была найдена асимптотически точная формула для $|M(n)|$. Эта формула выглядит достаточно громоздко, и поэтому здесь не приводится. Вместо этого для $|M(n)|$ ниже устанавливаются более грубые, но в тоже время значительно более простые неравенства.

7.2 Монотонные булевы функции

Покажем, что $|M(n)| \ge 2^{\binom{n}{\lfloor n/2 \rfloor}}$. Для доказательства этого неравенства достаточно рассмотреть множество монотонных булевых функций, каждая из которых удовлетворяет условию

$$f(x_1, \ldots, x_n) = \begin{cases} 1, & \text{если } \sum_{i=1}^n x_i > \lfloor \frac{n}{2} \rfloor, \\ 0, & \text{если } \sum_{i=1}^n x_i < \lfloor \frac{n}{2} \rfloor. \end{cases}$$

Так как каждая такая функция однозначно определяется своими значениями на наборах веса $\lfloor \frac{n}{2} \rfloor$, то очевидно, что таких функций ровно $2^{\binom{n}{\lfloor n/2 \rfloor}}$.

Верхняя оценка числа монотонных функций, доказываемая в следующей теореме, была установлена Ж. Анселем в 1966 году.

Теорема 7.1.

$$|M(n)| \le 3^{\binom{n}{\lfloor n/2 \rfloor}}.$$

Доказательство. Рассмотрим цепь $\boldsymbol{v}_1 \prec \boldsymbol{v}_2 \prec \boldsymbol{v}_3$. Нетрудно видеть, что существует единственная вершина $\boldsymbol{u}$ такая, что $\boldsymbol{v}_1 \prec \boldsymbol{u} \prec \boldsymbol{v}_3$. Вершину $\boldsymbol{u}$ назовем дополнительной для вершин $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$. Для доказательства теоремы потребуется следующее вспомогательное утверждение.

Лемма 7.1. *Булев куб $\mathbb{B}^n$ можно покрыть $\binom{n}{\lfloor n/2 \rfloor}$ непересекающимися цепями C_i так, что:*

(i) число цепей длины $n-2p+1$, где $p = 0, 1, \ldots, \lfloor n/2 \rfloor$, равно $\binom{n}{p} - \binom{n}{p-1}$;

(ii) для любых трех вершин $\boldsymbol{v}_1 \prec \boldsymbol{v}_2 \prec \boldsymbol{v}_3$, образующих цепь и принадлежащих одной и той же цепи C_i, дополнительная вершина принадлежит цепи C_j меньшей длины.

Доказательство. Лемму докажем индукцией по n. При $n = 2$ утверждение леммы очевидно — куб $\mathbb{B}^2$ можно покрыть одной цепью из трех вершин ($p = 0$) и одной цепью из одной вершины ($p = 1$), при этом свойство (ii) очевидно выполняется.

Допустим, что лемма верна для $n = k$. В кубе $\mathbb{B}^{k+1}$ рассмотрим два подмножества B_0 и B_1, состоящие из всех наборов, у которых $(k+1)$-е разряды равны, соответственно, нулю и единице. Каждое из множеств B_i изоморфно кубу $\mathbb{B}^k$ и поэтому в силу предположения индукции может быть покрыто непересекающимися цепями так, что число цепей длины $k-2p+1$, где $p = 0, 1, \ldots, \lfloor k/2 \rfloor$, равно $\binom{k}{p} - \binom{k}{p-1}$.

Рассмотрим одинаковые покрытия множеств B_1 и B_0, каждое из которых удовлетворяет условиям теоремы в кубе размерности k. Очевидно, что цепи этих покрытий не пересекаются и полностью покрывают куб $\mathbb{B}^{k+1}$. Пусть C_1 и C_0 — одинаковые цепи в рассматриваемых покрытиях множеств

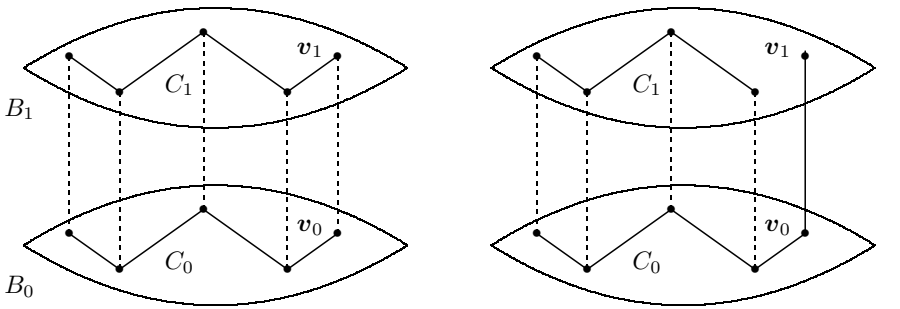


Рис. 7.1

B_1 и B_0, $\boldsymbol{v}_1$ и $\boldsymbol{v}_0$ — наибольшие элементы этих цепей. Очевидно, что $\boldsymbol{v}_0 \preceq \boldsymbol{v}_1$ и в наборах $\boldsymbol{v}_1$ и $\boldsymbol{v}_0$ все разряды кроме последнего совпадают. Рассматриваемые объекты представлены в левой части рисунка 7.1. На этом рисунке

штриховыми линиями изображены ребра, идущие в $(k+1)$-м направлении. Преобразуем цепи C_1 и C_0 следующим образом: набор $\boldsymbol{v}_1$ удалим из цепи C_1 и добавим к цепи C_0. В результате из двух цепей длины $k-2p+1$ получим цепь длины $k-2p=(k+1)-2p-1$ и цепь длины $k-2p+2=(k+1)-2(p-1)-1$. Новые цепи изображены в правой части рисунка 7.1. Выполним подобные преобразования над всеми парами одинаковых цепей в покрытиях B_1 и B_0.

Очевидно, что новые цепи покрывают $\mathbb{B}^{k+1}$ и не пересекаются. При этом цепь длины $(k+1)-2p-1$ получается либо удалением наибольшего набора из цепи длины $k-2(p-1)-1$, либо добавлением нового набора в цепь длины $k-2p-1$. Поэтому в преобразованном покрытии число цепей длины $(k+1)-2p-1$ равно

$$\left(\binom{k}{p}-\binom{k}{p-1}\right)+\left(\binom{k}{p-1}-\binom{k}{p-2}\right)=$$
$$=\left(\binom{k}{p}+\binom{k}{p-1}\right)-\left(\binom{k}{p-1}+\binom{k}{p-2}\right)=\binom{k+1}{p}-\binom{k+1}{p-1}.$$

Утверждение (i) доказано.

Докажем второе утверждение леммы. Прежде всего заметим, что если вершины $\boldsymbol{v}_1 \prec \boldsymbol{v}_2 \prec \boldsymbol{v}_3$ лежат в одном и том же подмножестве B_i, то до описанных выше преобразований они лежали на одной и той же цепи C и, в силу предположения индукции, их дополнительная вершина $\boldsymbol{u}$ лежала в B_i на цепи, длина которой меньше длины C не менее чем на два. Поэтому очевидно, что после преобразования цепей вершина $\boldsymbol{u}$ по прежнему будет лежать на цепи меньшей длины. Если же вершины $\boldsymbol{v}_1 \prec \boldsymbol{v}_2 \prec \boldsymbol{v}_3$ лежат в

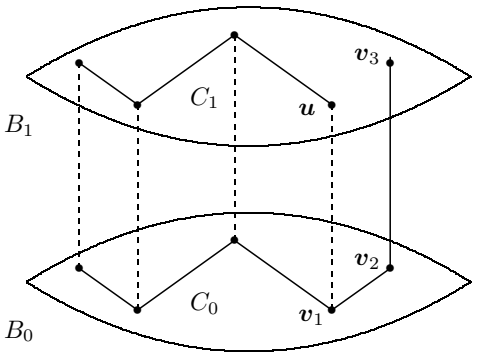


Рис. 7.2

разных подмножествах B_i, то (см. рис. 7.2) $\boldsymbol{v}_1$ и $\boldsymbol{v}_2$ принадлежат B_0, а $\boldsymbol{v}_3$ принадлежит B_1. В этом случае $\boldsymbol{v}_2$ и $\boldsymbol{v}_3$ отличаются в последней координате, и поэтому вершина $\boldsymbol{u}$ отличается от $\boldsymbol{v}_2$ также в последней координате, и, следовательно, лежит на цепи меньшей длины. Лемма доказана.

Теперь можно доказать неравенство теоремы 7.2. Сделаем это, оценив число различных способов, которыми можно задать значения n-местной монотонной булевой функции f на вершинах из $\mathbb{B}^n$. Сначала зададим значения функции на всех вершинах, которые принадлежат построенным в доказательстве леммы 7.1 цепям минимальной длины. Так как в силу леммы 7.1 длина минимальных цепей не превосходит двух, то на вершинах одной цепи значения монотонной функции можно задать не более чем тремя разными способами. Далее предположим, что значения f определены на всех вершинах, которые принадлежат цепям из k или меньшего числа вершин.

Рассмотрим цепь $\boldsymbol{v}_0 \prec \boldsymbol{v}_1 \prec \ldots \prec \boldsymbol{v}_{k+1}$ из $k+2$ вершин и множество вершин $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$, в котором каждая вершина $\boldsymbol{u}_i$ является дополнительной вершиной для $\boldsymbol{v}_{i-1}, \boldsymbol{v}_i, \boldsymbol{v}_{i+1}$, т. е. $\boldsymbol{v}_{i-1} \prec \boldsymbol{u}_i \prec \boldsymbol{v}_{i+1}$. Из леммы 7.1 и сделанного предположения следует, что значения функции f на вершинах $\boldsymbol{u}_i$ определены. Пусть s — максимальный индекс, для которого $f(\boldsymbol{u}_s)=0$, t — минимальный индекс, для которого $f(\boldsymbol{u}_t)=1$. Тогда $f(\boldsymbol{u}_{s+1})=1$, и, следовательно, $s+1 \geq t$. Поэтому

$$f(\boldsymbol{v}_{t-2}) \leq f(\boldsymbol{v}_{s-1}) \leq f(\boldsymbol{u}_s) = 0 < 1 = f(\boldsymbol{u}_t) \leq f(\boldsymbol{v}_{t+1}).$$

Таким образом, $f(\boldsymbol{v}_{t-2})=0$ и $f(\boldsymbol{v}_{t+1})=1$, и для того чтобы полностью определить значения функции f на вершинах рассматриваемой цепи, достаточно определить значения $f(\boldsymbol{v}_{t-1})$ и $f(\boldsymbol{v}_t)$. Очевидно, что сделать это можно тремя способами.

Так как общее число цепей равно $\binom{n}{\lfloor n/2 \rfloor}$ и значения функции f на вершинах каждой цепи можно задать не более чем тремя разными способами, то число монотонных n-местных булевых функций не превосходит $3^{\binom{n}{\lfloor n/2 \rfloor}}$. Теорема доказана.

В заключение приведем без доказательства формулу для асимптотики логарифма числа монотонных функций в $P_2(n)$

$$\log_2 |M(n)| \sim \binom{n}{\lfloor \frac{n}{2} \rfloor} \sim \frac{2^n}{\sqrt{\pi n/2}},$$

которая была установлена Д. Клейтменом в 1969 г.

7.3 Критерий полноты системы булевых функций

Система булевых функций $F=\{f_1, f_2, \ldots, f_i, \ldots\}$ называется *функционально полной*, если любая булева функция может быть реализована формулой в базисе F. Так как каждая булева функция реализуется своими совершенной дизъюнктивной нормальной формой и алгебраической нормальной формой, то системы функций $\{\&, \vee, \neg\}$ и $\{\&, \oplus, 1\}$ будут полными. Для установления полноты других систем можно воспользоваться следующей теоремой.

Теорема 7.2. *Пусть F и G — системы булевых функций, система F полная и каждая ее функция реализуется формулой в базисе G. Тогда G — полная система.*

Доказательство. Покажем, что произвольная булева функция h может быть реализована формулой в базисе G. Сделаем это индукцией по сложности формул, реализующих булевы функции в базисе F. В основание индукции положим формулы нулевой сложности, т. е. переменные. Далее предположим, что любая функция, реализуемая в базисе F формулой сложности $L-1$, может быть реализована формулой в базисе G. Пусть h — произвольная булева функция, для которой существует реализующая ее в базисе F формула H сложности L. Тогда $\mathsf{H} = f(\mathsf{H}_1, \ldots, \mathsf{H}_k)$, где $f \in F$ и $\mathsf{H}_1, \ldots, \mathsf{H}_k$ формулы в базисе F. Очевидно, что сложность каждой формулы H_i меньше сложности формулы H. Поэтому по предположению индукции реализуемые формулами $\mathsf{H}_1, \ldots, \mathsf{H}_k$ функции $h_1, \ldots, h_k$ реализуются также формулами $\mathsf{H}'_1, \ldots, \mathsf{H}'_k$ в базисе G. По условию теоремы функция $f(y_1, \ldots, y_k)$ реализуется формулой F в базисе G. В этой формуле каждую переменную y_i заменим формулой H'_i. Полученная формула H' реализует функцию h и является формулой в базисе G. Теорема доказана.

Из доказанной теоремы и законов двойственности немедленно следует полнота систем $\{\&, \neg\}$ и $\{\vee, \neg\}$. Также нетрудно видеть, что полной будет система $\{x \to y, \overline{x}\}$. Так как $x \to y = \overline{x} \vee y$, то $\overline{x} \to y = x \vee y$, т. е. дизъюнкция двух переменных реализуется формулой в базисе из импликации и отрицания. Следовательно, в силу предыдущей теоремы система $\{x \to y, \overline{x}\}$ является полной.

Следующая теорема, которая была доказана Э. Постом в первой половине двадцатого века, называется критерием Поста.

Теорема 7.3. *Для того чтобы система функций F была полной, необходимо и достаточно, чтобы она не содержалась целиком ни в одном из пяти замкнутых классов T_0, T_1, S, M и L.*

Доказательство. Необходимость. Пусть система функций F полна в P_2. Предположим, что F целиком содержится в замкнутом классе $R \in \{T_0, T_1, L, S, M\}$. Тогда из свойств операции замыкания, включения $F \subseteq R$ и равенства $[F] = P_2$ следует, что $P_2 = [F] = R$. Поэтому $R = P_2$. С другой стороны каждый из классов T_0, T_1, L, S, M отличен от P_2, т. е. $R \neq P_2$. Противоречие. Необходимость доказана.

Достаточность. Так как система функций F не содержится целиком ни в одном из пяти замкнутых классов перечисленных в условии теоремы, то в этой системе найдутся пять (не обязательно различных) функций f_{T_0}, f_{T_1}, f_S, f_M и f_L таких, что

$$f_{T_0} \notin T_0, \quad f_{T_1} \notin T_1, \quad f_S \notin S, \quad f_M \notin M, \quad f_L \notin L.$$

Если $f_{T_0}(1, \ldots, 1) = 0$, то $f_{T_0}(x, \ldots, x) = \overline{x}$. Если $f_{T_1}(0, \ldots, 0) = 1$, то $f_{T_1}(x, \ldots, x) = \overline{x}$. Если же $f_{T_0}(1, \ldots, 1) = 1$ и $f_{T_1}(0, \ldots, 0) = 0$, то

$$f_{T_0}(x, \ldots, x) = 1, \quad f_{T_1}(x, \ldots, x) = 0.$$

Следовательно, после отождествления переменных у функций f_{T_0} и f_{T_1} получаем либо (i) отрицание, либо (ii) две тождественные константы 0 и 1.

Последовательно рассмотрим эти возможности.

(i) Так как f_S не является самодвойственной, то найдутся такие противоположные наборы $\alpha_1, \ldots, \alpha_k$ и $\overline{\alpha}_1, \ldots, \overline{\alpha}_k$, что

$$f_S(\alpha_1, \ldots, \alpha_k) = f_S(\overline{\alpha}_1, \ldots, \overline{\alpha}_k).$$

Рассмотрим функцию $\varphi_1(x) = f_S(x^{\alpha_1}, \ldots, x^{\alpha_k})$. Легко видеть, что

$$\begin{aligned}\varphi_1(0) &= f_S(0^{\alpha_1}, \ldots, 0^{\alpha_k}) = f_S(\overline{\alpha}_1, \ldots, \overline{\alpha}_k) = \\ &= f_S(\alpha_1, \ldots, \alpha_k) = f_S(1^{\alpha_1}, \ldots, 1^{\alpha_k}) = \varphi_1(1),\end{aligned}$$

т. е. функция φ_1 является константой. Вторая константа получается из φ_1 и отрицания.

(ii) Так как f_M не является монотонной, то найдутся такие соседние наборы $\boldsymbol{a} = (\alpha_1, \ldots, \alpha_k)$ и $\boldsymbol{b} = (\beta_1, \ldots, \beta_k)$, что $\boldsymbol{a} \prec \boldsymbol{b}$ и

$$f_M(\alpha_1, \ldots, \alpha_k) > f_M(\alpha_1, \ldots, \alpha_k).$$

При каждом $1 \leq i \leq k$ определим функцию

$$g_i(x) = \begin{cases} 0, & \text{если } \alpha_i = \beta_i = 0; \\ 1, & \text{если } \alpha_i = \beta_i = 1; \\ x, & \text{если } \alpha_i = 0,\ \beta_i = 1. \end{cases}$$

Рассмотрим функцию $\varphi_2(x) = f_M(g_1(x), \ldots, g_k(x))$. Легко видеть, что

$$\varphi_2(0) = f_M(\alpha_1, \ldots, \alpha_k) > f_M(\beta_1, \ldots, \beta_k) = \varphi_2(1),$$

т. е. φ_2 — отрицание.

Таким образом, из функций f_{T_0}, f_{T_1}, f_S и f_M получены константы и отрицание.

Теперь, рассмотрим нелинейную функцию f_L. В многочлене Жегалкина этой функции найдется слагаемое, содержащее не менее двух переменных. Без ограничения общности будем полагать, что это x_1 и x_2. Группируя слагаемые, содержащие x_1, x_2 и x_1x_2, преобразуем многочлен Жегалкина функции f_L к виду

$$\begin{aligned}f_L(x_1, \ldots, x_k) =& x_1x_2f_0(x_3, \ldots, x_k) \oplus \\ &\oplus x_1f_1(x_3, \ldots, x_k) \oplus x_2f_2(x_2, \ldots, x_k) \oplus f_3(x_2, \ldots, x_k),\end{aligned}$$

где функция f_0 отлична от тождественного нуля. Далее воспользуемся равенством $x \oplus 1 = \overline{x}$. Пусть набор $(\alpha_3, \ldots, \alpha_k)$ таков, что $f_0(\alpha_3, \ldots, \alpha_k) = 1$. Введем функцию

$$\psi(x_1, x_2) = f_L(x_1, x_2, \alpha_3, \ldots, \alpha_k) = x_1x_2 \oplus \gamma_1x_1 \oplus \gamma_2x_2 \oplus \gamma_3,$$

где $\gamma_i = f_i(\alpha_3, \ldots, \alpha_k)$ — константы. Положим

$$\varphi_3(x_1, x_2) = \psi(x_1 \oplus \gamma_2, x_2 \oplus \gamma_1) \oplus \gamma_1\gamma_2 \oplus \gamma_3.$$

Легко видеть, что

$$\begin{aligned}\varphi_3(x_1,x_2) &= \psi(x_1\oplus\gamma_2, x_2\oplus\gamma_1)\oplus\gamma_1\gamma_2\oplus\gamma_3 = \\ &= (x_1\oplus\gamma_2)(x_2\oplus\gamma_1)\oplus\gamma_1(x_1\oplus\gamma_2)\oplus \\ &\quad\oplus\gamma_2(x_2\oplus\gamma_1)\oplus\gamma_3\oplus\gamma_1\gamma_2\oplus\gamma_3 = x_1x_2\end{aligned}$$

Таким образом, $\varphi_3(x_1,x_2)=x_1x_2$. Теперь утверждение теоремы следует из теоремы 7.2 и полноты системы функций $\{\&,\neg\}$.

Используя теорему 7.3, исследуем полноту системы $\{x\to y, xy\}$. Так как $1\to 1=1$ и $1\cdot 1=1$, то очевидно, что импликация и конъюнкция сохраняют единицу. Следовательно, система $\{x\to y, xy\}$ содержится в классе T_1 и в силу теоремы 7.3 не является полной.

Пусть F — замкнутый класс в P_2. Система булевых функций $\{f_i\}$ называется *базисом* в F, если ее замыкание совпадает с F, а любая ее собственая подсистема не является полной в F. Нетрудно видеть, что система функций $\{x\to y, \overline{x}\}$ является базисом в P_2. Эта система полна в P_2, так как ее первая функция нелинейная и несамодвойственная, а вторая не является монотонной и не сохраняет константы. Другая полная в P_2 система функций $\{\&,\vee,\neg\}$ базисом не является, так как две ее собственные подсистемы $\{\&,\neg\}$ и $\{\vee,\neg\}$ полны в P_2.

Булева функция f называется *шефферовой*, если $[f]=P_2$. Очевидно, что каждая шефферова функция является базисом в P_2.

Найдем все шефферовы функции в $P_2(2)$. Допустим, что f — шефферова, и $\mathbf{f}=(f_1,f_2,f_3,f_4)$ — вектор ее значений. Так как $f\notin T_0$ и $f\notin T_1$, то $f_1=1$ и $f_4=0$. Отсюда немедленно следует, что $f\notin M$. Теперь посмотрим, какие значения могут принимать вторая и третья компоненты вектора $\mathbf{f}$. Если $f_2\neq f_3$, то либо $\mathbf{f}=(1100)$, либо $\mathbf{f}=(1010)$. В обоих случаях f будет линейной самодвойственной функцией. Если $f_2=f_3$, то либо $\mathbf{f}=(1000)$, либо $\mathbf{f}=(1110)$. Непосредственной проверкой легко убедиться, что каждый из этих векторов задает нелинейную и несамодвойственную функцию. Таким образом, в $P_2(2)$ содержится ровно две шефферовы функции: стрелка Пирса и штрих Шеффера.

7.4 Задачи

7.1. Найти число монотонных функций в $P_2(3)$.

7.2. Показать, что функция двойственная к монотонной также будет монотонной.

7.3. Найти число монотонных булевых функций, зависящих от n переменных и принимающих единичное значение ровно на 7 наборах.

7.4. Пусть функция $f(x_1,\dots,x_n)$ принадлежит множеству $[x\to y]$ и существенно зависит не менее чем от двух переменных. Доказать, что она принимает единичные значения более чем на 2^{n-1} наборах.

7.5. Будет ли множество симметрических булевых функций замкнутым?

7.6. Найти

a) $|L(n)\cup T_0(n)\cup T_1(n)|$; b) $|L(n)\cap S(n)\cap T_1(n)|$;
c) $|L(n)\cap M(n)\cup T_0(n)|$; d) $|(S(n)\cup T_0(n))\setminus L(n)|$;
e) $|(S(n)\cup T_0(n)\cup T_1(n))\cap L(n)|$; f) $|S(n)\setminus(T_0(n)\setminus L(n))|$;
g) $|S(n)\cup L(n)\cup T_0(n)\cup T_1(n)|$; h) $|(L(n)\setminus S(n))\setminus M(n)|$.

7.7. Выяснить, является ли множество A базисом в B:

a) $A=\{xy\sim z\},\ B=T_1$;
b) $A=\{xy\vee z\},\ B=T_0$;
c) $A=\{x\sim y, x\oplus y\},\ B=L$;
d) $A=\{x_1\oplus x_2\oplus\dots\oplus x_k,1\}, k$ — константа, $B=L$.

7.8. Выяснить, при каких n функция f является шефферовой:

a) $f(x_1,\dots,x_n)=1\oplus x_1x_2\oplus\dots\oplus x_ix_{i+1}\oplus\dots\oplus x_{n-1}x_n\oplus x_nx_1$;
b) $f(x_1,\dots,x_n)=1\oplus x_1x_2\oplus\dots\oplus x_ix_{i+1}\oplus\dots\oplus x_{n-1}x_n$;
c) $f(x_1,\dots,x_n)=1\oplus\bigoplus_{1\le i<j\le n}x_ix_j$;
d) $f(x_1,\dots,x_n)=1\oplus(x_1|x_2)\oplus\dots\oplus(x_i|x_{i+1})\oplus\dots\oplus(x_{n-1}|x_n)\oplus(x_n|x_1)$;
e) $f(x_1,\dots,x_n)=1\oplus(x_1|x_2)\oplus\dots\oplus(x_i|x_{i+1})\oplus\dots\oplus(x_{n-1}|x_n)$;
f) $f(x_1,\dots,x_n)=1\oplus(x_1\to x_2)\oplus\dots\oplus(x_i\to x_{i+1})\oplus\dots\oplus(x_{n-1}\to x_n)$.

7.9. Указать какой-либо базис множества самодвойственных функций. Существует ли такая функция f, что $[f]=S$?

7.10. Доказать, что если f монотонна и существенно зависит не менее чем от двух переменных, то система $\{0,\overline{f}\}$ полна в P_2.

7.11. Набор $\boldsymbol{a}$ из B^n назовем нижней единицей монотонной функции f, если $f(\boldsymbol{a})=1$ и $f(\boldsymbol{b})=0$ для каждого $\boldsymbol{b}\prec\boldsymbol{a}$. Пусть монотонная функция f имеет ровно две нижние единицы. Доказать, что $\overline{f}$ — шефферова.

7.12. Пусть $\boldsymbol{a}$ — цепь в $\mathbb{B}^n$, $\boldsymbol{b}$ — цепь в $\mathbb{B}^m$. Показать, что на прямом произведении $\boldsymbol{a}\times\boldsymbol{b}$ можно определить ровно $\binom{|\boldsymbol{a}|+|\boldsymbol{b}|}{|\boldsymbol{a}|}$, где $|\boldsymbol{a}|$ и $|\boldsymbol{b}|$ — длины цепей, различных частичных монотонных булевых функций из $P_2(n+m)$.

7.13. Используя предыдущую задачу показать, что $|M(n)|<8^{\binom{n}{\lfloor n/2\rfloor}}$.

7.14. Найти число шефферовых функций в:

a) $P_2(3)$; b) $P_2(4)$; c) $P_2(5)$; d) $P_2(n)$.

Лекция 8

Сложность булевых функций

Любое вычисление можно представить в виде последовательности шагов, каждый из которых состоит в выполнении некоторого простого действия над исходными данными или над величинами, полученными на предыдущих шагах. Список команд, описывающих эти шаги и определяющих порядок их выполнения, обычно называется программой или схемой вычисления. Как правило, программы состоят из команд двух видов: вычислительных и управляющих. Вычислительные команды производят некоторые безусловные действия, например, складывают числа. Управляющие команды определяют порядок выполнения вычислительных команд. К таким командам, в частности, относится команда условного перехода. Число шагов, выполняемых в процессе вычисления, называется его сложностью. Если один и тот же объект, например булева функция, может быть вычислен различными способами, то его сложностью называется сложность самого простого вычисления.

Вычисления, программы которых состоят только из вычислительных команд, называются неветвящимися. В любом неветвящемся вычислении команды выполняются последовательно одна за другой в том порядке, в котором они расположены в программе. Основными математическими моделями неветвящихся вычислений являются неветвящиеся программы и схемы из функциональных элементов. Неветвящиеся программы моделируют работу универсальных вычислительных устройств, способных решать различные задачи в зависимости от программ, под управлением которых эти устройства работают в данный момент времени. Схемы из функциональных элементов являются моделями электронных схем, программы работы которых заложены в их конструкции и никогда не меняются. Несмотря на большие различия моделируемых объектов неветвящиеся программы и схемы из функциональных элементов очень похожи друг на друга и многие результаты, полученные для неветвящихся программ, легко переносятся на схемы из функциональных элементов и наоборот.

Далее рассматривается сложность вычисления значений булевых функций при помощи неветвящихся программ и схем из функциональных элементов.

8.1 Программы и схемы

Неветвящейся программой P над базисом B и множеством независимых переменных $\{x_1, \ldots, x_n\}$ называется последовательность равенств

$$y_i = f_i(a, b), \quad i = 1, \ldots, r, \tag{8.1}$$

где $f_i \in B$, $a, b \in \{x_1, \ldots, x_n, y_i, \ldots, y_{i-1}\}$. Величины y_i называются внутренними переменными программы P, индекс i — номером переменной y_i. Переменные a и b назовем, соответственно, первым и вторым предками переменной y_i, а переменную y_i — потомком переменных a и b. Число r назовем сложностью программы P. Индукцией по номеру внутренней переменной определим ее значение на наборе значений $\sigma_1, \ldots, \sigma_n$ независимых переменных $x_1, \ldots, x_n$. Для этого положим

$$y_i(\sigma_1, \ldots, \sigma_n) = f_i(a(\sigma_1, \ldots, \sigma_n), b(\sigma_1, \ldots, \sigma_n)).$$

Так как $a, b \in \{x_1, \ldots, x_n, y_i, \ldots, y_{i-1}\}$, то значение переменной y_i определено корректно. Будем говорить, что переменная y_i вычисляет функцию h, если равенство

$$y_i(\sigma_1, \ldots, \sigma_n) = h(\sigma_1, \ldots, \sigma_n)$$

справедливо для любого набора $\sigma_1, \ldots, \sigma_n$. В программе P выделим упорядоченный набор $z_1, \ldots, z_m$, состоящий из внутренних и независимых переменных. Будем говорить, что программа P вычисляет систему функций $h_1, \ldots, h_m$, если $z_i \equiv h_i$ для каждого $i = 1, \ldots, m$. Переменные $z_1, \ldots, z_m$ будем называть выходами программы.

По неветвящейся программе P можно построить ориентированный граф G, описывающий работу этой программы. Для этого каждой переменной программы поставим в соответствие собственную вершину графа. Затем проведем ребра, выходящие из вершин, соответствующих переменным–предкам, и входящие в вершины, соответствующие переменным–потомкам. Ребру, соединяющему i-го ($i = 1, 2$) предка с его потомком, припишем число i. Каждой вершине, соответствующей независимой переменной x_i, припишем символ x_i, а вершине, соответствующей внутренней переменной y_j, припишем символ f_j. Наконец каждой вершине, соответствующей какой-либо переменной z_k из выделенного набора $\{z_i\}$, припишем число k. Получившийся граф называется схемой программы P. В качестве примера приведем программу

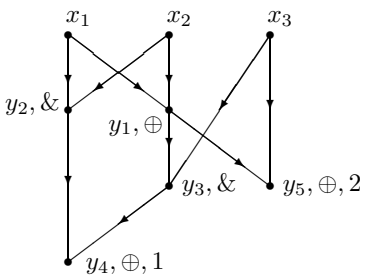


Рис. 8.1

$$\mathsf{P}: \quad y_1 = x_1 \oplus x_2, \quad y_2 = x_1 \& x_2, \quad y_3 = y_1 \& x_3, \quad y_4 = y_2 \& y_3, \quad y_5 = y_1 \oplus x_3,$$

в которой $z_1 = y_4$ и $z_2 = y_5$, и которая, как нетрудно проверить, вычисляет функции $x_1x_2 \oplus x_1x_3 \oplus x_2x_3$ и $x_1 \oplus x_2 \oplus x_3$. Построенная по программе P схема изображена на рис. 8.1, где предполагается, что первый предок каждой вершины находится левее второго предка, и поэтому отсутствует нумерация ребер, входящих в одну вершину.

Нетрудно видеть, что схема программы полностью описывает все вычисления, выполняемые программой. Тем не менее, говорить о полном соответствии между программами и их схемами нельзя, так как переменные в неветвящейся программе упорядочены линейно, а на множестве соответствующих им вершин схемы определен только частичный порядок.

Описание вычислений при помощи схем представляет независимый (от соответствующих им неветвящихся программ) интерес по ряду причин. Одна из главных заключается в тех возможностях, которые развитый язык теории графов представляет для анализа вычислений. Поэтому дадим независимое определение схемы.

Схемой из функциональных элементов (или *булевой схемой*) с n входами и m выходами называется ориентированный граф S, который не содержит контуров и обладает следующими свойствами:

– S содержит n вершин с входной степенью равной нулю. Такие вершины называются входами схемы.

– Входные степени остальных вершин не превосходят двух. Эти вершины называются элементами схемы. Ребра, входящие в один элемент, нумеруются числами от 1 до 2. Каждому элементу схемы с входной степенью равной j ($j = 0, 1, 2$) приписана j-местная функция из B. Множество B называется базисом схемы.

– m вершин помечены целыми числами от 1 до m. Эти вершины называются выходами схемы.

Если вершины w и u схемы S связаны ребром, ориентированным от u к w, то будем говорить, что вершина w подключена к вершине u. Вершину u будем называть предком вершины w, а вершину w — потомком вершины u. Если вершина w имеет двух предков и номер ребра (uw) равен единице, то вершину u будем называть первым предком вершины w, если номер (uw) равен двойке — вторым предком этой вершины.

Если вершине u схемы S приписана функция h, то будем говорить, что вершина u *реализует* функцию h. Для каждой вершины w этой схемы определим функцию $S(\mathsf{w})$, *вычисляемую* в вершине w. Сделаем это индуктивно в соответствии со следующими правилами:

– если w — вход схемы, которому приписана независимая переменная x_i, то $S(\mathsf{w}) = x_i$;

– если вершине w приписана нульместная функция f, то $S(\mathsf{w}) = f$;

– если вершине w приписана одноместная функция f и в вершине u — предке вершины w, вычисляется функция $S(\mathsf{u})$, то $S(\mathsf{w}) = f(S(\mathsf{u}))$;

– если вершине w приписана двухместная функция f и в вершине u — первом предке вершины w, вычисляется функция $S(\mathsf{u})$, а в вершине v — втором предке вершины w, вычисляется функция $S(\mathsf{v})$, то $S(\mathsf{w}) = f(S(\mathsf{u}), S(\mathsf{v}))$.

Будем говорить, что схема S вычисляет систему булевых функций

$$h_i(x_1, \ldots, x_n), \quad i = 1, \ldots, m,$$

если входы схемы S реализуют переменные $x_1, \ldots, x_n$, а в выходах u_i ($i = 1, \ldots, m$) вычисляются функции h_i, т. е. $S(\mathsf{u}_i) = h_i$.

При изображении схем как правило будем придерживаться следующих правил. 1) Предок всегда располагается выше потомка, т. е. все ребра ориентированы сверху вниз. Поэтому стрелки, указывающие ориентацию ребер, отсутствуют. 2) Элементы схем изображены треугольниками, в середине каждого треугольника помещен символ функции, реализуемой элементом. Если входная степень элемента, изображенного треугольником, равна двум, то ребра, связывающие элемент с его предками, присоединяются к разным точкам одной стороны треугольника — точка присоединения ребра от первого предка располагается левее точки присоединения ребра от второго предка. При таком изображении элементов отпадает необходимость в расстановке номеров ребер — нумерация однозначно восстанавливается по изображению. 3) Каждая вершина, являющаяся выходом схемы, отмечается полуребром, выходящим из этой вершины, и не входящим ни в какую другую вершину.

В соответствии с этими правилами на рисунке 8.2 изображены схемы S_1 и S_2, каждая из которых вычисляет функцию $x_1 \oplus x_2$. Вершины обеих

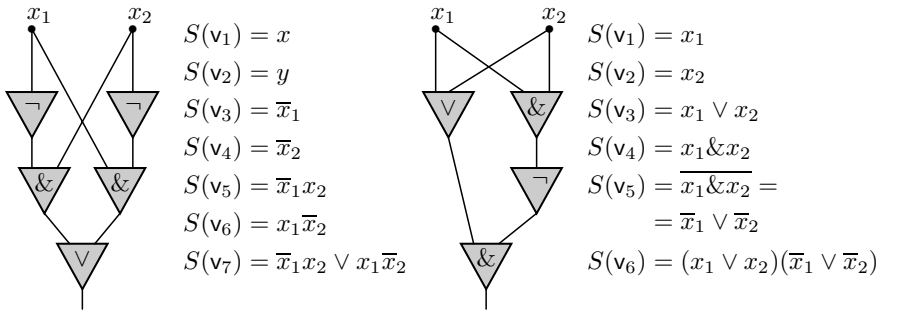


Рис. 8.2

схем перенумерованы слева направо–сверху вниз целыми числами, начиная с единицы. Функции, вычисляемые в вершинах каждой из этих схем, перечислены в столбце, стоящем рядом со схемой.

Выше было отмечено, что единственное отличие неветвящихся программ от схем заключается в том, что внутренние переменные программ линейно упорядочены, а на множестве соответствующих им элементов схем определен лишь частичный порядок. Поэтому разным программам может соответствовать одна схема. Так, например, изображенная в левой части рисунка 8.2 схема S_1 соответствует не только программе

$$\mathsf{P}_1: \quad y_1 = \overline{x}_1, \quad y_2 = \overline{x}_2, \quad y_3 = y_1 \& x_2, \quad y_4 = y_2 \& x_1, \quad y_5 = y_3 \vee y_4,$$

в которой внутренние переменные следуют друг за другом в том же порядке, что и вершины схемы S_1, но и программам

$$\mathsf{P}_2: \quad y_1 = \overline{x}_2, \quad y_2 = \overline{x}_1, \quad y_3 = y_1 \& x_1, \quad y_4 = y_2 \& x_2, \quad y_5 = y_3 \vee y_4,$$
$$\mathsf{P}_3: \quad y_1 = \overline{x}_2, \quad y_2 = y_1 \& x_1, \quad y_3 = \overline{x}_1, \quad y_4 = y_2 \& x_2, \quad y_5 = y_3 \vee y_4,$$

которые отличаются от P_1 только порядком выполнения операций.

Далее, учитывая существующее соответствие между схемами и программами, будем рассматривать в основном схемы. При этом будем полагать, что элементы каждой схемы линейно упорядочены, считая, что порядок элементов совпадает с порядком внутренних переменных одной из соответствующих данной схеме программ.

8.2 Схемы

Важнейшими характеристиками любой схемы являются ее сложность и глубина. *Сложностью* $L(\mathsf{S})$ схемы S называется число элементов этой схемы. Для определения глубины схемы S рассмотрим различные ориентированные цепи, связывающие ее входы и выходы. Длиной цепи называется число элементов, через которые проходит эта цепь. Цепь, проходящую через максимальное число элементов, назовем максимальной цепью схемы S, а ее длину (число элементов) назовем *глубиной* $D(\mathsf{S})$ схемы S.

Если среди всех схем, имеющих базис B и вычисляющих систему f, схема S содержит наименьшее число элементов, то S называется *минимальной* (по сложности) схемой системы f. Число элементов в минимальной (по сложности) схеме системы f называется *сложностью системы функций* f и обозначается через $L_B(f)$.

Аналогичным образом для произвольной системы булевых функций определяются ее минимальная (по глубине) схема и ее глубина. Если среди всех схем, имеющих базис B и вычисляющих систему f, схема S имеет наименьшую глубину D, то S называется минимальной. Глубина минимальной (по глубине) схемы системы f называется *глубиной системы функций* f и обозначается символом $D_B(f)$.

Снова рассмотрим изображенные рисунке 8.2 схемы, вычисляющие функцию $x_1 \oplus x_2$ в базисе $\{\&, \vee, \neg\}$. Сложность левой схемы S_1 равна пяти, сложность правой схемы S_2 — четырем. Глубина обеих схем равна трем. В левой схеме есть две цепи длины три. Одна из них начинается во входе x_1 и проходит через левый элемент отрицания, левый элемент дизъюнкции и левый элемент конъюнкции[1)]. В правой схеме цепь длины три начинается во входе x_2 и проходит через два конъюнктора и элемент отрицания. Очевидно, что схема S_1 не является минимальной по сложности схемой для функции $x_1 \oplus x_2$ среди схем с базисом $\{\&, \vee, \neg\}$ так как схема S_2 имеет такой же базис, вычисляет такую же функцию что и S_1 и при этом состоит из меньшего числа элементов.

[1)] Часто элемент дизъюнкции называется дизъюнктором, элемент конъюнкции — конъюнктором, элемент отрицания — инвертором.

Любой подграф G схемы S, являющийся схемой[2)], будем называть подсхемой схемы S. При построении больших схем, вычисляющих сложные функции, часто бывает удобно сначала построить схемы, вычисляющие некоторые вспомогательные функции, а потом, превратив их в подсхемы, собрать из них требуемую схему S. Подсхемы с более чем одним выходом будем изображать прямоугольниками, а подсхемы с одним выходом — треугольниками большого размера. Внутрь фигуры, изображающей подсхему, будем помещать либо символ функции, вычисляемой подсхемой, либо символ–имя подсхемы.

Используя понятие подсхемы, покажем, что любая булева функция n переменных может быть вычислена такой схемой S в базисе $\{\vee, \&, \neg\}$, что

$$L(\mathsf{S}) \le 3 \cdot 2^n - 4, \quad D(\mathsf{S}) \le 2n. \tag{8.2}$$

Сделаем это индукцией по числу переменных n. В основание индукции положим схемы, вычисляющие функции одной переменной. Самыми сложными и глубокими будут вычисляющие константы 0 и 1 схемы S_0 и S_1. Так как $0 = x\&\overline{x}$ и $1 = x \vee \overline{x}$, то легко видеть, что $L(\mathsf{S}_i) = 2$ и $D(\mathsf{S}_i) = 2$, т. е. неравенства (8.2) справедливы при $n = 1$. Теперь допустим, что эти неравенства справедливы при всех n, не превосходящих некоторое целое $k \ge 1$. Покажем, что тогда неравенства (8.2) имеют место и при $n = k + 1$. Для этого построим схему, вычисляющую произвольную булеву функцию $f(x_1, \ldots, x_n)$, и оценим ее сложность и глубину. Функцию f разложим по последней переменной:

$$f(x_1, \ldots, x_n) = \overline{x}_n f(x_1, ..., x_{n-1}, 0) \vee x_n f(x_1, ..., x_{n-1}, 1).$$

В соответствием с этим разложением построена изображенная на рисунке 8.3 схема S. Эта схема вычисляет функцию f и состоит из двух подсхем A и B, одного элемента отрицания, двух конъюнкторов и одного дизъюнктора. Подсхема A вычисляет функцию $f_1 = f(x_1, ..., x_{n-1}, 1)$, подсхема B — функцию $f_0 = f(x_1, ..., x_{n-1}, 0)$, элемент отрицания вычисляет $\overline{x}_n$. Далее два конъюнктора умножают функции f_0 и f_1, соответственно на $\overline{x}_n$ и x_n. Затем дизъюнктор вычисляет дизъюнкцию двух произведений. Легко видеть, что сложность и глубина схемы S выражаются через сложности и глубины подсхем A и B следующим образом:

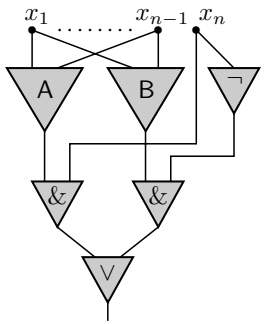


Рис. 8.3

$$L(\mathsf{S}) = L(\mathsf{A}) + L(\mathsf{B}) + 4,$$
$$D(\mathsf{S}) = \max(D(\mathsf{A}), D(\mathsf{B})) + 2.$$

[2)] Возможно в G придется добавить новые входы, если предки каких-либо элементов из G отсутствуют в G, и определить новые выходы, которыми станут те элементы, потомки которых отсутствуют в G.

По предположению индукции каждая из подсхем A и B состоит не более чем из $3\cdot 2^{n-1}-4$ элементов, а их глубина не превосходит $2(n-1)$. Поэтому,

$$L(\mathsf{S}) \le (3\cdot 2^{n-1}-4)+(3\cdot 2^{n-1}-4)+4 = 3\cdot 2^n-4,$$
$$D(\mathsf{S}) \le 2(n-1)+2 = 2n.$$

Неравенства (8.2) доказаны.

Так как система функций $\{\&,\vee,\neg\}$ полна в P_2, то из (8.2) следует существование таких констант c_1 и c_2, что

$$L_B(f) \le c_1\cdot 2^n, \quad D_B(f) \le c_2 n \tag{8.3}$$

для каждой n-местной функции f и любого базиса B. Поэтому существует простой алгоритм, позволяющий для произвольной булевой функции находить ее сложность, глубину и соответствующие минимальные схемы. Достаточно перебрать все схемы, сложность и глубина которых не превосходят правых частей неравенств (8.3), и выбрать среди них те, которые вычисляют f и имеют минимальные значения сложности и глубины. Однако число различных схем с увеличением числа входов растет так стремительно, что даже использование вычислительной техники не позволяет надеяться на успешное применение какого-либо переборного алгоритма уже при $n=8$. Поэтому при изучении сложности булевых функций используются иные методы.

Установим простейшие нижние оценки сложности и глубины булевых функций, которые справедливы при их вычислении схемами в любом базисе.

Теорема 8.1. *Если булева функция f существенно зависит от n переменных, то $L(f)\ge n-1$.*

Доказательство. Пусть S — минимальная схема функции f. Число элементов схемы S обозначим через L, а число ребер — через N. Величину N оценим двумя способами. С одной стороны, каждое ребро входит в какой-нибудь элемент. Поэтому $N\le 2L$. С другой стороны, из каждой вершины схемы S, кроме последней, выходит хотя бы одно ребро. Так как в S содержится $L+n$ вершин, то $N\ge L+n-1$. Следовательно, $2L\ge L+n-1$ или $L\ge n-1$. Теорема доказана.

Для доказательства нижней оценки глубины потребуется следующее вспомогательное утверждение.

Лемма 8.1. *Если схема S имеет один выход и любой ее элемент связан ориентированной цепью с этим выходом, то $D(\mathsf{S})\ge \log_2(L(\mathsf{S})+1)$.*

Доказательство. Лемму докажем индукцией по числу вершин схемы. В основание индукции положим очевидный случай схем, состоящих из одного элемента. Допустим, что для любой схемы с одним выходом и не более чем $L-1$ элементами теорема верна. Пусть S — схема из L элементов, v — последний элемент этой схемы, а u_1 и u_2 — предки v[3]. Функции, вычисляемые в вершинах u_1 и u_2, обозначим через φ_1 и φ_2. Пусть S_i — подсхема схемы S, которая вычисляет функцию φ_i. Нетрудно видеть, что S_1 и S_2 являются схемами, каждая из которых содержит не более $L-1$ элементов, и по крайней мере одна из этих частей, например S_1, состоит не менее чем из $\frac{1}{2}(L-1)$ элементов. По предположению индукции

$$D(\mathsf{S}_1)\ge \log_2(L(\mathsf{S}_1)+1)\ge \log_2\Big(\frac{1}{2}(L-1)+1\Big)=\log_2(L+1)-1.$$

Теперь осталось заметить, что глубина схемы S на единицу больше глубины схемы S_1. Лемма доказана.

Теперь из теоремы 8.1 и леммы 8.1 легко следует справедливость следующей теоремы.

Теорема 8.2. *Если булева функция f существенно зависит от n переменных, то $D(f)\ge \log_2 n$.*

Несмотря на свою простоту, и почти очевидность, неравенства теорем 8.1 и 8.2 в общем случае неулучшаемы. Для того, чтобы убедиться в этом достаточно рассмотреть например функцию $x_1\&\cdots\&x_n$, сложность и глубина которой в базисе $\{\vee,\&,\neg\}$ равны соответственно $n-1$ и $\lceil\log_2 n\rceil$.

Следует отметить, что доказательство нетривиальных нижних оценок сложности и глубины является очень трудной задачей для схем в полных базисах[4]. Так, например, лучшая из известных в настоящее время нижняя оценка сложности n-местных функций для базиса состоящего из всех не более чем двухместных булевых функций не превосходит $4n$. Более высокие нижние оценки удается доказывать только тогда, когда на базис или структуру схем накладываются существенные ограничения.

8.3 Свойства минимальных схем

Сформулируем и докажем несколько лемм о свойствах минимальных схем в базисе B_0, состоящем из всех не более чем двухместных булевых функций. Аналогичные утверждения нетрудно доказать и для других базисов. Первую очевидную лемму приведем без доказательства.

Лемма 8.2. *Если в схеме S есть две вершины, в которых вычисляется одна и та же функция, то одна из этих вершин может быть удалена из схемы, а сама схема S преобразована таким образом, что ее сложность уменьшится, вычисляемые ею функции не изменятся.*

Из леммы 8.2 следует, что в любой минимальной схеме в разных элементах вычисляются разные функции. Далее будем пользоваться этим свойством минимальных схем без ссылок на лемму 8.2.

[3] Случай, когда элемент v имеет одного предка рассматривается аналогично.

[4] Базис называется полным, если его замыкание совпадает с P_2.

Лемма 8.3. *Пусть* S *— минимальная схема в базисе* B_0*, у которой ни в одном из выходов не вычисляется тождественная константа. Тогда в* S *нет вершины, в которой реализуется тождественная константа.*

Доказательство. Пусть S — удовлетворяющая условиям леммы схема. Допустим, что в схеме S есть вершины, в которых вычисляется константы. Пусть s — самая нижняя такая вершина, и пусть в этой вершине вычисляется постоянная α. Рассмотрим фрагмент схемы содержащий s, элементы $\mathsf{s}_1, \ldots, \mathsf{s}_p$, — потомки вершины s, и вершины $\mathsf{u}_1, \ldots, \mathsf{u}_p$, — предки элементов $\mathsf{s}_1, ..., \mathsf{s}_p$. Функции, реализуемые элементами $\mathsf{s}_1, \ldots, \mathsf{s}_p$ будем обозначать через $v_1, \ldots, v_p$. Заметим, что в силу выбора вершины s, среди элементов $\mathsf{s}_1, \ldots, \mathsf{s}_p$ нет элементов, реализующих одноместные функции. Рассматриваемый фрагмент изображен в левой части рисунка 8.4. Преобразуем этот фрагмент следующим образом. Удалим элемент s, а элементы $\mathsf{s}_1, \ldots, \mathsf{s}_p$ заменим элементами $\mathsf{s}'_1, \ldots, \mathsf{s}'_p$, реализующими такие одноместные функции $v'_1, \ldots, v'_p$, что $v'_i(x) = v_i(\alpha, S(\mathsf{u}_i))$ (здесь полагаем, что в схеме S вершина s является первым предком элемента s_i). Преобразованный фрагмент показан в правой части рисунка 8.4. Легко видеть, что в элементах $\mathsf{s}'_1, \ldots, \mathsf{s}'_p$ вычисляются те же функции, что и ранее в элементах $\mathsf{s}_1, \ldots, \mathsf{s}_p$, и при этом сложность преобразованной схемы меньше сложности исходной. Пришли к противоречию с минимальностью схемы S. Лемма доказана.

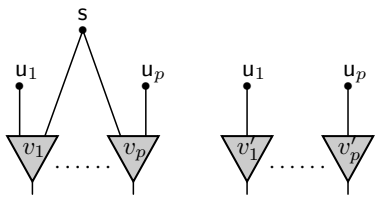


Рис. 8.4

Лемма 8.3 используется в часто применяемом преобразовании схем — подстановке константы вместо какого-либо входа. Рассмотрим это преобразование на примере схемы, вычисляющей функцию $x \oplus y$. Эта схема изображена в левой части рисунка 8.5. В схеме вместо переменной y подставим единицу. После такой подстановки правый элемент отрицания превратится в элемент, реализующий тождественный нуль, а левый конъюнктор — в тождественный элемент, который вычисляет такую же функцию, как и

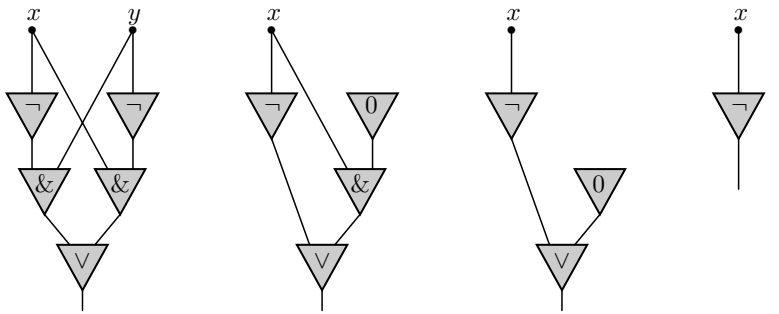


Рис. 8.5

левый инвертор, и поэтому может быть удален из схемы. Преобразованная схема изображена справа от исходной схемы. В новой схеме первый вход дизъюнктора подключен к инвертору вместо удаленного конъюнктора. Второй вход оставшегося конъюнктора подключен к тождественному нулю, и, следовательно, сам вычисляет тождественный нуль. Поэтому во второй схеме удалим элемент, вычисляющий тождественный нуль, а конъюнктор заменим элементом, вычисляющим тождественный нуль. В результате получится третья слева схема рисунка 8.5. В этой схеме второй вход дизъюнктора подключен к тождественному нулю, и, следовательно, вычисляет функцию, к которой подключен его первый вход. Поэтому конъюнктор вместе с тождественным нулем можно удалить из схемы, а элемент отрицания объявить выходом схемы. Получившаяся схема изображена в правой части рисунка 8.5. Эта схема вычисляет отрицание переменной x.

Лемма 8.4. *Пусть* S *— минимальная схема в базисе* B_0*, которая вычисляет такую систему функций* $f_1, \ldots, f_m$*, зависящих от переменных* $x_1, \ldots, x_n$*, что* $f_i \neq \overline{f}_j$ *и* $f_i \neq \overline{x}_k$ *для всех возможных значений индексов* i, j, k*. Тогда в* S *нет вершины, в которой реализуется отрицание.*

Доказательство. Пусть S — удовлетворяющая условиям леммы схема. Допустим, что в этой схеме есть элементы, в которых реализуется отрицание. Пусть s — самый нижний такой элемент и s не является выходом схемы. Рассмотрим фрагмент схемы, содержащий s, элементы $\mathsf{s}_1, \ldots, \mathsf{s}_p$, — потомки вершины s, вершину s_0 — предка s, и вершины $\mathsf{u}_1, \ldots, \mathsf{u}_p$, — предки элементов $\mathsf{s}_1, \ldots, \mathsf{s}_p$. Такой фрагмент изображен в левой части рис. 8.6. Преобразуем его следующим образом. Удалим элемент s. Элементы $\mathsf{s}_1, \ldots, \mathsf{s}_p$, реализующие функции $v_1, \ldots, v_p$, подключим к вершине s_0, после чего заменим элементами $\mathsf{s}'_1, \ldots, \mathsf{s}'_p$, реализующими такие функции $v'_1, \ldots, v'_p$, что $v'_i(x, y) = v_i(\overline{x}, y)$ (здесь как и ранее полагаем, что в схеме S вершина s_0 является первым предком элемента s_i). Преобразованный фрагмент схемы S изображен в правой части рис. 8.6. Легко видеть, что в элементах $\mathsf{s}'_1, \ldots, \mathsf{s}'_p$ вычисляются те же функции, что и ранее в элементах $\mathsf{s}_1, \ldots, \mathsf{s}_p$.

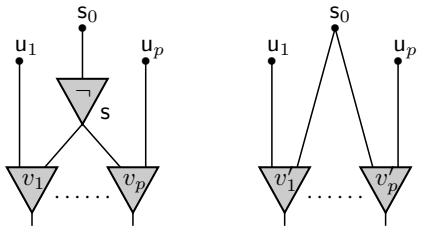


Рис. 8.6

Теперь рассмотрим случай, когда элемент s является одним из выходов схемы. В этом случае предок элемента s не может быть входом схемы, а обязательно будет элементом не являющимся выходом схемы. Пусть s_0 — элемент к которому подключена вершина s, $\mathsf{v}_1, \ldots, \mathsf{v}_p$ — элементы, подключенные к s, $\mathsf{r}_1, \ldots, \mathsf{r}_q$ — элементы, подключенные к s_0 и отличные от s. Реализуемую элементом s_0 функцию обозначим через φ. Такой фрагмент изображен в левой части рис. 8.7. В рассматриваемом фрагменте удалим элемент s, элемент s_0 заменим элементом s'_0, реализующим функцию $\overline{\varphi}$, объявим элемент s'_0 выходом схемы вместо элемента s, эле-

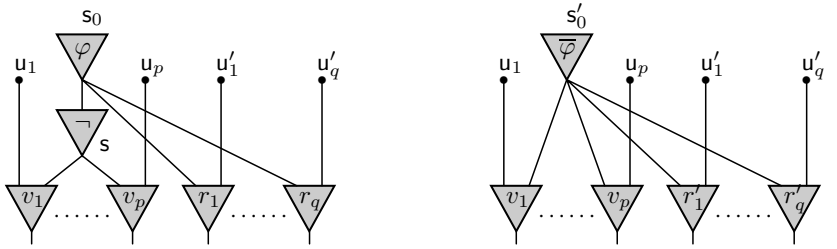


Рис. 8.7

менты $\mathsf{v}_1, \ldots, \mathsf{v}_p$ подключим к элементу s'_0, а элементы $\mathsf{r}_1, \ldots, \mathsf{r}_q$, реализующие функции $r_1, \ldots, r_q$, заменим элементами $\mathsf{r}'_1, \ldots, \mathsf{r}'_q$, реализующими такие функции $r'_1, \ldots, r'_q$, что $r'_i(x, y) = r_i(\overline{x}, y)$ (здесь как и ранее полагаем, что в схеме S вершина s_0 является первым предком элемента r_i). Преобразованный фрагмент схемы S изображен в правой части рис. 8.7. Легко видеть, что в элементах $\mathsf{r}'_1, \ldots, \mathsf{r}'_q$ вычисляются те же функции, что и ранее в элементах $\mathsf{r}_1, \ldots, \mathsf{r}_q$, и при этом сложность преобразованной схемы меньше сложности исходной. Пришли к противоречию с минимальностью схемы S. Лемма доказана.

Пусть схема S в базисе B_0 вычисляет систему функций из леммы 8.4. Применяя лемму 8.2 и рассуждения, использованные при доказательстве лемм 8.3 и 8.4, нетрудно показать, что схему S можно без увеличения сложности преобразовать так, что преобразованная схема будет состоять только из тех элементов, которые реализуют двуместные функции существенно зависящие от обоих своих аргументов. Пример такого преобразования по-

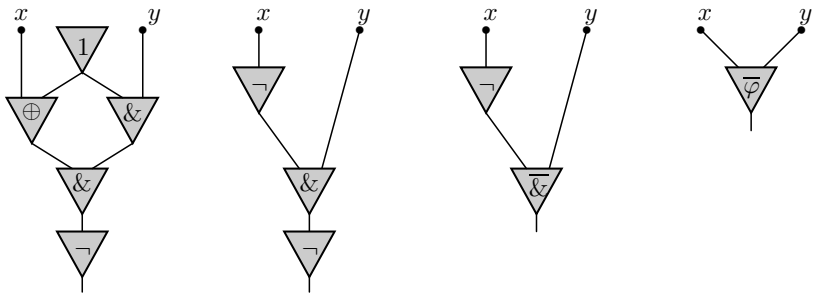


Рис. 8.8

казан на рисунке 8.8. Исходная схема в базисе B_0 изображена слева. Вершина, в которой вычисляется тождественная единица, удаляется так, как это было сделано в лемме 8.3. При этом элемент сложения заменяется элементом отрицания, а верхняя конъюнкция — тождественным элементом, который в свою очередь удаляется в силу леммы 8.2. Преобразованная схема изображена на втором слева фрагменте рисунка. Затем, в соответствии с леммой 8.4, удаляется нижний элемент отрицания и конъюнкция заменяется элементом, реализующим $\overline{\&}$. Наконец, удаляется последний элемент

отрицания, при этом элемент, реализующий $\overline{\&}$, заменяется элементом, реализующим функцию $\overline{\varphi}$. Результатом всех преобразований является одноэлементная схема, изображенная на правом фрагменте рисунка 8.8.

Доказываемое в следующей лемме свойство схем, часто используется при их анализе и позволяет устанавливать нетривиальные нижние оценки сложности булевых функций.

Лемма 8.5. *Пусть в схеме* S : (1) *к входу* x_i *подключен только один элемент* s *и этот элемент реализует нелинейную функцию*; (2) *второй вход* s *подключен к вершине* v, *в которой вычисляется функция, существенно зависящая только от переменных* $x_{i_1}, \ldots, x_{i_k}$. *Тогда найдутся такие постоянные* $\alpha_{i_1}, \ldots, \alpha_{i_k}$, *что после подстановки их вместо переменных* $x_{i_1}, \ldots, x_{i_k}$ *выход схемы* S *не зависит от* x_i.

Доказательство. Допустим, что к входу x_i подключен элемент s. Этот элемент реализует двухместную функцию $h(u, v) = (u \oplus \alpha)(v \oplus \beta) \oplus \gamma$, существенно зависящую от двух своих аргументов. Будем полагать, что к переменной x_i подключен второй вход элемента s. Тогда первый вход s подключен к некоторой вершине s_1, в которой вычисляется функция $h_1(x_{i_1}, \ldots, x_{i_k})$, не зависящая от x_i, и, в силу леммы 8.3, не равная тождественной постоянной. Подобная схема изображена на рисунке 8.9. В этой схеме условиям, наложенным на вход x_i, удовлетворяют входы x_2 и x_3. Если в качестве x_i взять вход x_2, то вершиной s_1 будет вход x_1. Подставляя вместо x_1 тождественную единицу видим, что функция, вычисляемая первым элементом схемы не зависит от x_2, так как $x_2 \vee 1 = 1$. Если в качестве x_i взять вход x_3, то вершиной s_1 будет вершина, реализующая дизъюнкцию. После подстановки $x_1 = 0$, $x_2 = 0$ первый элемент схемы будет вычислять тождественный нуль, и, следовательно, тождественный нуль так же будет вычислять конъюнктор, т. е. вычисляемая им функция не зависит от x_3. Легко видеть, что и в общем случае найдутся значения $\alpha_{i_1}, \ldots, \alpha_{i_k}$, переменных $x_{i_1}, \ldots, x_{i_k}$, при которых $h_1(\alpha_{i_1}, \ldots, \alpha_{i_k}) = \alpha$. Тогда

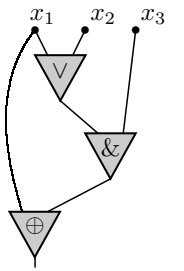


Рис. 8.9

$$\mathsf{S}(s) = h(h_1(\alpha_{i_1}, ..., \alpha_{i_k}), x_i) = (\alpha \oplus \alpha)(x_i \oplus \beta) \oplus \gamma = \gamma,$$

т. е. элемент s вычисляет функцию независящую от x_i. Следовательно, от x_i так же не зависит и выход схемы S. Лемма доказана.

8.4 Примеры

Рассмотрим три примера, в которых оценивается сложность и глубина булевых функций. В первых двух примерах для сложности и глубины рассматриваемых функций устанавливаются точные значения.

1. Покажем, что для сложности каждой булевой функции f, зависящей от трех переменных, справедливо неравенство

$$L_{B_0}(f) \le 4.$$

Для этого в многочлене Жегалкина булевой функции $f(x,y,z)$ соберем вместе все одночлены содержащие и все одночлены не содержащие переменную x. В результате получим следующее равенство:

$$\begin{aligned} f(x,y,z) &= a_0 \oplus a_1 x \oplus a_2 y \oplus a_3 z \oplus a_4 xy \oplus a_5 xz \oplus a_6 yz \oplus a_7 xyz = \\ &= x(a_1 \oplus a_4 y \oplus a_5 z \oplus a_7 yz) \oplus (a_0 \oplus a_2 y \oplus a_3 z \oplus a_6 yz) = \\ &= xh(y,z) \oplus g(y,z). \end{aligned}$$

Схема, построенная в соответствии с этим представлением, состоит из элемента реализующего функцию h, элемента реализующего функцию g, конъюнкции и элемента сложения. Легко видеть, что сложность такой схемы равна четырем, а глубина — двум.

Теперь покажем, что среди трехместных булевых функций есть функция, сложность которой в базисе из всех двухместных булевых функций не меньше четырех. Такой функцией является функция голосования

$$\tau_2(x,y,z) = xy \oplus xz \oplus yz \tag{8.4}$$

Прежде всего отметим, что подстановка произвольной константы на место любого аргумента функции голосования преобразует ее в функцию, существенно зависящую от двух оставшихся аргументов. Подставляя константы вместо переменной x в (8.4), легко видеть, что

$$\tau_2(0,y,z) = yz, \quad \tau_2(1,y,z) = y \oplus z \oplus yz. \tag{8.5}$$

Так как функция голосования является симметрической функцией, то равенства, аналогичные (8.5), справедливы и при подстановке констант вместо переменных y и z.

Пусть S — минимальная схема, вычисляющая функцию голосования. Перенумеруем элементы схемы S так, чтобы первый номер получил элемент, входы которого подключены только к независимым переменным, а последний номер получил выход схемы.

Допустим, что минимальная схема S состоит из двух элементов. Так как функция голосования является симметрической функцией, то без ограничения общности будем полагать, что входы первого элемента схемы подключены к переменным x и y. В этом случае схема S выглядит так, как это изображено на рис. 8.10. Пусть первый элемент реализует функцию f, а второй — функцию g. Покажем, что ни f, ни g не могут быть нелинейными функциями. Действительно, если f является нелинейной функцией, то из леммы 8.2 следует существование такой постоянной α, что после ее

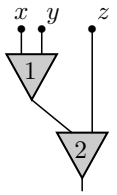


Рис. 8.10

подстановки вместо переменной y выход схемы не зависит от переменной x. Получили противоречие с (8.5). Если нелинейной функцией является функция $g = (u \oplus \alpha)(z \oplus \beta) \oplus \gamma$, то подставляя вместо переменной z константу β убеждаемся, что вычисляемая схемой S функция будет константой. Снова противоречие с (8.5).

Таким образом, функции f и g должны быть линейными функциями. Следовательно, вычисляемая схемой S функция должна быть линейной. Противоречие с (8.4).

Теперь предположим, что вычисляющая функцию голосования минимальная схема S состоит из трех элементов. Снова без ограничения общности будем полагать, что входы первого элемента схемы подключены к переменным x и y.

Общее число входов второго и третьего элементов схемы равно четырем. Три из этих четырех входов должны быть обязательно подключены:

1) к первому элементу;
2) к второму элементу;
3) к переменной z.

Следовательно, оставшийся четвертый вход может быть подключен либо к одной из переменных x и y (без ограничения общности будем полагать, что такой переменной будет y), либо к переменной z, либо к первому элементу схемы. Таким образом в схеме S два элемента могут подключаться либо к переменной y, либо к переменной z, либо к первому элементу. К переменной x обязательно подключен только первый элемент схемы.

Далее рассмотрим два случая:

(i) к первому элементу схемы S подключен элемент, который не является выходом схемы;

(ii) к первому элементу схемы S не подключен элемент, который не является выходом схемы.

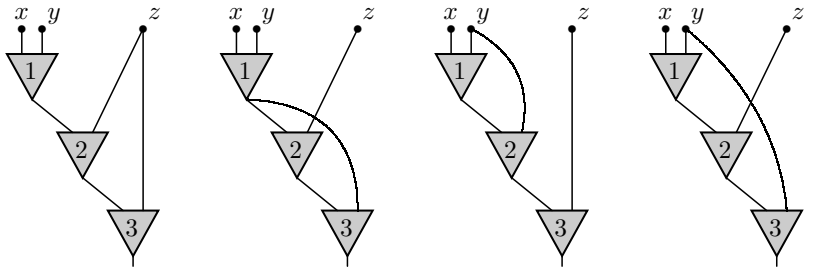


Рис. 8.11

Рассмотрим первый случай. В этом случае глубина схемы равна трем. С учетом сделанных выше предположений, в первом случае возможны только четыре различных конструкции схемы S. Все они представлены на рисунке 8.11. Из рисунка видно, что три левые схемы не являются минимальными. Каждая из этих схем может быть преобразована в схему из двух элементов так, что старая и новая схемы будут вычислять одну и туже функцию.

В качестве примера рассмотрим самую левую схему. В этой схеме второй и третий элементы вычисляют функцию, зависящую от двух аргументов — переменной z и функции, вычисляемой первым элементом. Поэтому второй и третий элементы можно заменить одним, реализующим подходящую двуместную функцию. Преобразованная схема будет выглядеть так, как это изображено на рис. 8.10.

Теперь рассмотрим правую схему. Как и при анализе двухэлементной схемы легко показать, что ни один из элементов не может быть нелинейным элементом. Подстановка вместо переменной y или z подходящей константы преобразует схему S в схему, вычисляющую константу или функцию одной переменной, что противоречит (8.5). Но если все элементы являются линейными элементами, то вычисляемая схемой S функция будет линейной. Противоречие с (8.4). Таким образом, ни одна из изображенных на рис. 8.11 схем не вычисляет функцию голосования. Случай (i) рассмотрен полностью.

Рассмотрим второй случай. С учетом сделанных предположений второй элемент схемы подключен к переменным y и z. Поэтому схема S выглядит так, как это изображено на рис. 8.12. Как и ранее, легко показать, что первый и второй элементы не могут быть нелинейными элементами. Подстановка вместо переменной y подходящей константы преобразует схему S в схему, вычисляющую константу или функцию одной переменной, что противоречит (8.5). Допустим, что первый и второй элементы являются линейными элементами, реализующими функции $x \oplus y \oplus \alpha$ и $y \oplus z \oplus \beta$. Преобразуем схему S, заменив переменные y и z переменной x. В преобразованной схеме первый элемент вычисляет константу α, второй элемент — константу β. Следовательно, сама схема также вычисляет тождественную константу. Пришли к противоречию с (8.4), так как подстановка $y = x$, $z = x$ преобразует функцию голосования в функцию x. Случай (ii) рассмотрен полностью.

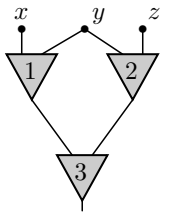


Рис. 8.12

Таким образом, минимальная схема функции голосования состоит не менее чем из четырех элементов, т. е. $L_{B_0}(\tau_2(x, y, z)) \geq 4$. Нижняя оценка глубины является тривиальным следствием теоремы 8.2.

2. Систему F из n линейных функций, зависящих от $2^n - 1$ переменных, назовем *универсальной* и обозначим через U_n, если j-й столбец матрицы $\mathbf{U}_n$ этой системы совпадает с двоичным представлением числа j. Например, универсальной системой U_3 с матрицей $\mathbf{U}_3$ будет система трех линейных функций $\{f_1, f_2, f_3\}$, если

$$\begin{aligned} f_1 &= x_4 \oplus x_5 \oplus x_6 \oplus x_7, \\ f_2 &= x_2 \oplus x_3 \oplus x_6 \oplus x_7, \\ f_3 &= x_1 \oplus x_3 \oplus x_5 \oplus x_7. \end{aligned} \qquad \mathbf{U}_3 = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Теорема 8.3. *Справедливы равенства*

$$L(U_n) = 2(2^n - n - 1), \qquad D(U_n) = n - 1,$$

которые достигаются на одной и той же схеме.

Доказательство. Верхние оценки. Для доказательства теоремы построим схему U_n, вычисляющую универсальную систему U_n, состоящую из $2(2^n - n - 1)$ элементов сложения и имеющую глубину $n - 1$. Схему U_n построим индукцией по n. На рисунке 8.13 представлена индуктивная процедура построения схемы U_n: в левой части изображен базис индукции, в правой — индуктивный переход. Индуктивный переход основан на следующем простом свойстве матрицы $\mathbf{U}_n$:

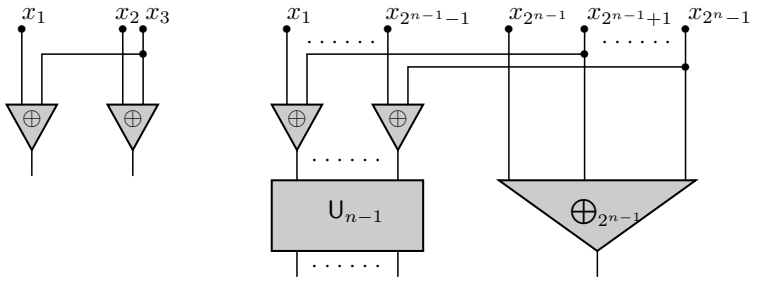


Рис. 8.13

$$u_{i,j} = u_{i,j+2^{n-1}} \quad \text{при всех } 2 \leq i \leq n \text{ и } 1 \leq j \leq 2^{n-1} - 1.$$

Найдем сложность $L(\mathsf{U}_n)$ и глубину $D(\mathsf{U}_n)$ схемы U_n. Легко видеть, что $L(\mathsf{U}_2) = 2$ и $D(\mathsf{U}_2) = 1$. Из конструкции схемы U_n видим, что

$$\begin{aligned} L(\mathsf{U}_n) &= L(\mathsf{U}_{n-1}) + L(\textstyle\bigoplus_{2^{n-1}}) + 2^{n-1} - 1 = L(\mathsf{U}_{n-1}) + 2(2^{n-1} - 1), \\ D(\mathsf{U}_n) &= \max(1 + D(\mathsf{U}_{n-1}), D(\textstyle\bigoplus_{2^{n-1}})) = \max(1 + D(\mathsf{U}_{n-1}), n - 1). \end{aligned}$$

По предположению индукции

$$L(\mathsf{U}_{n-1}) = 2(2^{n-1} - (n-1) - 1), \quad D(\mathsf{U}_{n-1}) = n - 2.$$

Следовательно,

$$\begin{aligned} L(\mathsf{U}_{n-1}) &= 2(2^{n-1} - (n-1) - 1) + 2(2^{n-1} - 1) = 2(2^n - n - 1), \\ D(\mathsf{U}_{n-1}) &= \max(1 + (n-2), n - 1) = n - 1. \end{aligned}$$

Верхние оценки сложности и глубины доказаны.

Прежде чем доказывать нижние оценки сложности системы U_n, дадим важное определение и докажем две необходимых леммы, имеющих также самостоятельный интерес.

Пусть S — произвольная схема. Вершину v схемы S назовем *вершиной ветвления* если эта вершина имеет не менее двух потомков, или является выходом схемы и имеет хотя бы одного потомка. Вершину ветвления v

назовем *вершиной первого ветвления* входа x_i, если v и вход x_i связаны цепью, проходящей через вершины $\mathsf{v}_0 = x_i, \mathsf{v}_1, \ldots, \mathsf{v}_{k-1}, \mathsf{v}_k = \mathsf{v}$, в которой ни одна из вершин v_j, где $j = 0, 1, \ldots, k-1$, не является вершиной ветвления. Для примера, рассмотрим первые две схемы, изображенные слева на рисунке 8.11 на странице 117. В первой схеме входы x и y не имеют вершин ветвления, а вход z сам является своей вершиной первого ветвления. На второй схеме входы x и y имеют общую вершину первого ветвления — элемент, помеченный единицей, а у входа z нет вершины первого ветвления.

Лемма 8.6. *В любой схеме S, вычисляющей систему линейных функций F, два входа имеют различные вершины первого ветвления, если в матрице системы соответствующие этим входам столбцы различны и каждый из этих столбцов содержит не менее двух единиц.*

Доказательство. Пусть схема S вычисляет систему линейных функций $F(x_1, \ldots, x_n)$, и пусть i-й и j-й входы S удовлетворяют условиям леммы. Допустим, что у i-го и j-го входов есть общая вершина первого ветвления v. В схеме S вместо всех переменных, исключая x_i и x_j, подставим тождественный нуль. После такой подстановки новая схема S' будет вычислять систему линейных функций, зависящих только от двух переменных x_i и x_j. Матрица этой системы будет состоять из i-го и j-го столбцов матрицы системы F. Далее рассмотрим два случая: первый, когда в новой матрице нет строки с двумя единицами; второй, когда такая строка есть.

В первом случае среди линейных функций обязательно найдется хотя бы одна, существенно зависящая только от x_i, и одна, существенно зависящая только от x_j. Во втором случае найдется функция, существенно зависящая только от одной переменной, например от x_i, и функция, существенно зависящая от x_i и x_j. В обоих случаях первую функцию обозначим через $f_1(x_i)$, вторую — через $f_2(x_i, x_j)$.

Теперь рассмотрим вершину v. После подстановки нулей в этой вершине будет вычисляться некоторая функция $h(x_i, x_j)$, зависящая только от x_i и x_j. Так как любая цепь, связывающая каждый из входов x_i или x_j с любым выходом схемы S', проходит через вершину v, то каждую функцию, вычисляемую схемой S', будем рассматривать как функцию, зависящую только от $h(x_i, x_j)$. Таким образом, $f_1(x_i) = g_1(h(x_i, x_j))$ и $f_2(x_i, x_j) = g_2(h(x_i, x_j))$. Очевидно, что ни g_1, ни g_2 не являются тождественными постоянными и что $g_1 \neq g_2$, т. е. одна из этих функций будет отрицанием, а вторая — тождественной функцией. Без ограничения общности будем полагать, что отрицанием будет g_1. Тогда $0 = f_1(0) = \overline{h}(0,0) \neq h(0,0) = f_2(0,0) = 0$. Противоречие. Лемма доказана.

Лемма 8.7. *Пусть система m линейных n-местных функций F такова, что среди столбцов ее (m,n)-матрицы: (i) нет ни одного нулевого столбца; (ii) есть k различных столбцов, каждый из которых содержит не менее двух единиц. Тогда*

$$L(F) \geq n + k - m.$$

Доказательство. Пусть S — минимальная схема, вычисляющая систему F. Из леммы 8.6 легко следует, что в S найдется не менее k входов, каждый из которых имеет собственную вершину первого ветвления, и, следовательно, в схеме S есть k вершин, каждая из которых имеет не менее двух потомков. Число элементов схемы S обозначим через L, число ребер — через N. Оценим N, подсчитывая ребра, выходящие из вершин схемы S.

1) Из каждого входа схемы обязательно выходит одно ребро — всего n ребер.

2) Из всех элементов схемы, не являющихся ее выходами, обязательно выходит по одному ребру — всего $L - m$ ребер.

3) Кроме этого в S есть не менее k вершин первого ветвления и из каждой выходит не менее одного ребра, не учтенного в пунктах 1) и 2).

Общее число выходящих ребер — $n + k + L - m$. Следовательно, $N \geq n + k + L - m$.

Теперь оценим N, подсчитывая ребра, входящие в вершины схемы S. Ребра входят только в элементы S, причем в каждый элемент входит не более двух ребер. Поэтому $N \leq 2L$. Следовательно, $n+k+L-m \leq N \leq 2L$. Из последнего неравенства немедленно получаем, что $L \geq n+k-m$. Лемма доказана.

Нижние оценки. Так как все $2^n - 1$ столбцов матрицы $\mathbf{U}_n$ различные, ненулевые и среди них есть $2^n - n - 1$ столбцов содержащих не менее чем по две единицы, то в силу леммы 8.7

$$L(U_n) \geq (2^n - 1) + (2^n - n - 1) - n = 2(2^n - n - 1).$$

Нижняя оценка глубины следует из теоремы 8.2. Теорема доказана.

3. Рассмотрим n-местную дизъюнкцию $x_1 \vee \cdots \vee x_n$. Легко видеть, что дизъюнкция сохраняет нуль, а функции $\&$ и $\oplus$ образуют базис в T_0. Поэтому дизъюнкцию $x_1 \vee \cdots \vee x_n$ можно вычислить схемой в базисе $\{\&, \oplus\}$. Далее найдем сложность такого вычисления, а затем покажем, что использование функций, не сохраняющих нуль, позволит уменьшить сложность вычисления дизъюнкции.

Утверждение 8.1. *Справедливо равенство*

$$L_{\{\&,\oplus\}}(x_1 \vee \ldots \vee x_n) = 3n - 3.$$

Доказательство. Верхняя оценка. Схема S_n, вычисляющая функцию $x_1 \vee \ldots \vee x_n$, легко строится индуктивно в соответствии с формулой

$$x_1 \vee \ldots \vee x_n = (x_1 \vee \ldots \vee x_{n-1})x_n \oplus (x_1 \vee \ldots \vee x_{n-1}) \oplus x_n.$$

Полагая, что $L(\mathsf{S}_{n-1}) = 3(n-1) - 3$ видим, что

$$L(\mathsf{S}_n) = L(\mathsf{S}_{n-1}) + 3 = 3n - 3.$$

Верхняя оценка доказана.

Нижняя оценка. Пусть S — минимальная схема в базисе $\{\&, \oplus\}$ для дизъюнкции n переменных. Покажем, что

$$L_{\{\&,\oplus\}}(x_1 \vee \ldots \vee x_n) \geq 3n - 3. \tag{8.6}$$

Прежде всего, покажем, что в S к каждому входу обязательно должен быть подключен хотя бы один элемент, реализующий функцию $\oplus$. Действительно, допустим, что к n-му входу подключены только элементы $\mathsf{s}_1, \ldots, \mathsf{s}_p$, реализующие конъюнкции. В S положим $x_1 = \ldots = x_{n-1} = 0$. Так как базис схемы S сохраняет нуль, то все ее вершины, находящиеся в схеме выше элементов $\mathsf{s}_1, \ldots, \mathsf{s}_p$ и не являющиеся n-м входом, будут вычислять тождественный нуль. Поэтому тождественный нуль будет вычисляться в вершинах, к которым подключены вторые входы элементов $\mathsf{s}_1, \ldots, \mathsf{s}_p$, а, следовательно, и в самих элементах $\mathsf{s}_1, \ldots, \mathsf{s}_p$. Последнее означает, что схема S вычисляет тождественный нуль при любом значении переменной x_n. Пришли к противоречию.

Теперь покажем, что в S обязательно найдется вход, к которому подключено не менее двух элементов. Предположим, что это не так. Тогда в схеме найдутся два входа, к которым подключен один и тот же элемент s. Без ограничения общности будем полагать, что такими входами будут $(n-1)$-й и n-й входы. В S положим $x_1 = \ldots = x_{n-2} = 0$, $x_{n-1} = x_n = x$. После такой подстановки все элементы S будут вычислять тождественный нуль при любом значении переменной x. Снова пришли к противоречию.

Пусть в S к n-му входу подключены два элемента v и u. Без ограничения общности будем полагать, что элемент u реализует $\oplus$. В S положим $x_n = 0$. Покажем, что после такой подстановки из S можно удалить не менее двух элементов, реализующих $\oplus$. Для этого рассмотрим всевозможные цепочки элементов $\mathsf{v} = \mathsf{v}_0, \mathsf{v}_1, \ldots, \mathsf{v}_k$ в которых каждый элемент v_{i+1} подключен к элементу v_i. Среди этих цепочек обязательно найдется такая, в которой первые k элементов реализуют конъюнкции, а последний — $\oplus$. (Если элемент v_0 реализует $\oplus$, то $k = 0$.) Слева на рисунке 8.14 изображена подобная цепочка из трех элементов. Если такой цепочки нет, то в S существует цепочка из одних конъюнкций, которая связывает n-й вход и выход схемы. В этом случае подстановка нуля вместо x_n приведет к тому, что в последнем элементе схемы будет вычисляться нуль независимо от значения первых $n-1$ переменных. Итак, пусть $\mathsf{w} = \mathsf{v}_k$ — элемент, реализующий $\oplus$ и связанный с v цепочкой из одних конъюнкций. Если w и u различные элементы, то именно они будут удалены после подстановки $x_n = 0$. Допустим теперь, w и u совпадают (подобная ситуация изображена в правой части рисунка 8.14). Сохраним за этим элементом обозначение w. Легко видеть, что после подстановки $x_n = 0$ вычисляемая в w функция будет тождественным нулем. Как и выше, рассматривая всевозможные цепочки конъюнкций, начинающиеся в w, найдем элемент z, реализующий $\oplus$.

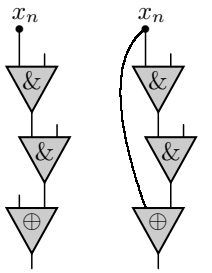


Рис. 8.14

Преобразование схемы, заключающееся в подстановке вместо x_i тождественного нуля и последующее удаление из схемы i-го входа и всех вершин, вычисляющих тождественный нуль, назовем операцией удаления i-го входа схемы.

Из доказанного выше следует, что из схемы S можно удалять входы до тех пор, пока в ней не останется единственный не удаленный вход. Следовательно, операция удаления входа может быть выполнена $n-1$ раз. При этом из S будет удалено не менее чем $2n-2$ элемента $\oplus$. Кроме того, легко видеть, что S содержит не менее чем $n-1$ элемент конъюнкции. Это следует из того, что степень дизъюнкции n переменных равна n. Таким образом, общее число элементов в схеме S не меньше чем $3n-3$. Неравенство (8.6) доказано. Утверждение доказано полностью.

Дизъюнкция принадлежит классу T_0, и как видно из доказанного утверждения, достаточно просто вычисляется схемами в базисе, целиком содержащемся в T_0. С другой стороны, легко видеть, что $x_1 \vee \cdots \vee x_n = (x_1 \oplus 1) \cdot \ldots \cdot (x_n \oplus 1) \oplus 1$. Откуда легко следует неравенство

$$L_{\{\&,\oplus,1\}}(x_1 \vee \cdots \vee x_n) \leq 2n + 1.$$

Таким образом, расширение базиса схемы до базиса полного в P_2 позволяет уменьшить сложность вычисления в полтора раза.

8.5 Задачи

8.1. Доказать равенства
a) $L_{\{\vee,\&,\neg\}}(x \sim y) = 4$; b) $L_{\{\vee,\&,\neg\}}(x \oplus y) = 4$.

8.2. Доказать равенства
a) $L_{\{\oplus,\&,1\}}(x \sim y) = 3$; b) $L_{\{\oplus,\&,1\}}(x \vee y) = 3$;
c) $L_{\{\oplus,\&,1\}}(x \to y) = 4$; d) $L_{\{\oplus,\&,1\}}(x \downarrow y) = 4$.

8.3. Доказать равенства
a) $L_{\downarrow}(x \to y) = 3$; b) $L_{\downarrow}(x\&y) = 3$; c) $L_{\downarrow}(1) = 3$;
d) $L_{\downarrow}(x|y) = 4$; e) $L_{\downarrow}(x \sim y) = 4$; f) $L_{\downarrow}(x \oplus y) = 5$.

8.4. Доказать равенства
a) $L_{|}(x\&y) = 2$; b) $L_{|}(x \vee y) = 3$; c) $L_{|}(0) = 3$;
d) $L_{|}(x \downarrow y) = 4$; e) $L_{|}(x \oplus y) = 4$; f) $L_{|}(x \sim y) = 5$.

8.5. Показать, что для системы дизъюнкций $\{x_1 \vee x_2 \vee x_3 \vee x_4,\ x_2 \vee x_3 \vee x_4\}$ не существует схемы, которая одновременно минимальна по сложности и глубине.

8.6. Найти $L_{B_0}(xyz \vee \overline{x}\,\overline{y}\,\overline{z})$.

8.7. Найти $L_{\{\&,\vee,\neg\}}(x \oplus y \oplus z)$.

8.8. Показать, что $L_{\{\&,\neg\}}(x_1 \vee \ldots \vee x_n) = 2n$.

8.9. Показать, что $L_{\{\vee,\neg\}}(x_1 \& \ldots \& x_n) = 2n$.

8.10. Пусть $B = P_2(2) \setminus \{\oplus, \sim\}$. Показать, что $L_B(x_1 \oplus \ldots \oplus x_n) = 3n - 3$.

8.11. Показать, что $L_{\{\&,\vee,\neg\}}(x_1 \oplus \ldots \oplus x_n) = 4n - 4$.

8.12. Найти $L_{\{\&,\oplus,1\}}(x_1 \vee \ldots \vee x_n)$.

8.13. Показать, что для любой булевой функции $f(x_1, \ldots, x_n)$ имеют место неравенства
а) $L_{\{\&,\neg\}} f(x_1, \ldots, x_n) \leq 2L_{\{\&,\vee,\neg\}}(f) + n$;
b) $L_{\{\vee,\neg\}} f(x_1, \ldots, x_n) \leq 2L_{\{\&,\vee,\neg\}}(f) + n$.

8.14. Показать, что в базисе $\{\downarrow\}$ можно построить схему из функциональных элементов, реализующую все булевы функции от n переменных и содержащую $2^{2^n} - n$ элементов. Доказать минимальность такой схемы.

8.15. Показать, что сложность реализации схемами в базисе $\{\vee\}$ системы из $2n$ дизъюнкций от 2^n переменных асимптотически равна 2^{n+1} при $n \to \infty$, если в матрице коэффициентов этой системы первые n разрядов j-го столбца являются двоичным разложением числа $j - 1$, а последние n разрядов являются двоичным разложением числа $2^n - j$.

8.16. Найти сложность реализации схемами в базисе $\{\oplus\}$ системы из n линейных n-местных булевых функций, если в матрице коэффициентов этой системы на главной диагонали стоят нули, а остальные элементы равны единице.

8.17. Найти сложность реализации схемами в базисе $\{\vee\}$ системы из n дизъюнкций, зависящих от n переменных, если в матрице коэффициентов этой системы на главной диагонали стоят нули, а остальные элементы равны единице.

8.18. Найти сложность реализации схемами в базисе $\{\oplus\}$ системы из 2^n линейных 2^n-местных булевых функций, если матрица коэффициентов этой системы определяется равенствами:

$$P_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad P_k = \begin{pmatrix} P_{k-1} & P_{k-1} \\ \overline{P}_{k-1} & P_{k-1} \end{pmatrix}.$$

8.19. Пусть f — такая система из $2n$ дизъюнкций, что матрица ее коэффициентов состоит из всех тех попарно различных векторов высоты $2n$, которые содержат ровно n единиц. Найти $L_{\{\vee\}}(f)$ при $n \to \infty$.

8.20. Пусть f — такая система из n дизъюнкций, что матрица ее коэффициентов состоит из всех тех попарно различных векторов высоты n, число единиц в которых делится на 3. Найти $L_{\{\vee\}}(f)$ при $n \to \infty$.

Лекция 9

Быстрые схемы

Рассматривая различные вычисления встречающиеся в реальной жизни, нетрудно заметить, что большинство таких вычислений состоит из большого числа часто повторяющихся однотипных операций. К таким операциям безусловно можно отнести операции сложения и умножения многоразрядных целых чисел, а также операцию сортировки наборов данных разной природы. Поэтому время выполнения этих операций оказывает существенное влияние на время вычисления в целом. При этом, как правило, условия реализации арифметических операций и сортировки допускают одновременное выполнение различных промежуточных действий, т. е. эффективность этих операций зависит в основном от их глубины, а не от сложности.

Ниже для операций сложения, умножения и сортировки рассматриваются быстрые схемы, т. е. схемы с небольшой глубиной.

9.1 Сложение

1. Рассмотрим булев оператор *сложения* $S_n : \{0,1\}^{2n} \to \{0,1\}^{n+1}$, вычисляющий сумму двух n-разрядных целых положительных чисел, представленных в двоичной системе счисления. Пусть

$$\boldsymbol{x} = \sum_{i=1}^{n} x_i 2^{i-1}, \quad \boldsymbol{y} = \sum_{i=1}^{n} y_i 2^{i-1}, \quad \boldsymbol{z} = \sum_{i=1}^{n+1} z_i 2^{i-1},$$

где $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{z}$. Тогда

$$S_n(x_1, \ldots, x_n, y_1, \ldots, y_n) = (z_1, \ldots, z_{n+1}).$$

Схему, вычисляющую оператор S_n, назовем n-разрядным *сумматором*.

В традиционном алгоритме сложения двух многоразрядных чисел j-й разряд суммы z_j равен сумме j-х разрядов слагаемых и переноса q_j из предыдущих $j - 1$ разрядов, т. е.

$$z_j = x_j \oplus y_j \oplus q_j. \tag{9.1}$$

При этом перенос в j-й разряд вычисляется по формуле

$$q_j = x_{j-1}y_{j-1} \oplus x_{j-1}q_{j-1} \oplus y_{j-1}q_{j-1} = x_{j-1}y_{j-1} \oplus q_{j-1}(x_{j-1} \oplus y_{j-1}) \quad (9.2)$$

после того, как вычислен перенос в $(j-1)$-й разряд. Нетрудно видеть, что глубина схемы, складывающей два числа и построенной в соответствии с формулами (9.1) и (9.2), будет линейной по числу разрядов складываемых чисел. Причина этого, очевидно, заключается в последовательном вычислении переносов. Поэтому для быстрого сложения многоразрядных чисел надо уметь вычислять переносы параллельно.

Докажем вспомогательное утверждение, которое позволит достаточно быстро и просто вычислить все переносы q_j, возникающие при сложении двух чисел.

Лемма 9.1. *Пусть зависящие от переменных $y_1, a_1, b_1, a_2, b_2, \ldots, a_{2^k}, b_{2^k}$ функции $y_1, y_2, \ldots, y_{2^k+1}$ такие, что $y_{j+1} = b_j \oplus a_j y_j$ при $j = 1, \ldots, 2^k$. Тогда существует вычисляющая функции $y_1, y_2, \ldots, y_{2^k+1}$ схема P_k, для сложности и глубины которой справедливы соотношения*

$$L(\mathsf{P}_k) \le 5 \cdot 2^k, \quad D(\mathsf{P}_k) \le 4k + 2.$$

Доказательство. Лемму докажем индукцией по k. При $k = 0$ нужно вычислить только одну функцию y_2. Это можно сделать схемой P_0, состоящей из одного элемента сложения и одного элемента умножения. Очевидно, что $L(\mathsf{P}_0) = 2$ и $D(\mathsf{P}_0) = 2$. Предположим, что при некотором $k \ge 0$ требуемая схема P_k существует. Используя эту схему, построим схему P_{k+1}.

Прежде всего заметим, что при каждом j, где $1 \le j \le 2^k$, справедливо равенство

$$\begin{aligned} y_{2j+1} &= b_{2j} \oplus a_{2j}y_{2j} = \\ &= b_{2j} \oplus a_{2j}(b_{2j-1} \oplus a_{2j-1}y_{2j-1}) = (b_{2j} \oplus a_{2j}b_{2j-1}) \oplus (a_{2j}a_{2j-1})y_{2j-1}. \end{aligned}$$

Для всех $j \in \{1, \ldots, 2^k\}$ введем новые функции

$$y'_{j+1} = y_{2j+1}, \quad b'_j = b_{2j} \oplus a_{2j}b_{2j-1}, \quad a'_j = a_{2j}a_{2j-1}.$$

Тогда новые функции y'_j и новые переменные a'_j и b'_j связаны следующими равенствами:

$$y'_{j+1} = b'_j \oplus a'_j y'_j, \quad \text{при } 1 \le j \le 2^k. \quad (9.3)$$

Воспользуемся этими равенствами для вычисления функций y_j. Сделаем это в три этапа. Сначала вычислим все новые переменные a'_j и b'_j. Затем вычислим что все функции y'_j, т. е. все функции y_j с нечетными индексами. Из равенств (9.3), условий леммы и предположения индукции следует, что это можно сделать при помощи схемы P_k, подключив ее входы к вычисленным ранее переменным a'_j и b'_j. Наконец, каждую функцию y_{2j} с четным индексом вычислим по формуле $y_{2j} = b_{2j-1} \oplus a_{2j-1}y_{2j-1}$, использующей вычисленную ранее функцию y_{2j-1}.

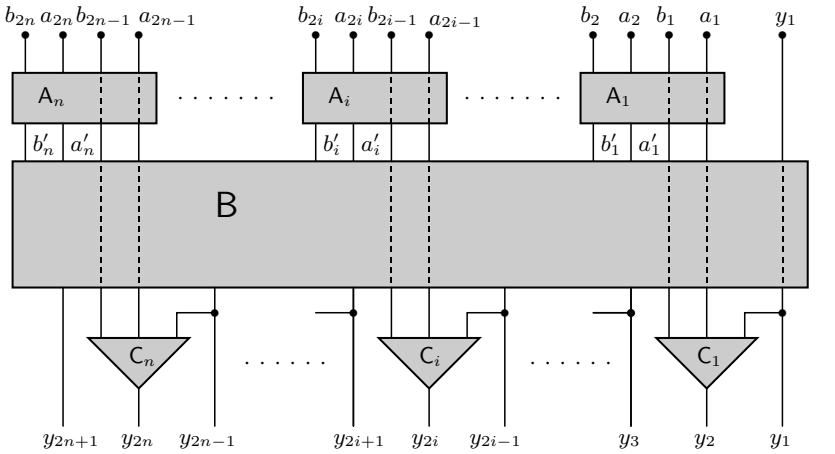


Рис. 9.1

Выполняющая указанные вычисления схема P_{k+1} изображена на рисунке 9.1, где $n = 2^k$. Эта схема состоит из 2^k подсхем A_j, $1 \le j \le 2^k$, подсхемы B и 2^k подсхем C_j, $1 \le j \le 2^k$. Кратко рассмотрим подсхемы P_{k+1} и оценим их сложности и глубины.

1. Подсхема A_j вычисляет функции a'_j и b'_j. Легко видеть, что $L(\mathsf{A}_j) = 3$ и $D(\mathsf{A}_j) = 2$.

2. Подсхема B является экземпляром схемы P_k. По предположению индукции $L(\mathsf{B}) \le 5 \cdot 2^k$ и $D(\mathsf{B}) \le 4k + 2$.

3. Подсхема C_j вычисляет функцию y_{2j} с четным индексом в соответствии с формулой $y_{2j} = b_{2j-1} \oplus a_{2j-1}y_{2j-1}$. Легко видеть, что $L(\mathsf{C}_j) = 2$ и $D(\mathsf{C}_j) = 2$.

Из конструкции схемы P_{k+1}, пп. 1–3 и предположения индукции легко получаем, что

$$\begin{aligned} L(\mathsf{P}_{k+1}) &\le L(\mathsf{P}_k) + 5 \cdot 2^k \le 5 \cdot 2^k + 5 \cdot 2^k = 5 \cdot 2^{k+1}, \\ D(\mathsf{P}_{k+1}) &\le D(\mathsf{P}_k) + 4 \le 4k + 2 + 4 = 4(k+1) + 2. \end{aligned}$$

Лемма доказана.

Теперь при помощи леммы 9.1 нетрудно доказать следующую теорему о существовании быстрых сумматоров.

Теорема 9.1. *Существует n-разрядный сумматор Σ_n, для сложности и глубины которого справедливы неравенства*

$$L(\Sigma_n) \le 13n, \quad D(\Sigma_n) \le 4\lceil \log_2 n \rceil + 4.$$

Доказательство. Рассмотрим сложение двух целых n-разрядных чисел $\boldsymbol{x}$ и $\boldsymbol{y}$. Для каждого $j \in \{1, \ldots, n\}$ определим функции

$$b_j = x_j y_j, \quad a_j = x_j \oplus y_j.$$

Тогда (см. (9.2) на стр. 126) для переноса q_{j+1} в $(j+1)$-й разряд суммы $\boldsymbol{x}+\boldsymbol{y}$ справедлива формула

$$q_{j+1} = x_j y_j \oplus (x_{j-1} \oplus y_{j-1}) q_j = b_j \oplus a_j q_j.$$

Вычислив величины b_j и a_j, для вычисления переносов q_{j+1} воспользуемся схемой $\mathsf{P}_{\lceil \log_2 n \rceil}$ из леммы 9.1. Легко видеть, для сложности и глубины схемы Q_n, производящей вычисления всех a_j, b_j и q_{j+1}, справедливы соотношения

$$L(\mathsf{Q}_n) \le 2n + 5 \cdot 2^{\lceil \log_2 n \rceil} \le 12n, \quad D(\mathsf{Q}_n) \le 4\lceil \log_2 n \rceil + 3.$$

Теперь для вычисления суммы $\boldsymbol{x}$ и $\boldsymbol{y}$ достаточно попарно сложить вычисленные схемой Q_n переносы q_j и суммы $x_j \oplus y_j$. Теорема доказана.

Более быстрые, асимптотически минимальные по глубине сумматоры построены в 1967 г. В. М. Храпченко, который в частности показал, что справедлива следующая теорема.

Теорема 9.2. *Существует n-разрядный сумматор Σ'_n, для сложности и глубины которого справедливы неравенства*

$$L(\Sigma'_n) \le 3n + 6 \cdot 2^{\lceil \log_2 n \rceil}, \quad D(\Sigma'_n) \le \lceil \log_2 n \rceil + 7\sqrt{2\lceil \log_2 n \rceil} + 14.$$

Сумматор Σ'_n устроен достаточно сложно, поэтому теорему 9.2 оставим без доказательства.

2. Далее покажем как можно использовать сумматоры для вычисления разности многоразрядных чисел. Разностью двух n-разрядных двоичных целых положительных чисел $\boldsymbol{x}$ и $\boldsymbol{y}$ назовем такой $(n+1)$-разрядный вектор $\boldsymbol{r}$, что его первые n разрядов образуют число $\boldsymbol{r}_1$, равное модулю разности $\boldsymbol{x}$ и $\boldsymbol{y}$,

$$\boldsymbol{r}_1 = (r_1, \ldots, r_n) = \sum_{i=1}^{n} r_i 2^{i-1} = |\boldsymbol{x} - \boldsymbol{y}|,$$

а его $(n+1)$-й разряд r_{n+1} равен знаку этой разности,

$$r_{n+1} = \begin{cases} 1, & \text{если } \boldsymbol{x} < \boldsymbol{y}, \\ 0, & \text{если } \boldsymbol{x} \ge \boldsymbol{y}. \end{cases}$$

Булев оператор $R_n : \{0,1\}^{2n} \to \{0,1\}^{n+1}$, вычисляющий разность двух n-разрядных целых положительных чисел, назовем оператором *вычитания*.

Опишем простой способ нахождения разности двух произвольных чисел $\boldsymbol{x}$ и $\boldsymbol{y}$. Вместе с двоичным n-разрядных числом $\boldsymbol{x} = \sum_{i=1}^{n} x_i 2^{i-1}$ рассмотрим его дополнение $\overline{\boldsymbol{x}} = \sum_{i=1}^{n} \overline{x}_i 2^{i-1}$. Очевидно, что $\boldsymbol{x} + \overline{\boldsymbol{x}} = 2^n - 1$. Следовательно, для оператора суммирования S_n и любых целых n-разрядных чисел $\boldsymbol{x}$ и $\boldsymbol{y}$ выполняется равенство

$$S_n(\overline{\boldsymbol{x}}, \boldsymbol{y}) = 2^n - 1 - \boldsymbol{x} + \boldsymbol{y}.$$

Далее через s_2 будем обозначать $(n+1)$-й разряд числа $S_n(\overline{\boldsymbol{x}}, \boldsymbol{y})$, а через s_1 — число, составленное из младших n разрядов числа $S_n(\overline{\boldsymbol{x}}, \boldsymbol{y})$, т. е. $S_n(\overline{\boldsymbol{x}}, \boldsymbol{y}) = s_2 2^n + \mathsf{s}_1$. Покажем, что

$$(\mathsf{s}_1 + s_2)^{s_2} = |\boldsymbol{x} - \boldsymbol{y}|.$$

Для этого рассмотрим два случая: $s_2 = 1$ и $s_2 = 0$. Если $s_2 = 1$, то очевидно $S_n(\overline{\boldsymbol{x}}, \boldsymbol{y}) \ge 2^n$, и, следовательно, $\boldsymbol{x} < \boldsymbol{y}$. В этом случае $\mathsf{s}_1 = -1 - \boldsymbol{x} + \boldsymbol{y}$. Тогда

$$(\mathsf{s}_1 + s_2)^{s_2} = \mathsf{s}_1 + s_2 = \mathsf{s}_1 + 1 = -\boldsymbol{x} + \boldsymbol{y} = |\boldsymbol{x} - \boldsymbol{y}|.$$

Если $s_2 = 0$, то $S_n(\overline{\boldsymbol{x}}, \boldsymbol{y}) < 2^n$, и, следовательно, $\boldsymbol{x} \ge \boldsymbol{y}$. В этом случае $\mathsf{s}_1 = 2^n - 1 - \boldsymbol{x} + \boldsymbol{y}$. Тогда

$$(\mathsf{s}_1 + s_2)^{s_2} = \overline{\mathsf{s}_1} = 2^n - 1 - (2^n - 1 - \boldsymbol{x} + \boldsymbol{y}) = |\boldsymbol{x} - \boldsymbol{y}|.$$

Таким образом, $(\mathsf{s}_1 + s_2)^{s_2} = |\boldsymbol{x} - \boldsymbol{y}|$ и число s_2 определяет знак разности $\boldsymbol{x} - \boldsymbol{y}$: разность отрицательна, если $s_2 = 1$, и неотрицательна, если $s_2 = 0$. Следовательно, пара (s_2, s_1) является разностью $\boldsymbol{x}$ и $\boldsymbol{y}$.

9.2 Вычисление суммы нескольких целых чисел

Ниже рассматриваются простые неглубокие схемы, вычисляющие сумму большого числа целых положительных чисел, заданных своими двоичными разложениями. Сначала докажем вспомогательное утверждение.

Лемма 9.2. *Пусть $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ — произвольные n-разрядные числа, $\mathbf{c}$ и $\boldsymbol{r}$ — такие $(n+1)$- и n-разрядное целые, что $\mathbf{c} - \boldsymbol{r} = \boldsymbol{x} + \boldsymbol{z} - \boldsymbol{y}$ и, более того,*

$$c_1 = 0, \quad 2c_{i+1} - r_i = x_i + z_i - y_i$$

для каждого $i \in \{1, \ldots, n\}$. Тогда существует вычисляющая $\mathbf{c}$ и $\boldsymbol{r}$ схема $\widetilde{\Sigma}_n$ для сложности и глубины которой справедливы равенства

$$L(\widetilde{\Sigma}_n) = 5n, \quad D(\widetilde{\Sigma}_n) = 3.$$

Доказательство. Так как $\mathbf{c} + \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{z} + \boldsymbol{r}$ и $c_1 = 0$ и $2c_{i+1} + y_i = x_i + z_i + r_i$ для каждого $i \in \{1, \ldots, n\}$, то легко видеть (см. (9.1) и (9.2)), что

$$c_{i+1} = x_i z_i \oplus r_i(x_i \oplus z_i), \quad y_i = x_i \oplus z_i \oplus r_i.$$

Из второго равенства находим $r_i = x_i \oplus y_i \oplus z_i$. Подставим r_i в первое равенство:

$$\begin{aligned} c_{i+1} &= x_i z_i \oplus r_i(x_i \oplus z_i) = x_i z_i \oplus (x_i \oplus y_i \oplus z_i)(x_i \oplus z_i) = \\ &= x_i z_i \oplus x_i \oplus z_i \oplus y_i(x_i \oplus z_i) = (x_i \vee z_i) \oplus y_i(x_i \oplus z_i). \end{aligned}$$

Тогда в качестве схемы $\widetilde{\Sigma}_n$ можно взять схему, изображенную на рисунке 9.2. Эта схема состоит из n независимо работающих одинаковых подсхем

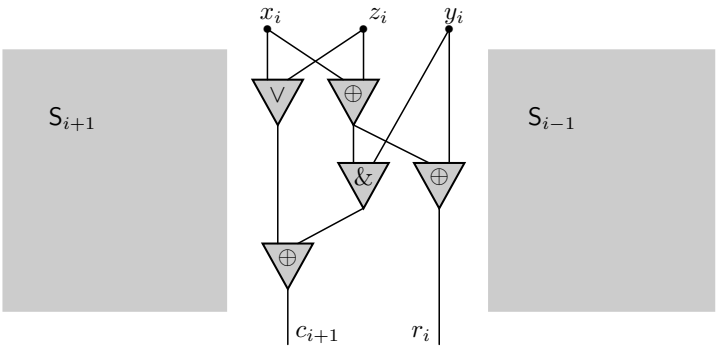


Рис. 9.2

S_i. Подсхема S_i имеет три входа и два выхода (входы и выходы S_i нумеруются слева направо). Входы S_i подключены к i-м разрядам чисел $\mathbf{x}$, $\boldsymbol{z}$ и $\boldsymbol{y}$. На первом выходе S_i вычисляется $(i+1)$-й разряд числа $\mathbf{c}$, на втором выходе — i-й разряд числа $\boldsymbol{r}$. Очевидно, что схема на рисунке 9.2 состоит из $5n$ элементов, а ее глубина равна трем. Лемма доказана.

Далее выходы $\widetilde{\Sigma}_n$ на которых вычисляются разряды числа $\mathbf{c}$ будем называть положительными, а выходы на которых вычисляются разряды числа $\boldsymbol{r}$ — отрицательными.

Пусть $\boldsymbol{x}_1, \boldsymbol{x}_2$ — n-разрядные двоичные числа. Пару $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ назовем n-разрядным двойным числом $\boldsymbol{x}$, значением $\boldsymbol{x}$ назовем разность $\boldsymbol{x}_1 - \boldsymbol{x}_2$. Число $\boldsymbol{x}_1$ называется положительной компонентой $\boldsymbol{x}$, а $\boldsymbol{x}_2$ — отрицательной компонентой $\boldsymbol{x}$. На множестве двойных чисел естественным образом определяются противоположное число и сумма двух чисел:

$$-(\boldsymbol{x}_1, \boldsymbol{x}_2) = (\boldsymbol{x}_2, \boldsymbol{x}_1),$$
$$(\boldsymbol{x}_1, \boldsymbol{x}_2) + (\boldsymbol{y}_1, \boldsymbol{y}_2) = (\boldsymbol{p}, \boldsymbol{q}),$$

где $\boldsymbol{p}$ и $\boldsymbol{q}$ такие, что $\boldsymbol{p} - \boldsymbol{q} = (\boldsymbol{x}_1 + \boldsymbol{y}_1) - (\boldsymbol{x}_2 + \boldsymbol{y}_2)$. Заметим, что хотя сумма двух двойных чисел определена не единственным образом (например $(0,0)+(1,0) = ((01),(00)) = ((10),(01))$), значение суммы всегда определено однозначно. Далее, говоря о сумме любого количества двойных чисел будем иметь ввиду не конкретную пару $(\boldsymbol{p}, \boldsymbol{q})$, а целый класс чисел имеющих одно и тоже значение. В частности под вычислением суммы двойных чисел понимается нахождение любого двойного числа значение которого равно сумме значений суммируемых чисел.

Прежде чем рассматривать вопросы о сложности и глубине сложения двойных чисел приведем одно полезное утверждение, являющееся тривиальным следствием доказательства леммы 9.2.

Лемма 9.3. *Пусть $\boldsymbol{x}, \boldsymbol{y}$ — произвольные n-разрядные числа, $\mathbf{c}$ и $\boldsymbol{r}$ — такие $(n+1)$- и n-разрядное целые, что $\mathbf{c} - \boldsymbol{r} = \boldsymbol{x} + \boldsymbol{y}$. Существует схема $\Sigma'_{2,n}$ вычисляющая $\mathbf{c}$ и $\boldsymbol{r}$ для сложности и глубины которой справедливы равенства*

$$L(\Sigma'_{2,n}) = 2n, \quad D(\Sigma'_{2,n}) = 1.$$

Лемму 9.3 можно рассматривать как утверждение о сложности и глубине сложения двух n-разрядных чисел, при условии, что результатом такого сложения будет двойное число.

Лемма 9.4. *Для любого $n \geq 1$ существует схема $\Sigma_{2,n}$, вычисляющая сумму двух n-разрядных двойных чисел, для сложности и глубины которой справедливы равенства*

$$L(\Sigma_{2,n}) = 10n - 3, \quad D(\Sigma_{2,n}) = 5.$$

Доказательство. Пусть $(\boldsymbol{x}, \boldsymbol{y})$, $(\boldsymbol{z}, \boldsymbol{w})$ — произвольные n-разрядные двойные числа. Положим $(\boldsymbol{x}, \boldsymbol{y}) + (\boldsymbol{z}, \boldsymbol{w}) = (\boldsymbol{p}, \boldsymbol{q})$. Пусть целые $\mathbf{c}$ и $\boldsymbol{r}$ вычислены схемой Σ_n из леммы 9.2 при условии, что на ее входы поданы числа $\boldsymbol{x}$, $\boldsymbol{z}$ и $\boldsymbol{y}$. Тогда $\boldsymbol{x} + \boldsymbol{z} - \boldsymbol{y} = \mathbf{c} - \boldsymbol{r}$ и

$$\begin{aligned}(\boldsymbol{x}, \boldsymbol{y}) + (\boldsymbol{z}, \boldsymbol{w}) &= (\boldsymbol{x} + \boldsymbol{z} - \boldsymbol{y}) - \boldsymbol{w} = \\ &= (\mathbf{c} - \boldsymbol{r}) - \boldsymbol{w} = -(\boldsymbol{r} + \boldsymbol{w} - \mathbf{c}).\end{aligned} \tag{9.4}$$

Из последнего равенства видно, что сумму $(\boldsymbol{x}, \boldsymbol{y}) + (\boldsymbol{z}, \boldsymbol{w})$ можно вычислить при помощи двух схем $\widetilde{\Sigma}_n$ и $\widetilde{\Sigma}_{n+1}$. Сначала схема $\widetilde{\Sigma}_n$ применяется к числам $\boldsymbol{x}$, $\boldsymbol{z}$ и $\boldsymbol{y}$. В результате получим два $(n+1)$-разрядное число $\mathbf{c}$ и n-разрядное число $\boldsymbol{r}$. Затем к числам $\boldsymbol{r}$, $\boldsymbol{w}$ и $\mathbf{c}$ применяется схема $\widetilde{\Sigma}_{n+1}$. Причем первые два входа j-й подсхемы схемы $\widetilde{\Sigma}_{n+1}$ подключаются к j-м разрядам чисел $\boldsymbol{r}$ и $\boldsymbol{w}$, а третий вход — к j-му разряду числа $\mathbf{c}$. В соответствии с (9.4), схема $\widetilde{\Sigma}_{n+1}$ вычислит число $\boldsymbol{r} + \boldsymbol{w} - \mathbf{c} = -(\boldsymbol{p}, \boldsymbol{q})$. Так как $-(\boldsymbol{p}, \boldsymbol{q}) = (\boldsymbol{q}, \boldsymbol{p})$, то очевидно, что сумма $(\boldsymbol{x}, \boldsymbol{y}) + (\boldsymbol{z}, \boldsymbol{w})$ вычислена: на отрицательных выходах $\widetilde{\Sigma}_{n+1}$ вычисляются разряды числа $\boldsymbol{p}$, на положительных — разряды числа $\boldsymbol{q}$. Схема $\Sigma_{2,n}$ построена.

Из леммы 9.2 легко следует, что $\Sigma_{2,n}$ содержит не более $10n+5$ элементов, а ее глубина не превосходит шести. В действительности рассматриваемая схема несколько проще. Заметим, что в схеме $\widetilde{\Sigma}_{n+1}$ в каждой подсхеме S_j глубина третьего входа равна двум, а в подсхеме S_{n+1} только один существенный вход. Поэтому схему $\Sigma_{2,n}$ можно представить в виде объединения n одинаковых подсхем S_j, каждая из которых содержит по десять элементов. Конструкция подсхемы S_j изображена на рисунке 9.3. Подсхема S_j имеет пять входов и три выхода. Входы S_j подключены к j-м разрядам чисел $\boldsymbol{x}$, $\boldsymbol{z}$, $\boldsymbol{y}$, $\boldsymbol{w}$, и к первому выходу подсхемы S_{j-1}. Второй и третий выходы S_j являются выходами схемы $\Sigma_{2,n}$. На втором выходе S_j вычисляется $(j+1)$-й разряд числа $\boldsymbol{q}$, на третьем выходе — j-й разряд числа $\boldsymbol{p}$. Первый выход подсхемы S_n также является выходом $\Sigma_{2,n}$ и на нем вычисляется $(n+1)$-й разряд числа $\boldsymbol{q}$.

При $j \in \{2, \ldots, n\}$ сложность каждой подсхемы S_j равна десяти, а глубина — пяти. Так как на пятый вход подсхемы S_1 подается тождественный нуль ($c_1 \equiv 0$), то легко видеть, что в S_1 три последних элемента можно

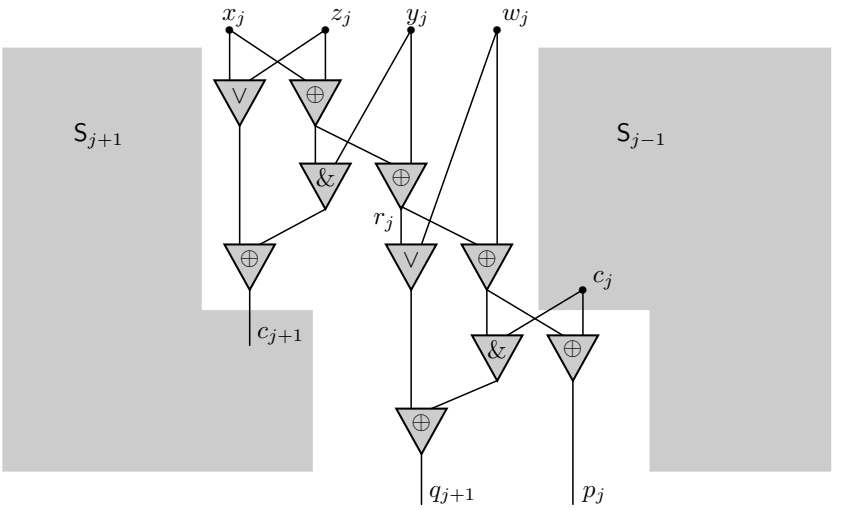


Рис. 9.3

удалить. Поэтому, сложность схемы $\Sigma_{2,n}$ равна $10n-3$, а глубина такая же как и у подсхем S_j, т. е. пять. Лемма доказана.

Теперь, применяя лемму 9.4, нетрудно установить следующий результат о глубине суммы большого числа двойных чисел.

Теорема 9.3. *Для любых $N, n \geq 1$ существует схема $\Sigma_{N,n}$, вычисляющая сумму N n-разрядных двойных чисел, для сложности и глубины которой при $N \to \infty$ справедливы неравенства*

$$L(\Sigma_{N,n}) \leq 10N(n+1)(1+o(1)), \qquad D(\Sigma_{N,n}) \leq 5\lceil \log_2 N \rceil.$$

Доказательство. Пусть $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ — произвольные n-разрядные двойные числа. Схему $\Sigma_{N,n}$ построим в соответствии со следующим алгоритмом. Числа $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ разобьем на пары и для каждой пары вычислим ее сумму используя построенные в лемме 9.4 схемы $\Sigma_{2,n}$. В результате получится примерно $\frac{1}{2}N$ $(n+1)$-разрядных двойных чисел. Новые числа снова разобьем на пары и для каждой пары вычислим ее сумму и т. д. Будем выполнять итерации до тех пор, пока не останется всего одно число.

Оценим глубину и сложность схемы $\Sigma_{N,n}$. Легко видеть, что число итераций не превосходит $\lceil \log_2 N \rceil$, а так как каждая итерация выполняется схемой глубины пять, то

$$D(\Sigma_{N,n}) \leq 5\lceil \log_2 N \rceil.$$

Теперь оценим сложность схемы $\Sigma_{N,n}$. Положим $R = \lceil \log_2 N \rceil$. Через N_i обозначим количество чисел, остающихся после i-й итерации. Легко видеть, что

$$N_i \leq \frac{1}{2}\Big(N_{i-1}+1\Big) < \left(\frac{1}{2}\right)^i N + 1.$$

На i-й итерации используется не более N_i схем $\Sigma_{2,n+i-1}$, поэтому

$$\begin{aligned} L(\Sigma_{N,n}) &\leq \sum_{i=1}^{R} 10(n+i-1)N_i \leq 10\sum_{i=1}^{R}(n+i-1)\Big(\Big(\frac{1}{2}\Big)^i N + 1\Big) \leq \\ &\leq 10\sum_{i=1}^{R}\Big(\Big(\frac{1}{2}\Big)^i N(n-1) + \Big(\frac{1}{2}\Big)^i Ni + (n+i-1)\Big) \leq \\ &\leq 10N(n-1) + 20N + 5R(2n+R). \end{aligned}$$

Следовательно, при $N \to \infty$, для сложности схемы $\Sigma_{N,n}$ справедливо неравенство

$$L(\Sigma_{N,n}) \leq \ 10N(n+1)(1+o(1)).$$

Теорема доказана.

9.3 Умножение

Из теорем 9.1 и 9.3 легко следует, что существует схема S, которая умножает два n-разрядных двоичных числа со сложностью $\mathcal{O}(n^2)$ и глубиной $\mathcal{O}(\log_2 n)$. Покажем, что существуют более простые схемы. Использованная в следующей теореме конструкция принадлежит А. А. Карацубе, который первым показал, что два n-разрядных числа можно умножить со сложностью меньшей чем n^2.

Теорема 9.4. *Существует схема M_n, вычисляющая произведение двух n-разрядных целых чисел, для сложности и глубины которой при $n \to \infty$ справедливы неравенства*

$$L(\mathsf{M}_n) = \mathcal{O}(n^{\log_2 3}), \quad D(\mathsf{M}_n) = \mathcal{O}(\log_2 n).$$

Доказательство. Пусть $n = 2^k + 2$, $\boldsymbol{x}$ и $\boldsymbol{y}$ — произвольные $(2n-2)$-разрядные двойные числа. Представим их в виде

$$\boldsymbol{x} = \boldsymbol{x}_2 2^{n-1} + \boldsymbol{x}_1, \quad \boldsymbol{y} = \boldsymbol{y}_2 2^{n-1} + \boldsymbol{y}_1,$$

где каждое из чисел $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{y}_1$, $\boldsymbol{y}_2$ состоит не более чем из $n-1$ разрядов. Тогда

$$\boldsymbol{x}\boldsymbol{y} = \boldsymbol{x}_2\boldsymbol{y}_2 2^{2n-2} + (\boldsymbol{x}_2\boldsymbol{y}_1 + \boldsymbol{x}_1\boldsymbol{y}_2)2^{n-1} + \boldsymbol{x}_1\boldsymbol{y}_1.$$

Откуда после несложных преобразований для произведения $\boldsymbol{x}\boldsymbol{y}$ получаем равенство

$$\boldsymbol{x}\boldsymbol{y} = \boldsymbol{x}_2\boldsymbol{y}_2 2^{2n-2} + (\boldsymbol{x}_2\boldsymbol{y}_2 + \boldsymbol{x}_1\boldsymbol{y}_1)2^{n-1} - (\boldsymbol{x}_2 - \boldsymbol{x}_1)(\boldsymbol{y}_2 - \boldsymbol{y}_1)2^{n-1} + \boldsymbol{x}_1\boldsymbol{y}_1. \quad (9.5)$$

Следовательно, умножение двух $(2n-2)$-разрядных чисел сводится к двум умножениям $(n-1)$-разрядных чисел, одному умножению n-разрядных чисел и нескольким сложениям.

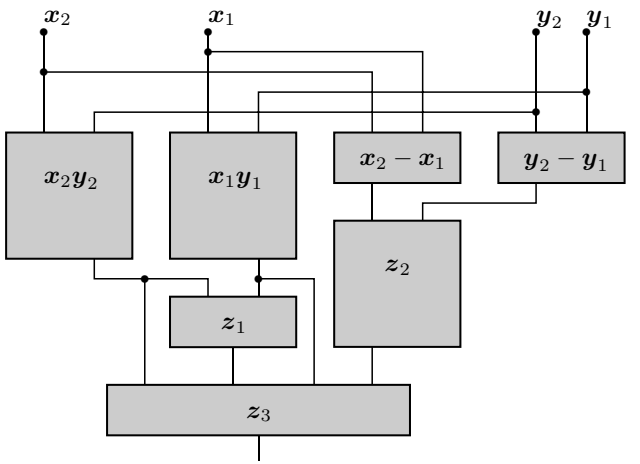


Рис. 9.4

Рекурсивная конструкция схемы M_{2n-2} показана на рисунке 9.4, где

$$\begin{aligned} z_1 &= x_2y_2 + x_1y_1, \\ z_2 &= (x_2 - x_1)(y_2 - y_1), \\ z_3 &= x_2y_2 2^{2n-2} + z_1 2^{n-1} - z_2 2^{n-1} + x_1y_1. \end{aligned}$$

Нетрудно видеть, что для сложности и глубины схемы M_{2n-2} имеют место рекуррентные соотношения

$$L(\mathsf{M}_{2n-2}) \le 3L(\mathsf{M}_n) + \mathcal{O}(n), \quad D(\mathsf{M}_{2n-2}) \le D(\mathsf{M}_n) + \mathcal{O}(1).$$

Так как $2n - 2 = 2^{k+1} + 2$ и $n = 2^k + 2$, то последние неравенства можно переписать в виде

$$L(\mathsf{M}_{2^{k+1}+2}) \le 3L(\mathsf{M}_{2^k+2}) + \mathcal{O}(2^k), \quad D(\mathsf{M}_{2^{k+1}+2}) \le D(\mathsf{M}_{2^k+2}) + \mathcal{O}(1).$$

Тогда справедливы неравенства

$$L(\mathsf{M}_{2^k+2}) \le \mathcal{O}(2^{k \log_2 3}), \quad D(\mathsf{M}_{2^k+2}) \le \mathcal{O}(k),$$

из которых легко следует утверждение теоремы. Теорема доказана.

В настоящее время разработаны различные алгоритмы умножения целых n-разрядных чисел, позволяющие строить при больших n значительно более экономные схемы. Наиболее простые из этих схем (см. [39]) состоят из $\mathcal{O}(n \log_2 n \cdot \log_2 \log_2 n)$ элементов, а их глубина равна $\mathcal{O}(\log_2 n)$. К сожалению константы, входящие в "$\mathcal{O}$" таковы, что эти схемы представляют в основном теоретический интерес.

9.4 Сортировка

Пусть $\boldsymbol{x} = (x_1, \ldots, x_n)$ — набор действительных чисел. *Сортировкой* набора $\boldsymbol{x}$ называется перестановка его разрядов в порядке невозрастания их величин. Сортировка встречается в качестве составной части большого числа разнообразных алгоритмов и является одной из наиболее важных комбинаторных задач. В этом параграфе будут построены схемы, сортирующие булевы наборы и имеющие небольшие сложности и глубины. Затем будет показано, что построенные схемы могут быть использованы для сортировки не только булевых наборов, но и наборов действительных чисел.

1. Набор $\boldsymbol{\alpha}$ из $\mathbb{B}^n$ называется *упорядоченным* если $\alpha_i \le \alpha_j$ для всех $1 \le i < j \le n$. Схема в базисе $\{\vee, \&\}$ с n входами и n выходами называется схемой *двоичной сортировки* или *сортирующей схемой*, если она преобразует произвольный набор в упорядоченный набор такого же веса. В следующей теореме приводится конструкция эффективных сортирующих схем. Эти схемы строятся из двухэлементных подсхем с двумя входами и двумя выходами. На первом выходе каждой подсхемы вычисляется конъюнкция ее входов, а на втором выходе — дизъюнкция входов. Такие двухэлементные схемы (см. рисунок 9.5) будем называть *булевыми компараторами*. Легко видеть, что компаратор является схемой, сортирующей наборы длины два.

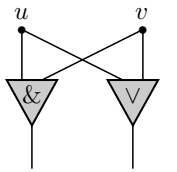


Рис. 9.5

Теорема 9.5. *Существует схема S_{2^k}, сортирующая наборы длины 2^k, для сложности и глубины которой справедливы равенства*

$$L(\mathsf{S}_{2^k}) = k(k-1)2^{k-1} + 2^{k+1} - 2, \qquad D(\mathsf{S}_{2^k}) = \frac{1}{2}k(k+1).$$

Доказательство. Сначала построим схему $\mathsf{S}_{2n,2n}$, объединяющую два упорядоченных набора $(u_1, \ldots, u_{2n})$ и $(v_1, \ldots, v_{2n})$ в один упорядоченный набор $(w_1, \ldots, w_{4n})$. Схема $\mathsf{S}_{2n,2n}$ называется $(2n, 2n)$-схемой нечетно-четного слияния и строится индуктивно. В основании индукции лежит схема $\mathsf{S}_{1,1}$, упорядочивающая два одноэлементных набора и состоящая из единственного компаратора. Очевидно, что

$$L(\mathsf{S}_{1,1}) = 2, \qquad D(\mathsf{S}_{1,1}) = 1. \tag{9.6}$$

Предположим, что схема $\mathsf{S}_{n,n}$ построена. Тогда схема $\mathsf{S}_{2n,2n}$, конструкция которой представлена на рисунке 9.6, строится следующим образом.

1. Из элементов с нечетными номерами составляются два упорядоченных набора $(u_1, u_3, \ldots, u_{2n-1})$ и $(v_1, v_3, \ldots, v_{2n-1})$, которые сливаются схемой $\mathsf{S}_{n,n}$ в упорядоченный набор $(p_1, \ldots, p_{2n})$.

2. Из элементов с четными номерами составляются два упорядоченных набора $(u_2, u_4, \ldots, u_{2n})$ и $(v_2, v_4, \ldots, v_{2n})$, которые сливаются в упорядоченный набор $(q_1, \ldots, q_{2n})$ схемой $\mathsf{S}_{n,n}$.

3. Наборы $(p_1, \ldots, p_{2n})$ и $(q_1, \ldots, q_{2n})$ преобразуются в упорядоченный набор $(w_1, \ldots, w_{4n})$ по формулам $w_1 = p_1$, $w_{2i} = p_{i+1} \& q_i$, $w_{2i+1} = p_{i+1} \vee q_i$ для $i = 1, 2, \ldots, 2n-1$, и $w_{4m} = q_{2m}$.

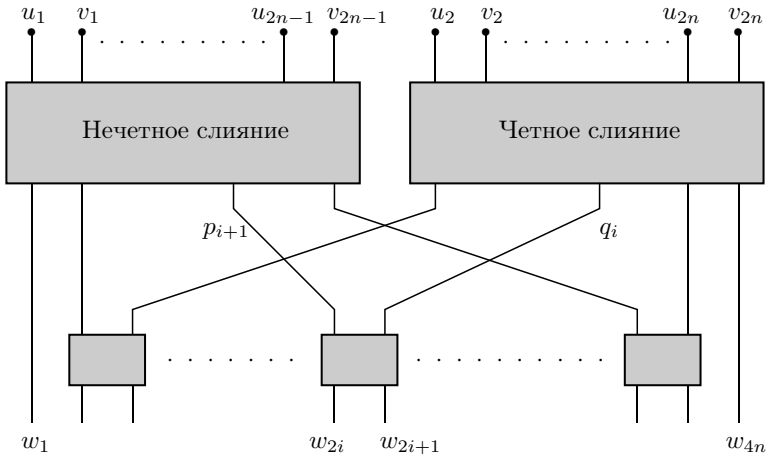


Рис. 9.6

Покажем, что схема $\mathsf{S}_{2n,2n}$ действительно преобразует два упорядоченных n-элементных набора в упорядоченный набор. Допустим, что набор $\boldsymbol{u} = (u_1, \ldots, u_{2n})$ состоит из k нулей и $2n-k$ единиц, а набор $\boldsymbol{v} = (v_1, \ldots, v_{2n})$ — из l нулей и $2n-l$ единиц. Тогда в наборе $(p_1, \ldots, p_{2n})$, получившемся после слияния нечетных элементов наборов $\boldsymbol{u}$ и $\boldsymbol{v}$, будет $t = \lceil k/2 \rceil + \lceil l/2 \rceil$ нулей и $2n - t$ единиц. Так же легко видеть, что набор $(q_1, \ldots, q_{2n})$, получившейся после слияния четных элементов наборов $\boldsymbol{u}$ и $\boldsymbol{v}$, будет состоять из $s = \lfloor k/2 \rfloor + \lfloor l/2 \rfloor$ нулей и $2n - s$ единиц. Так как для любого x разность $\lceil x \rceil - \lfloor x \rfloor$ может быть равна только нулю или единице, то величина

$$R = (\lceil k/2 \rceil + \lceil l/2 \rceil) - (\lfloor k/2 \rfloor + \lfloor l/2 \rfloor)$$

может принимать только три значения: 0, 1 и 2.

Если $R = 0$ или $R = 1$, то набор $(p_1, q_1, \ldots, p_{2n}, q_{2n})$ будет упорядоченным. Если $R = 2$, то в этом наборе на первых $k + l - 1$ местах будут стоять нули, на $(k+l)$-м месте будет стоять единица, на $(k+l+1)$-м месте будет стоять последний нуль, и на оставшихся местах — единицы. Из равенства $R = 2$ легко следует, что числа k и l нечетные, и, следовательно, $k + l = 2h$ — четное число. Тогда,

$$w_{2h} = p_{h+1} \& q_h = q_h = 0, \qquad w_{2h+1} = p_{h+1} \vee q_h = p_{h+1} = 1,$$

т. е. один из компараторов, находящихся в последнем ряду схемы, поменяет местами последний нуль и первую единицу набора $(p_1, q_1, \ldots, p_{2n}, q_{2n})$. Легко видеть, что после этого набор станет упорядоченным. Следовательно, схема $\mathsf{S}_{2n,2n}$ действительно объединяет два упорядоченных набора в упорядоченный набор.

Оценим сложность и глубину этой схемы. Из конструкции схемы имеем

$$L(\mathsf{S}_{2n,2n}) = 2L(\mathsf{S}_{n,n}) + (4n-2), \tag{9.7}$$
$$D(\mathsf{S}_{2n,2n}) = D(\mathsf{S}_{n,n}) + 1. \tag{9.8}$$

Индукцией по k покажем, что при $k \geq 0$ для сложности схемы $\mathsf{S}_{2^k,2^k}$ справедливо равенство

$$L(\mathsf{S}_{2^k,2^k}) = k2^{k+1} + 2. \tag{9.9}$$

Действительно, при $k = 0$ равенство (9.9) следует из (9.6). Допустим, что (9.9) верно при всех $k \leq m-1$. Тогда из этого предположения и равенства (9.7) имеем

$$\begin{aligned} L(\mathsf{S}_{2^m,2^m}) &= 2L(\mathsf{S}_{2^{m-1},2^{m-1}}) + 2 \cdot 2^m - 2 = \\ &= 2((m-1)2^m + 2) + 2 \cdot 2^m - 2 = m2^{m+1} + 2. \end{aligned}$$

Следовательно, (9.9) справедливо при всех целых $k \geq 0$.

Аналогичным образом, из (9.8) и (9.6) при всех целых $k \geq 0$ для глубины $\mathsf{S}_{2^k,2^k}$ имеем

$$D(\mathsf{S}_{2^k,2^k}) = k + 1. \tag{9.10}$$

Теперь, также индуктивно, построим схему S_{4n}, сортирующую наборы из $4n$ элементов. В основание индукции положим схему S_2, сортирующую двухэлементные наборы и состоящую из одного компаратора. Очевидно, что

$$L(\mathsf{S}_2) = 2, \qquad D(\mathsf{S}_2) = 1. \tag{9.11}$$

Допустим, что схема S_{2n} построена. Тогда схему S_{4n} составим из двух схем,

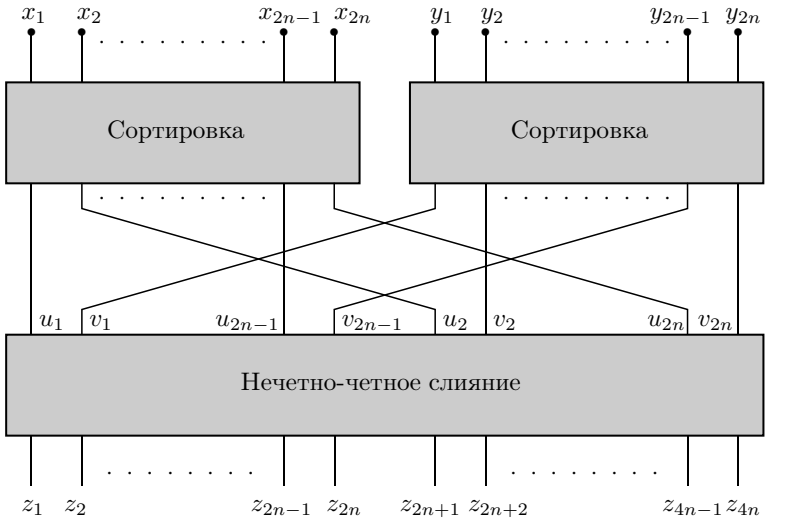


Рис. 9.7

сортирующих $2n$-элементные наборы, и одной схемы нечетно-четного слияния двух $2n$-элементных наборов. Конструкция схемы представлена на рисунке 9.7.

Оценим сложность и глубину этой схемы. Из конструкции схемы имеем

$$L(\mathsf{S}_{4n}) = 2L(\mathsf{S}_{2n}) + L(\mathsf{S}_{2n,2n}), \tag{9.12}$$
$$D(\mathsf{S}_{4n}) = D(\mathsf{S}_{2n}) + D(\mathsf{S}_{2n,2n}). \tag{9.13}$$

Индукцией по k покажем, что для сложности схемы S_{2^k} при всех $k \geq 1$ справедливо равенство

$$L(\mathsf{S}_{2^k}) = k(k-1)2^{k-1} + 2^{k+1} - 2. \tag{9.14}$$

При $k = 1$ равенство (9.14) следует из (9.11). Допустим, что оно верно при всех $k \leq m-1$. Тогда из предположения индукции, равенства (9.12) и равенства (9.9) имеем

$$\begin{aligned} L(\mathsf{S}_{2^m}) &= 2L(\mathsf{S}_{2^{m-1}}) + L(\mathsf{S}_{2^{m-1},2^{m-1}}) = \\ &= 2((m-1)(m-2)2^{m-2} + 2^m - 2) + (m-1)2^m + 2 = \\ &= m(m-1)2^{m-1} + 2^{m+1} - 2. \end{aligned}$$

Следовательно, (9.14) справедливо при всех целых $k \geq 1$. Из (9.10) и (9.13) для глубины $\mathsf{S}_{2^k,2^k}$ имеем

$$D(\mathsf{S}_{2^k}) = D(\mathsf{S}_{2^{k-1}}) + k = \sum_{j=1}^{k} j = \frac{1}{2}k(k+1).$$

Теорема доказана.

Построенные в доказательстве теоремы 9.5 сортирующие схемы были предложены в 1968 году Бэтчером, и называются теперь *схемами Бэтчера*.

2. Схемы из теоремы 9.5 можно использовать для сортировки наборов действительных чисел. Для этого в схеме S_{2^k} элементы дизъюнкции надо заменить элементами вычисления максимума, а элементы конъюнкции — элементами вычисления минимума. То, что преобразованная схема будет сортировать наборы действительных чисел вытекает из следующей теоремы.

Теорема 9.6. *Схема S с n входами не является сортирующей схемой только в том случае, когда существует булев набор длины n, который не может быть отсортирован этой схемой.*

Доказательство. Допустим, что схема S преобразует последовательность $u_1, \ldots, u_n$ в последовательность $v_1, \ldots, v_n$. Пусть $f : \mathbb{R} \to \mathbb{R}$ — произвольная монотонная функция. Индукцией по числу компараторов легко показать, что схема S преобразует последовательность $f(u_1), \ldots, f(u_n)$ в последовательность $f(v_1), \ldots, f(v_n)$. Предположим, что последовательность

$v_1, \ldots, v_n$ не является упорядоченной. Тогда найдется такое i, что $v_{i+1} < v_i$. Функцию f определим следующим образом:

$$f(x) = \begin{cases} 0, & \text{если } x \leq v_{i+1}, \\ 1, & \text{если } x > v_{i+1}. \end{cases}$$

В этом случае двоичная последовательность $f(u_1), \ldots, f(u_n)$ будет преобразована схемой S в неупорядоченную двоичную последовательность $f(v_1), \ldots, f(v_n)$. Теорема доказана.

В 1983 г. М. Айтаи, Я. Комлош и Е. Семереди показали, что существуют сортирующие схемы с n входами, которые состоят из $\mathcal{O}(n \log_2 n)$ компараторов и глубина которых есть $\mathcal{O}(\log_2 n)$. По имени авторов эти схемы сейчас называются AKS-схемами. Однако как и в случае схем, умножающих целые числа, константы, входящие в "$\mathcal{O}$", слишком велики для того, чтобы AKS-схемы использовались в реальных вычислениях.

9.5 Задачи

9.1. Функции $y_j = \bigoplus_{i=1}^{j} x_i$, где $j = 1, \ldots, n$, называются префиксными суммами переменных $x_1, \ldots, x_n$. Показать, что при $n = 2^k$ существует вычисляющая префиксные суммы схема S такая, что $L(\mathsf{S}) \leq 2n$ и $D(\mathsf{S}) = 2\log_2 n$.

9.2. Показать, что при $n = 2^k$ существует вычисляющая префиксные суммы схема S такая, что $L(\mathsf{S}) = \mathcal{O}(n \log_2 n)$ и $D(\mathsf{S}) = \log_2 n$.

9.3. Показать, что при $n \to \infty$ существует вычисляющая префиксные суммы схема S такая, что $L(\mathsf{S}) = \mathcal{O}(n)$ и $D(\mathsf{S}) = \log_2 n + \mathcal{O}(1)$.

9.4. Показать, что при $n \to \infty$ существует схема, вычисляющая сумму двух двоичных n-разрядных чисел, сложность которой есть $\mathcal{O}(n)$, а глубина асимптотически не больше $2\log_2 n$.

9.5. Построить схему, преобразующую три целых n-разрядных числа в два числа с такой же суммой.

9.6. Построить схему, преобразующую четыре целых n-разрядных числа в два числа с такой же суммой.

9.7. Построить схему, вычисляющую сумму трех n-разрядных двойных чисел, глубина которой равна семи.

9.8. Показать, что при $n \to \infty$ существует схема, преобразующая n целых двоичных чисел в два числа с такой же суммой, глубина которой асимптотически не больше $4\log_2 n$.

9.9. Показать, что при $n \to \infty$ существует схема, преобразующая n целых двоичных чисел в два числа с такой же суммой, глубина которой асимптотически не больше $3,71\log_2 n$.

9.10. Показать, что при $n \to \infty$ существует схема, вычисляющая сумму n двоичных n-разрядных чисел, глубина которой асимптотически не больше $6\log_2 n$.

9.11. Пусть $f_1, \ldots, f_k$ — булевы функции, $f = f_1 \vee \cdots \vee f_k$. Показать, что

$$D(f) \leq \left\lceil \log_2 \sum_{i=1}^{k} 2^{D(f_i)} \right\rceil.$$

9.12. Пусть $\mathbf{f}$ — вектор значений n-местной булевой функции f, $\widehat{\mathbf{f}}$ — вектор коэффициентов многочлена Жегалкина функции f. Показать, что существует схема S_n, состоящая из элементов умножения и сложения по модулю два, вычисляющая вектор $\widehat{\mathbf{f}}$ по вектору $\mathbf{f}$, сложность и глубина которой удовлетворяют равенствам:

$$L(\mathsf{S}_n) = n2^{n-1}, \qquad D(\mathsf{S}_n) = n.$$

9.13. Пусть f и g — булевы функции n переменных. Показать, что сложность и глубина вычисления коэффициентов многочлена Жегалкина функции fg по известным коэффициентам многочленов Жегалкина функций f и g есть $\mathcal{O}(n2^{n-1})$ и $\mathcal{O}(n)$.

9.14. Показать, что любая сортирующая схема с n входами состоит не менее чем из $n \log_2 n - 2n$ компараторов.

Лекция 10

Универсальные методы синтеза схем

В середине сороковых годов двадцатого века в математике появились задачи, связанные со сложностью булевых функций, при помощи которых описывалось функционирование узлов создаваемых в то время первых компьютеров. Одной из таких задач стала задача К. Шеннона о сложности произвольной n-местной булевой функции. Для вычисления булевых функций Шеннон использовал контактные схемы — математическую модель, наиболее адекватно описывающую элементную базу вычислительной техники того времени. Он показал, что при $n \to \infty$ сложность каждой n-местной булевой функции есть $\mathcal{O}(2^n/n)$, и при этом сложность почти всех таких функций по порядку величины не меньше чем $2^n/n$. Для доказательства верхней оценки сложности произвольной функции им был создан универсальный метод построения контактных схем, позволяющий единым для всех булевых функций образом строить вычисляющие эти функции схемы. Для доказательства нижней оценки Шенноном совместно с Дж. Риорданом был предложен мощностной метод, основанный на простом наблюдении, что различные функции должны вычислятся разными схемами, и заключающийся в сравнении числа различных схем определенной сложности и числа всех n-местных булевых функций.

Предложенный Шенноном подход к изучению сложности булевых функций оказался очень плодотворным, был перенесен на другие модели вычислений, в том числе и на схемы из функциональных элементов, и впоследствии был развит О. Б. Лупановым, из результатов которого в частности следует, что при $n \to \infty$ сложность почти всех n-местных булевых функций асимптотически равна $2^n/n$.

10.1 Метод Шеннона

Основная идея метода Шеннона состоит в предварительном вычислении большого количества вспомогательных величин, каждая из которых многократно используется для получения окончательного результата. Идея Шен-

нона проста и эффективна и может быть успешно использована в различных задачах. Ниже эта идея используется в двух случаях: для построения схем, вычисляющих произвольные булевы функции, и для построения схем, вычисляющих системы линейных булевых функций.

Следующая теорема является простым следствием известного результата Шеннона 1949 г. о сложности реализации булевых функций контактными схемами.

Теорема 10.1. *Пусть $n \to \infty$. Тогда для каждой булевой функции f, зависящей от n переменных,*

$$L(f) \le \frac{6 \cdot 2^n}{n}(1 + o(1)).$$

Доказательство. Функцию f разложим по первым k переменным:

$$f(x_1, \ldots, x_n) = \bigvee_{\sigma_1 \ldots \sigma_k} f(\sigma_1, ..., \sigma_k, x_{k+1}, \ldots, x_n) \cdot x_1^{\sigma_1} \cdots x_k^{\sigma_k}. \qquad (10.1)$$

Опишем конструкцию схемы S, вычисляющей функцию f. Эта схема состоит из трех подсхем S_1–S_3, и ее конструкция основана на формуле (10.1). Подсхема S_1 вычисляет все элементарные конъюнкции первых k переменных. Подсхема S_2 вычисляет все булевы функции, зависящие от последних $n-k$ переменных. Подсхема S_3 вычисляет функцию f по формуле (10.1), используя функции, вычисленные подсхемами S_1 и S_2. Нетрудно видеть, что подсхема S_3 состоит из не более чем 2^k дизъюнкторов и 2^k конъюнкторов.

Далее в двух леммах оценим сложности подсхем S_1 и S_2. Сначала оценим сложность множества K_k, состоящего из всех элементарных конъюнкций вида $x_1^{\sigma_1} \& \ldots \& x_k^{\sigma_k}$.

Лемма 10.1. *Существует схема K_k, вычисляющая все элементарные конъюнкции переменных $x_1, \ldots, x_k$, и для сложности которой справедливо равенство*

$$L(\mathsf{K}_k) = 2^k + \mathcal{O}\left(k 2^{k/2}\right).$$

Доказательство. Положим $m = \lceil k/2 \rceil$, $l = \lfloor k/2 \rfloor$. Опишем схему K_n, удовлетворяющую равенству леммы. Эта схема состоит из трех подсхем A, B и C. Подсхема A вычисляет все элементарные конъюнкции переменных $x_1, \ldots, x_m$. Подсхема B вычисляет все элементарные конъюнкции переменных $x_{m+1}, \ldots, x_{m+l}$. Подсхема C состоит из $2^m \cdot 2^l = 2^k$ конъюнкторов. Каждый из этих конъюнкторов умножает один выход подсхемы A на один выход подсхемы B. Выходами схемы K_n являются все элементы подсхемы C. Нетрудно видеть, что $L(\mathsf{A}) = \mathcal{O}(m2^m)$ и $L(\mathsf{B}) = \mathcal{O}(l2^l)$. Следовательно,

$$L(\mathsf{K}_n) = L(\mathsf{A}) + L(\mathsf{B}) + L(\mathsf{C}) \le \mathcal{O}(k2^{k/2}) + 2^k. \qquad (10.2)$$

Лемма доказана.

Теперь оценим сложность множества U_{n-k}, состоящего из всех $(n-k)$-местных булевых функций.

Лемма 10.2. *Существует схема U_{n-k}, вычисляющая все $(n-k)$-местные булевы функции, для сложности которой справедливо неравенство*

$$L(\mathsf{U}_{n-k}) \le 3 \cdot 2^{n-k} 2^{2^{n-k}}.$$

Доказательство. Утверждение леммы следует из того, что существует ровно $2^{2^{n-k}}$ различных $(n-k)$-местных булевых функций, и в силу первого из неравенств (8.2) на стр. 109 сложность каждой из этих функций не превосходит $3 \cdot 2^{n-k}$. Лемма доказана.

Продолжим доказательство теоремы. Из конструкции схемы S и лемм 10.1 и 10.2 легко следует неравенство

$$L(\mathsf{S}) \le 3 \cdot 2^k (1 + o(1)) + \mathcal{O}\left(2^{n-k} 2^{2^{n-k}}\right). \qquad (10.3)$$

Положим $k = \lceil n - \log_2(n - 3\log_2 n) \rceil$. Тогда $n - k \le \log_2(n - 3\log_2 n)$. Подставляя выбранное значение k в (10.3), видим, что

$$L(\mathsf{S}) \le \frac{3 \cdot 2^{n+1}}{n}(1 + o(1)) + \mathcal{O}\left(n \cdot \frac{2^n}{n^3}\right) = \frac{6 \cdot 2^n}{n}(1 + o(1)).$$

Теорема доказана.

В доказательстве теоремы 10.1 параметр k выбран так, что число всех $(n-k)$-местных булевых функций существенно меньше (примерно в n^2 раз) числа элементарных дизъюнкций k переменных. Поэтому каждая функция $f(\sigma_1, ..., \sigma_k, x_{k+1}, \ldots, x_n)$ в формуле (10.1) умножается в среднем на n^2 элементарных дизъюнкций, т. е. эти функции и есть те вспомогательные величины, которые упоминались перед формулировкой теоремы 10.1.

Теперь воспользуемся идеей Шеннона в несколько более сложной ситуации — рассмотрим задачу о вычислении произвольной системы линейных булевых функций. В этой задаче идея Шеннона применяется для вычисления систем, в которых число переменных существенно больше числа функций.

Лемма 10.3. *Пусть $m \le \log_2 n - 2\log_2 \log_2 n$. Тогда при $n \to \infty$ произвольная система $F(x_1, \ldots, x_n)$, состоящая из m линейных булевых функций, может быть вычислена такой схемой S, что*

$$L(\mathsf{S}) = n + \mathcal{O}\left(\frac{n}{\log_2 n}\right).$$

Доказательство. Пусть $\mathbf{F}$ — матрица системы F. Если система F симметрична относительно переменных $x_{i_1}, \ldots, x_{i_k}$, то столбцы матрицы $\mathbf{F}$ с номерами $i_1, \ldots, i_k$ будут одинаковыми. Параметры леммы m и n таковы ($2^m \ll n$), что число различных видов столбцов высоты m много меньше числа столбцов в $\mathbf{F}$. Поэтому в матрице $\mathbf{F}$ будет много одинаковых столбцов, и, следовательно, система F симметрична относительно некоторых подмножеств своих переменных. Используем это свойство рассматриваемой системы. Для каждого $j \in \{1, 2, \ldots, 2^m - 1\}$ сформируем множество переменных

M_j такое, что переменная x_i принадлежит M_j тогда и только тогда, когда i-й столбец матрицы $\mathbf{F}$ совпадает с двоичным представлением числа j. Далее для каждого непустого множества M_j определим сумму входящих в это множество переменных:

$$y_j = \bigoplus_{x_i \in M_j} x_i. \tag{10.4}$$

Легко видеть, что для k-й функции f_k, $1 \le k \le m$, системы F справедливо равенство

$$f_k = \bigoplus y_j, \tag{10.5}$$

в котором суммирование производится по всем тем целым j, $1 \le j \le 2^m - 1$, в двоичном представлении которых коэффициент при $(m-k)$-й степени двойки равен единице.

Каждую функцию y_j вычислим отдельной схемой S_j. Так как каждая переменная x_i входит не более чем в одну сумму y_j, то, очевидно, что

$$\sum_{j=1}^{2^m-1} L(\mathsf{S}_j) < n. \tag{10.6}$$

Каждую функцию f_k вычислим отдельной схемой S'_k. Схема S'_k вычисляет f_k в соответствии с формулой (10.5) и ее входы подключены к выходам соответствующих схем S_j. Из (10.5) немедленно следует, что

$$\sum_{k=1}^{m} L(\mathsf{S}'_k) \le m \cdot 2^m \le \frac{n}{\log_2 n}. \tag{10.7}$$

Таким образом, из (10.6) и (10.7) следует, что $L(\mathsf{S}) \le n + \frac{n}{\log_2 n}$. Лемма доказана.

Нетрудно видеть, что упоминавшимися выше вспомогательными величинами в доказательстве леммы 10.3 являются функции y_j. каждая из которых в общем случае входит примерно в половину сумм в (10.5).

Теперь рассмотрим общую ситуацию, когда двоичный логарифм числа переменных системы может быть больше числа ее функций. В этом случае в матрице системы может не найтись одинаковых столбцов, и поэтому напрямую воспользоваться разработанной выше техникой построения схем не удастся. Расширить границы применимости этой техники можно при помощи очень простого приема — разделения системы на несколько подсистем, размеры каждой из которых удовлетворяют условиям леммы 10.3. Именно такой способ использован в доказательстве следующей леммы.

Лемма 10.4. *При $n \to \infty$ произвольная система $F(x_1, \ldots, x_n)$, состоящая из m линейных функций, может быть вычислена такой схемой S, что*

$$L(\mathsf{S}) \le \frac{n(m + \log_2 n)}{\log_2 n} \left(1 + \mathcal{O}\left(\frac{\log_2 \log_2 n}{\log_2 n}\right)\right).$$

Доказательство. Положим $q = \lfloor \log_2 n - 2\log_2 \log_2 n \rfloor$, $k = \lceil m/q \rceil$. Без ограничения общности полагаем, что $m \ge q$, так как при $m < q$ настоящая лемма следует из леммы 10.3. Из системы F сформируем k новых систем $F_1, \ldots, F_k$ так, что

$$F_j = \{f_{(j-1)q+1}, \ldots, f_{jq}\} \text{ при } j = 1, 2, \ldots, k-1,$$
$$F_k = \{f_{(k-1)q+1}, \ldots, f_m\}.$$

Каждую из систем F_i вычислим собственной схемой S_i, сложность которой в силу леммы 10.3 удовлетворяет неравенству

$$L(\mathsf{S}_i) \le n + \frac{n}{\log_2 n}.$$

Очевидно, что схема $\mathsf{S} = \cup_{i=1}^{k} \mathsf{S}_i$ вычисляет систему $F = \cup_{i=1}^{k} F_i$, и для ее сложности справедливы неравенства

$$L(\mathsf{S}) = \sum_{i=1}^{k} L(\mathsf{S}_i) \le \left(n + \frac{n}{\log_2 n}\right) \left\lceil \frac{m}{\lfloor \log_2 n - 2\log_2 \log_2 n \rfloor} \right\rceil \le$$
$$\le \frac{n(m + \log_2 n)}{\log_2 n} \left(1 + \mathcal{O}\left(\frac{\log_2 \log_2 n}{\log_2 n}\right)\right).$$

Лемма доказана.

10.2 Метод Лупанова

В 1958 г. О. Б. Лупанов предложил новый метод синтеза схем для произвольных булевых функций. Метод Лупанова менее универсален чем метод Шеннона, так как в большей мере использует особенности решаемой задачи. Учет этих особенностей позволил усилить результат теоремы 10.1 и получить (см. теоремы 10.2 и 10.5) в некотором смысле окончательное решение задачи о сложности произвольной булевой функции.

Пусть $H_{n,m}(i) = \{h\}$ — множество всех таких n-местных булевых функций, что $h(\alpha) = 0$, если $|\alpha| \le m(i-1)$ или $|\alpha| > mi$. Нетрудно видеть, что каждое множество $H_{n,m}(i)$, где $i = 1, \ldots, \lceil 2^n/m \rceil$, состоит не более чем из 2^m различных функций, а вместе все множества $H_{n,m}(i)$ содержат не более $2^{n+m+1}/m$ различных функций. Также нетрудно видеть, что любая n-местная булева функция f может быть представлена в виде дизъюнкции

$$f(x_1, \ldots, x_n) = \bigvee_{i=1}^{\lceil 2^n/m \rceil} f_i(x_1, \ldots, x_n) \tag{10.8}$$

функций из $H_{n,m}(i)$, где $f_i \in H_{n,m}(i)$ и $f_i(\alpha) = f(\alpha)$ для каждого α такого, что $m(i-1) < |\alpha| \le mi$. Используем равенство (10.8) для доказательства следующей теоремы Лупанова.

Теорема 10.2. *Пусть $n \to \infty$. Тогда для каждой булевой функции f, зависящей от n переменных,*

$$L(f) \le \frac{2^n}{n}\left(1 + \mathcal{O}\left(\frac{\log_2 n}{n}\right)\right).$$

Доказательство. Функцию f разложим по первым k переменным:

$$\begin{aligned} f(x_1,\dots,x_n) &= \bigvee_{\sigma_1\dots\sigma_k} f(\sigma_1,...,\sigma_k,x_{k+1},\dots,x_n)\cdot x_1^{\sigma_1}\cdots x_k^{\sigma_k} = \\ &= \bigvee_{\sigma_1\dots\sigma_k} f_{\sigma_1,\dots,\sigma_k}(x_{k+1},\dots,x_n)\cdot x_1^{\sigma_1}\cdots x_k^{\sigma_k}. \end{aligned} \tag{10.9}$$

Затем каждую функцию $f_{\sigma_1,\dots,\sigma_k}(x_{k+1},\dots,x_n)$ разложим в дизъюнкцию

$$f_{\sigma_1,\dots,\sigma_k}(x_{k+1},\dots,x_n) = \bigvee_{i=1}^{\lceil 2^{n-k}/m\rceil} f_{\boldsymbol{\sigma},i}(x_{k+1},\dots,x_n) \tag{10.10}$$

функций из $H_{n-k,m}(i)$, где $\boldsymbol{\sigma} = (\sigma_1\dots\sigma_k)$ и $f_{\boldsymbol{\sigma},i} \in H_{n-k,m}(i)$. Тогда

$$f(x_1,\dots,x_n) = \bigvee_{\sigma_1\dots\sigma_k} x_1^{\sigma_1}\cdots x_k^{\sigma_k}\cdot\Big(\bigvee_{i=1}^{\lceil 2^{n-k}/m\rceil} f_{\boldsymbol{\sigma},i}(x_{k+1},\dots,x_n)\Big). \tag{10.11}$$

Воспользуемся последним равенством для построения вычисляющей функцию f схемы S. Эта схема состоит из пяти подсхем (соединение подсхем изображено на рисунке 10.1), устроенных следующим образом.

1. Подсхема S_1 является экземпляром схемы K_k, построенной в лемме 10.1, и вычисляет все элементарные конъюнкции вида $x_1^{\sigma_1}\cdots x_k^{\sigma_k}$. Очевидно, что

$$L(\mathsf{S}_1) \le 2^k + \mathcal{O}\left(k2^{k/2}\right).$$

2. Подсхема S_2 также является экземпляром схемы K_{n-k}, построенной в лемме 10.1, и вычисляет все элементарные конъюнкции вида $x_{k+1}^{\sigma_{k+1}}\cdots x_n^{\sigma_n}$. Очевидно, что

$$L(\mathsf{S}_2) \le 2^{n-k} + \mathcal{O}\left((n-k)2^{(n-k)/2}\right).$$

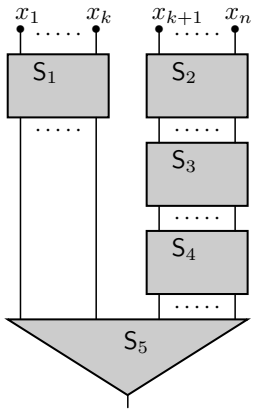


Рис. 10.1

3. Подсхема S_3 вычисляет все функции множеств $H_{n-k,m}(i)$, используя элементарные конъюнкции, вычисленные подсхемой S_2. Так как каждая функция из $H_{n-k,m}(i)$ является дизъюнкцией не более чем m элементарных конъюнкций степени $n-k$, то очевидно, что сложность схемы S_3 удовлетворяют неравенству

$$L(\mathsf{S}_3) \le m2^{n-k+m+1}/m = 2^{n-k+m+1}.$$

4. Подсхема S_4 вычисляет все функции $f_{\sigma_1,\dots,\sigma_k}(x_{k+1},\dots,x_n)$. Так как для вычисления одной функции требуется не более $\lceil 2^{n-k}/m\rceil$ дизъюнкторов, то

$$L(\mathsf{S}_4) \le 2^k\lceil 2^{n-k}/m\rceil \le \frac{2^n}{m} + 2^k.$$

5. Подсхема S_5 умножает функции, вычисленные подсхемой S_3, на элементарные конъюнкции, вычисленные подсхемой S_1, и вычисляет дизъюнкцию этих произведений. Легко видеть, что

$$L(\mathsf{S}_5) = 2\cdot 2^k - 1.$$

Таким образом,

$$L(\mathsf{S}) \le 4\cdot 2^k + 2^{n-k} + \mathcal{O}\left(n\left(2^{k/2} + 2^{n-k}\right)\right) + \frac{2^n}{m} + 2^{n-k+m+1}. \tag{10.12}$$

Положим $k = \lfloor n - 2\log_2 n\rfloor$ и $m = \lfloor n - 4\log_2 n\rfloor$. Легко убедиться, что при выбранных значениях параметров k и m и $n \to \infty$ все слагаемые в (10.12) кроме предпоследнего есть $\mathcal{O}\left(\frac{2^n}{n^2}\right)$. Следовательно, $L(\mathsf{S}) \le \frac{2^n}{m} + \mathcal{O}\left(\frac{2^n}{n^2}\right)$, и после несложных преобразований имеем

$$L(\mathsf{S}) \le \frac{2^n}{n}\left(1 + \mathcal{O}\left(\frac{\log_2 n}{n}\right)\right).$$

Теорема доказана.

Теорема 10.2 легко обобщается на случай вычисления вектор-функций. Соответствующий результат, также как и теорема 10.2 принадлежащий Лупанову, приведем без доказательства.

Теорема 10.3. *Пусть $n \to \infty$ и $\log_2 m = o(2^n)$. Тогда для каждой n-местной булевой вектор-функции f_m с m компонентами*

$$L(f_m) \le \frac{m2^n}{n + \log_2 m}(1 + o(1)).$$

10.3 Нижние мощностные оценки

Развивая мощностной метод Шеннона–Риордана, Лупанов установил следующий общий результат о нижних оценках сложности.

Теорема 10.4. *Пусть $F_{n,m}$ — такое множество n-местных булевых вектор-функций с m компонентами, что $n + m = o\big(\frac{\log_2|F_{n,m}|}{\log_2\log_2|F_{n,m}|}\big)$ при $n \to \infty$. Тогда для любой постоянной $\varepsilon > 0$ доля функций f из $F_{n,m}$, для которых*

$$L(f) \ge (1-\varepsilon)\frac{\log_2|F_{n,m}|}{\log_2\log_2|F_{n,m}|},$$

стремится к единице при $n \to \infty$.

Доказательство. Пусть $N(n,m,L)$ — число различных минимальных схем сложности L с n входами и m выходами.

Лемма 10.5. *Существует такая постоянная c, что*

$$N(n,m,L) \le \big(c(L+n)\big)^{L+n+m}.$$

Доказательство. Пусть схема S_σ состоит из L элементов, причем на множестве ее вершин определена нумерация σ, в которой вершины перенумерованы числами от 1 до $L+n$ так, что j-й вход имеет номер j для каждого j из $\{1,\dots,n\}$. Такую схему S_σ можно описать при помощи пары $(\mathsf{T}_\sigma, \boldsymbol{t}_\sigma)$, где T_σ — таблица из L строк и трех столбцов, а $\boldsymbol{t}_\sigma$ — набор длины m. В таблице i-я строка соответствует $(n+i)$-й вершине схемы, и в ее трех ячейках находятся: символ базисной функции, реализуемой в этой вершине; номер первого предка $(n+i)$-й вершины; номер второго предка этой вершины. Набор $\boldsymbol{t}_\sigma$ состоит из номеров вершин, которые являются выходами схемы. Очевидно, что число различных пар $(\mathsf{T}_\sigma, \boldsymbol{t}_\sigma)$, соответствующих схемам из L элементов, не превосходит $\big(16(L+n)^2\big)^L (L+n)^m$.

Пусть теперь S — минимальная схема, S_{σ_1} и S_{σ_2} — два экземпляра схемы S с различными нумерациями вершин, T_{σ_1} и T_{σ_2} — соответствующие этим схемам таблицы, π — подстановка на множестве $\{n+1,\dots,L+n\}$, преобразующая нумерацию σ_1 в σ_2. Нетрудно видеть, что если в таблице T_{σ_1} каждый элемент t из второго и третьего столбцов заменить на $\pi(t)$ и затем переставить строки так, чтобы i-я строка превратилась в $(\pi(i+n)-n)$-ю, то получится таблица T_{σ_2}. Поэтому для каждого j в j-й вершине схемы S_{σ_1} и в $\pi(j)$-й вершине схемы S_{σ_2} вычисляется одна и та же функция.

Покажем теперь, что таблицы T_{σ_1} и T_{σ_2} различны. Допустим, что это не так, и пусть k — такой индекс, что $\pi(k) \ne k$. В этом случае из равенства таблиц следует, что для каждого i, в том числе и для $i=k$, в i-х вершинах схем S_{σ_1} и S_{σ_2} вычисляются одинаковые функции. Следовательно, в k-й и $\pi(k)$-й вершинах схемы S_{σ_2} вычисляются одинаковые функции, что, очевидно, противоречит минимальности схемы S.

Таким образом для числа $N'(n,m,L)$ различных минимальных схем с n входами и m выходами, и состоящих ровно из L элементов, справедливо неравенство

$$N'(n,m,L) \le \big(16(L+n)^2\big)^L (L+n)^m \Big/ L!.$$

Так как $(1+\frac{1}{x})^x \le 4$ при $x>0$ и $k! \ge (\frac{1}{4}k)^k$, то

$$\begin{aligned} N'(n,m,L) &\le \big(16(L+n)\big)^{L+m} 4^n \Big(1+\frac{n}{L}\Big)^L \le \\ &\le \big(16(L+n)\big)^{L+m} 4^n 4^n \le \big(16(L+n)\big)^{L+m+n}. \end{aligned}$$

Теперь, учитывая равенство $N(n,m,L) = \sum_{k=1}^{L} N'(n,m,k)$, легко видеть, что

$$N(n,m,L) \le \sum_{k=1}^{L} \big(16(k+n)\big)^{k+m+n} \le \sum_{k=0}^{L+n+m} \big(16(L+n)\big)^k \le$$

$$\le \frac{\big(16(L+n)\big)^{L+n+m+1}}{16(L+n)-1} \le \big(32(L+n)\big)^{L+n+m}.$$

Лемма доказана.

Допустим, что R различных вектор-функций могут быть вычислены схемами сложности не более L. Так как для каждой функции найдется по крайней мере одна минимальная схема, и все эти схемы различны, то

$$R \le \big(c(L+n)\big)^{L+n+m},$$

или после логарифмирования

$$\log_2 R \le (L+n+m)\log_2(c(L+n)). \tag{10.13}$$

Положим

$$L = \frac{(1-2\varepsilon)\log_2 |F_{n,m}|}{\log_2\log_2 |F_{n,m}|}, \tag{10.14}$$

где ε — положительная постоянная. Тогда, подставляя выбранное значение L в (10.13) и учитывая, что $n+m = o\big(\frac{\log_2 |F_{n,m}|}{\log_2\log_2 |F_{n,m}|}\big)$, для любой постоянной $\varepsilon > 0$ при $n \to \infty$ имеем

$$\begin{aligned} \log_2 R &\le \Big(\frac{(1-2\varepsilon)\log_2 |F_{n,m}|}{\log_2\log_2 |F_{n,m}|} + n + m\Big) \log_2\Big(c\Big(\frac{(1-2\varepsilon)\log_2 |F_{n,m}|}{\log_2\log_2 |F_{n,m}|} + n\Big)\Big) \le \\ &\le \frac{(1-\varepsilon)\log_2 |F_{n,m}|}{\log_2\log_2 |F_{n,m}|} \cdot \log_2\big((1-\varepsilon)\log_2 |F_{n,m}|\big) < (1-\varepsilon)\log_2 |F_{n,m}|. \end{aligned}$$

Из последнего неравенства следует, что при помощи схем, сложность которых не превосходит (10.14), можно вычислить не более $|F_{n,m}|^{1-\varepsilon} = o(|F_{n,m}|)$ различных вектор-функций. Теорема доказана.

Применяя теорему 10.4 к множеству всех n-местных булевых функций, получаем следующий результат.

Теорема 10.5. *Для любой постоянной $\varepsilon > 0$ доля n-местных булевых функций f, для которых*

$$L(f) \ge (1-\varepsilon)\frac{2^n}{n},$$

стремится к единице при $n \to \infty$.

Сравнивая утверждения теорем 10.2 и 10.5, нетрудно видеть, что при $n \to \infty$ сложность почти всех n-местных булевых функций асимптотически равна $\frac{2^n}{n}$. Таким образом, метод Лупанова позволяет для почти всех булевых функций строить асимптотически минимальные схемы.

Применяя теорему 10.4 к множеству всех n-местных булевых вектор-функций с m компонентами, получаем следующий результат.

Теорема 10.6. *Для любой постоянной $\varepsilon > 0$ доля n-местных булевых вектор-функций f_m, для которых*

$$L(f_m) \geq (1-\varepsilon)\frac{m2^n}{n+\log_2 m},$$

стремится к единице при $n \to \infty$.

Сравнивая утверждения теорем 10.2 и 10.5, видим, что при $n \to \infty$ и $\log_2 m = o(2^n)$ сложность почти всех n-местных булевых вектор-функций асимптотически равна $\frac{m2^n}{n+\log_2 m}$.

10.4 Частичные функции

Частичные булевы функции образуют важный класс функций, естественным образом появляющихся в различных задачах. Следующая теорема легко следует из теоремы Л. А. Шоломова, доказанной в 1967 году на основе ранних работ Э. И. Нечипорука.

Теорема 10.7. *Пусть $n \to \infty$, $D \subseteq \{0,1\}^n$. Тогда для каждой частичной булевой функции f, определенной на области D,*

$$L(f) \leq \frac{|D|}{\log_2 |D|}(1+o(1)) + \mathcal{O}(n). \tag{10.15}$$

Доказательство теоремы 10.7 проведем в два этапа. На первом этапе установим справедливость неравенства (10.15) для областей большой мощности. На втором этапе, используя линейные булевы операторы для вложения областей n-мерного пространства в пространство меньшей размерности, сведем задачу вычисления произвольной частичной функции к вычислению частичной функции, определенной на области большой мощности, т. е. к задаче, решенной на первом этапе.

1. Набор $\boldsymbol{\alpha} \in \{0,1\}^m$ назовем доопределением набора $\boldsymbol{\beta} \in \{0,1,*\}^m$, если $\alpha_i = \beta_i$ для всех тех i, для которых $\beta_i \in \{0,1\}$. Множество $B \subseteq \{0,1\}^m$ назовем доопределением множества $A \subseteq \{0,1,*\}^m$, если для каждого элемента $\boldsymbol{\alpha}$ из A в B найдется элемент $\boldsymbol{\beta}$, являющийся доопределением $\boldsymbol{\alpha}$. Справедливо следующее утверждение о мощности доопределения.

Лемма 10.6. *Пусть $A = \{\boldsymbol{\alpha}\}$ — множество наборов из $\{0,1,*\}^m$, каждый из которых содержит ровно k булевых компонент. Тогда существует доопределение множества A, состоящее не более чем из $m2^{k+1}$ наборов.*

Доказательство. Допустим, что любое N-элементное подмножество множества $\{0,1\}^m$ не является доопределением множества A. Тогда для каждого такого подмножества можно указать хотя бы один набор из A, для которого в этом подмножестве нет доопределения. Поэтому число пар $(\boldsymbol{\alpha}, B)$, где $\boldsymbol{\alpha} \in A$, а B — N-элементное подмножество множества $\{0,1\}^m$, таких, что в B нет допределения $\boldsymbol{\alpha}$, не меньше, чем $\binom{2^m}{N}$. Так как A состоит

из $\binom{m}{k}2^k$ элементов, то в A найдется такой набор $\boldsymbol{\alpha}$, что по крайней мере $\binom{2^m}{N}\Big/\binom{m}{k}2^k$ N-элементных подмножеств множества $\{0,1\}^m$ не содержат доопределение $\boldsymbol{\alpha}$. С другой стороны, легко видеть, что для любого набора из A ровно $\binom{2^m-2^{m-k}}{N}$ N-элементных подмножеств множества $\{0,1\}^m$ не содержат его доопределение. Поэтому должно выполняться неравенство

$$\binom{2^m}{N}\Big/\binom{m}{k}2^k \leq \binom{2^m-2^{m-k}}{N}. \tag{10.16}$$

Преобразуя (10.16), видим, что

$$\binom{m}{k}2^k \geq \frac{2^m(2^m-1)\cdots(2^m-N+1)}{(2^m-2^{m-k})\cdots(2^m-2^{m-k}-N+1)} > \Big(\frac{2^m}{2^m-2^{m-k}}\Big)^N =$$
$$= \Big(\frac{1}{1-2^{-k}}\Big)^N \geq \Big(1+2^{-k}\Big)^N = \Big(1+2^{-k}\Big)^{2^k\cdot N2^{-k}} \geq 2^{N2^{-k}}.$$

Так как $\binom{m}{k}2^k < 2^{2m}$, то, логарифмируя крайние части последнего выключенного неравенства, заключаем, что $N < 2m2^k$. Таким образом, из предположения, что любое N-элементное подмножество множества $\{0,1\}^m$ не является доопределением множества A, следует неравенство $N < 2m2^k$. Поэтому при N больших или равных $m2^{k+1}$ среди N-элементных подмножеств множества $\{0,1\}^m$ найдется хотя бы одно доопределение множества A. Лемма доказана.

Лемма 10.7. *Пусть $D \subseteq \{0,1\}^n$, D состоит из N наборов. Если $\log_2 N \sim n$ при $n \to \infty$, то для любой частичной булевой функции $f : D \to \{0,1\}$*

$$L(f) \leq \frac{N}{n}(1+o(1)).$$

Доказательство. Введем параметры R и k, значения которых определим позднее. Значения частичной n-местной булевой функции f запишем в таблице T_f из 2^k столбцов и 2^{n-k} строк, поставив в соответствие i-му столбцу таблицы двоичный набор $(\sigma_1,\ldots,\sigma_k)$, являющийся двоичным представлением числа $i-1$, а j-й строке — набор $(\sigma_{k+1},\ldots,\sigma_n)$, являющийся двоичным представлением числа $j-1$. В таблице на пересечении i-го столбца и j-й строки поставим значение $f(\sigma_1,\ldots,\sigma_k,\sigma_{k+1},\ldots,\sigma_n)$. Каждую строку таблицы представим в виде следующих друг за другом элементарных наборов $\boldsymbol{\alpha}$, каждый из которых, кроме быть может последнего, содержит R булевых компонент. Множество таких наборов разобьем на классы, поместив в класс P_{ij} наборы, начинающиеся в i-й и заканчивающиеся в j-й позициях. Нетрудно видеть, что число различных классов не превосходит величины 2^{2k-1}.

Из леммы 10.6 следует, что для множества элементарных наборов класса P_{ij} существует множество их доопределений, которое состоит не более

чем из $2^k 2^{R+1}$ наборов длины 2^k, в каждом из которых первые $i-1$ и последние $2^k - j$ компонент равны нулю. Следовательно, для множества всех элементарных наборов существует множество их доопределений H, которое состоит не более чем из $2^{3k}2^R$ наборов длины 2^k.

Преобразуем таблицу T_f, заменив в ней каждый элементарный набор каким-либо его доопределением из H. Нетрудно видеть, что преобразованная таблица будет таблицей значений некоторой n-местной булевой функции h, являющейся доопределением функции f.

Номером двоичного набора $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_m)$ назовем величину $|\boldsymbol{\sigma}| = 1 + \sum_{i=1}^{m} \sigma_i 2^{i-1}$. Пусть $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_{2^k}) \in H$. Введем множество G, состоящее из функций

$$g_{\boldsymbol{\gamma}}(x_1, \ldots, x_k) = \bigvee_{\boldsymbol{\sigma}=(\sigma_1,\ldots,\sigma_k)} x_1^{\sigma_1} \cdot \ldots \cdot x_k^{\sigma_k} \cdot \boldsymbol{\gamma}_{|\sigma|},$$

вектора значений которых, как нетрудно видеть, являются элементами множества H. Очевидно, что $|G| \le 2^{3k}2^R$, и функция h может быть выражена через функции системы G следующим образом:

$$h(x_1, \ldots, x_n) = \bigvee_{\boldsymbol{\sigma}=(\sigma_{k+1},\ldots,\sigma_n)} \Big(\bigvee_{g \in G} g(x_1, \ldots, x_k) \Big) x_{k+1}^{\sigma_{k+1}} \cdot \ldots \cdot x_n^{\sigma_n}. \qquad (10.17)$$

Оценим число функций g в (10.17). Прежде всего заметим, что число функций, соответствующих наборам $\boldsymbol{\alpha}$, содержащим менее R булевых компонент, не превосходит числа строк таблицы, т. е. не больше чем 2^{n-k}. Число остальных функций очевидно не превосходит N/R. Поэтому общее число элементарных наборов в T_f, а, следовательно, и функций g в (10.17) не превосходит

$$N \Big/ R + 2^{n-k}. \qquad (10.18)$$

Опишем схему S, вычисляющую функцию f и удовлетворяющую требованиям леммы. Эта схема состоит из трех подсхем S_1–S_3, и ее конструкция основана на формуле (10.17). Подсхема S_1 вычисляет все элементарные конъюнкции первых k переменных и, используя эти конъюнкции, все функции из G. Учитывая, что $|G| \le 2^{3k}2^R$ и каждая функция из G является дизъюнкцией не более чем 2^k элементарных конъюнкций, имеем

$$L(\mathsf{S}_1) \le 2^{5k}2^R. \qquad (10.19)$$

Подсхема S_2 вычисляет все элементарные конъюнкции последних $n-k$ переменных. Очевидно, что

$$L(\mathsf{S}_2) \le 2^{n-k}(1 + o(1)). \qquad (10.20)$$

Подсхема S_3 подключена к выходам подсхем S_1 и S_2 и вычисляет полностью определенную функцию $h(x_1, \ldots, x_n)$ в соответствии с равенством (10.17). Из (10.18) следует, что

$$L(\mathsf{S}_3) \le 2^{n-k+1} + N \Big/ R. \qquad (10.21)$$

Суммируя неравенства (10.19)–(10.21), видим, что

$$L(\mathsf{S}) \le N \Big/ R + 2^{5k}2^R + 2^{n-k}(1 + o(1)). \qquad (10.22)$$

Положим

$$k = \lceil n - \log_2 N + 2\log_2 n \rceil, \qquad R = \lfloor \log_2 N - 5k - 2\log_2 n \rfloor. \qquad (10.23)$$

Тогда, подставляя равенства (10.23) в неравенство (10.22) и учитывая условие $\log_2 N \sim n$, имеем

$$L(\mathsf{S}) \le \frac{N}{n}(1 + o(1)).$$

Лемма доказана.

2. Докажем несколько утверждений об инъективных и "почти" инъективных линейных булевых операторах. Затем, используя эти утверждения, установим окончательный результат о сложности частичных функций.

Лемма 10.8. *Пусть $A, B \subseteq \{0,1\}^n$, $A \cap B = \varnothing$, m — целое. Тогда существует линейный оператор $\mathcal{L} : \{0,1\}^n \to \{0,1\}^m$ такой, что*

$$|\{(\boldsymbol{x}, \boldsymbol{y}) \,|\, \boldsymbol{x} \in A, \boldsymbol{y} \in B, \mathcal{L}(\boldsymbol{x}) = \mathcal{L}(\boldsymbol{y})\}| \le 2^{-m}|A||B|.$$

Доказательство. Обозначим через $F(n,m)$ множество всех линейных операторов $\mathcal{L} : \{0,1\}^n \to \{0,1\}^m$. Очевидно, что $|F(n,m)| = 2^{nm}$. Так как для любых двух различных наборов $\boldsymbol{x}$ и $\boldsymbol{y}$ из $\{0,1\}^n$ имеется ровно 2^{n-1} n-местных линейных функций f с нулевым свободным членом, значения которых на этих наборах совпадают, т. е. $f(\boldsymbol{x}) = f(\boldsymbol{y})$, то поэтому в $F(n,m)$ имеется 2^{nm-m} различных операторов $\mathcal{L}$ таких, что $\mathcal{L}(\boldsymbol{x}) = \mathcal{L}(\boldsymbol{y})$. Следовательно, величина

$$2^{-nm} \sum_{\boldsymbol{x} \in A, \boldsymbol{y} \in B} 2^{nm-m} = 2^{-m}|A||B|$$

является средним значением для числа таких пар $(\boldsymbol{x}, \boldsymbol{y})$, на которых значения оператора из $F(n,m)$ одинаковы. Поэтому в $F(n,m)$ найдется оператор $\mathcal{L}$, для которого равенство $\mathcal{L}(\boldsymbol{x}) = \mathcal{L}(\boldsymbol{y})$ выполняется не более чем на $2^{-m}|A||B|$ парах $(\boldsymbol{x}, \boldsymbol{y})$, $\boldsymbol{x} \in A$, $\boldsymbol{y} \in B$. Лемма доказана.

Лемма 10.9. *Пусть $A, B \subseteq \{0,1\}^n$, $A \cap B = \varnothing$, $m = \lceil \log_2 |B| + k \rceil$, где $k \ge 0$. Тогда существует линейный оператор $\mathcal{L} : \{0,1\}^n \to \{0,1\}^m$ такой, что*

$$|\{(\boldsymbol{x}, \boldsymbol{y}) \,|\, \boldsymbol{x} \in A, \boldsymbol{y} \in B, \mathcal{L}(\boldsymbol{x}) = \mathcal{L}(\boldsymbol{y})\}| \le \frac{1}{2^k}|A|.$$

Доказательство. Из леммы 10.8 следует, что найдется такой линейный (n,m)-оператор $\mathcal{L}$, для которого равенство $\mathcal{L}(\boldsymbol{x}) = \mathcal{L}(\boldsymbol{y})$ выполняется не более чем на $2^{-m}|A||B|$ парах $(\boldsymbol{x}, \boldsymbol{y})$, где $\boldsymbol{x} \in A$, $\boldsymbol{y} \in B$. Так как $2^m \ge 2^k|B|$, то $2^{-m}|A||B| \le \frac{1}{2^k}|A|$. Лемма доказана.

Лемма 10.10. *Пусть $D \subseteq \{0,1\}^n$ состоит из N наборов, $\log_2 N \geq \frac{1}{3}n$, функция f определена на D. Тогда*

$$L(f(x_1,\ldots,x_n)) \leq \frac{N}{\log_2 N}(1+o(1)).$$

Доказательство. Положим $k = 3\log_2 n$, $D_0 = \{\boldsymbol{x} \in D \mid f(\boldsymbol{x}) = 0\}$, $D_1 = \{\boldsymbol{x} \in D \mid f(\boldsymbol{x}) = 1\}$. Без ограничения общности будем полагать, что $|D_0| \geq |D_1|$. К областям D_0 и D_1 применим лемму 10.9, полагая, что $A = D_1$ и $B = D_0$. В результате найдется такой линейный оператор $\mathcal{L} : \{0,1\}^n \to \{0,1\}^m$, что $m = \lceil \log_2 |D_0| + 3\log_2 n \rceil$ и множество

$$D' = \{(\boldsymbol{x},\boldsymbol{y}) \mid \boldsymbol{x} \in D_0, \boldsymbol{y} \in D_1, \mathcal{L}(\boldsymbol{x}) = \mathcal{L}(\boldsymbol{y})\}$$

состоит не более чем из N/n^3 наборов. Далее введем определенную на области $\mathcal{L}(D)$ частичную m-местную булеву функцию

$$g(\boldsymbol{y}) = \begin{cases} 0, & \text{если } \exists\, \boldsymbol{x} \in D_0 \text{ такой, что } \boldsymbol{y} = \mathcal{L}(\boldsymbol{x}), \\ 1, & \text{в противном случае,} \end{cases}$$

и определенную на области D частичную функцию $h(\boldsymbol{x}) = f(\boldsymbol{x}) \oplus g(\mathcal{L}(\boldsymbol{x}))$. Нетрудно видеть, что $h(\boldsymbol{x})$ равна единице не более чем на N/n^3 наборах из D. Так как $f(\boldsymbol{x}) = h(\boldsymbol{x}) \oplus g(\mathcal{L}(\boldsymbol{x}))$, то

$$L(f) \leq L(g) + L(h) + L(\mathcal{L}) + 1.$$

Нетрудно видеть, что $\log_2 N \sim m$, и поэтому для оценки сложности функции g можно воспользоваться леммой 10.7. Из этой леммы следует, что

$$L(g) \leq \frac{N}{\log_2 N}(1+o(1)).$$

Для вычисления функции h воспользуемся ее совершенной дизъюнктивной формой, полагая, что вне области D эта функция равна нулю. Так как h равна единице не более чем на N/n^3 наборах, то

$$L(h) \leq \frac{Nn}{n^3} = o\Big(\frac{N}{n}\Big).$$

Очевидно, что сложность оператора $\mathcal{L}$ не превосходит n^2. Поэтому из условий леммы следует, что $L(\mathcal{L}) = o(g)$. Лемма доказана.

3. Доказательство теоремы 10.7. Если $\log_2 N \geq \frac{1}{3}N$, то утверждение теоремы следует из леммы 10.10. Поэтому далее полагаем, что $\log_2 N \leq \frac{1}{3}N$.

Введем области $D_0 = \{\boldsymbol{x} \in D \mid f(\boldsymbol{x}) = 0\}$, $D_1 = \{\boldsymbol{x} \in D \mid f(\boldsymbol{x}) = 1\}$ и положим $k = \log_2 |D_1| + 1$. К областям D_0 и D_1 применим лемму 10.9, полагая, что $A = D_1$ и $B = D_0$. В результате найдется такой линейный оператор $\mathcal{L} : \{0,1\}^n \to \{0,1\}^m$, что $m = \lceil \log_2 |D_0| + \log_2 |D_1| + 1 \rceil$ и при

этом нет таких элементов $\boldsymbol{x} \in D_0$ и $\boldsymbol{y} \in D_1$, что $\mathcal{L}(\boldsymbol{x}) = \mathcal{L}(\boldsymbol{y})$. Далее введем определенную на области $\mathcal{L}(D)$ частичную m-местную булеву функцию

$$g(\boldsymbol{y}) = \begin{cases} 0, & \text{если } \exists\, \boldsymbol{x} \in D_0 \text{ такой, что } \boldsymbol{y} = \mathcal{L}(\boldsymbol{x}), \\ 1, & \text{если } \exists\, \boldsymbol{x} \in D_1 \text{ такой, что } \boldsymbol{y} = \mathcal{L}(\boldsymbol{x}). \end{cases}$$

Нетрудно видеть, что $f(\boldsymbol{x}) = g(\mathcal{L}(\boldsymbol{y}))$. Так как $m \leq \lceil 2\log_2 N \rceil$, то можно воспользоваться леммой 10.10. В силу этой леммы

$$L(g) \leq \frac{N}{\log_2 N}(1+o(1)). \tag{10.24}$$

Наконец заметим, что в силу леммы 10.4

$$L(\mathcal{L}) = \mathcal{O}\left(\frac{n(\log_2 N + \log_2 n)}{\log_2 n}\right). \tag{10.25}$$

Нетрудно видеть, что при $n \to \infty$ для любого N сумма правых частей неравенств (10.24) и (10.25) не превосходит $\frac{|D|}{\log_2 |D|}(1+o(1)) + \mathcal{O}(n)$. Теорема доказана.

Применяя теорему 10.4 к множеству всех частичных n-местных булевых функций, определенных на одной и той же области, получаем следующую теорему.

Теорема 10.8. *Пусть $n \to \infty$, $D \subseteq \{0,1\}^n$. Для любой постоянной $\varepsilon > 0$ доля n-местных булевых функций $f : D \to \{0,1\}$, для которых*

$$L(f) \geq (1-\varepsilon)\frac{|D|}{\log_2 |D|},$$

стремится к единице при $n \to \infty$.

Сравнивая утверждения теорем 10.7 и 10.8, нетрудно видеть, что при $n \to \infty$ для любой области $D \subseteq \{0,1\}^n$ такой, что $|D| \gg n\log_2 n$ сложность почти всех частичных n-местных булевых функций, определенных на области D, асимптотически равна $\frac{|D|}{\log_2 |D|}$.

10.5 Монотонные функции

Из теоремы 10.5 и нижней оценки числа монотонных функций (см. стр. 96) легко следует, что доля n-местных монотонных булевых функций f, для которых

$$L(f) \geq (1-\varepsilon)\frac{2^n}{n\sqrt{\pi n/2}},$$

стремится к единице при $n \to \infty$. Справедлива и аналогичная верхняя оценка. Точнее, при $n \to \infty$ сложность каждой n-местной монотонной функции удовлетворяет неравенству

$$L(f) \leq \frac{2^n}{n\sqrt{\pi n/2}}(1+o(1)), \tag{10.26}$$

которое было установлено А. Б. Угольниковым в 1976 г. Доказательство неравенства (10.26) длинное и технически сложное. Поэтому далее вместо (10.26) докажем более слабое неравенство.

Теорема 10.9. *Для сложности любой монотонной n-местной булевой функции f справедливо неравенство*

$$L(f) \le \frac{2^{n+1}}{n\sqrt{\pi n/2}}(1+o(1)).$$

Напомним, что пара вершин n-мерного булева куба $\{0,1\}^n$ называется ребром, если эти вершины различаются ровно в одном разряде. Пусть $(\boldsymbol{x},\boldsymbol{y})$ и $(\boldsymbol{u},\boldsymbol{v})$ — такие ребра n-мерного булева куба, что $\boldsymbol{y} \preccurlyeq \boldsymbol{x}$ и $\boldsymbol{v} \preccurlyeq \boldsymbol{u}$. Эти ребра назовем несравнимыми, если $\boldsymbol{y}$ несравнима с $\boldsymbol{u}$ или $\boldsymbol{x}$ несравнима с $\boldsymbol{v}$. Последовательность вершин $\boldsymbol{\alpha} = (\boldsymbol{x}_0,\ldots,\boldsymbol{x}_n)$ такую, что $\boldsymbol{x}_0 \prec \cdots \prec \boldsymbol{x}_n$, будем называть максимальной цепью n-мерного булева куба. Будем говорить, что цепь $\boldsymbol{\alpha}$ проходит через ребро $(\boldsymbol{u},\boldsymbol{v})$, если вершины $\boldsymbol{u}$ и $\boldsymbol{v}$ принадлежат $\boldsymbol{\alpha}$.

Следующая лемма является реберным аналогом известной оценки числа элементов антицепи в булевом кубе. Далее полагаем, что n — четное.

Лемма 10.11. *Любое множество попарно несравнимых ребер n-мерного булева куба состоит не более чем из $\frac{n}{2}\binom{n}{n/2}$ элементов.*

Доказательство. Пусть $\mathcal{N}$ — произвольное множество попарно несравнимых ребер n-мерного булева куба, а $\mathcal{N}_k$ — его подмножество, состоящее из всех тех ребер множества $\mathcal{N}$, которые в булевом кубе соединяют вершины k-го и $(k+1)$-го слоев. Если $(\alpha,\beta) \in \mathcal{N}_k$, то через это ребро проходит ровно $k!(n-k-1)!$ максимальных цепей. Так как каждая максимальная цепь проходит ровно через одно ребро множества $\mathcal{N}$, то через все ребра множества $\mathcal{N}$ проходит ровно $\sum_{k=0}^{n-1} k!(n-k-1)!|\mathcal{N}_k|$ максимальных цепей. Так как в n-мерном булевом кубе существует только $n!$ различных максимальных цепей, то

$$n! \ge \sum_{k=0}^{n-1} k!(n-k-1)!|\mathcal{N}_k| = (n-1)!\sum_{k=0}^{n-1}|\mathcal{N}_k|\Big/\binom{n-1}{k} \ge$$
$$\ge (n-1)!\Big(\sum_{k=0}^{n-1}|\mathcal{N}_k|\Big)\Big/\binom{n-1}{n/2} = (n-1)!|\mathcal{N}|\Big/\binom{n-1}{n/2}.$$

Преобразуя предыдущее неравенство, получаем требуемую оценку:

$$|\mathcal{N}| \le n\binom{n-1}{n/2} \le \frac{n\cdot(n-1)!}{(n/2)!(n/2-1)!} =$$
$$= \frac{n!}{(n/2-1)!(n/2)!} = \frac{n/2\cdot n!}{n/2\cdot(n/2-1)!(n/2)!} = \frac{n/2\cdot n!}{(n/2)!(n/2)!} = \frac{n}{2}\binom{n}{n/2}.$$

Лемма доказана.

Пусть f — монотонная булева функция. Ребро $(\boldsymbol{x},\boldsymbol{y})$ назовем непостоянным, если $f(\boldsymbol{x}) \ne f(\boldsymbol{y})$. Так как любые два непостоянных ребра несравнимы, то из леммы 10.11 вытекает следующее утверждение.

Лемма 10.12. *У любой n-местной монотонной булевой функции число непостоянных ребер не превосходит $\frac{n}{2}\binom{n}{n/2}$.*

Будем говорить, что ребро $(\boldsymbol{x},\boldsymbol{y})$ проходит в i-м направлении, если наборы $\boldsymbol{x}$ и $\boldsymbol{y}$ различаются в i-м разряде.

Лемма 10.13. *Для любой n-местной монотонной булевой функции найдутся такие направления i и j, что число непостоянных ребер, проходящих в этих направлениях, не превосходит $\binom{n}{n/2}$.*

Доказательство. Допустим, что в любых двух направлениях i и j проходит в совокупности больше чем $\binom{n}{n/2}$ непостоянных ребер. Тогда сумма S количества непостоянных ребер, взятых по всем парам направлений, больше $\binom{n}{n/2}\binom{n}{2}$. При этом непостоянные ребра каждого направления будут посчитаны ровно $n-1$ раз. Поэтому в силу леммы 10.12 сумма S не превосходит величины $(n-1)\frac{n}{2}\binom{n}{n/2}$. Таким образом,

$$\binom{n}{n/2}\binom{n}{2} < S \le (n-1)\frac{n}{2}\cdot\binom{n}{n/2}.$$

Перенося множитель $\binom{n}{2}$ из левой части полученного неравенства в правую, получим противоречие. Лемма доказана.

Символом $f_{ij}^{\alpha\beta}(\boldsymbol{x})$ обозначим $(n-2)$-местную функцию, получающуюся из n-местной булевой функции f подстановкой констант α и β вместо ее i-го и j-го аргументов, а символом $\boldsymbol{x}_{ij}^{\alpha\beta}$ — булев набор длины n, у которого i-й и j-й разряды равны α и β и который после удаления этих разрядов превращается булев набор $\boldsymbol{x}$ длины $n-2$.

Лемма 10.14. *Для любой n-местной монотонной булевой функции f найдутся такие i и j, что $f_{ij}^{11}(\boldsymbol{x}) \ne f_{ij}^{00}(\boldsymbol{x})$ не более чем для $\frac{1}{2}\binom{n}{n/2}$ различных наборов $\boldsymbol{x}$ длины $n-2$.*

Доказательство. Рассмотрим такие i и j, для которых у функции f в i-м и j-м направлениях в совокупности проходит не более $\binom{n}{n/2}$ непостоянных ребер. Если $f_{ij}^{11}(\boldsymbol{x}) \ne f_{ij}^{00}(\boldsymbol{x})$, то $f_{ij}^{11}(\boldsymbol{x}) = 1$ и $f_{ij}^{00}(\boldsymbol{x}) = 0$ и, следовательно, одно из ребер $(\boldsymbol{x}_{ij}^{00},\boldsymbol{x}_{ij}^{01})$ или $(\boldsymbol{x}_{ij}^{01},\boldsymbol{x}_{ij}^{11})$ будет непостоянным. Точно также непостоянным будет и одно из ребер $(\boldsymbol{x}_{ij}^{00},\boldsymbol{x}_{ij}^{10})$ или $(\boldsymbol{x}_{ij}^{10},\boldsymbol{x}_{ij}^{11})$. Поэтому каждому неравенству $f_{ij}^{11}(\boldsymbol{x}) \ne f_{ij}^{00}(\boldsymbol{x})$ соответствует два непостоянных ребра, проходящих в i-м или j-м направлениях. Так как число непостоянных ребер i-го и j-го направлений ровно в два раза больше числа наборов $\boldsymbol{x}$, на которых $f_{ij}^{11}(\boldsymbol{x}) \ne f_{ij}^{00}(\boldsymbol{x})$, то утверждение леммы следует из выбора i и j. Лемма доказана.

Символом $\boldsymbol{x}_{ij}$ обозначим булев набор длины $n-2$, получающийся из булева набора $\boldsymbol{x}$ длины n удалением i-го и j-го разрядов.

Лемма 10.15. *Для любой n-местной монотонной булевой функции f найдутся такие целые i и j, $(n-2)$-местные монотонные функции p и q и область $D \subseteq \{0,1\}^n$, что $|D| \le \binom{n}{n/2}$ и*

$$f(\boldsymbol{x}) = p(\boldsymbol{x}_{ij})x_i x_j \vee q(\boldsymbol{x}_{ij})f_D(\boldsymbol{x}). \tag{10.27}$$

Доказательство. Пусть i и j такие, как в лемме 10.14. В $\{0,1\}^n$ определим множество $D = \{\boldsymbol{x} \mid f_{ij}^{11}(\boldsymbol{x}_{ij}) \ne f_{ij}^{00}(\boldsymbol{x}_{ij})\}$. Из леммы 10.14 следует, что это множество состоит не более чем из $\binom{n}{n/2}$ элементов. Легко видеть, что $f(\boldsymbol{x}) = f_{ij}^{11}(\boldsymbol{x}_{ij})$, если $\boldsymbol{x} \notin D$ и $f(\boldsymbol{x}) = f_D(\boldsymbol{x})$, если $\boldsymbol{x} \in D$. Так как функции $f_{ij}^{11}(\boldsymbol{x}_{ij}) \oplus f_{ij}^{00}(\boldsymbol{x}_{ij}) \oplus 1$ является характеристической функцией множества D, то для f имеет место равенство

$$f(\boldsymbol{x}) = f_{ij}^{11}(\boldsymbol{x}_{ij})(f_{ij}^{11}(\boldsymbol{x}_{ij}) \oplus f_{ij}^{00}(\boldsymbol{x}_{ij}) \oplus 1) \vee (f_{ij}^{11}(\mathbf{x}_{ij}) \oplus f_{ij}^{00}(\boldsymbol{x}_{ij}))f_D(\boldsymbol{x}),$$

которое после элементарных преобразований превращается в равенство

$$f(\boldsymbol{x}) = f_{ij}^{11}(\boldsymbol{x}_{ij})f_{ij}^{00}(\boldsymbol{x}_{ij}) \vee (f_{ij}^{11}(\boldsymbol{x}_{ij}) \vee f_{ij}^{00}(\boldsymbol{x}_{ij}))f_D(\boldsymbol{x}).$$

Положив $p(\boldsymbol{x}_{ij}) = f_{ij}^{11}(\boldsymbol{x}_{ij})f_{ij}^{00}(\boldsymbol{x}_{ij})$ и $q(\boldsymbol{x}_{ij}) = f_{ij}^{11}(\boldsymbol{x}_{ij}) \vee f_{ij}^{00}(\boldsymbol{x}_{ij})$, получим равенство (10.27). Лемма доказана.

Из леммы 10.15 следует, что вычисление значения n-местной монотонной функции можно свести к вычислению значений двух $(n-2)$-местных монотонных функций и одной n-местной частичной функции, определенной на области, состоящей не более чем из $\binom{n}{n/2}$ наборов.

Допустим, что любая n-местная монотонная булева функция может быть вычислена схемой, сложность которой ограничена величиной $L_M(n)$. Тогда из теоремы 10.7 и леммы 10.15 следует, что для $L_M(n)$ справедливо рекуррентное неравенство

$$L_M(n) \le 2L_M(n-2) + \frac{2^n}{n\sqrt{\pi n/2}}(1+o(n)) + 2. \tag{10.28}$$

Положим $k = \lceil \log_2 n \rceil$. Для оценки величины $L_M(n)$ применим k раз соответствующее неравенство из (10.28). В результате имеем[1)]

$$\begin{aligned} L_M(n) &\lesssim 2L_M(n-2) + \frac{2^n}{n\sqrt{\pi n/2}} \lesssim \\ &\lesssim 4L_M(n-4) + \frac{2\cdot 2^{(n-2)}}{(n-2)\sqrt{\pi n/2(n-2)}} + \frac{2^{n+1}}{n\sqrt{\pi n/2}} \lesssim \dots \\ &\lesssim 2^k L_M(n-2k) + \frac{2^n}{n\sqrt{\pi n/2}} \sum_{i=1}^{k} \frac{1}{2^{i-1}} \lesssim 2^k L_M(n-2k) + \frac{2\cdot 2^n}{n\sqrt{\pi n/2}}. \end{aligned} \tag{10.29}$$

[1)] Неравенство $a(n) \lesssim b(n)$ означает, что $\lim_{n\to\infty} a(n)/b(n) \le 1$.

Очевидно, что $L_M(n-2k) \lesssim \frac{2^{n-2k}}{n-2k}$ в силу теоремы 10.2. Поэтому из (10.29) и цепочки неравенств

$$L_M(n) \lesssim \frac{2^k \cdot 2^{n-2k}}{n-2k} + \frac{2\cdot 2^n}{n\sqrt{2\pi n}} \lesssim \frac{2^n}{n^2} + \frac{2\cdot 2^n}{n\sqrt{\pi n/2}} \sim \frac{2\cdot 2^n}{n\sqrt{\pi n/2}}$$

следует справедливость теоремы для четных n. Так как равенство

$$f(x_1,\dots,x_n,x_{n+1}) = x_{n+1}f(x_1,\dots,x_n,1) \vee f(x_1,\dots,x_n,0)$$

имеет место для любой монотонной функции f, то, очевидно, что утверждение теоремы справедливо и для нечетного числа аргументов. Теорема доказана.

10.6 Задачи

10.1. Найти сложность и глубину системы всех одночленов переменных $x_1,\dots,x_n$. Показать, что для этой системы существует схема, являющаяся одновременно минимальной по сложности и глубине.

10.2. Пусть $B \subseteq P_2(2)$ и $[B] = P_2$. Показать, что при $n \to \infty$ для каждой функции f из $P_2(n)$ справедливо неравенство $L_B(f) \lesssim \frac{2^n}{n}$.

10.3. Пусть $\mathbb{V}$ — подпространство в $\{0,1\}^n$ размерности k, $F_{\mathbb{V}}$ — подмножество $P_2(n)$, состоящее из всех функций постоянных на смежных классах пространства $\{0,1\}^n$ по $\mathbb{V}$. Оценить $\max L(f)$, где максимум берется по всем функциям из $F_{\mathbb{V}}$ при условии, что $n-k \to \infty$.

10.4. Пусть V — подпространство в $\{0,1\}^n$ размерности k, F_V — подмножество $P_2(n)$, состоящее из всех функций, равных единице на наборах не принадлежащих V. Оценить $\max L(f)$, где максимум берется по всем функциям из F_V при условии, что $k \to \infty$.

10.5. Положим $|\alpha| = \sum_{i=1}^{n} \alpha_i 2^{i-1}$ для каждого α из $\{0,1\}^n$. Пусть A состоит из всех таких булевых функций f, что $f(x) = f(y)$, если $|x| = |y| \pmod{n^2}$. Оценить $\max_{f\in A} L(f)$.

10.6. Оценить $\max L(f)$, где максимум берется по всем булевым операторам $f : \{0,1\}^n \to \{0,1\}^n$ таким, что $\|f(x)\| = 1$ для всех $x \in \{0,1\}^n$

10.7. Показать, что для каждого линейного булева (n,n)-оператора f при $n \to \infty$: а) $L(f) \lesssim \frac{n^2}{\log_2 n}$; б) $L(f) \lesssim \frac{n^2}{2\log_2 n}$.

10.8. Показать, что для каждой квадратичной булевой функции f при $n \to \infty$: а) $L(f) \lesssim \frac{n^2}{2\log_2 n}$; б) $L(f) \lesssim \frac{n^2}{4\log_2 n}$.

10.9. Пусть $P_{n,m} = \{f \in P_2(n), \deg f \le m\}$. Показать, что при любой постоянной m и $n \to \infty$ для почти всех функций f из $P_{n,m}$ справедливо равенство $L(f) = \Theta(\frac{n^m}{\log_2 n})$.

10.10. Пусть $N_{n,m} = \{f \in P_2(n), ||f|| \le n^m\}$. Показать, что при любой постоянной m и $n \to \infty$ для почти всех функций f из $N_{n,m}$ справедливо равенство $L(f) = \Theta(\frac{n^{m+1}}{\log_2 n})$.

10.11. Пусть A — множество всех булевых функций $f(x_1,\dots,x_n,y_1,\dots,y_n)$ симметричных относительно переменных x_i и относительно переменных y_j. Оценить $\max_{f\in A} L(f)$.

10.12. Пусть A — множество всех булевых функций $f(x_1,\dots,x_n,y_1,\dots,y_n)$ симметричных относительно всех пар переменных x_i, y_i. Оценить $\max_{f\in A} L(f)$.

10.13. Пусть A — множество, состоящее из всех булевых функций, удовлетворяющих равенству $f(x_1,\dots,x_{2n-1}) = 0$, если $\sum_{i=1}^{2n-1} x_i \geq n$. Найти $\max_{f\in A} L(f)$.

10.14. Пусть A — множество, состоящее из всех булевых функций, удовлетворяющих равенству $f(x_1,\dots,x_n) = f(\overline{x}_1,\dots,\overline{x}_n)$. Найти $\max_{f\in A} L(f)$.

10.15. Пусть A — множество всех булевых функций $f(x_1,\dots,x_{2n})$ симметричных относительно первых n переменных. Оценить $\max_{f\in A} L(f)$.

10.16. Показать, что $L(f) = \mathcal{O}(n)$ для любой симметрической функции $f(x_1,\dots,x_n)$.

10.17. Доказать теорему 10.3.

10.18. Пусть $f \in P_2(n)$, $\|f\| \leq 2^{n-1}$ и $\log_2\log_2\|f\| \sim \log_2 n$. Показать, что $L(f) \lesssim \frac{\log_2\binom{2^n}{\|f\|}}{\log_2\log_2\binom{2^n}{\|f\|}}$.

10.19. Используя задачу 7.13, показать, что $L_{\{\vee,\&\}}(f) = \mathcal{O}\left(\binom{n}{\lfloor n/2\rfloor}\Big/ n\right)$ для любой монотонной булевой функции $f(x_1,\dots,x_n)$.

10.20. Показать, что при $n\to\infty$ для вычисления n-й степени действительного числа a достаточно выполнить $(1+o(1))\log_2 n$ умножений действительных чисел.

Лекция 11

Средняя сложность булевых функций

Важная особенность неветвящихся программ и схем из функциональных элементов состоит в том, что каждая программа и каждая схема на любом наборе аргументов выполняет одно и то же число элементарных операций, т. е. при помощи неветвящихся программ и схем изучается сложность функций в "худшем" случае. Однако не всегда число элементарных операций, выполненных в "худшем" случае, является реалистичной мерой сложности. Часто более важно знать среднее, по всем возможным аргументам, число выполненных операций, которое может значительно отличатся числа операций, выполненных в "худшем" случае. Существует большое число различных задач для которых не известны алгоритмы, решающие эти задачи достаточно быстро при произвольном выборе их аргументов. В то же время многие такие задачи могут быть быстро решены "в среднем", т. е. для каждой задачи найдется алгоритм, время работы которого на почти всех аргументах мало, и только для незначительной доли аргументов алгоритм работает долго.

Далее рассматривается средняя сложность булевых функций. Функции вычисляются при помощи неветвящихся программ с условной остановкой. Эти программы отличаются от рассматривавшихся выше неветвящихся программ наличием управляющей команды — команды условной остановки, и являются естественной моделью вычислений, в которых нет условного перехода и косвенной адресации, но есть возможность досрочного прекращения работы при выполнении определенного условия. Такие вычисления можно представить следующим образом. Вычисления выполняет процессор, снабженный памятью, состоящей из отдельных ячеек. Процессор способен вычислять некоторое количество двухместных элементарных функций, составляющих базис вычислений. Каждая ячейка памяти в любой момент времени доступна процессору как для чтения, так и для записи информации. Процессор работает под управление программы, являющейся последовательностью элементарных команд двух видов. Каждая команда первого вида вычисляется значение некоторой базисной функции, аргументами которой являются величины, содержащиеся в определенных ячейках памяти.

Вычисленный результат также помещается в одну из ячеек памяти. Команда второго вида может прекратить выполнение программы. Каждая такая команда имеет единственный аргумент — содержимое некоторой ячейки памяти. Если значение аргумента равно единице, то выполнение программы прекращается; если значение аргумента равно нулю, то выполняется следующая команда программы. В памяти выделяется множество особых ячеек содержимое которых после прекращения работы объявляется результатом работы программы.

11.1 Неветвящиеся программы с условной остановкой

Пусть $X = \{x_1, \dots, x_n\}$ — множество независимых булевых переменных. Введем множество переменных $Y = \{\mathtt{y}_1, \dots, \mathtt{y}_l\}$ и множество переменных $Z = \{\mathtt{z}_1, \dots, \mathtt{z}_m\}$. Переменные из множества Y назовем внутренними, а переменные из множества Z — выходными переменными. Пусть, далее, $\mathtt{a} \in Y \cup Z$, $\mathtt{b}, \mathtt{c} \in X \cup Y \cup Z$, f — булева функция, зависящая не более чем двух переменных. *Вычислительной командой* $\mathtt{p}$ назовем выражение

$$\mathtt{p}: \quad \mathtt{a} = f(\mathtt{b}, \mathtt{c}).$$

Переменную $\mathtt{a}$ назовем *выходом* вычислительной команды $\mathtt{p}$, а переменные $\mathtt{b}$, $\mathtt{c}$ — *входами* этой команды. Если переменная $\mathtt{a}$ является выходом команды $\mathtt{p}$, то будем говорить, что команда $\mathtt{p}$ изменяет значение этой переменной, а если $\mathtt{a}$ не является выходом $\mathtt{p}$, то будем говорить, что команда $\mathtt{p}$ не изменяет ее значение.

Пусть теперь $\mathtt{a} \in X \cup Y \cup Z$. *Командой остановки* $\mathtt{p}$ назовем выражение

$$\mathtt{p}: \quad \mathtt{Stop}(\mathtt{a}).$$

Переменную $\mathtt{a}$ назовем входом команды остановки $\mathtt{p}$.

Последовательность $\mathsf{P} = \mathtt{p}_1 \dots \mathtt{p}_i \dots \mathtt{p}_L$, состоящая из вычислительных команд и команд остановки, называется *неветвящейся программой с условной остановкой*, если при любом $j \in \{1, 2, \dots, L\}$ каждый вход команды $\mathtt{p}_j$ есть либо независимая переменная, либо выход некоторой вычислительной команды $\mathtt{p}_i$, где $i < j$.

Неветвящаяся программа работает в дискретные моменты времени $t = 0, 1, 2, \dots$, не изменяет значения независимых переменных и изменяет значения внутренних и выходных переменных. Значения $\mathtt{y}_i(\boldsymbol{x}; t)$ внутренних переменных $\mathtt{y}_i$ и значения $\mathtt{z}_j(\boldsymbol{x}; t)$ выходных переменных $\mathtt{z}_j$ программы P в произвольный момент времени t на наборе независимых переменных $\boldsymbol{x} = (x_1, \dots, x_n)$ определим индуктивно:

• В начальный момент времени $t = 0$ значения всех внутренних и выходных переменных считаем неопределенными;

• Если команда $\mathtt{p}_t$ не изменяет значения внутренней переменной $\mathtt{y}_i$ (или выходной переменной $\mathtt{z}_j$), то положим

$$\mathtt{y}_i(\boldsymbol{x}; t) = \mathtt{y}_i(\boldsymbol{x}; t-1), \qquad \mathtt{z}_j(\boldsymbol{x}; t) = \mathtt{z}_j(\boldsymbol{x}; t-1);$$

• Если команда $\mathtt{p}_t$ изменяет значения внутренней переменной $\mathtt{y}_i$ (или выходной переменной $\mathtt{z}_j$), и значения первого и второго входов команды $\mathtt{p}_t$ в момент времени $t-1$ равны соответственно $\mathtt{b}(\boldsymbol{x}; t-1)$ и $\mathtt{c}(\boldsymbol{x}; t-1)$, то положим

$$\mathtt{y}_i(\boldsymbol{x}; t) = f_t(\mathtt{b}(\boldsymbol{x}; t-1), \mathtt{c}(\boldsymbol{x}; t-1)),$$
$$\mathtt{z}_j(\boldsymbol{x}; t) = f_t(\mathtt{b}(\boldsymbol{x}; t-1), \mathtt{c}(\boldsymbol{x}; t-1)).$$

Значением команды $\mathtt{p}_t$ программы P на наборе независимых переменных $\boldsymbol{x} = (x_1, ..., x_n)$ назовем значение ее выхода в момент времени t и обозначим через $\mathtt{p}_t(\boldsymbol{x})$.

Через $n(\mathtt{p})$ обозначим номер команды $\mathtt{p}$ в программе P, т. е. $n(\mathtt{p}_i) = i$. Пусть $\mathtt{p}_{t_1}, \dots, \mathtt{p}_{t_r}$ — все команды остановки из P, причем $t_1 < \dots < t_r$. Тогда через $\mathtt{s}_j$ будем обозначать j-ю команду остановки программы P, т. е. $\mathtt{s}_j \equiv \mathtt{p}_{t_j}$.

Вычислительную команду $\mathtt{p}_i$ (переменную x_l) назовем *нулевым аргументом* команды остановки $\mathtt{s}_j$, $n(\mathtt{s}_j) = r$, и обозначим через $\mathtt{q}_j$, если:

(i) выход команды $\mathtt{p}_i$ (переменная x_l) является входом команды $\mathtt{s}_j$.

(ii) среди команд $\mathtt{p}_t$, $i < t < r$, нет команды, выход которой совпадает с выходом команды $\mathtt{p}_i$.

Будем говорить, что k-я команда остановки $\mathtt{s}_k$ прекращает вычисления программы P на наборе $\boldsymbol{x}$, если

$$\mathtt{q}_1(\boldsymbol{x}) = \dots = \mathtt{q}_{k-1}(\boldsymbol{x}) = 0, \quad \mathtt{q}_k(\boldsymbol{x}) = 1.$$

Результат действия программы P на наборе $\boldsymbol{x}$ обозначим через $\mathsf{P}(\boldsymbol{x})$ и его l-ю компоненту $\mathsf{P}_l(\boldsymbol{x})$ определим следующим образом:

$$\mathsf{P}_l(\boldsymbol{x}) = \begin{cases} \mathtt{z}_l(\boldsymbol{x}; t_k), & \text{если } \mathtt{q}_1(\boldsymbol{x}) = \dots = \mathtt{q}_{k-1}(\boldsymbol{x}) = 0, \ \ \mathtt{q}_k(\boldsymbol{x}) = 1, \\ \mathtt{z}_l(\boldsymbol{x}; L), & \text{если } \mathtt{q}_1(\boldsymbol{x}) = \dots = \mathtt{q}_r(\boldsymbol{x}) = 0, \end{cases}$$

т. е. $\mathsf{P}_l(\boldsymbol{x})$ равно значению l-й выходной переменной $\mathtt{z}_l$ в момент остановки программы. Легко видеть, что

$$\begin{aligned} \mathsf{P}_l(\boldsymbol{x}) = {} & \mathtt{q}_1(\boldsymbol{x})\mathtt{z}_l(\boldsymbol{x}; t_1) \vee \overline{\mathtt{q}}_1(\boldsymbol{x})\mathtt{q}_2(\boldsymbol{x})\mathtt{z}_l(\boldsymbol{x}; t_2) \vee \dots \\ & \dots \vee \overline{\mathtt{q}}_1(\boldsymbol{x})\overline{\mathtt{q}}_2(\boldsymbol{x}) \cdots \overline{\mathtt{q}}_{k-1}(\boldsymbol{x})\mathtt{q}_k(\boldsymbol{x})\mathtt{z}_l(\boldsymbol{x}; t_k) \vee \dots \\ & \dots \vee \overline{\mathtt{q}}_1(\boldsymbol{x})\overline{\mathtt{q}}_2(\boldsymbol{x}) \cdots \overline{\mathtt{q}}_{r-1}(\boldsymbol{x})\mathtt{q}_r(\boldsymbol{x})\mathtt{z}_l(\boldsymbol{x}; t_r) \\ & \vee \overline{\mathtt{q}}_1(\boldsymbol{x})\overline{\mathtt{q}}_2(\boldsymbol{x}) \cdots \overline{\mathtt{q}}_r(\boldsymbol{x})\mathtt{z}_l(\boldsymbol{x}; L). \end{aligned} \tag{11.1}$$

Иногда формулу (11.1) удобнее использовать в преобразованном виде

$$\begin{aligned} \mathsf{P}_l(\boldsymbol{x}) = {} & \mathtt{q}_1(\boldsymbol{x})\mathtt{z}_l(\boldsymbol{x}; t_1) \vee \overline{\mathtt{q}}_1(\boldsymbol{x})\big(\mathtt{q}_2(\boldsymbol{x})\mathtt{z}_l(\boldsymbol{x}; t_2) \vee \overline{\mathtt{q}}_2(\boldsymbol{x})\big(\dots \\ & \dots \big(\mathtt{q}_{k-1}(\boldsymbol{x})\mathtt{z}_l(\boldsymbol{x}; t_{k-1}) \vee \overline{\mathtt{q}}_{k-1}(\boldsymbol{x})\big(\dots \\ & \cdots \vee \overline{\mathtt{q}}_{r-1}\big(\mathtt{q}_r(\boldsymbol{x})\mathtt{z}_l(\boldsymbol{x}; t_r) \vee \overline{\mathtt{q}}_r\mathtt{z}_l(\boldsymbol{x}; L)\big)\dots\big)\dots\big)\dots\big). \end{aligned} \tag{11.2}$$

Будем говорить, что программа P вычисляет n-местную булеву функцию f, если $\mathsf{P}(\boldsymbol{x}) = f(\boldsymbol{x})$ для любого $\boldsymbol{x}$ из $\{0,1\}^n$.

Рассмотрим три разных программы, вычисляющих дизъюнкцию четырех переменных. Команды каждой из этих программ расположены в одном из следующих вертикальных столбцов:

$\mathsf{p}_1:\ \mathtt{z} = 1$	$\mathtt{z} = x_1 \vee x_2$	$\mathtt{y}_1 = x_1 \vee x_2$
$\mathsf{p}_2:\ \mathtt{Stop}(x_1)$	$\mathtt{Stop}(\mathtt{z})$	$\mathtt{y}_2 = x_3 \vee x_4$
$\mathsf{p}_3:\ \mathtt{Stop}(x_2)$	$\mathtt{z} = x_3 \vee x_4$	$\mathtt{z} = \mathtt{y}_1 \vee \mathtt{y}_2$
$\mathsf{p}_4:\ \mathtt{Stop}(x_3)$		
$\mathsf{p}_5:\ \mathtt{Stop}(x_4)$		
$\mathsf{p}_6:\ \mathtt{z} = 0$		

Первая программа состоит из шести команд и работает следующим образом. Сначала выходной переменной присваивается значение единица. Затем последовательно проверяются условия равенства единице переменных x_i. Если первая переменная равна единице, то первая команда остановки прекращает работу программы. Если $x_1 = 0$, то начинает работу вторая команда остановки, которая в свою очередь прекращает работу программы, если $x_2 = 1$. Если $x_2 = 0$, то аналогичным образом работает третья команда остановки, а затем, если $x_3 = 0$, — четвертая. Если ни одна из команд остановки не прекратила работу программы, т. е. если все переменные равны нулю, то выполняется последняя команда программы, которая присваивает выходной переменной нулевое значение. Используя (11.2) и тождество $x \vee \overline{x}\, y = x \vee y$, убеждаемся, что первая программа действительно вычисляет дизъюнкцию четырех переменных:

$$\begin{aligned}\mathsf{P}(\boldsymbol{x}) &= x_1 \cdot 1 \vee \overline{x}_1(x_2 \cdot 1 \vee \overline{x}_2(x_3 \cdot 1 \vee \overline{x}_3(x_4 \cdot 1 \vee \overline{x}_4 \cdot 0))) = \\ &= x_1 \vee \overline{x}_1(x_2 \vee \overline{x}_2(x_3 \vee \overline{x}_3 x_4)) = x_1 \vee \overline{x}_1(x_2 \vee \overline{x}_2(x_3 \vee x_4)) = \\ &= x_1 \vee \overline{x}_1(x_2 \vee x_3 \vee x_4) = x_1 \vee x_2 \vee x_3 \vee x_4.\end{aligned}$$

Вторая программа состоит из трех команд. Аналогичным образом для ее значения имеем

$$\mathsf{P}(\boldsymbol{x}) = (x_1 \vee x_2)(x_1 \vee x_2) \vee \overline{(x_1 \vee x_2)}(x_3 \vee x_4) = x_1 \vee x_2 \vee x_3 \vee x_4.$$

Третья программа состоит только из вычислительных команд и, поэтому, по существу является схемой из функциональных элементов. Очевидно, что она так же вычисляет дизъюнкцию переменных x_1, x_2, x_3 и x_4.

Сложностью $C(\mathsf{P})$ программы P назовем число команд этой программы. *Временем работы* $T_{\mathsf{P}}(\boldsymbol{x})$ программы P на наборе переменных $\boldsymbol{x}$ назовем минимальное $n(\mathsf{s}_j)$ такое, что $\mathsf{q}_j(\boldsymbol{x}) = 1$, т. е. это число команд, выполненных до остановки программы. Если все $\mathsf{q}_j(\boldsymbol{x}) = 0$, то выполняются все команды программы и в этом случае $T_{\mathsf{P}}(\mathbf{x}) = C(\mathsf{P})$. Величину

$$T(\mathsf{P}) = 2^{-n} \sum T_{\mathsf{P}}(\boldsymbol{x}),$$

где суммирование производится по всем двоичным наборам длины n, назовем *средним временем работы* программы P. Если для некоторого булевой функции f и любого двоичного набора $\boldsymbol{x}$ справедливо равенство $f(\boldsymbol{x}) = \mathsf{P}(\boldsymbol{x})$, то будем говорить, что программа P вычисляет функцию f. Величину

$$T(f) = \min T(\mathsf{P}),$$

где минимум берется по всем программам, вычисляющим f, назовем *средним временем вычисления* (*средней сложностью*) функции f. Программу P, вычисляющую функцию f, для которой справедливо равенство $T(\mathsf{P}) = T(f)$, назовем *минимальной* программой. Величину

$$C(f) = \min C(\mathsf{P}),$$

где минимум берется по всем программам, вычисляющим f, назовем *программной сложностью* функции f. Величина $C(f)$ характеризует время, необходимое для вычисления f в худшем случае, поэтому $C(f)$ так же будем называть сложностью в худшем случае. Отметим, что неветвящаяся программа, не содержащая команды остановки и вычисляющая функцию, отличную от независимой переменной, является обычной схемой из функциональных элементов, базис которой состоит из всех не более чем двухместных булевых функций. Поэтому средняя сложность любой булевой функции $f(x_1, \ldots, x_n)$, существенно зависящей не менее чем от двух переменных, не меньше ее схемной сложности, т. е.

$$T(f(x_1, \ldots, x_n)) \le L(f(x_1, \ldots, x_n)).$$

Рассмотрим две программы P_1 и P_2, вычисляющие систему из двух функций — дизъюнкции и конъюнкции четырех переменных:

$\mathsf{p}_1:\ \mathtt{z}_1 = x_1 \oplus x_2$	$\mathtt{z}_1 = x_1 \vee x_2$
$\mathsf{p}_2:\ \mathtt{z}_2 = 0$	$\mathtt{z}_1 = \mathtt{z}_1 \vee x_3$
$\mathsf{p}_3:\ \mathtt{Stop}(\mathtt{z}_1)$	$\mathtt{z}_1 = \mathtt{z}_1 \vee x_4$
$\mathsf{p}_4:\ \mathtt{z}_1 = x_3 \oplus x_4$	$\mathtt{z}_2 = x_1 \,\&\, x_2$
$\mathsf{p}_5:\ \mathtt{Stop}(\mathtt{z}_1)$	$\mathtt{z}_2 = \mathtt{z}_2 \,\&\, x_3$
$\mathsf{p}_6:\ \mathtt{z}_1 = x_1 \vee x_3$	$\mathtt{z}_2 = \mathtt{z}_2 \,\&\, x_4$
$\mathsf{p}_7:\ \mathtt{z}_2 = x_1 \,\&\, x_3$	

В этих программах дизъюнкция вычисляется переменной $\mathtt{z}_1$, а конъюнкция — переменной $\mathtt{z}_2$. Легко видеть, что сложности первой и второй программ равны, соответственно, семи и шести. Найдем их средние времена работы. В первой программе первая команда остановки прекращает ее работу на восьми наборах: (0100), (0101), (0110), (0111), (1000), (1001), (1010), (1011); вторая команда остановки — на четырех наборах: (0001), (0001), (1101), (1111); наконец, на оставшихся четырех наборах (0000), (0011), (1100) и

(1111) выполняются все команды программы. Поэтому для среднего времени работы программы P_1 имеем

$$T(\mathsf{P}_1) = \frac{1}{16}\Big(3\cdot 8 + 5\cdot 4 + 7\cdot 5\Big) = \frac{9}{2}.$$

Вторая программа состоит только из вычислительных команд, и, следовательно, среднее время работы этой программы совпадает с ее сложностью, т. е. $T(\mathsf{P}_2) = 6$.

Пусть переменная $\mathtt{a}$ является входом команды p_i программы P, а переменная $\mathtt{b}$ — выходом вычислительной команды p_j этой программы. Будем говорить, что на входе команды p_i вычисляется булева функция $f(\boldsymbol{x})$, если

$$\mathtt{a}(\boldsymbol{x}; i-1) = f(\boldsymbol{x})$$

при всех значениях независимых переменных, при которых значение переменной $\mathtt{a}$ определено. Аналогичным образом скажем, что на выходе команды p_j вычисляется булева функция $h(\boldsymbol{x})$, если

$$\mathtt{b}(\boldsymbol{x}; j) = h(\boldsymbol{x})$$

при всех значениях независимых переменных, при которых значение переменной $\mathtt{b}$ определено.

Будем говорить, что две команды p_i и p_j одной программы имеют общий вход, если на их входах вычисляются одинаковые функции.

Среди всех программ выделим множество приведенных программы и покажем, что любая программа без увеличения сложности и среднего времени работы может быть преобразована в приведенную программу.

Программу P назовем *приведенной*, если:

• вход каждой команды программы P не является тождественной постоянной;

• никакие две команды остановки программы P не имеют общего входа.

Лемма 11.1. *Произвольная программа* P *может быть преобразована в приведенную программу* P' *так, что*:

(*i*) $T(\mathsf{P}') \le T(\mathsf{P}), \quad C(\mathsf{P}') \le C(\mathsf{P})$;

(*ii*) $\mathsf{P}'(x_1,\ldots,x_n) = \mathsf{P}(x_1,\ldots,x_n)$ *при всех* $(x_1,\ldots,x_n)$.

Доказательство. Пусть P — произвольная программа, $\boldsymbol{x} = (x_1,\ldots,x_n)$ — набор независимых переменных, $\mathsf{p}_t:\ \mathtt{a} = g(\mathtt{b},\mathtt{c})$ — команда, на первом входе которой вычисляется постоянная функция, т. е. $\mathtt{b}(\boldsymbol{x}; t-1) = \text{const}$. В этом случае $\mathtt{a}(\boldsymbol{x}; t)$ зависит только от $\mathtt{c}(\boldsymbol{x}; t-1)$, и, следовательно, существует функция h такая, что $h(\mathtt{c}(\boldsymbol{x}; t-1)) = g(\mathtt{b}(\boldsymbol{x}; t-1), \mathtt{c}(\boldsymbol{x}; t-1))$. Заменяя в P команду p_t командой $\boldsymbol{p}'_t:\ \mathtt{a} = h(\mathtt{c})$, получаем программу, удовлетворяющую условиям (*i*) и (*ii*).

Допустим теперь, что в программе P найдутся две различных команды остановки p_i и p_j, где $i < j$, с одним и тем же входом $\mathtt{a}$. Если $\mathtt{a}(\boldsymbol{x}; i-1) = \mathtt{a}(\boldsymbol{x}; j-1) = 1$, то команда p_i останавливает выполнение программы, и команда p_j не выполняется. Если $\mathtt{a}(\boldsymbol{x}; j-1) = 0$, то команда p_j не останавливает выполнение программы. Следовательно, команда p_j может быть удалена из программы P, и это ни как не скажется на ее работе. Лемма доказана.

11.2 Примеры

Рассмотрим три примера, в которых оценивается средняя сложность конкретных булевых функций.

1. Эффекты, связанные с возможностью досрочного прекращения вычислений, начинают проявляться уже при вычислении булевых функций трех переменных. Выше было показано (см. стр. 116), что схемная сложность каждой булевой функции трех переменных не превосходит четырех, причем существуют функции, например функция голосования τ_2, сложность которых равна четырем. Покажем, что средняя сложность любой булевой функции трех переменных не превосходит $2\frac{1}{2}$.

Пусть f — произвольная булева функция трех переменных. Разложим f по первой переменной

$$f(x_1,x_2,x_3) = x_1 f_1(x_2,x_3) \vee \overline{x}_1 f_2(x_2,x_3).$$

Легко видеть, что программа P

$$\begin{aligned}
&\mathsf{p}_1:\ \mathtt{z} = f_1(x_2,x_3)\\
&\mathsf{p}_2:\ \mathtt{Stop}(x_1)\\
&\mathsf{p}_3:\ \mathtt{z} = f_2(x_2,x_3)
\end{aligned}$$

вычисляет функцию f. Действительно, в соответствии с определением функции, вычисляемой неветвящейся программой с условной остановкой, для $\mathsf{P}(\boldsymbol{x})$ справедливы равенства

$$\mathsf{P}(\boldsymbol{x}) = \mathsf{q}_1(\boldsymbol{x})\mathtt{z}(\boldsymbol{x};1) \vee \overline{\mathsf{q}}_1(\boldsymbol{x})\mathtt{z}(\boldsymbol{x};3) = x_1 f_1(x_2,x_3) \vee \overline{x}_1 f_2(x_2,x_3).$$

На четырех наборах — $(1,0,0)$, $(1,0,1)$, $(1,1,0)$, $(1,1,1)$ — программа P выполняет два действия, на остальных четырех наборах — три. Поэтому

$$T(f) \le T(\mathsf{P}) = \frac{1}{8}\,(2\cdot 4 + 3\cdot 4) = 2\frac{1}{2}.$$

Теперь покажем, что средняя сложность функции голосования τ_2 равна $2\frac{1}{2}$. Пусть P — программа, вычисляющая τ_2 с минимальным средним временем. Если первая команды остановки s_1 является третьей командой P, то, очевидно, что $T(\mathsf{P}) \ge 3$. Так как в любой программе на первом месте должна стоять вычислительная команда, то далее полагаем, что s_1 является второй командой P. Теперь для доказательства неравенства $T(g) \ge 2\frac{1}{2}$ достаточно показать, что в P, команда остановки прекращает вычисления не более чем на четырех наборах из восьми, так как в этом случае $T(\mathsf{P}) \ge \frac{1}{8}\,(4\cdot 2 + 4\cdot 3) = 2\frac{1}{2}$.

Легко видеть, что нулевым аргументом команды остановки может быть либо выходная, либо независимая переменная, т. е. начало программы P имеет следующий вид:

$$\begin{array}{lll} \mathsf{p}_1: & \mathtt{z} = f(x_1, x_2) & \mathtt{z} = f(x_1, x_2) \\ \mathsf{p}_2: & \mathtt{Stop}(\mathtt{z}) & \mathtt{Stop}(x_i) \end{array}$$

где $i \in \{1, 2, 3\}$. В первом случае если $T_{\mathsf{P}}(\boldsymbol{x}) = 2$, то $\mathsf{P}(\boldsymbol{x}) = 1$. Следовательно, команда p_2 прекращает вычисления не более чем на четырех наборах. Во втором случае очевидно, что p_2 прекращает вычисления ровно на четырех наборах.

2. Симметрической пороговой функцией n переменных с порогом m называется такая функция $\tau_{n,m}$, что

$$\tau_{n,m}(x_1, \ldots, x_n) = \begin{cases} 1, & \text{если } \sum_{i=1}^n x_i \geq m, \\ 0, & \text{если } \sum_{i=1}^n x_i < m. \end{cases}$$

Нетрудно показать (см. теорему 8.1 и доказательство леммы 9.3), что $L(\tau_{n,m}) \asymp n$, т. е. с точностью до постоянного множителя сложность симметрических пороговых функций пропорциональна числу аргументов и не зависит от величины порога. В случае средней сложности ситуация обратная — средняя сложность (при $m \leq \frac{1}{2}n$) не зависит от числа аргументов функции и с точностью до постоянного множителя пропорциональна величине порога.

Теорема 11.1.

$$T(\tau_{n,m}) \asymp \min(m, n - m).$$

Доказательство. Без ограничения общности будем полагать, что $m \leq \frac{1}{2}n$. Случай $m > \frac{1}{2}n$ рассматривается аналогично.

Сначала покажем, что $T(\tau_{n,m}) \geq m$. Пусть P — минимальная программа, вычисляющая $\tau_{n,m}$, s_1 — первая команда остановки этой программы, $\alpha = (\alpha_1, \ldots \alpha_n)$ — набор на котором команда s_1 останавливает вычисления. Если $n(\mathsf{s}_1) < m$, то система функций $\{\mathsf{q}_1(\boldsymbol{x}), \mathtt{z}(\boldsymbol{x}; n(\mathsf{s}_1) - 1)\}$ существенно зависит не более чем от $m - 1$ переменных. Без ограничения общности полагаем, что это $x_1, \ldots x_m$. Тогда $\mathsf{P}(\alpha_1, \ldots \alpha_m, 0, \ldots, 0) = \mathsf{P}(\alpha_1, \ldots \alpha_m, 1, \ldots, 1)$, что противоречит очевидным равенствам

$$\tau_{n,m}(\alpha_1, \ldots \alpha_m, 0, \ldots, 0) = 0, \quad \tau_{n,m}(\alpha_1, \ldots \alpha_m, 1, \ldots, 1) = 1.$$

Следовательно, $T(\tau_{n,m}) \geq m$.

Теперь покажем, что $T(\tau_{n,m}) = \mathcal{O}(m)$. Для этого опишем программу $\mathsf{P}_{n,m}$, вычисляющую функцию $\tau_{n,m}$, и оценим среднее время работы этой программы. Пусть $n = (2m-1)t + k$, где $0 \leq k < 2m - 1$. Программу $\mathsf{P}_{n,m}$ представим в виде последовательности независимых частей—подпрограмм:

$$\mathsf{P}_{n,m} = \mathsf{P}_1 \ldots \mathsf{P}_j \ldots \mathsf{P}_t \mathsf{P}_{t+1}.$$

Подпрограмма P_1 присваивает выходной переменной значение 1, вычисляет сумму $S_1 = \sum_{i=1}^{(2m-1)} x_i$ и останавливает вычисления если $S_1 \geq m$. При каждом $j \in \{2, 3, \ldots, t\}$ подпрограмма P_j вычисляет сумму $S_j = \sum_{i=(2m-1)(j-1)+1}^{(2m-1)j} x_i$ и останавливает вычисления если $S_j \geq m$. Подпрограма P_{t+1} вычисляет сумму $S = \sum_{i=1}^n x_i$, и присваивает выходной переменной значение 1 если $S_j \geq m$, и значение 0 если $S < m$. Легко видеть, что сложность каждой подпрограммы $C(\mathsf{P}_j)$ есть $\mathcal{O}(m)$.

Нетрудно убедиться в том, что подпрограмма P_1 останавливает вычисления на $2^{n-(2m-1)} \sum_{i=m}^{2m-1} \binom{2m-1}{i} = 2^{n-1}$ наборах, а при $j \leq t$ каждая подпрограмма P_j — на $2^{n-j(2m-1)} \left(\sum_{i=m}^{2m-1} \binom{2m-1}{i}\right)^j = 2^{n-j}$ наборах. Таким образом вместе подпрограммы $\mathsf{P}_1, \ldots, \mathsf{P}_t$ прекращают вычисления на $2^n(1 - 2^{-t})$ наборах, и поэтому подпрограмма P_{t+1} работает только на 2^{n-t} наборах. Следовательно,

$$\begin{aligned} T(\mathsf{P}) &\asymp \frac{1}{2^n} \left(\sum_{j=1}^{t} \left(2^{n-j} \sum_{i=1}^{j} C(\mathsf{P}_i) \right) + 2^{n-t} C(\mathsf{P}) \right) \asymp \\ &\asymp \frac{m}{2^n} \left(2^n \sum_{j=1}^{t} \sum_{i=1}^{j} 2^{-j} + t \cdot 2^{n-t} \right) \asymp \frac{m}{2^n} \left(2^n \sum_{j=1}^{\infty} j \cdot 2^{-j} + t \cdot 2^{n-t} \right) = \\ &= m \left(\sum_{j=1}^{\infty} \sum_{i=j}^{\infty} 2^{-j} + t \cdot 2^{-t} \right) \leq m \left(\sum_{j=1}^{\infty} 2^{1-j} + 1 \right) = 3m. \end{aligned}$$

Теорема доказана.

3. Рассмотрим две простейшие симметрические пороговые функции — дизъюнкцию и конъюнкции растущего числа аргументов. Покажем, что для средней сложности этих функций справедлива следующая теорема.

Теорема 11.2. *Пусть $n \to \infty$. Тогда*

$$T(x_1 \vee x_2 \vee \cdots \vee x_n) \sim \frac{8}{3}, \qquad T(x_1 \,\&\, x_2 \,\&\, \ldots \,\&\, x_n) \sim \frac{11}{3}.$$

Доказательство. Верхние оценки. Верхние оценки средней сложности получим оценив среднее время работы приводимых ниже программ $\mathsf{P}_\vee$ и $\mathsf{P}_\&$, вычисляющих соответственно дизъюнкцию (второй столбец) и конъюнкцию (третий столбец) n переменных. Без ограничения общности полагаем, что n — четное:

$$\begin{array}{lll} \mathsf{p}_1: & \mathtt{z} = x_1 \vee x_2 & \mathtt{y} = \overline{x_1 \& x_2} \\ \mathsf{p}_2: & \mathtt{Stop}(\mathtt{z}) & \mathtt{z} = 0 \\ \mathsf{p}_3: & \mathtt{z} = x_3 \vee x_4 & \mathtt{Stop}(\mathtt{y}) \\ \mathsf{p}_4: & \mathtt{Stop}(\mathtt{z}) & \mathtt{y} = \overline{x_3 \& x_4} \\ \mathsf{p}_5: & \mathtt{z} = x_5 \vee x_6 & \mathtt{Stop}(\mathtt{y}) \end{array}$$

.....	.....	.....
p_j :	$\mathtt{z} = x_j \vee x_{j+1}$	$\mathtt{Stop(y)}$
p_{j+1} :	$\mathtt{Stop(z)}$	$\mathtt{y} = \overline{x_j \& x_{j+1}}$
p_{j+2} :	$\mathtt{z} = x_{j+2} \vee x_{j+3}$	$\mathtt{Stop(y)}$
.....	.....	.....
p_{n-3} :	$\mathtt{z} = x_{n-3} \vee x_{n-2}$	$\mathtt{Stop(y)}$
p_{n-2} :	$\mathtt{Stop(z)}$	$\mathtt{y} = \overline{x_{n-3} \& x_{n-2}}$
p_{n-1} :	$\mathtt{z} = x_{n-1} \vee x_n$	$\mathtt{Stop(y)}$
p_n :		$\mathtt{z} = x_{n-1} \& x_n$

Легко видеть, что

$$T(\mathsf{P}_\vee) \le \frac{1}{2^n}\left(\sum_{j=1}^{n/2} 2j \cdot 3 \cdot 2^{n-2j}\right) < 6\left(\sum_{j=1}^{\infty} \frac{j}{4^j}\right) =$$
$$= 6\left(\sum_{j=1}^{\infty}\sum_{i=j}^{\infty} \frac{1}{4^i}\right) = 6\left(\sum_{j=1}^{\infty} \frac{1}{4^j}\frac{1}{1-\frac{1}{4}}\right) = 6 \cdot \frac{4}{3} \cdot \frac{1}{4} \cdot \frac{4}{3} = \frac{8}{3}.$$

Сравнивая программу $\mathsf{P}_\&$ с программой $\mathsf{P}_\vee$, легко видеть, что справедливо равенство $T(\mathsf{P}_\&) = 1 + T(\mathsf{P}_\vee)$. Следовательно,

$$T(\mathsf{P}_\vee) \le 1 + \frac{8}{3} = \frac{11}{3}.$$

Верхние оценки доказаны.

Нижние оценки. Нижнюю оценку докажем только для конъюнкции. Доказательство для дизъюнкции аналогично. Прежде всего отметим следующие факты:

(i) среди первых $n-1$ команд минимальной программы P_n, вычисляющей конъюнкцию n переменных, нет команды остановки, входом которой является выходная переменная.

(ii) среди первых $n-2$ команд минимальной программы P_n, вычисляющей конъюнкцию n переменных, нет команды остановки, входом которой является независимая переменная.

Докажем (i). Предположим, что такая команда $\mathsf{p}_t : \mathtt{Stop(z)}$ в минимальной программе есть. Тогда на некотором наборе $(\sigma_1, \ldots, \sigma_n)$ эта команда прекращает выполнение программы. Легко видеть, что выходная переменная $\mathtt{z}(\boldsymbol{x}; t-1)$ существенно зависит не более чем от $n-1$ переменных. Без ограничения общности полагаем, что $\mathtt{z}(\boldsymbol{x}; t-1)$ не зависит от x_n. Тогда $\mathtt{z}(\boldsymbol{x}; t-1)$ равна единице на наборе $(\sigma_1, \ldots, \sigma_{n-1}, 0)$. Противоречие.

Докажем (ii). Предположим, что такая команда $\mathsf{p}_t : \mathtt{Stop}(x_l)$ в минимальной программе есть. Легко видеть, что выходная переменная $\mathtt{z}(\boldsymbol{x}; t-1)$ существенно зависит не более чем от $n-2$ переменных. Без ограничения общности полагаем, что $\mathtt{z}(\boldsymbol{x}; t-1)$ не зависит от x_{n-1} и x_n и $n \ne l$. Тогда программа прекращает работу на наборе $x_1 = x_2 = \cdots = x_{n-1} = 1$ и результат работы не зависит от x_n. Противоречие.

Отметим также, что первой командой любой программы, вычисляющей полностью определенную функцию, не может быть команда остановки; перед командой остановки необходима хотя бы одна вычислительная команда, выходом которой является выходная переменная.

Пусть P_n — минимальная программа, вычисляющая функцию $\&_n = x_1 \& x_2\& \ldots \& x_n$. Если первая команда остановки программы P_n является четвертой командой P_n, то очевидно, что $T(\mathsf{P}_n) > 4$. Поэтому достаточно рассмотреть случай, когда в программе P_n первая команда остановки стоит на втором или третьем месте.

Так как из свойств (i) и (ii) следует, что независимая и выходная переменные не могут быть входом команды остановки, то очевидно, что в программе P_n первая команда остановки не может быть второй командой программы. Поэтому достаточно рассмотреть ситуацию, когда первая команда остановки s_1 стоит в P_n на третьем месте. Из свойств (i) и (ii) легко следует, что возможны только три принципиально различных, с точностью до переименования переменных, случая для выбора первых трех команд этой программы:

p_1 :	$\mathtt{y} = g(x_1, x_2)$	$\mathtt{y} = g(x_1, x_2)$	$\mathtt{z} = g(x_1, x_2)$
p_2 :	$\mathtt{z} = f(\mathtt{y}, x_3)$	$\mathtt{z} = f(x_3, x_4)$	$\mathtt{y} = f(\mathtt{z}, x_3)$
p_3 :	$\mathtt{Stop(y)}$	$\mathtt{Stop(y)}$	$\mathtt{Stop(y)}$

Так как вычисляемая первыми двумя командами программы система функций $\{\mathtt{y}(\boldsymbol{x}; 2), \mathtt{z}(\boldsymbol{x}; 2)\}$ существенно зависит не более чем от четырех переменных, то при $n > 4$, найдется переменная, например x_n, не являющаяся существенной для системы $\{\mathtt{y}(\boldsymbol{x}; 2), \mathtt{z}(\boldsymbol{x}; 2)\}$. Так как на любом наборе с $x_n = 0$ конъюнкция n переменных равна нулю, то очевидно, что

$$\mathtt{z}(x_1, x_2, \ldots, x_n; 2) \& \mathtt{y}(x_1, x_2, \ldots, x_n; 2) = 0, \tag{11.3}$$

при всех возможных значениях переменных $x_1, \ldots, x_n$.

Далее рассмотрим два первых случая, когда функция $\mathtt{y}(\boldsymbol{x}; 2)$ существенно зависит только от двух переменных. Покажем, что при всех $(\sigma_3, \ldots, \sigma_n)$

$$T_{P_n}(1, 1, \sigma_3, \ldots, \sigma_n) > 3, \tag{11.4}$$

т. е. при $x_1 = x_2 = 1$ первая команда остановки программы P_n не останавливает вычисления. Сделаем это методом от противного. Допустим, что найдется набор $(\sigma'_3, \ldots, \sigma'_n)$ такой, что $T_{\mathsf{P}_n}(1, 1, \sigma'_3, \ldots, \sigma'_n) = 3$. Следовательно, $\mathtt{y}(1, 1; 2) = 1$. Так как $\mathtt{y}(\boldsymbol{x}; 2)$ существенно зависит только от первых двух переменных, то в силу (11.3) имеем

$$\mathtt{z}(1, 1, \ldots, 1; 2) \& \mathtt{y}(1, 1, \ldots, 1; 2) = 0.$$

Поэтому, $\mathtt{z}(1, 1, \ldots, 1; 2) = 0$. Противоречие.

В программе P_n вместо первых двух переменных подставим единицы и преобразуем ее в приведенную программу. Новая программа P_{n-2} вычисляет конъюнкцию $n-2$ переменных и содержит по крайней мере на две команды меньше чем программа P_n. Из (11.4) следует, что первую команду остановки программы P_n, т. е. команду p_3, можно удалить, так как после подстановки единиц вместо x_1 и x_2 она не останавливает вычисления. После этой подстановки первая команда программы P_n вычисляет тождественный нуль и его выход — внутренняя переменная. Из леммы 11.1 следует, что выход этой команды не является входом никакой другой команды P_{n-2} и поэтому ее можно удалить. Следовательно,

$$\sum_{\boldsymbol{x}\in\{0,1\}^n, x_1=x_2=1} T_{\mathsf{P}_n}(x_1,\ldots,x_n) \geq \sum_{\boldsymbol{x}\in\{0,1\}^{n-2}} \left(T_{\mathsf{P}_{n-2}}(x_3,\ldots,x_n)+2\right).$$

Так как

$$\frac{1}{2^{n-2}} \sum_{\boldsymbol{x}\in\{0,1\}^{n-2}} T_{\mathsf{P}_{n-2}}(x_3,\ldots,x_n) \geq T(\&_{n-2}(x_3,\ldots,x_n)),$$

то

$$\begin{aligned} T(\&_n) &= \frac{1}{2^n}\left(\sum_{x\in\{0,1\}^n, x_1\&x_2=0} 3 + \sum_{x\in\{0,1\}^n, x_1=x_2=1} T_{P_n}(x_1,\ldots,x_n)\right) \geq \\ &\geq \frac{1}{2^n}\left(3\cdot 3\cdot 2^{n-2} + (2+T(\&_{n-2}(x_3,\ldots,x_n)))\,2^{n-2}\right) \geq \\ &\geq \frac{9}{4}+\frac{2}{4}+\frac{T(\&_{n-2})}{4} = \frac{11}{4}+\frac{T(\&_{n-2})}{4}. \end{aligned}$$

Таким образом

$$\begin{aligned} T(\&_n) &\geq \frac{11}{4}+\frac{T(\&_{n-2})}{4} \geq \frac{11}{4}+\frac{11}{16}+\frac{T(\&_{n-4})}{16} \geq \cdots \geq \\ &\geq \frac{11}{4}\left(1+\frac{1}{4}+\frac{1}{16}+\ldots\right) \sim \frac{11}{4}\cdot\frac{1}{1-\frac{1}{4}} \sim \frac{11}{3}. \end{aligned}$$

Рассмотрим третий случай. Если $\mathsf{y}(\boldsymbol{x};2)$ существенно зависит только от двух переменных, то доказательство полностью аналогично доказательству в первых двух случаях. Поэтому далее полагаем, что $\mathsf{y}(\boldsymbol{x};2)$ существенно зависит от трех переменных, и, следовательно, $\mathsf{z}(\boldsymbol{x};1)$ существенно зависит от двух переменных. Разложим $\mathsf{y}(\boldsymbol{x};2)$ по аргументу z:

$$\mathsf{y}(\boldsymbol{x};2) = \mathsf{z}(\boldsymbol{x};1)f_1(x_3) \vee \overline{\mathsf{z}}(\boldsymbol{x};1)f_2(x_3).$$

Так как $\mathsf{z}(\boldsymbol{x};1) = \mathsf{z}(\boldsymbol{x};2)$, то из (11.3) легко видеть, что

$$0 = \mathsf{z}(\boldsymbol{x};2)\mathsf{y}(\boldsymbol{x};2) = \mathsf{z}(\boldsymbol{x};2)f_1(x_3) \vee \mathsf{z}(\boldsymbol{x};2)\overline{\mathsf{z}}(\boldsymbol{x};2)f_2(x_3) = \mathsf{z}(\boldsymbol{x};2)f_1(x_3).$$

Следовательно, $f_1(x_3)\equiv 0$. Поэтому,

$$\mathsf{y}(\boldsymbol{x};2) = \overline{\mathsf{z}}(\boldsymbol{x};1)f_2(x_3).$$

Но тогда $\mathsf{y}(x_1,x_2,x_3;2)$ равна единице не более чем на трех наборах значений переменных x_1,x_2,x_3. Поэтому, первая команда остановки прекращает вычисления не более чем на $3\cdot 2^{n-3}$ наборах значений переменных $x_1,\ldots,x_n$. На остальных $5\cdot 2^{n-3}$ наборах должны выполниться еще хотя бы две команды: команда остановки и ее аргумент — вычислительная команда (см. свойства (i) и (ii)). Следовательно,

$$T(\mathsf{P}) \geq \frac{1}{2^n}\left(3\cdot 3\cdot 2^{n-3} + 5\cdot 5\cdot 2^{n-3}\right) > 4.$$

Теорема доказана.

11.3 Средняя сложность почти всех функций

В лекции 10 было установлено, что при $n\to\infty$ почти все n-местные булевы функции имеют экспоненциальную, относительно n, сложность. Покажем, что аналогичный эффект имеет место и для средней сложности. Кроме того, покажем, что средняя сложность почти каждой булевой функции с точностью до постоянного множителя совпадает с ее обычной сложностью.

С каждой программой P, вычисляющей n-местную булеву функцию, свяжем линейный порядок на множестве двоичных наборов длины n. Сделаем это следующим образом. Каждому двоичному набору $\boldsymbol{x}$ длины n, рассматриваемому как двоичная запись натурального числа, поставим в соответствие его номер $N_{\mathsf{P}}(\boldsymbol{x})$ такой, что $1\leq N_{\mathsf{P}}(\boldsymbol{x})\leq 2^n$; $N_{\mathsf{P}}(\boldsymbol{x}) < N_{\mathsf{P}}(\boldsymbol{y})$, если $T_{\mathsf{P}}(\boldsymbol{x}) < T_{\mathsf{P}}(\boldsymbol{y})$; $N_{\mathsf{P}}(\boldsymbol{x}) < N_{\mathsf{P}}(\boldsymbol{y})$, если $T_{\mathsf{P}}(\boldsymbol{x}) = T_{\mathsf{P}}(\boldsymbol{y})$ и $\boldsymbol{x} < \boldsymbol{y}$.

Теорема 11.3. *Пусть $n\to\infty$. Тогда:*

(i) Для любой постоянной $\varepsilon>0$ доля n-местных булевых функций f, для которых

$$T(f) \geq (1-\varepsilon)\frac{2^{n-3}}{n},$$

стремится к единице;

(ii) для каждой булевой функции f, зависящей от n переменных,

$$T(f) \lesssim \frac{2^{n-1}}{n}.$$

Доказательство. (i) Оценим число булевых функций, средняя сложность каждой из которых не превосходит величины $(1-\varepsilon)\frac{2^{n-3}}{n}$. Пусть f — одна из таких функций, P — минимальная программа, вычисляющая f. Рассмотрим набор $\boldsymbol{x}_0$ такой, что $N_{\mathsf{P}}(\boldsymbol{x}_0) = 2^{n-1}$. Тогда из определения средней сложности следует, что

$$T(\mathsf{P}) = 2^{-n}\sum_{\boldsymbol{y}} T_{\mathsf{P}}(\boldsymbol{y}) > 2^{-n} \sum_{\boldsymbol{y}\ |\ N(\boldsymbol{y})>N(\boldsymbol{x}_0)} T_{\mathsf{P}}(\boldsymbol{y}) \geq \frac{1}{2}T_{\mathsf{P}}(\boldsymbol{x}_0). \qquad (11.5)$$

Поэтому, $T_{\mathsf{P}}(\boldsymbol{x}_0) < 2T(f)$. Так как $T(f) \le (1-\varepsilon)\frac{2^{n-3}}{n}$, то легко видеть, что

$$T_{\mathsf{P}}(\boldsymbol{x}_0) < 2(1-\varepsilon)\frac{2^{n-3}}{n} = (1-\varepsilon)\frac{2^{n-2}}{n}. \tag{11.6}$$

Каждая функция однозначно определяется первыми $T_{\mathsf{P}}(\boldsymbol{x}_0)$ командами своей минимальной программы P и двоичным вектором длины не более чем 2^{n-1}, состоящим из значений функции f на тех аргументах, время работы P на которых больше времени работы этой программы на $\boldsymbol{x}_0$. Обозначим через N_0 число различных программ, состоящих не более чем из $T_{\mathsf{P}}(\boldsymbol{x}_0)$ команд. Тогда число функций, средняя сложность которых не превосходит $(1-\varepsilon)\frac{2^{n-3}}{n}$, ограничена сверху величиной $N_0 2^{2^{n-1}}$. Оценим N_0.

Любая программа P определяется списком своих команд p_i, каждая из которых однозначно задается следующими данными:

• типом команды — возможны всего два варианта, команда может быть либо вычислительной, либо командой остановки;

• двуместной булевой функцией f_i, вычисляемой вычислительной командой (для команды остановки эта информация опускается) — существует всего 16 различных двуместных булевых функций;

• номером переменной, выходной или внутренней, являющейся выходом вычислительной команды (для команд остановки эта информация опускается) — если программа P состоит из L команд, то общее число внутренних и выходной переменных не превосходит L и без ограничения общности полагаем, что внутренние переменные нумеруются числами от 1 до $L-1$, а выходной переменной присваивается номер L;

• номерами переменных, независимых или внутренних, являющихся входами команды — полагаем, что независимые переменные нумеруются числами от $L+1$ до $L+n$, таким образом общее число пар номеров не превосходит $(L+n)^2$.

Поэтому для числа N, равного числу различных программ, состоящих из L команд, справедливо неравенство

$$N \le \left(2 \cdot 16 \cdot L \cdot (L+n)^2\right)^L \le (4(L+n))^{3L}. \tag{11.7}$$

Подставляя в (11.7) вместо L величину $T_{\mathsf{P}}(\boldsymbol{x}_0)$ и учитывая неравенство (11.6), получаем, что при начиная с некоторого n имеет место неравенство

$$N_0 \le \left(4\left(T_{\mathsf{P}}(\boldsymbol{x}_0)+n\right)\right)^{3T_{\mathsf{P}}(\boldsymbol{x}_0)} \le$$
$$\le \left(4\left((1-\varepsilon)\frac{2^{n-2}}{n}+n\right)\right)^{3(1-\varepsilon)\cdot 2^{n-2}/n} \le 2^{3(1-\varepsilon)\cdot 2^{n-2}}.$$

Следовательно, число функций, средняя сложность которых не превосходит $(1-\varepsilon)\frac{2^{n-4}}{n}$, не больше чем

$$2^{3\cdot(1-\varepsilon)\cdot 2^{n-2}} 2^{2^{n-1}} = 2^{2^n(1-3\varepsilon/4)} = o\left(2^{2^n}\right).$$

Таким образом, при $n \to \infty$ для любой положительной постоянной ε средняя сложность почти каждой булевой функции, зависящей от n переменных, не меньше чем $(1-\varepsilon)\frac{2^{n-3}}{n}$. Первое неравенство теоремы доказано.

(ii) Каждому двоичному набору $(\sigma_1 \dots \sigma_k)$ поставим в соответствие его номер $N(\sigma_1 \dots \sigma_k) = \sum_{i=1}^{k} \sigma_i 2^{i-1}$.

Положим $s = \lfloor n - \log_2 n \rfloor$. Функцию f разложим по первым $n-s$ переменным:

$$f(x_1, \dots, x_n) =$$
$$= \bigvee_{\sigma_1 \dots \sigma_{n-s}} f(\sigma_1, \dots, \sigma_{n-s}, x_{n-s+1}, \dots, x_n) \& x_1^{\sigma_1} \& \dots \& x_{n-s}^{\sigma_{n-s}}.$$

Программу, вычисляющую функцию f, представим в следующем виде

$$\mathsf{P} = \mathsf{P}_0 \dots \mathsf{P}_j \dots \mathsf{P}_{2^{n-s}-1},$$

где $j = N(\sigma_1 \dots \sigma_{n-s})$, P_j — программа, вычисляющая функцию

$$f_j(x_{n-s+1}, \dots, x_n) = f(\sigma_1, \dots, \sigma_{n-s}, x_{n-s+1}, \dots, x_n)$$

и прекращающая работу программы P, если $x_1^{\sigma_1} \& \dots \& x_{n-s}^{\sigma_{n-s}} = 1$. Так как сложность произвольной булевой функции, зависящей от s переменных, асимптотически не превосходит $2^s/s$, то $C(\mathsf{P}_j) \lesssim 2^s/s$. Поэтому

$$T(\mathsf{P}) \sim \frac{1}{2^n} \sum_{j=0}^{2^{n-s}-1} \left(2^s \sum_{i=1}^{j} C(\mathsf{P}_i)\right) = \frac{1}{2^n} 2^s \sum_{j=0}^{2^{n-s}-1} \sum_{i=1}^{j} C(\mathsf{P}_i) \lesssim$$
$$\lesssim \frac{1}{2^n} 2^s \frac{2^s}{s} \sum_{j=0}^{2^{n-s}-1} j \lesssim \frac{1}{2^n} \frac{2^{2s}}{s} \frac{2^{2(n-s)}}{2} \sim \frac{2^{n-1}}{s} \sim \frac{2^{n-1}}{n}.$$

Теорема доказана.

Без доказательства приведем обобщение теоремы 11.3 для булевых вектор-функций с растущим числом компонент.

Теорема 11.4. *Пусть $n, m \to \infty$ и $m = n^{\mathcal{O}(1)}$. Тогда:*

(i) Для любой постоянной $\varepsilon > 0$ доля n-местных булевых вектор-функций f с m компонентами, для которых

$$T(f) \ge (1-\varepsilon)\frac{2^{n-2}m}{n},$$

стремится к единице;

(ii) для каждой n-местной булевой вектор-функций f с m компонентами

$$T(f) \lesssim \frac{2^{n-2}m}{n}.$$

11.4 Средняя vs. максимальная сложность

Определим насколько сильно могут различаться среднее время вычисления конкретной булевой функции и ее сложность, т. е. время вычисления в худшем случае. Положим

$$m(f) = L(f)\Big/T(f), \quad m(n) = \max m(f),$$

где максимум берется по всем функциям, зависящим от n переменных. Прежде всего отметим, что в силу теорем 10.2, 10.5 и 11.3 при $n \to \infty$ для почти каждой n-местной булевой функции f справедливы неравенства

$$2 \lesssim m(f) \lesssim 8,$$

т. е. для почти каждой булевой функции досрочное прекращение вычислений позволяет уменьшить объем вычислений не более чем в конечное число раз. С другой стороны, пример симметрических пороговых функций показывает, что такое уменьшение может быть более значительным. Максимально возможное значение такого уменьшения оценивается в следующей теореме, где при $n \to \infty$ устанавливается порядок роста функции $m(n)$.

Теорема 11.5. *Существуют такие постоянные c_1 и c_2, что*

$$c_1 \left(\frac{2^n}{n}\right)^{1/2} \le m(n) \le c_2 \left(\frac{2^n}{n}\right)^{1/2}.$$

Доказательство. Пусть $k = \lceil (n+\log n)/2 \rceil$ и g — самая сложная булева функция от k переменных, т. е. $L(g) = \Theta(2^n/n)^{1/2}$. Рассмотрим функцию

$$f(x_1, \ldots, x_n) = \overline{x}_{k+1} \& \cdots \& \overline{x}_n \& g(x_1, \ldots, x_k).$$

Пусть минимальная программа $\mathsf{P}(g)$ вычисляет функцию g. Тогда программа P, начинающаяся с команд

$$\begin{array}{ll} \mathsf{p}_1: & \mathtt{z} = 0 \\ \mathsf{p}_2: & \mathtt{Stop}(x_{k+1}) \\ \mathsf{p}_3: & \mathtt{Stop}(x_{k+2}) \\ \ldots\ldots & \ldots\ldots\ldots\ldots \\ \mathsf{p}_{n-k+1}: & \mathtt{Stop}(x_n), \end{array}$$

после которых идут команды программы $\mathsf{P}(g)$, вычисляет f. Легко видеть, что

$$T(\mathsf{P}) \le 2^{-n} \left(\sum_{j=1}^{n-k} (j+1) 2^{n-j} + (n-k+1+C(g)) 2^k \right) = \mathcal{O}(1),$$

и $L(f) = \Theta(2^n/n)^{1/2}$. Следовательно, $m(n) \ge c_1 (2^n/n)^{1/2}$.

Теперь покажем, что $m(n) \le c_2 (2^n/n)^{1/2}$. Пусть f — произвольная булева функция от n переменных, P — программа, которая вычисляет f, и среднее время ее работы минимально.

Положим $k = \lfloor (n + \log n)/2 \rfloor$. Рассмотрим набор $\boldsymbol{x}$ такой, что $N_{\mathsf{P}}(\boldsymbol{x}) = 2^n - 2^k$. Так как

$$T(\mathsf{P}) = 2^{-n} \sum_{\boldsymbol{y}} T_{\mathsf{P}}(\boldsymbol{y}) > 2^{-n} \sum_{\boldsymbol{y} \mid N(\boldsymbol{y}) > N(\boldsymbol{x})} T_{\mathsf{P}}(\boldsymbol{y}) \ge 2^{-n} 2^k T_{\mathsf{P}}(\boldsymbol{x}),$$

то легко видеть, что

$$2^{k-n} T_{\mathsf{P}}(\boldsymbol{x}) < T(f). \tag{11.8}$$

Далее, пусть $\tilde{f}$ — частичная булева функция, определенная на всех таких наборах $\boldsymbol{y}_i$, что $N_{\mathsf{P}}(\boldsymbol{y}_i) > N_{\mathsf{P}}(\boldsymbol{x})$, и совпадающая на этих наборах с f. Так как $2^k = \Theta(2^n n)^{1/2}$, то из теоремы 10.7 следует существование схемы S, вычисляющей $\tilde{f}$, и такой, что

$$L(\mathsf{S}) = \mathcal{O}\left(\frac{2^n}{n}\right)^{1/2}. \tag{11.9}$$

Теперь опишем программу P', вычисляющую функцию f. Сначала воспользуемся программой P, которая за минимальное среднее время вычисляет f. С ее помощью будем вычислять значения функции f на наборах $\boldsymbol{y}$ таких, что $N_{\mathsf{P}}(\boldsymbol{y}) \le N_{\mathsf{P}}(\boldsymbol{x})$. Так как $2^{n-k} = \mathcal{O}(2^n/n)^{1/2}$, то из (11.8) следует, что для вычисления функции f на этих наборах потребуется выполнить не более H_1 команд программы P, где

$$H_1 = \mathcal{O}\left(\left(\frac{2^n}{n}\right)^{1/2} T(f)\right). \tag{11.10}$$

Для вычисления функции f на оставшихся наборах воспользуемся неветвящейся программой—схемой, реализующей функцию $\tilde{f}$. Очевидно, что эта программа S содержит не более H_2 команд, где

$$H_2 = \mathcal{O}\left(\left(\frac{2^n}{n}\right)^{1/2}\right), \tag{11.11}$$

что следует из (11.9).

Таким образом, из (11.10) и (11.11) следует, что сложность $C(\mathsf{P}')$ программы P' по порядку не превосходит величины

$$\left(\frac{2^n}{n}\right)^{1/2} T(f) + \left(\frac{2^n}{n}\right)^{1/2} \le 2 \left(\frac{2^n}{n}\right)^{1/2} T(f).$$

Так как $L(f) = \mathcal{O}(C(\mathsf{P}'))$ (это неравенство является простым следствием равенства (11.2)), то найдется константа c_2 такая, что

$$L(f)\Big/T(f) \le c_2 \left(\frac{2^n}{n}\right)^{1/2}.$$

Теорема доказана.

Теперь покажем, что для каждой булевой функции f существует программа, сложность и среднее время работы которой одновременно близки, соответственно, к сложности в худшем случае и к средней сложности этой функции.

Теорема 11.6. *Для любой булевой функции f найдется такая вычисляющая ее программа* P*, что*

$$T(\mathsf{P}) \le 2T(f), \quad C(\mathsf{P}) \le 4C(f).$$

Доказательство. Пусть P — программа, вычисляющая функцию f за минимальное среднее время. Пусть $\boldsymbol{x}$ — набор с минимальным номером такой, что $T_{\mathsf{P}}(\boldsymbol{x}) \ge 2C(f)$ (если такого набора нет, то утверждение теоремы тривиально, так как тогда $T(f) = T(\mathsf{P}) \le C(\mathsf{P}) \le 2C(f)$). Ясно, также что $T_{\mathsf{P}}(\boldsymbol{x}) \le 3C(f)$, поскольку невыполнение этого неравенства влечет существование в программе P расположенных друг за другом $C(f)+1$ вычислительных команд, что противоречит минимальности программы P. Среднее время работы программы P можно представить следующим образом:

$$\begin{aligned} T(\mathsf{P}) &= 2^{-n}\Big(\sum_{N_{\mathsf{P}}(\boldsymbol{y})<N_{\mathsf{P}}(\boldsymbol{x})} T_{\mathsf{P}}(\boldsymbol{y}) + \sum_{N_{\mathsf{P}}(\boldsymbol{y})\ge N_{\mathsf{P}}(\boldsymbol{x})} T_{\mathsf{P}}(\boldsymbol{y})\Big) = \\ &= T_1 + T_2 \ge T_1 + 2C(f)(2^n - N_{\mathsf{P}}(\boldsymbol{x}) + 1)2^{-n}. \end{aligned}$$

Преобразуем программу P, заменив команды с номерами большими $T_{\mathsf{P}}(\boldsymbol{x})$ программой без команд остановки — минимальной схемой из функциональных элементов вычисляющей f. Легко видеть, для сложности новой программы P' справедливо неравенство

$$C(\mathsf{P}') \le T_{\mathsf{P}}(\boldsymbol{x}) + C(f) \le 4C(f),$$

а для среднего времени работы этой программы — неравенство:

$$\begin{aligned} T(\mathsf{P}') &\le 2^{-n}\Big(\sum_{N_{\mathsf{P}}(\boldsymbol{y})<N_{\mathsf{P}}(\boldsymbol{x})} T_{\mathsf{P}'}(\boldsymbol{y}) + \sum_{N_{\mathsf{P}}(\boldsymbol{y})\ge N_{\mathsf{P}}(\boldsymbol{x})} T_{\mathsf{P}'}(\boldsymbol{y})\Big) = \\ &= T_1' + T_2' \le T_1' + 4C(f)(2^n - N_{\mathsf{P}}(\boldsymbol{x}) + 1)2^{-n} \le \\ &\le 2(T_1' + 2C(f)(2^n - N_{\mathsf{P}}(\boldsymbol{x}) + 1)2^{-n}). \end{aligned}$$

Так как $T_1' = T_1$, то $T(\mathsf{P}') \le 2T(\mathsf{P})$. Теорема доказана.

11.5 Задачи

11.1. Показать, что средняя сложность системы, состоящей из n-местных дизъюнкции и конъюнкции, зависящих от одних и тех же переменных, не превосходит пяти.

11.2. Показать, что в любой приведенной программе с n входами число команд остановки не превосходит $\frac{1}{2}(L+n)$, где L — сложность программы.

11.3. Показать, что программа P, вычисляющая булеву функцию f, может быть преобразована в вычисляющую f схему из функциональных элементов S так, что $L(\mathsf{S}) \le 2C(\mathsf{P})$.

11.4. Найти все возможные значения, которые может принимать средняя сложность функций трех переменных.

11.5. Найти все такие функции $f(x,y,z)$, существенно зависящие от всех своих переменных, для которых $T(f) = L(f)$.

11.6. Пусть $\mu(f)$ равно максимальному числу последовательных слоев, на которых симметрическая n-местная функция f принимает одинаковые значения. Показать, что $T(f) \asymp n - \mu(f) + 2$.

11.7. Пусть $n \to \infty$. Показать, что для всех $\sigma_1, \ldots, \sigma_n$ справедливы асимптотические равенства:
a) $T(x_1^{(\sigma_1)} \vee \ldots \vee x_n^{(\sigma_n)}) \sim 8/3$, b) $T(x_1^{(\sigma_1)} \& \ldots \& x_n^{(\sigma_n)}) \sim 11/3$.

11.8. Показать, что $T(x_1 \oplus x_2 \oplus \cdots \oplus x_n) = n - 1$.

11.9. Показать, что при $n \to \infty$ найдется такой линейный булев (n,n)-оператор f, что $T(f) = \Theta(\frac{n^2}{\log_2 n})$.

11.10. Показать, что при $n \to \infty$ найдется такая квадратичная n-местная булева функция f, что $T(f) = \Theta(\frac{n^2}{\log_2 n})$.

11.11. Показать, что при $n \to \infty$ неравенство $T(f) = \mathcal{O}\left(\frac{\|f\|}{\log_2 \|f\|}\right)$ справедливо для любой n-местной булевой функции f такой, что $n^3 \le \|f\| \le 2^{n-1}$.

11.12. Показать, что при $n \to \infty$ найдется такая n-местная булева функция f, что $n^3 \le \|f\| \le 2^{n-1}$ и $T(f) = \Theta\left(\frac{\|f\|}{\log_2 \|f\|}\right)$.

11.13. Оценить $\max L(f)/T(f)$, где максимум берется по всем n-местным функциям одинаковой сложности L.

11.14. Доказать теорему 11.4.

Лекция 12

Алфавитное кодирование

Множества $A = \{a_1, \ldots, a_n\}$ и $B = \{b_1, \ldots, b_m\}$ будем называть алфавитами, элементы этих множеств — буквами, а элементы из A^* и B^* — словами над алфавитами A и B, соответственно. Отображение $\varphi : A \to B^*$ будем называть *побуквенным кодированием* алфавита A словами над алфавитом B. Отображение φ естественным образом продолжается до отображения $\varphi^* : A^* \to B^*$ так, что $\varphi^*(a_{i_1} \ldots a_{i_k}) = \varphi(a_{i_1}) \ldots \varphi(a_{i_k})$. Множество $V = \{v_1, \ldots, v_n\}$, где $v_i = \varphi(a_i)$, будем называть m-ичным *кодом алфавита* A, порожденным отображением φ. Будем полагать, что каждой букве a_i приписана вероятность $p(a_i) = p_i$ так, что $p_1 + \cdots + p_n = 1$, т. е. на алфавите A задано распределение вероятностей $P = \{p_1, \ldots, p_n\}$. *Стоимостью кода* $V = \{v_1, \ldots, v_n\}$ при распределении вероятностей $P = \{p_1, \ldots, p_n\}$ назовем величину

$$C(V, P) = \sum_{i=1}^{n} p_i l(v_i),$$

где $l(v_i)$ — число букв в слове v_i. Стоимость $C(V, P)$ можно рассматривать как среднее число символов алфавита B, приходящихся в коде V на одну букву алфавита A. Ниже рассматриваются вопросы построения кодов, имеющих при данном распределении вероятностей минимальную стоимость. Далее везде будем рассматривать только двоичные коды, т. е. в качестве алфавита B будем использовать двоичный алфавит $\{0, 1\}$.

12.1 Разделимые и префиксные коды

Порожденный отображением φ код V назовем *разделимым*[1)], если отображение φ^* инъективно. Для каждого разделимого кода V определено обратное к φ^* отображение $(\varphi^*)^{-1} : \varphi^*(A^*) \to A$, которое будем называть декодированием кода V.

В следующей теореме устанавливается необходимое условие разделимости кода. Неравенство (12.1) называется неравенством Крафта–Макмиллана.

[1)] Разделимые коды также называются однозначно декодируемыми.

Теорема 12.1. *Пусть $V = \{v_1, \ldots, v_n\}$, где $n \geq 2$, — разделимый код. Тогда*

$$\sum_{i=1}^{n} 2^{-l(v_i)} \leq 1. \tag{12.1}$$

Доказательство. Обозначим через D множество двоичных наборов, каждый из которых является последовательностью слов из V. Пусть D_k равно числу двоичных наборов длины k в D, а w — весовая функция на D, равная длине набора. Нетрудно видеть, что

$$F_D^w(x) = \sum_{k=0}^{\infty} D_k x^k = \sum_{k=0}^{\infty} \Big(\sum_{v_i \in V} x^{l(v_i)} \Big)^k = \frac{1}{1 - \sum_{v_i \in V} x^{l(v_i)}}.$$

Так как код V разделимый, то любой двоичный набор разлагается в последовательность слов из V не более чем одним способом. Поэтому $D_k \leq 2^k$ для любого k, и, следовательно, для любой постоянной ε из полуинтервала $(0, \frac{1}{2}]$ и для любого $x \in [0, \frac{1}{2} - \varepsilon]$ имеет место неравенство

$$\frac{1}{1 - \sum_{v_i \in V} x^{l(v_i)}} \leq \frac{1}{1 - 2x},$$

где $\frac{1}{1-2x}$ — производящая функция множества всех двоичных наборов конечной длины. Последнее неравенство легко преобразуется в неравенство

$$\sum_{v_i \in V} x^{l(v_i)} \leq 2x,$$

которое, в свою очередь, при $x = \frac{1}{2} - \varepsilon$ превращается в неравенство

$$\sum_{v_i \in V} 2^{-l(v_i)} (1 - 2\varepsilon)^{l(v_i)} \leq 1 - 2\varepsilon. \tag{12.2}$$

Переходя в (12.2) к пределу при $\varepsilon \to 0$, получаем неравенство теоремы. Теорема доказана.

Код $V = \{v_1, \ldots, v_n\}$ называется *префиксным*, если ни одно его слово не является началом другого. Префиксные коды удобно представлять при помощи плоских корневых бинарных деревьев, в которых листья соответствуют кодовым словам, а пути от корня к листьям задают эти слова. При этом

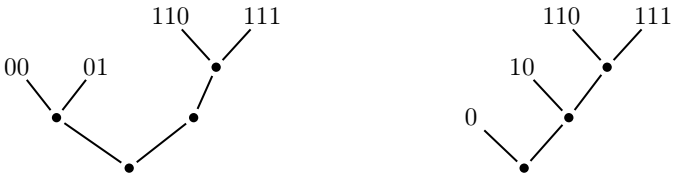


Рис. 12.1

ребро, направленное вверх налево, соответствует нулю, а ребро, направленное вверх направо, соответствует единице. На рис. 12.1 изображены два дерева, представляющие префиксные коды $\{00, 01, 110, 111\}$ и $\{0, 10, 110, 111\}$. Префиксный код называется полным, если в соответствующем ему дереве нет отличных от корня внутренних вершин степени два. В правом дереве на рис. 12.1 таких вершин нет, поэтому код $\{0, 10, 110, 111\}$ полный в отличие от кода $\{00, 01, 110, 111\}$.

Представляя префиксные коды при помощи деревьев, нетрудно убедиться в том, что любой префиксный код разделим. Более того, используя представляющее префиксный код $V = \{v_1, \ldots, v_n\}$ дерево T_V, можно быстро и легко выполнить декодирование, т. е. по последовательности $v_{i_1} \ldots v_{i_k}$ определить последовательность $a_{i_1} \ldots a_{i_k}$. Для декодирования последовательности $v_{i_1} \ldots v_{i_k}$ достаточно посмотреть на эту последовательность как на последовательность описаний путей, начинающихся в корне дерева T_V.

Теорема 12.2. *Пусть для целых чисел $l_1 \le \cdots \le l_n$ справедливо неравенство*

$$\sum_{i=1}^{n} 2^{-l_i} \le 1.$$

Тогда найдется такой префиксный код $V = \{v_1, \ldots, v_n\}$, что $l(v_i) = l_i$ для каждого $i \in \{1, \ldots, n\}$.

Доказательство. Положим $q_1 = 0$ и $q_m = \sum_{i=1}^{m-1} 2^{-l_i}$ для $m = 2, \ldots, n$. Из теоремы 12.1 следует, что каждое из чисел q_i меньше единицы. Из первых l_i знаков после запятой числа q_i составим слово v_i. Пусть $j > i$. Тогда

$$q_j - q_i = \sum_{m=i}^{j-1} 2^{-l_m} \ge 2^{-l_i}.$$

Следовательно числа q_j и q_i отличаются по крайней мере в одном из первых l_i знаков после запятой, и так как каждое из слов v_i и v_j состоит не менее чем из l_i знаков, то ни одно из этих слов не является началом другого. Таким образом $V = \{v_1, \ldots, v_n\}$ — префиксный код, в котором $l(v_i) = l_i$ для каждого $i \in \{1, \ldots, n\}$. Теорема доказана.

Теорема 12.2 показывает, что неравенство Крафта–Макмиллана является не только необходимым, но и достаточным условием существования разделимого кода с заданными длинами слов.

Из теорем 12.2 и 12.1 следует простая, но важная теорема, которую приведем без доказательства.

Теорема 12.3. *Для любого разделимого кода $W = \{w_1, \ldots, w_n\}$ найдется такой префиксный код $V = \{v_1, \ldots, v_n\}$, что $l(w_i) = l(v_i)$ для каждого $i \in \{1, \ldots, n\}$.*

Теорема 12.3 показывает, что при изучении стоимости разделимых кодов можно рассматривать только префиксные коды, которые в отличие от произвольных разделимых кодов имеют простую процедуру декодирования.

12.2 Оптимальные коды

Код $V = \{v_1, \ldots, v_n\}$ называется *оптимальным* для распределения вероятностей $P = \{p_1, \ldots, p_n\}$, если для данного P его стоимость $C(V, P)$ минимальна среди всех n-элементных кодов. Существование оптимального кода легко следует из теоремы 12.3 и того факта, что при поиске оптимального кода можно ограничиться только полными префиксными кодами, максимальная длина слов в которых не превосходит $n - 1$. Далее через $C(P)$ будем обозначать стоимость оптимального кода для распределения P.

Легко видеть, что в оптимальном коде $l(v_i) \ge l(v_j)$ если $p(a_i) < p(a_j)$. Также заметим, что в оптимальном префиксном коде всегда найдется по крайней мере два слова максимальной длины, причем эти слова будут отличаться только в последнем (правом) символе. Допустим, что это не так, и найдется такой код $V = \{v_1, \ldots, v_n\}$, что слово v, отличающееся от максимально длинного слова v_n только последним символом, не принадлежит V. Тогда, удалив последний символ в слове v_n, получим новый префиксный код $V' = \{v_1, \ldots, v_{n-1}, v'_n\}$, стоимость которого меньше стоимости кода V.

Отмеченные свойства оптимальных префиксных кодов позволяют доказать следующую теорему.

Теорема 12.4. *Пусть $V = \{v_1, \ldots, v_n\}$ — оптимальный префиксный код для распределения $P = \{p_1, \ldots, p_n\}$, где $p_1 \ge \cdots \ge p_n$, и пусть q_0, q_1 — такие положительные, что $q_0 + q_1 = p_i$ для некоторого $i \in \{1, \ldots, n\}$ и $p_n \ge q_0, q_1$. Тогда префиксный код*

$$V' = \{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n, v_i 0, v_i 1\}$$

будет оптимальным для распределения

$$P' = \{p_1, \ldots, p_{i-1}, p_{i+1}, \ldots, p_n, q_0, q_1\}.$$

Доказательство. Допустим, что код V' не является оптимальным для распределения P'. Тогда существует код $W' = \{w_1, \ldots, w_{i-1}, w_{i+1}, \ldots, w_{n+2}\}$ такой, что $l(w_1) \le \cdots \le l(w_{n+1}) = l(w_{n+2})$ и $C(W', P') < C(V', P')$. Учитывая отмеченное выше свойство оптимальных префиксных кодов, без ограничения общности будем считать, что в коде W' вероятностям q_0 и q_1 соответствуют слова w_{n+1} и w_{n+2}, и эти слова отличаются только в последнем символе, т. е. $w_{n+1} = w0$ и $w_{n+2} = w1$. Пусть $W = \{w_1, \ldots, w_i, \ldots, w_n\}$, где $w_i = w$. Тогда

$$C(V', P') - C(V, P) = l(v_i 0)q_0 + l(v_i 1)q_1 - l(v_i)p_i = p_i,$$
$$C(W', P') - C(W, P) = l(w0)q_0 + l(w1)q_1 - l(w)p_i = p_i.$$

Следовательно, в силу сделанного предположения,

$$C(V, P) - C(W, P) = C(V', P') - C(W', P') > 0.$$

Пришли к противоречию с оптимальностью кода V. Теорема доказана.

Формулировка теоремы 12.4 фактически заключает в себе простой индуктивный алгоритм построения оптимального префиксного кода для произвольного распределения вероятностей $P_n = \{p_1, \ldots, p_n\}$. Алгоритм состоит в следующем. Сначала исходя из распределения P_n строится последовательность распределений $P_n, P_{n-1}, \ldots, P_2$, в которой распределение P_{i-1} получается из P_i заменой двух самых маленьких вероятностей на одну, равную их сумме. Префиксный код для распределения P_2 строится тривиально — таким кодом всегда будет код $V_2 = \{0, 1\}$. Затем исходя из кода V_2 при помощи теоремы 12.4 строится последовательность оптимальных кодов $V_2, V_3, \ldots, V_n$. Построенный при помощи этого алгоритма код V_n называется

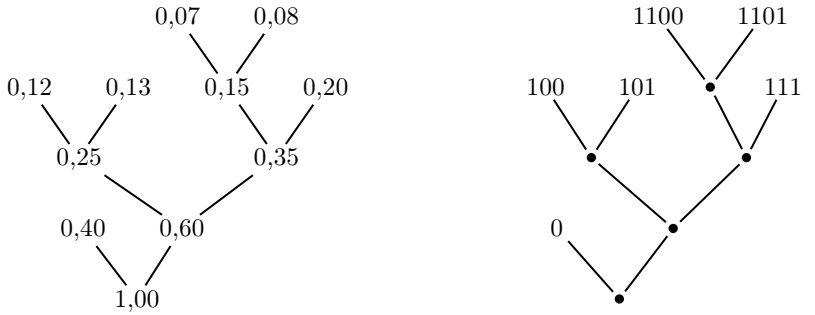


Рис. 12.2

кодом Хаффмана.

Процесс построения кода Хаффмана для распределения вероятностей $\{0,07;\ 0,08;\ 0,12;\ 0,13;\ 0,20;\ 0,40\}$ представлен на рис. 12.2.

12.3 Стоимость кодирования

Метод Хаффмана позволяет строить оптимальные коды для любого распределения вероятностей, но к сожалению не дает явной информации о стоимости построенного кода. Поэтому для оценки величины $C(P)$ приходится привлекать дополнительные соображения.

Пусть $P = \{p_1, \ldots, p_n\}$ — распределение вероятностей. Функция[2)]

$$H(p_1, \ldots, p_n) = -\sum_{i=1}^{n} p_i \log_2 p_i \tag{12.3}$$

называется энтропией распределения P. Нетрудно видеть, что использованная в лекции 3 для оценки биномиальных коэффициентов функция $H(x) = -x\log_2 x - (1-x)\log_2(1-x)$ является частным случаем функции (12.3). В следующей теореме, доказанной Шенноном в середине 40-х годов 20 века, энтропия используется для оценки стоимости алфавитного кодирования.

[2)] Здесь и далее полагаем, что функция $x \log_2 x$ равна нулю при $x = 0$.

Теорема 12.5. *Для любого распределения вероятностей* $\{p_1, \ldots, p_n\}$

$$H(p_1, \ldots, p_n) \le C(p_1, \ldots, p_n) \le H(p_1, \ldots, p_n) + 1. \tag{12.4}$$

Доказательство. Нижняя оценка. Покажем, что левое неравенство (12.4) справедливо для кода Хаффмана. Сделаем это индукцией по n. В основание индукции положим случай $n = 2$. Прежде всего заметим, что функция $-x\log_2 x$ выпукла вверх при $x \in (0, 1)$. Поэтому

$$\frac{1}{2}(-x_1 \log_2 x_1 - x_2 \log_2 x_2) \le -\frac{x_1 + x_2}{2} \log_2 \frac{x_1 + x_2}{2}$$

для любых $x_1, x_2 \in (0, 1)$. Следовательно,

$$H(x, 1-x) = 2(-x\log_2 x - (1-x)\log_2(1-x)) \le -2 \cdot \frac{1}{2}\log_2 \frac{1}{2} = 1$$

для любого $x \in (0, 1)$. Таким образом при $n = 2$ для любого распределения вероятностей $\{p_1, p_2\}$ имеет место неравенство $H(p_1, p_2) \le 1$. С другой стороны, очевидно, что при $n = 2$ независимо от распределения вероятностей стоимость кодирования всегда равна единице. Следовательно, $C(p_1, p_2) \ge H(p_1, p_2)$.

Теперь допустим, что нижняя оценка теоремы справедлива для любого n не превосходящего $k - 1$, где $k \ge 3$. Установим ее справедливость для $n = k$. Пусть $P = \{p_1, \ldots, p_k\}$ — произвольное распределение вероятностей, в котором вероятности p_i упорядочены по убыванию. Рассмотрим новое распределение $P' = \{p_1, \ldots, p_{k-2}, p'_{k-1}\}$, где $p'_{k-1} = p_{k-1} + p_k$. Распределение P' возникает на первом шаге алгоритма Хаффмана, и поэтому очевидно, что

$$C(P) - C(P') = p_{k-1}l_{k-1} + p_k l_k - p'_{k-1}l'_{k-1} = p_{k-1} + p_k.$$

Также легко видеть, что

$$H(P) - H(P') = -p_{k-1}\log_2 p_{k-1} - p_k \log_2 p_k + p'_{k-1}\log_2 p'_{k-1}.$$

По предположению индукции $C(P') - H(P') \ge 0$. Тогда

$$\begin{aligned} C(P) - H(P) &\ge (C(P) - H(P)) - (C(P') - H(P')) = \\ &= (C(P) - C(P')) - (H(P) - H(P')). \end{aligned}$$

Следовательно,

$$\begin{aligned} C(P) - H(P) &\ge p_{k-1} + p_k + p_{k-1}\log_2 p_{k-1} + p_k \log_2 p_k - p'_{k-1}\log_2 p'_{k-1} = \\ &= \frac{1}{2}\Big(2p_{k-1}\log_2 2p_{k-1} + 2p_k \log_2 2p_k\Big) - p'_{k-1}\log_2 p'_{k-1} \ge \\ &\ge \frac{2p_{k-1} + 2p_k}{2}\log_2 \frac{2p_{k-1} + 2p_k}{2} - p'_{k-1}\log_2 p'_{k-1} = 0. \end{aligned}$$

Нижняя оценка доказана.

Верхняя оценка. Пусть $P = \{p_1, \ldots, p_n\}$ — произвольное распределение вероятностей, в котором вероятности p_i упорядочены по убыванию. Без ограничения общности будем считать, что все вероятности положительны. Так как

$$\sum_{i=1}^{n} 2^{-(-\lfloor \log_2 p_i \rfloor)} \le \sum_{i=1}^{n} p_i = 1,$$

то в силу теоремы 12.2 существует префиксный код $V = \{v_1, \ldots, v_n\}$, в котором $l(v_i) = -\lfloor \log_2 p_i \rfloor$ для каждого $i \in \{1, \ldots, n\}$. Такой код называется *кодом Шеннона*, и нетрудно видеть, что для его стоимости справедливо неравенство

$$\begin{aligned} C(p_1, \ldots, p_n) &= \sum_{i=1}^{n} p_i(-\lfloor \log_2 p_i \rfloor) \le \\ &\le \sum_{i=1}^{n} p_i(-\log_2 p_i + 1) = H(p_1, \ldots, p_n) + 1. \end{aligned}$$

Теорема доказана.

12.4 Блочное кодирование

Пусть $P = \{p_1, \ldots, p_n\}$ — распределение вероятностей на алфавите A. Далее будем рассматривать кодирование блоками — будем кодировать слова длины k, составленные из букв алфавита A, полагая, что вероятности всех букв в каждом слове независимы, т. е. что равенство

$$p\,(a_{i_1} \cdots a_{i_k}) = p\,(a_{i_1}) \cdots p\,(a_{i_k}) = p_{i_1} \cdots p_{i_k}$$

справедливо для любого слова $a_{i_1} \cdots a_{i_k}$ из A^k. Распределение вероятностей на множестве A^k будем обозначать через P^k. Пусть V_k — код множества A^k. Величину

$$C_k(V_k, P) = \frac{1}{k} C(V_k, P^k)$$

назовем средней стоимостью буквы алфавита A для распределения P при использовании кода V_k, а величину

$$C_k(P) = \min_{V_k} C_k(V_k, P) = \frac{1}{k} C(P^k)$$

— средней стоимостью буквы для распределения P при кодировании блоками длины k. Ниже в теореме 12.6 устанавливаются оценки Шеннона для величины C_k. Для доказательства этой теоремы потребуется имеющая самостоятельный интерес лемма об аддитивности энтропии независимых распределений.

Лемма 12.1. *Пусть $P = \{p_1, \ldots, p_n\}$ и $Q = \{q_1, \ldots, q_m\}$ — независимые распределения вероятностей на алфавитах $A = \{a_1, \ldots, a_n\}$ и $B = \{b_1, \ldots, b_m\}$. Тогда для распределения вероятностей PQ на алфавите $AB = \{a_i b_j\}$ справедливо равенство*

$$H(PQ) = H(P) + H(Q).$$

Доказательство.

$$\begin{aligned} H(PQ) &= -\sum_{ij} p_i q_j \log_2 p_i q_j = -\sum_i \sum_j p_i q_j (\log_2 p_i + \log_2 q_j) = \\ &= -\sum_i \Big(p_i \log_2 p_i \sum_j q_j \Big) - \sum_i \Big(p_i \sum_j q_j \log_2 q_j \Big) = \\ &= -\Big(\sum_i p_i \log_2 p_i \Big) \Big(\sum_j q_j \Big) - \Big(\sum_i p_i \Big) \Big(\sum_j q_j \log_2 q_j \Big) = H(P) + H(Q). \end{aligned}$$

Лемма доказана.

Из леммы 12.1 легко следует, что

$$H(P^k) = kH(P) \tag{12.5}$$

для любого распределения P и любого натурального k.

Теорема 12.6. *Для любого распределения вероятностей $P = \{p_1, \ldots, p_n\}$*

$$H(p_1, \ldots, p_n) \le C_k(p_1, \ldots, p_n) \le H(p_1, \ldots, p_n) + \frac{1}{k}.$$

Доказательство. Из равенства (12.5) и теоремы 12.5 следует, что

$$kH(P) = H(P^k) \le C(P^k) \le H(P^k) + 1 = kH(P) + 1. \tag{12.6}$$

Так как $C(P^k) = kC_k(P)$, то разделив все члены в (12.6) на k, получим требуемые неравенства для $C_k(P)$. Теорема доказана.

Из теоремы 12.6 следует, что выбирая длину блока достаточно большой, среднюю стоимость буквы можно сделать сколь угодно близкой к энтропии распределения P.

12.5 Универсальное блочное кодирование

Доказательство теоремы 12.6 существенным образом опирается на знание распределения вероятностей P. Б. М. Фитингоф в 1966 году предложил независящий от распределения вероятностей P универсальный метод кодирования блоками и показал, что для любого распределения P среднюю стоимость буквы при таком кодировании можно сделать сколь угодно близкой к энтропии этого неизвестного распределения P. *Универсальный метод кодирования Фитингофа* описывается ниже в теореме 12.7. Для оценки стоимости этого метода потребуются два вспомогательных утверждения о свойствах энтропии, доказываемые ниже в леммах 12.2 и 12.3.

Лемма 12.2. *Пусть $m_1+m_2+\cdots+m_n=m$ и все m_i — целые неотрицательные. Тогда*

$$\frac{m!}{m_1!\cdots m_n!}\le 2^{mH(\frac{m_1}{m},\ldots,\frac{m_n}{m})}.$$

Доказательство. Неравенство леммы докажем индукцией по n. При этом без ограничения общности будем считать, что все m_i положительны. В основание индукции положим приведенное на стр. 32 неравенство (2.13), которое является частным случаем доказываемого неравенства при $n=2$. Пусть $n>2$. Тогда по предположению индукции

$$\frac{m!}{m_1!\cdots m_n!}=\frac{m!}{m_1!\cdots m_{n-2}!(m_{n-1}+m_n)!}\cdot\frac{(m_{n-1}+m_n)!}{m_{n-1}!m_n!}\le$$
$$\le 2^{mH(\frac{m_1}{m},\ldots,\frac{m_{n-2}}{m},\frac{m_{n-1}+m_n}{m})}\cdot 2^{(m_{n-1}+m_n)H(\frac{m_{n-1}}{m_{n-1}+m_n},\frac{m_n}{m_{n-1}+m_n})}.$$

Для доказательства леммы достаточно показать, что

$$mH\Big(\frac{m_1}{m},\ldots,\frac{m_n}{m}\Big)=mH\Big(\frac{m_1}{m},\ldots,\frac{m_{n-2}}{m},\frac{m_{n-1}+m_n}{m}\Big)+$$
$$+(m_{n-1}+m_n)H\Big(\frac{m_{n-1}}{m_{n-1}+m_n},\frac{m_n}{m_{n-1}+m_n}\Big) \qquad (12.7)$$

Нетрудно видеть, что разность правой и левой частей равенства (12.7) равна

$$-m_{n-1}\log_2\frac{m_{n-1}}{m}-m_n\log_2\frac{m_n}{m}+(m_{n-1}+m_n)\log_2\frac{m_{n-1}+m_n}{m}+$$
$$+m_{n-1}\log_2\frac{m_{n-1}}{m_{n-1}+m_n}+m_n\log_2\frac{m_n}{m_{n-1}+m_n}.$$

Приводя в последней формуле подобные члены, видим, что

$$-m_{n-1}\log_2\Big(\frac{m_{n-1}}{m}\cdot\frac{m}{m_{n-1}+m_n}\cdot\frac{m_{n-1}+m_n}{m_{n-1}}\Big)-$$
$$-m_n\log_2\Big(\frac{m_n}{m}\cdot\frac{m}{m_{n-1}+m_n}\cdot\frac{m_{n-1}+m_n}{m_n}\Big)=0.$$

Лемма доказана.

Лемма 12.3. *Пусть $p_1+p_2+\cdots+p_n=1$ и все $p_i\ge 0$. Тогда для любой выпуклой вверх функции f и любых $x_1,x_2,\ldots,x_n$*

$$p_1f(x_1)+p_2f(x_2)+\cdots+p_nf(x_n)\le f(p_1x_1+p_2x_2+\cdots+p_nx_n). \qquad (12.8)$$

Доказательство. Без ограничения общности будем считать, что все p_i положительные. На плоскости xy рассмотрим график функции $y=f(x)$ и лежащие на этом графике точки M_i с координатами $(x_i,f(x_i))$. Так как функция f выпукла вверх, то многоугольник с вершинами в точках M_i лежит ниже графика функции f так, как это изображено на рис. 12.3. Покажем, что точка N с координатами (u,v), где $u=p_1x_1+p_2x_2+\cdots+p_nx_n$

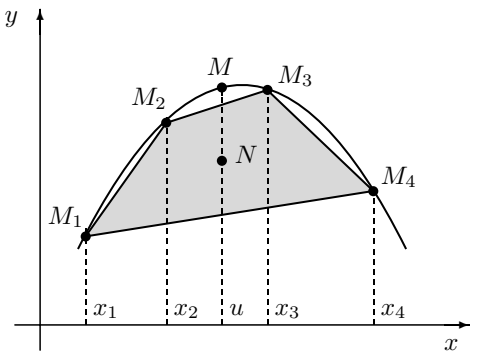


Рис. 12.3

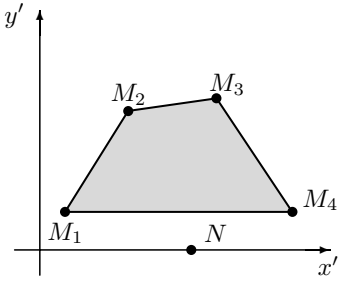


Рис. 12.4

и $v=p_1f(x_1)+p_2f(x_2)+\cdots+p_nf(x_n)$, лежит внутри многоугольника. Допустим, что это не так. Тогда на плоскости xy введем новые прямоугольные координаты x' и y' так, что бы ось x' была параллельна ближайшей стороне многоугольника и проходила через точку N. Подобная ситуация изображена на рис. 12.4. Тогда, с одной стороны, координата y' точки N равна нулю, а с другой стороны, эта координата больше нуля, так как она является суммой положительных вторых координат точек M_i, умноженных на положительные числа p_i. Полученное противоречие доказывает, что N лежит внутри многоугольника с вершинами в точках M_i.

Теперь осталось заметить, что точка M (см. рис. 12.3), вторая координата которой равна $f(p_1x_1+p_2x_2+\cdots+p_nx_n)$, лежит выше точки N, вторая координата которой равна $p_1f(x_1)+p_2f(x_2)+\cdots+p_nf(x_n)$. Лемма доказана.

Так как при $x\ge 0$ функция $-x\log_2 x$ выпукла вверх, то для любых неотрицательных $p_1,\ldots,p_n$ таких, что $p_1+p_2+\cdots+p_n=1$, и любых неотрицательных $x_1,x_2,\ldots,x_n$

$$-\sum_{i=1}^n p_ix_i\log_2 x_i\le-\Big(\sum_{i=1}^n p_ix_i\Big)\log_2\Big(\sum_{i=1}^n p_ix_i\Big). \qquad (12.9)$$

Теорема 12.7. *Для любого k существует такой универсальный код V_k, что неравенство*

$$C_k(V_k,P)\le H(P)+\frac{(n-1)\log_2(k+1)+2}{k}.$$

справедливо для любого распределения вероятностей P на множестве букв алфавита $A=\{a_1,\ldots,a_n\}$.

Доказательство. Через $A(m_1,\ldots,m_n)$, где $m_1+\cdots+m_n=k$, обозначим подмножество множества A^k, в котором каждое слово содержит ровно m_i букв a_i. Число различных подмножеств $A(m_1,\ldots,m_n)$ равно числу упорядоченных разбиений числа k не более чем на n слагаемых, т. е. $\binom{k+n-1}{n-1}$. Подмножества $A(m_1,\ldots,m_n)$ перенумеруем числами от единицы

до $\binom{k+n-1}{n-1}$, а слова в каждом подмножестве $A(m_1,\dots,m_n)$ — числами от единицы до $\frac{k!}{m_1!\cdots m_n!}$. В результате каждое k-буквенное слово из A^k будет однозначно определятся по двум номерам — номеру подмножества, которому оно принадлежит, и своему номеру в этом подмножестве. Слову $\boldsymbol{a}$ из $A(m_1,\dots,m_n)$ поставим в соответствие двоичное слово $v(\boldsymbol{a})$ длины

$$\Big\lceil \log_2 \binom{k+n-1}{n-1} \Big\rceil + \Big\lceil \log_2 \frac{k!}{m_1!\cdots m_n!} \Big\rceil,$$

составленное из номера $n(m_1,\dots,m_n)$ множества $A(m_1,\dots,m_n)$ и приписанному к нему справа номеру $n(\boldsymbol{a})$ слова $\boldsymbol{a}$. Первый их этих номеров является двоичным $\lceil \log_2 \binom{k+n-1}{n-1} \rceil$-разрядным числом, а второй — двоичным $\lceil \log_2 \frac{k!}{m_1!\cdots m_n!} \rceil$-разрядным числом[3)].

Нетрудно видеть, что составленный из слов $v(\boldsymbol{a})$ код V_k будет префиксным, а для его стоимости будет справедливо равенство

$$C(V_k,P^k) = \sum_{\substack{m_1,\dots,m_n \\ v\in A(m_1,\dots,m_n)}} p_1^{m_1}\cdots p_n^{m_n}\Big(\Big\lceil \log_2 \binom{k+n-1}{n-1}\Big\rceil + \Big\lceil \log_2 \frac{k!}{m_1!\cdots m_n!}\Big\rceil\Big),$$

правую часть которого очевидно можно оценить следующим образом:

$$\sum_{\substack{m_1,\dots,m_n \\ v\in A(m_1,\dots,m_n)}} p_1^{m_1}\cdots p_n^{m_n}\Big(2+\log_2\binom{k+n-1}{n-1} + \log_2\frac{k!}{m_1!\cdots m_n!}\Big) \le$$
$$\le 2+\log_2\binom{k+n-1}{n-1} + \sum_{\substack{m_1,\dots,m_n \\ v\in A(m_1,\dots,m_n)}} p_1^{m_1}\cdots p_n^{m_n}\log_2\frac{k!}{m_1!\cdots m_n!}. \quad (12.10)$$

Оценим разность R между последним слагаемым в правой части (12.10) и величиной $H(P^k)$. Нетрудно видеть, что

$$R = \sum_{\substack{m_1,\dots,m_n \\ v\in A(m_1,\dots,m_n)}} p_1^{m_1}\cdots p_n^{m_n}\log_2\frac{k!}{m_1!\cdots m_n!} -$$
$$-\Big(-\sum_{\substack{m_1,\dots,m_n \\ v\in A(m_1,\dots,m_n)}} p_1^{m_1}\cdots p_n^{m_n}\log_2(p_1^{m_1}\cdots p_n^{m_n})\Big) =$$
$$= \sum_{\substack{m_1,\dots,m_n \\ v\in A(m_1,\dots,m_n)}} p_1^{m_1}\cdots p_n^{m_n}\Big(\log_2\frac{k!}{m_1!\cdots m_n!} + \log_2(p_1^{m_1}\cdots p_n^{m_n})\Big).$$

Покажем что каждая сумма, стоящая в скобках в последнем равенстве, не больше нуля для всех распределений P. Так как в силу леммы 12.2

$$\log_2\frac{k!}{m_1!\cdots m_n!} \le kH(\frac{m_1}{k},\dots,\frac{m_n}{k}) = -k\sum_{i=1}^n \frac{m_i}{k}\log_2\frac{m_i}{k},$$

[3)] Старшие разряды чисел $n(m_1,\dots,m_n)$ и $n(\boldsymbol{a})$ могут быть нулями.

то применяя (12.9) видим, что

$$\log_2\frac{k!}{m_1!\cdots m_n!} + \log_2(p_1^{m_1}\cdots p_n^{m_n}) \le -k\sum_{i=1}^n\frac{m_i}{k}\log_2\frac{m_i}{k} + \sum_{i=1}^n m_i\log_2 p_i =$$
$$= -k\sum_{i=1}^n p_i\cdot\frac{m_i}{kp_i}\log_2\frac{m_i}{kp_i} \le -k\Big(\sum_{i=1}^n p_i\frac{m_i}{kp_i}\Big)\log_2\Big(\sum_{i=1}^n p_i\frac{m_i}{kp_i}\Big) =$$
$$= -k\Big(\sum_{i=1}^n\frac{m_i}{k}\Big)\log_2\Big(\sum_{i=1}^n\frac{m_i}{k}\Big) = -k\log_2 1 = 0.$$

Таким образом $R\le 0$, и, следовательно,

$$\sum_{\substack{m_1,\dots,m_n \\ v\in A(m_1,\dots,m_n)}} p_1^{m_1}\cdots p_n^{m_n}\log_2\frac{k!}{m_1!\cdots m_n!} \le H(P^k). \quad (12.11)$$

Наконец заметим, что

$$\log_2\binom{k+n-1}{n-1} \le \log_2(k+1)^{n-1} \le (n-1)\log_2(k+1). \quad (12.12)$$

Подставляя (12.11) и (12.12) в (12.10) видим, что

$$C(V_k,P^k) \le H(P^k) + (n-1)\log_2(k+1) + 2.$$

Теперь утверждение теоремы легко следует из равенств $H(P^k) = kH(P)$ и $C_k(V_k,P) = \frac{1}{k}C(V_k,P^k)$. Теорема доказана.

Из (12.1) видно, что выбирая k достаточно большим, величину $C_k(V_k,P)$ можно сделать сколь угодно близкой к $H(P)$.

12.6 Задачи

12.1. Показать, что сумма длин всех слов любого n-элементного оптимального кода не превосходит $\frac{1}{2}(n+2)(n-1)$.

12.2. Показать, что сумма длин всех слов любого разделимого n-элементного кода не меньше чем $n\log_2 n$.

12.3. Показать, что для любого n найдется такое распределение вероятностей $P=\{p_1,\dots,p_n\}$, где все $p_i>0$, что $C(P)=H(P)$.

12.4. Показать, что для любого n и любого $\varepsilon>0$ найдется такое распределение вероятностей $P=\{p_1,\dots,p_n\}$, где все $p_i>0$, что $C(P)\ge H(P)+1-\varepsilon$.

12.5. Указать минимальное n и соответствующее ему распределение вероятностей $P=\{p_1,\dots,p_n\}$, для которых найдутся два оптимальных кода с разными длинами слов.

12.6. Привести пример распределения вероятностей, при котором код Шеннона не является оптимальным.

12.7. Множество N состоит из всех двоичных последовательностей длины n, каждая из которых содержит ровно k единиц. Каждая последовательности $\boldsymbol{a}$ из N делится на блоки по m символов, после чего каждый блок кодируется при помощи некоторого кода V над $\{0,1\}$, и последовательность кодов этих блоков образует код последовательности $\boldsymbol{a}$. Указать код для множества $\{0,1\}^m$, при котором средняя длина кода последовательности из N была бы как можно меньше. Оценить эту длину.

12.8. Показать, что для каждого m-ичного разделимого кода $\{\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n\}$ выполняется неравенство Крафта–Макмиллана

$$\sum_{i=1}^{n} m^{-l(v_i)} \le 1.$$

12.9. Показать, что в m-ичном n-элементном коде найдется слово, длина которого не меньше $\log_m(1+n(m-1))$.

12.10. Показать, что при любом распределении вероятностей P для стоимости $C(V,P)$ соответствующего этому распределению m-ичного оптимального кода V имеет место неравенство $C(V,P) \le H_m(P)+1$, где $H_m(p_1,\ldots,p_n) = -\sum_{i=1}^{n} p_i \log_m p_i$.

Лекция 13

Коды, исправляющие ошибки

Рассмотрим канал передачи информации, в котором под воздействием внешних причин передаваемая информация может искажаться. Будем полагать, что по каналу связи из пункта A в пункт B передаются двоичные последовательности, и что каждый символ передаваемой последовательности независимо от других символов с вероятностью p, где $p < 1/2$, превращается в противоположный, т. е. единица превращается в нуль, а нуль в единицу. Такой канал связи называется *двоичным симметричным каналом с вероятностью ошибки* p, а ошибки называются *аддитивными*, так как ошибку в любом разряде передаваемой последовательности можно рассматривать как прибавление к этому разряду единицы по модулю два.

Для борьбы с искажениями обычно поступают следующим образом. Передаваемое сообщение длины K делят на слова одинаковой длины k, и каждое слово $\boldsymbol{a}$ длины k преобразуют по определенным правилам в слово $\mathbf{g}$ большей длины n — это преобразование называется кодированием, а множество G слов $\mathbf{g}$, получаемых в результате кодирования всевозможных слов $\boldsymbol{a}$ длины k, называется блочным кодом длины n[1)]. Затем преобразованные слова передают по каналу связи. Принимающая сторона делит принятое сообщение на слова длины n и для каждого такого слова $\boldsymbol{v}$ выполняет "обратное" преобразование, которое, как правило, состоит из двух этапов. Сначала определяют такое слово $\mathbf{g}$ из G, для которого условная вероятность $P(\boldsymbol{v}|\mathbf{g})$ получить слово $\boldsymbol{v}$ при условии, что было передано слово $\mathbf{g}$, максимальна. Такой способ определения $\mathbf{g}$ называется *декодированием методом максимального правдоподобия*. После этого для полученного слова $\mathbf{g}$ находится его прообраз $\boldsymbol{a}$.

Параметры k, n и процедуры кодирования и декодирования выбираются так, чтобы при данном p уменьшить до приемлемого значения вероятность неправильного декодирования, сделать как можно больше отношение k/n, называемое скоростью кода, и по-возможности уменьшить сложность выполнения процедур кодирования и декодирования. Далее будем рассматривать различные вопросы, связанные с существованием, мощностью, построением, кодированием и декодированием кодов.

[1)] Помимо блочных кодов используются также не рассматриваемые здесь сверточные коды.

13.1 Двоичный симметричный канал

Нетрудно видеть, что для любой положительной постоянной ε и любого натурального k любую двоичную последовательность длины k можно передать по двоичному симметричному каналу так, что несмотря на появляющиеся в процессе передачи ошибки, получатель информации может правильно определить посланную ему последовательность с вероятность не меньшей чем $1-\varepsilon$. Эта вероятность называется *надежностью* передачи. Для достижения требуемой надежности достаточно каждый передаваемый символ повторить определенное количество раз, и его значением в B считать то, которое встретится чаще. Если при данном k для достижения надежности $1-\varepsilon$ каждый символ необходимо повторить m раз, то скорость передачи информации будет равна $1/m$. Нетрудно видеть, что с ростом k и/или уменьшением ε скорость передачи будет уменьшаться. Поэтому возникает естественный вопрос: Какой максимальной скорости передачи информации можно достичь в двоичном симметричном канале при условии, что надежность должна быть близка к единице? Не вводя строгих определений и рассуждая неформально, попробуем ответить на этот вопрос.

Допустим, что для любого достаточно большого k двоичный набор длины k можно преобразовать в двоичный набор большей длины n так, что невзирая на возникающие во время передачи ошибки, из принятого сообщения можно извлечь исходный набор длины k. В этом случае вместе с переданным набором длины k станет известен и набор длины n, содержащий единицы в тех разрядах, где произошли ошибки. Из теоремы 12.6 следует, что при больших n для записи такого набора потребуется не меньше $nH(p,1-p) = nH(p)$ двоичных разрядов. Поэтому k не превосходит $n(1-H(p))$, и, следовательно, для скорости k/n передачи информации справедливо неравенство

$$\frac{k}{n} \le 1 - H(p). \tag{13.1}$$

Величина $1-H(p)$ называется пропускной способностью двоичного симметричного канала. Далее пропускную способность будем использовать в качестве критерия при оценке эффективности различных методов исправления ошибок.

13.2 Параметры и простейшие свойства кодов

Произвольное подмножество $G = \{\mathbf{g}_1, \ldots, \mathbf{g}_m\}$ множества $\mathbb{B}^n$ называется *двоичным кодом длины* n, а его элементы кодовыми словами или кодовыми векторами. Допустим, что при передаче по двоичному симметричному каналу кодовый вектор $\mathbf{g}$ длины n превратился в вектор $\boldsymbol{v}$ такой же длины. В этом случае ненулевые компоненты вектора $\mathbf{c} = \mathbf{g} \oplus \boldsymbol{v}$ указывают положение ошибок, а сам вектор $\mathbf{c}$ называется *вектором ошибок*. Нетрудно видеть, что вероятность $p(\mathbf{c})$ появления вектора ошибок $\mathbf{c}$ в двоичном симметричном канале не зависит от передаваемого кодового слова, и для нее

справедливо равенство

$$p(\mathbf{c}) = p^{\|\mathbf{c}\|}(1-p)^{n-\|\mathbf{c}\|}, \tag{13.2}$$

где $||\mathbf{c}||$ — вес вектора $\mathbf{c}$. Так как ошибки происходят независимо и вероятность одной ошибки меньше $1/2$, то из (13.2) легко следует, что максимум условных вероятностей $P(\boldsymbol{v}|\mathbf{g}_i)$ получить слово $\boldsymbol{v}$ при условии, что было передано кодовое слово $\mathbf{g}_i$, достигается на том слове $\mathbf{g}$, расстояние Хемминга до которого от $\boldsymbol{v}$ минимально. Если максимум достигается на единственном элементе кода G, то говорят об исправлении ошибки в $\boldsymbol{v}$. Определение для произвольного вектора $\boldsymbol{v} \in \mathbb{B}^n$ такого вектора $\mathbf{g} \in G$, что расстояние от вектора $\boldsymbol{v}$ до вектора $\mathbf{g}$ меньше чем расстояние от $\boldsymbol{v}$ до любого другого элемента G называется *декодированием* вектора $\boldsymbol{v}$, или *исправлением ошибок* в векторе $\boldsymbol{v}$. Если полученный в результате декодирования вектор $\mathbf{g}$ совпадает с переданным вектором $\mathbf{g}'$, то декодирование называется правильным, если не совпадает — неправильным. Если же максимум условных вероятностей достигается более чем на одном кодовом слове, то говорят об обнаружении ошибки, исправить которую нельзя.

Каждому элементу $\mathbf{g}$ кода G поставим в соответствие множество $V(\mathbf{g})$ всех тех наборов $\boldsymbol{v}$ из $\mathbb{B}^n$, для каждого из которых расстояние до $\mathbf{g}$ строго меньше, чем расстояние до любого другого элемента кода G. Очевидно, что, если в результате ошибок передачи кодовое слово $\mathbf{g}$ превращается в какой-либо элемент $\boldsymbol{v}$ из $V(\mathbf{g})$, то ошибки в слове $\boldsymbol{v}$ можно исправить. Поэтому множество $C(\mathbf{g}) = \{\, \mathbf{c} \mid \mathbf{c} = \mathbf{g} \oplus \boldsymbol{v}, \text{где } \boldsymbol{v} \in V(\mathbf{g})\}$ называется *множеством исправляемых ошибок кодового слова* $\mathbf{g}$ кода G. Нетрудно видеть, что сумма

$$\sum_{\boldsymbol{c} \in C(\mathbf{g})} p^{||\mathbf{c}||}(1-p)^{n-||\mathbf{c}||} \tag{13.3}$$

равна вероятности исправления ошибки при передаче кодового слова $\mathbf{g}$. Минимум величин (13.3), взятый по всем кодовым словам $\mathbf{g}$, называется *вероятностью исправления ошибки* кодом G.

Как правило, множества $C(\mathbf{g})$ имеют слишком сложную структуру, создающую значительные трудности при анализе кодов и их свойств. Большинство этих трудностей можно избежать, если, рассматривая коды, заменить множества $C(\mathbf{g})$ на их подмножества, являющиеся шарами одинакового для всех $\mathbf{g}$ радиуса. Такая замена приводит к понятиям кодового расстояния и кода с данным кодовым расстоянием.

Подмножество $G = \{\mathbf{g}_1, \ldots, \mathbf{g}_m\}$ множества $\mathbb{B}^n$ называется *кодом длины* n *с кодовым расстоянием* d, если для любых двух его элементов $\mathbf{g}_i$ и $\mathbf{g}_j$ расстояние Хемминга между ними не меньше d. Также говорят, что код G исправляет t независимых ошибок, если его кодовое расстояние не меньше чем $2t+1$. Нетрудно видеть, что в коде, исправляющем t ошибок, каждый элемент кода является центром шара радиуса t, и шары с центрами в разных элементах кода не пересекаются. Поэтому для такого кода длины n

вероятность P_e неправильного декодирования удовлетворяет неравенству

$$P_e \le \sum_{\mathbf{c}\,|\,||\mathbf{c}||>t} p^{\|\mathbf{c}\|}(1-p)^{n-\|\mathbf{c}\|} = \sum_{i=t+1}^{n} \binom{n}{i} p^i (1-p)^{n-i}. \qquad (13.4)$$

При заданной вероятности ошибки p неравенство (13.4) позволяет определить значения параметров n и t, при которых вероятность неправильного декодирования P_e принимает допустимые (с точки зрения решаемой при помощи кодирования задачи) значения. Например, используя неравенство Чебышева, можно показать, что для любого положительного числа δ

$$P_e \to 0 \ \text{ при } \ n \to \infty \ \text{ и } \ t = (1+\delta)pn. \qquad (13.5)$$

Также можно показать[2)], что для любого положительного числа δ

$$P_e \to 1 \ \text{ при } \ n \to \infty \ \text{ и } \ t = (1-\delta)pn. \qquad (13.6)$$

Для кода длины n с кодовым расстоянием d величину d/n назовем *относительным кодовым расстоянием*. Из соотношений (13.5) и (13.6) можно сделать следующий качественный вывод о кодах, обеспечивающих надежную передачу информации по двоичному симметричному каналу: такие коды должны иметь большую длину, а их относительное кодовое расстояние должно быть константой.

Код G, исправляющий t ошибок, называется *максимальным* для данного t, если мощность G максимальна среди всех кодов, исправляющих t ошибок. Покажем, что для мощности любого максимального кода G длины n справедливы неравенства

$$\frac{2^n}{\Sigma_{i=0}^{2t}\binom{n}{i}} \le |G| \le \frac{2^n}{\Sigma_{i=0}^{t}\binom{n}{i}}. \qquad (13.7)$$

Так как шары радиуса t с центрами в различных кодовых словах не пересекаются, то, очевидно, что для любого кода длины n произведение числа кодовых слов и мощности шара радиуса t не превосходит 2^n, и, следовательно, $|G| \le 2^n/\Sigma_{i=0}^{t}\binom{n}{i}$. С другой стороны, для любого элемента $\mathbf{g}$ кода G в шаре радиуса $2t$ с центром в этом элементе есть ровно одно кодовое слово — элемент $\mathbf{g}$. Поэтому, если набор $\boldsymbol{a}$ не принадлежит ни одному из шаров радиуса $2t$ с центрами в элементах кода, то этот набор можно добавить в G, и дополненное множество по прежнему будет исправлять t ошибок. Поэтому, если $|G|(\Sigma_{i=0}^{2t}\binom{n}{i}) < 2^n$, то код G не является максимальным. Следовательно, неравенство $|G|(\Sigma_{i=0}^{2t}\binom{n}{i}) \ge 2^n$ является необходимым условием максимальности кода G.

Для кода G длины n величина $R = \frac{\log_2|G|}{n}$ называется его *скоростью*. Выражая скорость максимального кода через его мощность из (13.7), получим, что для скорости максимального кода длины n с исправлением t ошибок справедливы неравенства

$$1 - \frac{\log_2 \Sigma_{i=0}^{2t}\binom{n}{i}}{n} \le R \le 1 - \frac{\log_2 \Sigma_{i=0}^{t}\binom{n}{i}}{n},$$

которые при больших n с помощью теорем 2.5 и 2.9 легко преобразуются к виду

$$1 - H\Big(\frac{2t}{n}\Big) \le R \le 1 - H\Big(\frac{t}{n}\Big) + \mathcal{O}\Big(\frac{\log_2 n}{n}\Big), \qquad (13.8)$$

где левое неравенство справедливо при $t \le n/4$, а правое при $t \to \infty$ и $t \le n/2$. Объединяя неравенства (13.5) и (13.8), заключаем, что при помощи кодов по двоичному симметричному каналу с вероятностью ошибки p можно передавать информацию с близкой к нулю вероятностью неправильного декодирования и скоростью

$$1 - H((1+\delta)2p) \le R \le 1 - H((1+\delta)p), \qquad (13.9)$$

где δ — сколь угодно малое положительное число, удовлетворяющее неравенству $(1+\delta)2p \le 1/2$.

Доказательство нижней оценки в (13.7) фактически содержит алгоритм построения кода G, на котором эта оценка достигается, но к сожалению ничего не говорит о каких-либо свойствах этого кода, знание которых могло бы помочь в построении хороших алгоритмов кодирования и декодирования. Поэтому кодирование G будем рассматривать как произвольную булеву функцию $f : \{0,1\}^{\lceil \log_2 |G| \rceil} \to \{0,1\}^n$. При фиксированном p и $n \to \infty$ сложность такой функции есть $\mathcal{O}(n|G|/\log_2 n|G|)$. Для декодирования кода G можно вычислить расстояния от принятого вектора до всех кодовых слов и выбрать слово с минимальным расстоянием. Сложность такого декодирования есть $\mathcal{O}(n|G|)$. Существуют другие, более экономные с точки зрения числа операций алгоритмы декодирования произвольных кодов. Однако все эти алгоритмы используют тот или иной перебор кодовых слов и требуют для своей работы память, способную вместить $|G|$ двоичных наборов длины n. Так как малая вероятность безошибочной передачи информации достигается при использовании кодов большой длины, то указанные методы кодирования и декодирования не представляют практического интереса.

Часто декодирование кода, исправляющего t ошибок, можно существенно упростить, если исправлять ошибки только в тех случаях, когда число ошибок не превосходит t. Если же декодируемое слово находится на расстоянии бóльшем чем t от любого кодового слова, то в этом случае говорят об обнаружении ошибки, которую нельзя исправить. Такое декодирование называется декодированием в пределах кодового расстояния и именно такому декодированию соответствует приведенная выше оценка (13.4) вероятности неправильного декодирования.

[2)] Соотношения (13.5) и (13.6) далее будут подробно рассмотрены в п. 13.4.

13.3 Линейные коды

Код G называется *линейным* (n,k)-кодом, если он является k-мерным линейным подпространством пространства $\mathbb{B}^n$. Справедлива следующая тео-

рема о кодовом расстоянии линейного кода.

Теорема 13.1. *В каждом линейном коде G кодовое расстояние d равно весу его минимального ненулевого элемента*:

$$d = \min_{\mathbf{g}\neq\mathbf{0},\mathbf{g}\in G} \|\mathbf{g}\|.$$

Доказательство. Так как нулевой набор всегда принадлежит линейному коду, то, очевидно, что кодовое расстояние не превосходит веса минимального ненулевого элемента. Допустим, что $d < \min\|\mathbf{g}\|$. В этом случае в G найдутся два элемента $\mathbf{g}_1$ и $\mathbf{g}_2$, расстояние между которыми меньше d. Следовательно,

$$\|\mathbf{g}_1 \oplus \mathbf{g}_2\| = d(\mathbf{g}_1, \mathbf{g}_2) < d.$$

С другой стороны, сумма $\mathbf{g}_1 \oplus \mathbf{g}_2$ обязательно принадлежит G. Поэтому $\|\mathbf{g}_1 \oplus \mathbf{g}_2\| \geq d$. Пришли к противоречию. Теорема доказана.

Теорема 13.1 является частным случаем следующего более общего утверждения об ошибках, исправляемых линейными кодами.

Теорема 13.2. *В линейном коде множества исправляемых ошибок всех элементов совпадают.*

Доказательство. Допустим, что теорема не верна, и в каком-нибудь линейном коде G найдутся такие два элемента $\mathbf{g}_1$ и $\mathbf{g}_2$, что вектор $\mathbf{c}$ принадлежит множеству исправляемых ошибок элемента $\mathbf{g}_1$ и не принадлежит множеству исправляемых ошибок элемента $\mathbf{g}_2$. Тогда для вектора $\mathbf{g}_1 \oplus \mathbf{c}$ ближайшим элементом кода будет $\mathbf{g}_1$, а для вектора $\mathbf{g}_2 \oplus \mathbf{c}$ найдется элемент кода $\mathbf{g}_3$, расстояние до которого не меньше, чем до $\mathbf{g}_2$. Следовательно, существует вектор $\mathbf{c}'$ такой, что $\mathbf{g}_2 \oplus \mathbf{c} \oplus \mathbf{c}' = \mathbf{g}_3$ и $\|\mathbf{c}'\| \leq \|\mathbf{c}\|$. Но тогда вектор $\mathbf{c} \oplus \mathbf{c}' = \mathbf{g}_2 \oplus \mathbf{g}_3 = \mathbf{g}_4$ является элементом кода G. Поэтому расстояние от вектора $\mathbf{g}_1 \oplus \mathbf{c} = \mathbf{g}_1 \oplus \mathbf{g}_4 \oplus \mathbf{c}'$ до элемента $\mathbf{g}_1 \oplus \mathbf{g}_4$ не меньше, чем расстояние до $\mathbf{g}_1$, т. е. вектор $\mathbf{c}$ не принадлежит множеству исправляемых ошибок элемента $\mathbf{g}_1$. Полученное противоречие показывает, что множества исправляемых ошибок всех элементов совпадают. Теорема доказана.

Теорема 13.2 позволяет говорить о множестве ошибок, исправляемых линейным кодом. Пусть линейный код G исправляет ошибки из множества C. Повторяя доказательство теоремы 13.2, нетрудно показать, что

$$\mathbf{c}_i \oplus \mathbf{c}_j \notin G \ \text{ для всех } \ \mathbf{c}_i, \mathbf{c}_j \ \text{ из } \ C. \tag{13.10}$$

Действительно, если в C найдутся такие $\mathbf{c}_i$ и $\mathbf{c}_j$, что $\mathbf{c}_i \oplus \mathbf{c}_j \in G$, то для каждого из этих векторов найдется по крайней мере два ближайших элемента кода — нулевой и $\mathbf{c}_i \oplus \mathbf{c}_j$.

Булева (k, n)-матрица $\mathbf{G}$ называется *порождающей* матрицей линейного кода G, если линейная оболочка $\langle \mathbf{g}_1, \ldots, \mathbf{g}_k \rangle$ строк матрицы $\mathbf{G}$ совпадает с G. При помощи порождающей матрицы $\mathbf{G}$ очень просто выполняется процедура кодирования: для преобразования информационного вектора $\boldsymbol{a}$ длины k в кодовое слово $\mathbf{g}$ длины n достаточно вычислить произведение $\boldsymbol{a}\mathbf{G}$.

Булева $(n-k, n)$-матрица $\mathbf{H}$ называется *проверочной* матрицей линейного кода G, если $\mathbf{Hg} = \mathbf{0}$ для каждого $\mathbf{g} \in G$ и $\mathbf{H}\boldsymbol{x} \neq \mathbf{0}$ для каждого $\boldsymbol{x} \notin G$. Вектор $\mathbf{H}\boldsymbol{x}$ называется *синдромом* вектора $\boldsymbol{x}$ и обозначается символом S. Нетрудно видеть, что $\mathbf{Hc}_i \neq \mathbf{Hc}_j$ для любых исправляемых кодом G ошибок $\mathbf{c}_i$ и $\mathbf{c}_j$, так как в противном случае $\mathbf{H}(\mathbf{c}_i \oplus \mathbf{c}_j) = \mathbf{0}$ и, следовательно, $\mathbf{c}_i \oplus \mathbf{c}_j \in G$, что, очевидно, невозможно. Таким образом, справедлива следующая теорема.

Теорема 13.3. *Для того, чтобы матрица $\mathbf{H}$ была проверочной матрицей кода, исправляющего ошибки из множества C, необходимо и достаточно, чтобы $\mathbf{Hc}_i \neq \mathbf{Hc}_j$ для любых ошибок $\mathbf{c}_i$ и $\mathbf{c}_j$ из C.*

Отметим, что

$$\mathbf{H}(\mathbf{g} \oplus \mathbf{c}) = \mathbf{H}(\mathbf{g}) \oplus \mathbf{H}(\mathbf{c}) = \mathbf{H}(\mathbf{c})$$

для любого элемента кода $\mathbf{g}$ и любого вектора ошибок $\mathbf{c}$. Поэтому вычисление синдрома может существенно упростить декодирование по сравнению с общим нелинейным случаем. Для декодирования набора $\boldsymbol{x}$ надо вычислить его синдром и затем сравнить полученный результат с заранее вычисленными синдромами векторов ошибок. Такое декодирование называется синдромным и его сложность (без учета сложности вычисления синдрома) есть $\mathcal{O}(n2^{n-k})$. Эту величину можно значительно уменьшить при помощи метода согласования, успешно работающего в различных ситуациях. Опишем этот метод.

Пусть линейный (n, k)-код G исправляет t ошибок. Допустим, что при передаче вектора $\mathbf{g}$ произошло не более t ошибок, и был получен вектор $\boldsymbol{x}$. Пусть $\mathbf{c}$ — вектор ошибок, т. е. $\mathbf{g} \oplus \mathbf{c} = \boldsymbol{x}$. Пусть A — множество синдромов $S(\mathbf{c}_l)$ всех векторов ошибок $\mathbf{c}_l$, вес которых не превосходит $\lceil t/2 \rceil$, B — множество попарных сумм синдрома $S(\boldsymbol{x})$ принятого вектора $\boldsymbol{x}$ и синдромов $S(\mathbf{c}_m)$ всех векторов ошибок $\mathbf{c}_m$, вес которых не превосходит $\lfloor t/2 \rfloor$. Так как любой вектор, вес которого не превосходит t, можно представить в виде суммы двух векторов, вес первого из которых не превосходит $\lceil t/2 \rceil$, а второго — $\lfloor t/2 \rfloor$, то, очевидно, что найдутся такие векторы $\mathbf{c}_i$ и $\mathbf{c}_j$, что $\mathbf{c} = \mathbf{c}_i \oplus \mathbf{c}_j$, где $\|\mathbf{c}_i\| \leq \lceil t/2 \rceil$ и $\|\mathbf{c}_j\| \leq \lfloor t/2 \rfloor$. Поэтому в силу линейности синдрома

$$S(\boldsymbol{x}) = S(\mathbf{c}) = S(\mathbf{c}_i \oplus \mathbf{c}_j) = S(\mathbf{c}_i) \oplus S(\mathbf{c}_j).$$

Переписав последнее равенство в виде $S(\boldsymbol{x}) \oplus S(\mathbf{c}_i) = S(\mathbf{c}_j)$, заключаем, что существует непустое пересечение множеств A и B, и если в этих множествах найти пару одинаковых элементов, то по этой паре можно будет восстановить вектор ошибок. Найти такую пару можно следующим образом. Сначала вычислим все синдромы из множества A и все суммы из множества B. Затем упорядочим множество A[3)4)]. После этого последовательно

[3)] Так как элементы множества A являются двоичными векторами длины $n-k$, то для упорядочивания проще всего использовать лексикографический порядок.

[4)] Вычисление множества A и его упорядочивание не зависят от декодируемого вектора $\boldsymbol{x}$. Поэтому будем считать, что эти действия уже выполнены до начала декодирования, и не будем их учитывать при оценке сложности декодирования.

для каждого элемента из B попробуем найти равный ему элемент из A. Если такой элемент есть, то его можно найти, выполнив не более $\lceil \log_2 |A| \rceil$ сравнений текущего элемента из B с элементами из A. Сначала элемент из B сравнивается со средним элементом из A. Если элемент из B окажется меньше, то далее поиск ведется в первой половине A, если больше — во второй половине A. Если в A есть элемент, равный текущему элементу из B, то он будет обнаружен во время одного из сравнений. Нетрудно видеть, что для декодирования вектора $\boldsymbol{x}$ достаточно выполнить

$$\mathcal{O}(|B| \log_2 |A|) = \mathcal{O}\Big(\Big(\sum_{i=0}^{\lfloor t/2 \rfloor} \binom{n}{i}\Big) \log_2 \sum_{i=0}^{\lceil t/2 \rceil} \binom{n}{i}\Big) = \mathcal{O}\Big(2^{nH(t/2n)} nH(t/2n)\Big) \tag{13.11}$$

операций над векторами длины $n-k$.

В следующей теореме устанавливается фундаментальное свойство линейных кодов, лежащее в основе подавляющего числа конструкций этих кодов.

Теорема 13.4. *Для того, чтобы матрица* $\mathbf{H}$ *была проверочной матрицей линейного кода с кодовым расстоянием не меньшим* d *необходимо и достаточно, чтобы любые* $d-1$ *столбцов матрицы* $\mathbf{H}$ *были линейно независимы.*

Доказательство. Установим необходимость. Пусть $\mathbf{H}$ — проверочная матрица кода G с расстоянием d. Если в матрице $\mathbf{H}$ сумма столбцов с номерами $i_1, \ldots, i_l$ равна нулевому вектору, то произведение $\mathbf{H}\boldsymbol{v}$ матрицы $\mathbf{H}$ и вектора $\boldsymbol{v}$, у которого единичные компоненты имеют номера $i_1, \ldots, i_l$, также будет равно нулевому вектору. Следовательно, вектор $\boldsymbol{v}$ принадлежит G, и, поэтому, $l \geq d$. С другой стороны, если любые $d-1$ столбцов матрицы $\mathbf{H}$ линейно независимы, то и произведение матрицы $\mathbf{H}$ и любого вектора $\boldsymbol{v}$ с не более чем $d-1$ единичными компонентами не равно нулевому вектору, и в силу теоремы 13.1 нулевое пространство матрицы $\mathbf{H}$ будет кодом с расстоянием не меньшим d. Теорема доказана.

Докажем нижнюю оценку для мощности максимальных линейных кодов, исправляющих данное число ошибок. Эта оценка называется неравенством Варшамова–Гилберта.

Теорема 13.5. *Если числа* n, m *и* d *удовлетворяют неравенству*

$$2^{n-m} > \sum_{i=0}^{d-2} \binom{n-1}{i},$$

то существует линейный (n, m)*-код с расстоянием* d.

Доказательство. Допустим, что найдется матрица $\mathbf{H}_k$ из $n-m$ строк и k столбцов, у которой любые $d-1$ столбцов линейно независимы. Тогда существует не более $\sum_{i=0}^{d-2} \binom{k}{i}$ различных линейных комбинаций столбцов этой матрицы, в каждую из которых входит не более чем $d-2$ ненулевых слагаемых. Если $2^{n-m} > \sum_{i=0}^{d-2} \binom{k}{i}$, то найдется хотя бы один ненулевой вектор $\boldsymbol{h}$ длины $n-m$, не совпадающий ни с одной из этих линейных комбинаций. Нетрудно видеть, что в матрице $\mathbf{H}_{k+1} = (\mathbf{H}_k \boldsymbol{h})$, составленной из столбцов матрицы $\mathbf{H}_k$ и вектора $\boldsymbol{h}$, любые $d-1$ столбцов линейно независимы, и в силу теоремы 13.4 эта матрица будет проверочной матрицей кода с расстоянием d. Теорема доказана.

Так как синдромы всех исправляемых линейным (n, m)-кодом ошибок различны, то неравенство $n - m \geq \lceil \log_2 \Sigma_{i=0}^{t} \binom{n}{i} \rceil$ справедливо для любого такого кода, исправляющего t ошибок. Объединив это неравенство с границей Варшамова–Гилберта, для мощности максимального линейного (n, m)-кода G, исправляющего t ошибок, получим двойное неравенство

$$\frac{2^n}{\Sigma_{i=0}^{2t-1} \binom{n}{i}} \leq |G| = 2^m \leq 2^{n - \lceil \log_2 \Sigma_{i=0}^{t} \binom{n}{i} \rceil} \leq \frac{2^n}{\Sigma_{i=0}^{t} \binom{n}{i}}, \tag{13.12}$$

являющееся аналогом неравенства (13.7) для произвольных кодов. Заметим, что нижняя оценка в (13.12) немного усиливает нижнюю оценку в (13.7). Однако это усиление не столь велико, чтобы существенно улучшить оценки скорости линейных кодов по сравнению с аналогичными оценками (13.8) и (13.9). Как и в случае произвольных кодов нетрудно показать, что для скорости максимального линейного кода длины n, исправляющего t ошибок, справедливы неравенства

$$1 - H\Big(\frac{2t}{n}\Big) \leq R \leq 1 - H\Big(\frac{t}{n}\Big) + \mathcal{O}\Big(\frac{\log_2 n}{n}\Big), \tag{13.13}$$

и что при помощи линейных кодов по двоичному симметричному каналу с вероятностью ошибки p можно передавать информацию с близкой к нулю вероятностью неправильного декодирования и скоростью

$$1 - H((1+\delta)2p) \leq R \leq 1 - H((1+\delta)p), \tag{13.14}$$

где δ — сколь угодно малое положительное число, удовлетворяющее неравенству $(1+\delta)2p \leq 1/2$.

13.4 Теорема Шеннона

Неравенства (13.14), так же как и неравенства (13.9), оценивают максимальную скорость надежной передачи информации по двоичному симметричному каналу, достижимую при помощи кодов, в которых множества исправляемых ошибок являются шарами. Далее в двух теоремах устанавливаются верхняя и нижняя оценки максимальной скорости передачи информации по двоичному симметричному каналу при помощи блочных кодов без каких либо ограничений на структуру множества исправляемых ошибок. Эти теоремы называются прямой и обратной теоремами Шеннона о кодировании в двоичном симметричном канале. Сначала докажем прямую теорему. В этой теореме устанавливается нижняя оценка скорости.

Теорема 13.6. *Пусть $p \in (0,1/2)$ — вероятность ошибки в одном символе. Для любого положительного ε существует линейный код D со скоростью не меньшей $1 - H(p) - \varepsilon$ и с вероятностью неправильного декодирования не большей ε.*

Перед доказательством теоремы установим несколько вспомогательных фактов. Прежде всего заметим, что функция $H(x)$ выпукла вверх, возрастает на отрезке от нуля до $1/2$ и $H(0) = 0$. Поэтому в силу неравенства леммы 12.3

$$\alpha H(x) \le H(\alpha x) \quad \text{при } x \in (0,1/2) \text{ и } \alpha \in (0,1). \tag{13.15}$$

Далее положим

$$T_p(x) = -x \log_2 p - (1-x)\log_2(1-p),$$

и рассмотрим функцию

$$\begin{aligned} H(x) - T_p(x) = &- x\log_2 x - (1-x)\log_2(1-x) + \\ &+ x\log_2 p + (1-x)\log_2(1-p), \end{aligned}$$

где $x \in (0,1)$, а $p \in (0,1/2)$. Нетрудно видеть, что эта функция равна нулю при $x = p$, а ее производная равна

$$\log_2 \frac{1-x}{x} - \log_2 \frac{1-p}{p},$$

имеет единственный корень $x = p$, который больше нуля, если $x < p$, и меньше нуля, если $x > p$. Следовательно,

$$H(x) - T_p(x) \quad \begin{cases} < 0 \text{ и возрастает} & \text{при } x \in (0,p); \\ = 0, & \text{при } x = p; \\ < 0 \text{ и убывает}, & \text{при } x \in (p,1); \end{cases}$$

для любого p из интервала $(0,1/2)$. Поэтому для любых положительных постоянных δ и p таких, что $0 < (1-\delta)p < (1+\delta)p < 1/2$, найдется такая положительная постоянная γ, что

$$H(x) - T_p(x) \le -\gamma, \quad \text{при } x \in (0,(1-\delta)p] \cup [(1+\delta)p, 1). \tag{13.16}$$

Воспользуемся этим свойством для доказательства следующего утверждения.

Лемма 13.1. *Пусть $C \subseteq \{0,1\}^n$, $p < 1/2$, $\delta > 0$, $|C| \le 2^{nH((1-\delta)p)}$ и $p(\mathbf{c}) = p^{\|\mathbf{c}\|}(1-p)^{n-\|\mathbf{c}\|}$. Тогда существуют такие n_0 и положительная постоянная γ, что при всех $n \ge n_0$*

$$\sum_{\mathbf{c}\in C} p(\mathbf{c}) \le n 2^{-\gamma n}.$$

Доказательство. Множество векторов из C упорядочим по возрастанию их весов и разобъем на подмножества $C_0, C_1, \ldots, C_i, \ldots$ так, что C_0 будет состоят из первого элемента, C_1 — из следующих $2^{nH(1/n)}$ элементов, C_i — из $2^{nH(i/n)}$ элементов минимального веса множества $C \setminus (C_0 \cup \cdots \cup C_{i-1})$. Из условий леммы следует, что максимальный индекс i подмножества A_i не превосходит $(1-\delta)pn + 1$. Пусть δ_0 — такая постоянная, что неравенство $(1-\delta)pn + 1 \le (1-\delta_0)pn$ выполняется при $n \ge n_0$. Так как $\binom{n}{i} \le 2^{nH(i/n)}$, то вес любого вектора $\mathbf{c}$ из C_i не меньше i, и так как $p < 1/2$, то $p(\mathbf{c}) \le p^i(1-p)^{n-i}$. Поэтому, учитывая (13.16), при всех $0 \le i \le (1-\delta_0)pn$ имеем

$$\sum_{\mathbf{c}\in C_i} p(\mathbf{c}) \le 2^{nH(i/n)} p^i (1-p)^{n-i} = 2^{n(H(i/n) - T_p(i/n))} \le 2^{-\gamma n},$$

и, следовательно,

$$\sum_{\mathbf{c}\in C} p(\mathbf{c}) \le \sum_{i=0}^{(1-\delta_0)pn} 2^{n(H(i/n) - T_p(i/n))} \le n 2^{-\gamma n}. \tag{13.17}$$

Лемма доказана.

Наконец докажем следующее утверждение о кодах, исправляющих почти все ошибки из данного множества.

Лемма 13.2. *Пусть $C \subseteq \{0,1\}^n$, $m = \lceil (1+\gamma)\log_2 |C| \rceil$, где $\gamma \ge 0$. Тогда существует такая матрица $\mathbf{H}$ из m строк и n столбцов, что*

$$|\{\boldsymbol{x} \in C \mid \mathbf{H}\boldsymbol{x} \ne \mathbf{H}\boldsymbol{y} \ \forall \boldsymbol{y} \in C\}| > |C| - |C|^{1-\gamma}.$$

Доказательство. Пусть $M(n,m)$ — множество всех булевых матриц из m строк и n столбцов, $\boldsymbol{z} = (z_1, \ldots, z_n)$ — ненулевой вектор длины n. Без ограничения общности будем полагать, что $z_n = 1$. Нетрудно видеть, что $\boldsymbol{z}$ принадлежит нулевому пространству матрицы $\mathbf{H} = (h_{ij})$ из $M(n,m)$ тогда и только тогда, когда $h_{in} = \bigoplus_{j=1}^{n-1} h_{ij} z_j$ для каждого $i \in \{1, \ldots, m\}$. Поэтому ненулевой вектор принадлежит нулевому пространству ровно 2^{nm-m} различных матриц из $M(n,m)$. Так как $|M(n,m)| = 2^{nm}$, то величина

$$2^{-nm} \sum_{\boldsymbol{x},\boldsymbol{y}\in C, \boldsymbol{x}\ne\boldsymbol{y}} 2^{nm-m} = 2^{-m} \frac{|C|(|C|-1)}{2}$$

является средним значением для числа таких пар $(\boldsymbol{x}, \boldsymbol{y})$, на которых значения матрицы из $M(n,m)$ одинаковы. Поэтому в $M(n,m)$ найдется матрица $\mathbf{H}$, для которой равенство $\mathbf{H}\boldsymbol{x} = \mathbf{H}\boldsymbol{y}$ выполняется менее чем на $2^{-m}|C|^2/2$ парах $(\boldsymbol{x}, \boldsymbol{y})$ из C. Так как $2^m \ge |C|^{1+\gamma}$, то $2^{-m-1}|C|^2 < |C|^{1-\gamma}/2$. Следовательно, менее чем $|C|^{1-\gamma}$ векторов множества C образуют пары, на элементах которых совпадают значения матрицы $\mathbf{H}$. Поэтому найдется более $C| - |C|^{1-\gamma}$ векторов, на которых значения матрицы $\mathbf{H}$ различны. Лемма доказана.

Доказательство теоремы. Рассмотрим в $\{0,1\}^n$ шар A радиуса $t = \lfloor (1+\delta)pn \rfloor$, где постоянная δ удовлетворяет неравенству $(1+2\delta)p < 1/2$. Пусть $\mathbf{H}$ — матрица из леммы 13.2 с $C = A$ и $\gamma = 2\delta$. В качестве кода D возьмем линейный код с проверочной матрицей $\mathbf{H}$. Из теоремы 13.3 и леммы 13.2 следует, что множество ошибок, исправляемых кодом D, содержит более $|A| - |A|^{1-2\delta}$ наборов шара A, при этом код D состоит из 2^{n-m} элементов, и его скорость равна $\frac{n-m}{n}$, где $m = \lceil (1+2\delta)\log_2 |A| \rceil$. Покажем, что параметры δ и n можно выбрать так, что код D будет удовлетворять заключению теоремы.

Сначала оценим разность $n - m$. Нетрудно видеть, что

$$\begin{aligned} n - m &\geq n - (1+2\delta)\log_2 |A| - 1 \geq n(1 - (1+2\delta)H(t/n) - 1/n) \geq \\ &\geq n(1 - (1+2\delta)H((1+\delta)p) - 1/n). \end{aligned} \tag{13.18}$$

Так как в силу свойства (13.15) функции $H(x)$

$$\frac{1}{1+\delta} H((1+\delta)p) \leq H\Big(\frac{1+\delta}{1+\delta} \cdot p\Big) \leq H(p),$$

то, продолжая неравенство (13.18), видим, что

$$\begin{aligned} n - m &\geq n(1 - (1+\delta)(1+2\delta)H(p) - 1/n) = \\ &= n(1 - H(p) - (3\delta + 2\delta^2)H(p) - 1/n). \end{aligned}$$

Выберем δ_1 и n_1 так, чтобы при данном p выполнялись неравенства

$$(3\delta_1 + 2\delta_1^2)H(p) < \varepsilon/2, \qquad 1/n_1 < \varepsilon/2.$$

Нетрудно видеть, что в этом случае при $n \geq n_1$ скорость R кода D удовлетворяет неравенству

$$R = \frac{n-m}{n} > 1 - H(p) - \varepsilon.$$

Теперь покажем, что вероятность P неправильного декодирования рассматриваемого кода не превосходит ε. Обозначим через A' подмножество множества ошибок кода D, в котором вес каждого вектора ошибок не превосходит $\lfloor (1+\delta)pn \rfloor$. Очевидно, что для величины P справедливо равенство

$$P = \sum_{\boldsymbol{c} \in A \setminus A'} p(\boldsymbol{c}) + \sum_{\boldsymbol{c} \notin A} p(\boldsymbol{c}), \tag{13.19}$$

где $p(\boldsymbol{c})$ — вероятность появления вектора ошибок $\boldsymbol{c}$.

Оценим первую сумму правой части (13.19). Из конструкции кода D имеем

$$|A \setminus A'| < |A|^{1-2\delta} = \Big(\sum_{i=0}^{\lfloor (1+\delta)pn \rfloor} \binom{n}{i}\Big)^{1-2\delta} \leq \Big(2^{nH((1+\delta)p)}\Big)^{1-2\delta}.$$

В силу неравенства (13.15)

$$\begin{aligned} (1-2\delta)nH((1+\delta)p) &\leq nH((1-2\delta)(1+\delta)p) = \\ &= nH((1+\delta-2\delta-2\delta^2)p) = nH((1-\delta)p - 2\delta^2 p) \leq nH((1-\delta)p). \end{aligned}$$

Следовательно, при $\delta = \delta_1$

$$|A \setminus A'| < 2^{nH((1-\delta_1)p)}.$$

Из леммы 13.1 следует, что найдется такое n_2, что при $n \geq n_2$

$$\sum_{\boldsymbol{c} \in A \setminus A'} p(\boldsymbol{c}) \leq n 2^{-\gamma n} \leq \varepsilon/2. \tag{13.20}$$

Теперь оценим вторую сумму правой части (13.19). Учитывая (13.16), видим, что для этой суммы справедлива оценка

$$\begin{aligned} \sum_{c \notin A} p(c) &= \sum_{i > \lfloor (1+\delta_1)pn \rfloor} \binom{n}{i} p^i (1-p)^{n-i} \leq \\ &\leq \sum_{i > (1+\delta_1)pn} 2^{n(H(i/n) - T(i/n,p))} \leq n 2^{-\gamma n}, \end{aligned} \tag{13.21}$$

с такой же постоянной γ как и в (13.20). Поэтому как и в предыдущем случае при $n \geq n_2$ величина в правой части (13.21) будет меньше $\varepsilon/2$. Таким образом, при $n \geq \max(n_1, n_2)$ вероятность P неправильного декодирования не превосходит ε. Теорема доказана.

Теперь докажем обратную теорему Шеннона. Эта теорема утверждает, что при передаче информации по двоичному симметричному каналу со скорость большей, чем $1 - H(p)$ вероятность ошибки не может быть сделана произвольно малой. Более того, из доказательства теоремы следует, что с ростом длины кода вероятность ошибки стремится к единице.

Теорема 13.7. *Пусть $p \in (0, 1/2)$ — вероятность ошибки в одном символе. Существует такая положительная постоянная ε_0, что для любого блочного кода D со скоростью большей $1 - H(p)$ вероятность неправильного декодирования не меньше ε_0.*

Доказательство. Пусть D — код длины n со скоростью $1 - H(p) + \delta$, где δ — положительная постоянная. Этот код состоит из $2^{n(1-H(p)+\delta)}$ элементов и исправляет не более $2^{n(H(p)-\delta)}$ ошибок. Так как

$$H(p) - \delta \leq (1-\delta)H(p) \leq H((1-\delta)p),$$

то число исправляемых этим кодом ошибок можно оценить сверху величиной

$$2^{nH((1-\delta)p)}. \tag{13.22}$$

Множество исправляемых кодом ошибок обозначим через A. Из леммы 13.1 следует существование такой положительной постоянной γ, что для вероятности P правильного исправления ошибки справедливо неравенство

$$P \le \sum_{c \in A} p(c) \le n 2^{-\gamma n}.$$

Очевидно, что при достаточно большом n, вероятность будет меньше любой положительной постоянной ε.

Пусть теперь D — код длины n_0 со скоростью $1-H(p)+\delta$ и вероятностью неправильного декодирования ε. Тогда код D_m, являющийся конкатенацией m кодов D, будет иметь такую же скорость $1 - H(p) + \delta$, а для вероятности P_m правильного исправления ошибки этим кодом справедливы неравенства

$$(1-\varepsilon)^m \le P_m \le nm 2^{-\gamma nm}.$$

Отсюда после извлечения корня m-й степени и перехода к пределу при $m \to \infty$, получаем, что

$$1 - \varepsilon \le \lim_{m \to \infty} \sqrt[m]{nm} \cdot 2^{-\gamma n} = 2^{-\gamma n},$$

или

$$\varepsilon \ge 1 - 2^{-\gamma n}.$$

Так как γ — постоянная, то минимум величин $1-2^{-\gamma n}$, взятый по всем возможным значениям n, является постоянной ε_0 из формулировки теоремы. Теорема доказана.

13.5 Хорошие коды

Определим условия, которым должен удовлетворять код для того, чтобы считаться хорошим при использовании в двоичном симметричном канале.

Прежде всего код должен быть надежным, т. е. вероятность неправильного декодирования должна быть маленькой. На стр. 196 был сделан вывод о том, что для этого код должен быть длинным и иметь относительное кодовое расстояние, равное константе. Из неравенств 13.13 и 13.14 следует, что существуют сколь угодно длинные линейные коды, у которых относительное расстояние и скорость равны константам. Поэтому к необходимым условиям хорошего кода отнесем большую длину и независящие от длины скорость и относительное расстояние. При использовании длинных кодов важное значение приобретает сложность процедур кодирования и декодирования, которая при больших значениях может оказаться главным фактором, ограничивающим скорость передачи информации. Простота этих процедур является по важности вторым, после надежности, условием хорошего кода. Обычно сложность считают хорошей, если она ограничена сверху полиномом небольшой степени от длины кода.

Таким образом код назовем хорошим если: 1) код имеет большую длину; 2) его относительное расстояние равно константе; 3) сложность его процедур кодирования и декодирования ограничена сверху полиномом небольшой степени от его длины; 4) его скорость равна константе.

Как уже было сказано выше существуют сколь угодно длинные линейные коды, у которых относительное расстояние и скорость равны константам. Кодирование линейного кода осуществляется умножением порождающей матрицы на информационный вектор. Поэтому с точки зрения скорости, относительного расстояния и сложности кодирования существуют хорошие линейные коды. С другой стороны, если относительное расстояние линейного кода равно константе, то формула 13.11 гарантирует лишь экспоненциальную относительно его длины сложность декодирования. Такая сложность декодирования является абсолютно неприемлемой для длинных кодов. Поэтому далее основные усилия будут направлены на построение линейных кодов с константным относительным расстоянием, простым декодированием и, по возможности, большой скоростью.

13.6 Задачи

13.1. Пусть G — линейный (n,k)-код. Доказать, что если код G имеет хотя бы одно слово нечетного веса, то все кодовые слова четного веса образуют $(n, k-1)$-код.

13.2. Показать, что если существует линейный (n,k)-код с минимальным расстоянием $2t+1$, то существует линейный $(n+1,k)$-код с расстоянием $2t+2$.

13.3. Пусть k — натуральное. Для заданного k найти максимальное n, для которого существует линейный $(n, n-k)$-код, исправляющий одну ошибку.

13.4. Показать, что любой код длины $n = 2k$ с минимальным расстоянием k содержит не более $2n$ слов.

13.5. Написать проверочную матрицу кода длины n с минимальным расстоянием d, если: *a*) $n = 16,\ d = 3$; *b*) $n = 9,\ d = 4$; *c*) $n = 11,\ d = 5$.

13.6. Линейный (n,k)-код G называется систематическим, если его проверочная матрица $\mathbf{H}$ имеет вид $(\mathbf{E}_{n-k}\mathbf{P})$, где $\mathbf{E}_{n-k}$ — единичная матрица порядка $n-k$. Найти порождающую матрицу систематического кода.

13.7. Показать, что если множество ошибок $C \subseteq \mathbb{B}^n$ состоит не более чем из $2^{n/2}$ наборов, то найдется такой линейный (n,m)-код, исправляющий ошибки из этого множества, что $n - m \le \lfloor 2 \log_2 |C| \rfloor - 1$.

13.8. Показать, что если $N/\log_2 n \to \infty$ при $n \to \infty$, то доля множеств ошибок $C \subseteq \mathbb{B}^n$ таких, что $|C| = N$ и для каждого из этих множеств существует исправляющий ошибки данного множества (n,m)-код с $n - m \ge (2-\varepsilon)\log_2 |C|$, стремится к нулю для любой положительной постоянной ε.

Лекция 14

Линейные коды

Ниже рассматриваются конструкции и методы декодирования трех самых известных линейных кодов: кодов Хемминга, кодов Рида–Маллера и кодов Боуза–Чоудхури–Хоквингема.

14.1 Коды Хемминга

В 1950 г. Р. Хемминг опубликовал первую нетривиальную конструкцию кодов, исправляющих ошибки. Предложенные Хеммингом коды исправляют одну и обнаруживают две ошибки в двоичных наборах, длина которых равна степени двойки, и, что особенно важно, имеют простой и эффективный алгоритм декодирования. Сейчас такие коды называются расширенными кодами Хемминга.

Пусть $\mathbf{H}_m = (h_{ij})$ — булева матрица из m строк и $2^m - 1$ столбцов, у которой j-й столбец $\mathbf{h}_j = (h_{1j}, \ldots, h_{mj})$ совпадает с двоичным разложением числа j. Например, матрица $\mathbf{H}_3$ выглядит следующим образом:

$$\mathbf{H}_3 = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Кодом Хемминга длины $n = 2^m - 1$ называется линейный $(2^m - 1, 2^m - m - 1)$-код с проверочной матрицей $\mathbf{H}_m$. Легко видеть, что любые два столбца матрицы $\mathbf{H}_m$ будут линейно независимыми. Поэтому из теоремы 13.4 следует, что матрица $\mathbf{H}_m$ является проверочной матрицей кода, исправляющего одну ошибку в наборах длины $2^m - 1$. Исправление ошибки в векторе $\boldsymbol{v}$ при использовании кода Хемминга выполняется очень просто. Нетрудно видеть, что если ошибки не было, то вектор $\mathbf{H}_m\boldsymbol{v}$ состоит из одних нулей, а если одиночная ошибка присутствует, то вектор $\mathbf{H}_m\boldsymbol{v}$ является двоичным представлением номера ошибочного разряда. Поэтому для исправления ошибки достаточно вычислить синдром, и если ошибка присутствует, то ненулевое значение синдрома укажет место этой ошибки.

Исправляющий t ошибок код G длины n называется *совершенным*, если любой набор длины n принадлежит шару радиуса t с центром в каком-либо

кодовом слове. Так как в $\mathbb{B}^n$ шар единичного радиуса состоит из $n+1$ наборов, а код Хемминга длины $n = 2^m - 1$ состоит из $2^n/(n+1)$ элементов, то нетрудно видеть, что при любом $n = 2^m - 1$ код Хемминга длины n является совершенным, а, следовательно, и максимальным среди всех исправляющих одну ошибку кодов длины n.

Воспользуемся кодом Хемминга для построения кода с кодовым расстоянием, равным четырем. Для этого к матрице $\mathbf{H}_m$ слева добавим нулевой столбец, а затем к новой матрице сверху добавим строку из одних единиц. Получившуюся матрицу обозначим через $\mathbf{H}'_m$. При $m = 3$ матрица $\mathbf{H}_m$ выглядит следующим образом:

$$\mathbf{H}_3 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Линейный $(2^m, 2^m - m - 1)$-код с проверочной матрицей $\mathbf{H}'_m$ называется *расширенным кодом Хемминга* длины $n = 2^m$. В матрице $\mathbf{H}'_m$ любые три столбца линейно независимы. Это следует из двух очевидных фактов: матрица $\mathbf{H}_m$ является подматрицей матрицы $\mathbf{H}'_m$; первая компонента суммы трех любых столбцов матрицы $\mathbf{H}'_m$ равна единице. Поэтому любой расширенный код Хемминга будет исправлять одну и обнаруживать две ошибки.

Для декодирования расширенного кода Хемминга нужно вычислить синдром и воспользоваться следующим простым правилом. Если все компоненты синдрома равны нулю, то ошибок нет. Если синдром содержит ненулевые компоненты, и его первая компонента равна единице, то произошла одна ошибка в позиции, номер которой равен двоичному числу, составленному из последних m компонент синдрома. Если синдром содержит ненулевые компоненты, и его первая компонента равна нулю, то произошло две ошибки.

Покажем, что расширенный код Хемминга является максимальным кодом с расстоянием четыре. Допустим, что это не так. Тогда для какого-нибудь $n = 2^m$ найдется состоящий более чем из $2^{n-1}/n$ элементов код G длины n с расстоянием четыре. Код G преобразуем в новый код G' длины $n - 1$, удалив из всех кодовых слов G последнюю компоненту. Нетрудно видеть, что новый код будет состоять более чем из $2^{n-1}/n$ элементов, а его кодовое расстояние будет не меньше трех, т. е. G' содержит больше элементов, чем код Хемминга такой же длины, что, очевидно, невозможно в силу максимальности последнего. Следовательно, расширенный код Хемминга действительно является максимальным линейным кодом с расстоянием четыре.

14.2 Коды Рида–Маллера

Коды Рида–Маллера образуют обширный класс линейных кодов длины 2^n с кодовым расстоянием 2^{n-k}, где k изменяется от единицы до $n - 2$. При

$k = n-2$ коды Рида–Маллера совпадают с расширенными кодами Хемминга. Эти коды были найдены Д. Маллером в 1953 г., а в 1954 г. И. Рид опубликовал эффективный метод для их декодирования. Предложенный Ридом метод не требует вычисления синдрома декодируемого набора и является первым представителем класса мажоритарных методов декодирования, в которых символы информационного (еще не закодированного) сообщения вычисляются голосованием среди определенным образом вычисленных линейных сумм разрядов декодируемого набора.

Кодом Рида–Маллера $\mathrm{PM}(n,k)$ длины 2^n порядка k называется множество векторов значений всех n-местных булевых функций, степени многочленов Жегалкина которых не превосходят k[1)]. Будем полагать, что наборы аргументов n-местных булевых функций упорядочены лексикографически так, что переменная x_1 соответствует старшему разряду набора, а переменная x_n — младшему (см. стр. 80). Так, например, множество $\mathrm{PM}(2,0)$ состоит из векторов значений двух булевых констант, т. е. $\mathrm{PM}(2,0) = \{(0000),(1111)\}$. Множество $\mathrm{PM}(2,1)$ состоит из векторов значений всех линейных булевых функций двух переменных x_1 и x_2:

$$0=(0000),\quad x_1=(0011),\quad 1\oplus x_1=(1100),\qquad 1\oplus x_1\oplus x_2=(1001),$$
$$1=(1111),\quad x_2=(0101),\quad 1\oplus x_2=(1010),\qquad\quad x_1\oplus x_2=(0110).$$

Наконец, множество $\mathrm{PM}(2,2)$ совпадает с $P_2(2)$. Нетрудно видеть, что рассмотренные множества являются линейными кодами длины четыре с кодовыми расстояниями равными, соответственно, четырем, двум и единице. Далее покажем, что при произвольных n и k код $\mathrm{PM}(n,k)$ будет линейным

Таблица 14.1

1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
x_1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
x_2	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
x_3	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
x_4	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
x_1x_2	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
x_1x_3	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1
x_1x_4	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	1
x_2x_3	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1
x_2x_4	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0	1
x_3x_4	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1
$x_1x_2x_3$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
$x_1x_2x_4$	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
$x_1x_3x_4$	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
$x_2x_3x_4$	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1

[1)] Здесь для краткости будем говорить не о векторах значений булевых функций, а просто о функциях.

пространством размерности $\sum_{i=0}^{k}\binom{n}{i}$, состоящим из векторов длины 2^n. В таблице 14.1 представлены базисы кодов $\mathrm{PM}(4,1)$, $\mathrm{PM}(4,2)$ и $\mathrm{PM}(4,3)$. Первые пять строк этой таблицы образуют базис кода $\mathrm{PM}(4,1)$, первые одиннадцать — базис кода $\mathrm{PM}(4,2)$, все строки — базис кода $\mathrm{PM}(4,3)$.

Теорема 14.1. *Код Рида–Маллера $PM(n,k)$ длины 2^n порядка k является линейным кодом длины 2^n с кодовым расстоянием 2^{n-k}.*

Доказательство. Теорему докажем индукцией по числу переменных n. При $n=2$ утверждение теоремы следует из рассмотренного выше примера.

Предположим, что теорема верна при всех n, не превосходящих некоторое $m-1\ge 2$. Покажем, что из этого предположения следует утверждение теоремы для $n=m$.

Так как сумма двух функций степени $\le k$ также будет функцией степени $\le k$, то очевидно, что множество $\mathrm{PM}(n,k)$ является линейным пространством. Поэтому в силу теоремы 13.1 достаточно показать, что вес каждой ненулевой функции из $\mathrm{PM}(n,k)$ не меньше, чем 2^{m-k}.

Пусть f — произвольная m-местная булева функция степени k. Если $k=m$, то утверждение теоремы очевидно, так как $\|f\|\ge 1$. Поэтому далее полагаем, что $k<m$. В многочлене Жегалкина функции f соберем вместе одночлены, содержащие и не содержащие x_m. В результате получим равенство

$$f(x_1,\dots,x_m) = x_m f_1(x_1,\dots,x_{m-1}) \oplus f_2(x_1,\dots,x_{m-1}), \tag{14.1}$$

где $f_1\in \mathrm{PM}(m-1,k-1)$ и $f_2\in \mathrm{PM}(m-1,k)$. Нетрудно видеть, что $f_1\oplus f_2\in \mathrm{PM}(m-1,k)$. Если функции f_1, f_2 и $f_1\oplus f_2$ отличны от тождественного нуля, то по предположению индукции для этих функций справедливы неравенства

$$\|f_2\|\ge 2^{m-k},\quad \|f_2\|\ge 2^{m-1-k},\quad \|f_1\oplus f_2\|\ge 2^{m-1-k}. \tag{14.2}$$

Для каждого $\boldsymbol{x}$ из $\mathbb{B}^m$ через $\boldsymbol{x}'$ обозначим первые $(m-1)$ координат набора $\boldsymbol{x}$. Булев куб $\mathbb{B}^m$ разобьем на два непересекающихся подмножества B_0^m и B_1^m, первое из которых состоит из всех наборов с последним разрядом равным нулю, а второе — из всех наборов с последним разрядом, равным единице. Оценим вес функции f. Из (14.1) имеем

$$\|f\| = \sum_{\boldsymbol{x}\in B_1^m}(x_m f_1(\boldsymbol{x}')\oplus f_2(\boldsymbol{x}')) + \sum_{\boldsymbol{x}\in B_0^m}(x_m f_1(\boldsymbol{x}')\oplus f_2(\boldsymbol{x}')) =$$
$$= \sum_{\boldsymbol{x}'\in\mathbb{B}^{m-1}}(f_1(\boldsymbol{x}')\oplus f_2(\boldsymbol{x}')) + \sum_{\boldsymbol{x}'\in\mathbb{B}^{m-1}} f_2(\boldsymbol{x}') = \|f_1\oplus f_2\| + \|f_2\|.$$

Далее рассмотрим три случая.

1. Функции $f_1\oplus f_2$ и f_2 отличны от тождественного нуля. Из (14.2) имеем

$$\|f\|\ge \|f_1\oplus f_2\| + \|f_2\| \ge 2^{m-1-k} + 2^{m-1-k} = 2^{m-k}.$$

2. Функция $f_1 \oplus f_2$ отлична от тождественного нуля и $f_2 \equiv 0$. Тогда $f_1 \oplus f_2 = f_1 \in \mathrm{PM}(m-1, k-1)$. Из (14.2) имеем

$$\|f\| = \|f_1 \oplus f_2\| = \|f_1\| \geq 2^{m-k}.$$

3. Функция f_2 отлична от тождественного нуля и $f_1 \oplus f_2 \equiv 0$. Тогда $f_2 \equiv f_1 \in \mathrm{PM}(m-1, k-1)$. Из (14.2) имеем

$$\|f\| = \|f_2\| = \|f_1\| \geq 2^{m-k}.$$

Теорема доказана.

Нетрудно видеть, что скорость R кода Рида–Маллера длины $N = 2^n$ порядка k равна $\left(\sum_{i=0}^{k} \binom{n}{i}\right)/2^n$. Поэтому, если $k \leq n/2 - \sqrt{n\varphi(n)}$, где $\varphi(n)$ неограниченно растет вместе с n, то $R \to 0$ при $n \to \infty$. Отсюда легко следует, что для любой положительной постоянной ε у кода Рида–Маллера длины N с кодовым расстоянием не меньшем $N^{1/2+\varepsilon}$ скорость стремится к нулю при $n \to \infty$. Таким образом, с точки зрения скорости коды Рида–Маллера являются плохими кодами. Тем не менее, несмотря на свою плохую скорость эти коды представляют значительный интерес по следующим причинам. Во-первых, для всех n при $k = 1$ и $k = n-2$ коды Рида–Маллера являются максимальными кодами с расстоянием 2^{n-1} и 4, соответственно. Во-вторых, для кодов Рида–Маллера существует простой алгоритм исправления ошибок.

14.3 Декодирование кодов Рида–Маллера

Алгоритм декодирования кодов Рида–Маллера основан на рассматриваемом ниже методе определения коэффициентов многочлена Жегалкина булевой функции. Для функции f, степень которой не превосходит k, где $1 \leq k \leq n-2$, рассматриваемый метод позволяет правильно определять коэффициенты многочлена Жегалкина даже в том случае, когда вместо вектора значений функции f известен вектор значений функции $f \oplus c$, где c — неизвестная булева функция, вес которой не превосходит $2^{n-k-1} - 1$. Опишем этот метод.

Сначала рассмотрим способ определения коэффициента при одночлене $x_1 \cdot \ldots \cdot x_k$. Для этого одночлена определим разбиение булева куба $\mathbb{B}^n$ на 2^{n-k} подмножеств B_i, $i = 0, 1, \ldots, 2^{n-k} - 1$, так, что подмножество B_i состоит из всех тех наборов, у которых на последних $n-k$ местах стоят такие константы $\beta_{k+1}, \ldots, \beta_n$, что $i = \sum_{j=1}^{n-k} \beta_{k+j} 2^{n-k-j}$. Очевидно, что каждое из этих подмножеств является подкубом размерности k. Например, для одночлена x_1x_2, рассматриваемого как функция переменных x_1, x_2, x_3 и x_4, описываемое разбиение куба состоит из подмножеств

$$B_0 = \{(0000), (0100), (1000), (1100)\}, \quad B_1 = \{(0001), (0101), (1001), (1101)\},$$
$$B_2 = \{(0010), (0110), (1010), (1110)\}, \quad B_3 = \{(0011), (0111), (1011), (1111)\}.$$

Ограничением булевой функции $f(x_1, \ldots, x_n)$ на множество B_i, где $i = \sum_{j=1}^{n-k} \beta_{k+j} 2^{n-k-j}$, называется функция $f(x_1, \ldots, x_k, \beta_{k+1}, \ldots, \beta_n)$, получающаяся из f подстановкой констант $\beta_{k+1}, \ldots, \beta_n$ вместо переменных $x_{k+1}, \ldots, x_n$. Очевидно, что ограничение одночлена $x_1 \cdot \ldots \cdot x_k$ на любое из множеств B_i получается из исходного одночлена удалением всех его фиктивных переменных, равно единице на единственном наборе аргументов $x_1 = 1, \ldots, x_k = 1$, и поэтому для каждого $i \in \{0, 1, \ldots, 2^{n-k} - 1\}$ справедливы равенства

$$\bigoplus_{\boldsymbol{x} \in B_i} x_1 \cdot \ldots \cdot x_k = \bigoplus_{x_1, \ldots, x_k \in \mathbb{B}^k} x_1 \cdot \ldots \cdot x_k = 1. \tag{14.3}$$

Легко видеть, что ограничение любого другого одночлена $x_{i_1} \cdot \ldots \cdot x_{i_m}$ степени не больше k на любое множество B_i будет одночленом степени строго меньшей k. Поэтому такое ограничение обязательно будет иметь хотя бы одну фиктивную переменную, его вес будет четным числом, и, следовательно, для каждого $i \in \{0, 1, \ldots, 2^{n-k} - 1\}$ будут справедливы равенства

$$\bigoplus_{\boldsymbol{x} \in B_i} x_{i_1} \cdot \ldots \cdot x_{i_m} = \bigoplus_{\substack{x_1, \ldots, x_k \in \mathbb{B}^k \\ x_{k+1} = \beta_{k+1}, \ldots, x_n = \beta_n}} x_{i_1} \cdot \ldots \cdot x_{i_m} = 0. \tag{14.4}$$

Для любой булевой функции $f(x_1, \ldots, x_k)$, степень которой не превосходит k, равенства (14.3) и (14.4) позволяют определить входит ли одночлен $x_1 \cdot \ldots \cdot x_k$ в многочлен Жегалкина этой функции. Действительно, для любой f и для любого $\boldsymbol{x}$ значение $f(\boldsymbol{x})$ равно взятой по модулю два сумме значений одночленов, входящих в многочлен Жегалкина f. Поэтому, сумма значений функции f на любом из множеств B_i, взятая по модулю два, равна единице только в том случае, когда одночлен $x_1 \cdot \ldots \cdot x_k$ входит в ее многочлен Жегалкина. Таким образом, если степень булевой функции f не превосходит k, то существуют 2^{n-k} независимых соотношений для определения вхождения одночлена $x_1 \cdot \ldots \cdot x_k$ в многочлен Жегалкина f.

Теперь рассмотрим сумму $f(\boldsymbol{x}) \oplus c(\boldsymbol{x})$ булевых функций $f(\boldsymbol{x})$ и $c(\boldsymbol{x})$ таких, что $\deg f \leq k$ и $\|c\| < 2^{n-k-1}$. Для каждого $i \in \{0, 1, \ldots, 2^{n-k} - 1\}$ вычислим две суммы

$$\bigoplus_{\boldsymbol{x} \in B_i} (f(\boldsymbol{x}) \oplus c(\boldsymbol{x})), \qquad \bigoplus_{\boldsymbol{x} \in B_i} f(\boldsymbol{x}), \tag{14.5}$$

которые, очевидно, различаются не более чем на $\|c\|$ множествах B_i. Поэтому для определения коэффициента при одночлене $x_1 \cdot \ldots \cdot x_k$ в многочлене Жегалкина функции f можно использовать вектор значений функции $f \oplus c$. Достаточно вычислить все суммы $\bigoplus_{\boldsymbol{x} \in B_i} (f \oplus c)(\boldsymbol{x})$. Эти суммы называются проверочными суммами для одночлена $x_1 \cdot \ldots \cdot x_k$. Если больше половины проверочных сумм равны единице, то одночлен $x_1 \cdot \ldots \cdot x_k$ входит в многочлен Жегалкина функции f, а если больше половины вычисленных сумм равны нулю, то не входит.

Определение коэффициента при произвольном одночлене $x_{i_1} \cdot \ldots \cdot x_{i_k}$ степени k отличается от приведенного выше метода только способом определения множеств B_i — в случае одночлена $x_{i_1} \cdot \ldots \cdot x_{i_k}$ каждое B_i состоит из всех тех наборов, у которых принимают фиксированные значения разряды с индексами, не принадлежащими множеству $\{i_1, \ldots, i_k\}$. Например, для одночлена x_1x_3, рассматриваемого как функция переменных x_1, x_2, x_3 и x_4, множества B_i определяются следующим образом:

$$B_0 = \{(0000), (0010), (1000), (1010)\}, \quad B_1 = \{(0001), (0011), (1001), (1011)\},$$
$$B_2 = \{(0100), (0110), (1100), (1110)\}, \quad B_3 = \{(0101), (0111), (1101), (1111)\}.$$

После того как для функции f определены коэффициенты многочлена Жегалкина при всех одночленах степени k, к функции $f \oplus c$ прибавим все одночлены степени k, коэффициенты при которых равны единице. В результате получим новую функцию $f' \oplus c$, где $\deg f' \le k-1$, а вес c по прежнему не превосходит $2^{n-k-1} - 1$. Теперь, используя вектор значений функции $f' \oplus c$, описанным выше способом определим в многочлене Жегалкина функции f коэффициенты при одночленах степени $k-1$. Затем одночлены степени $k-1$, коэффициенты при которых равны единице, прибавим к функции $f' \oplus c$. Очевидно, что степень новой функции $f'' \oplus c$ не превосходит $k-2$.

Нетрудно убедиться в том, что повторяя приведенные вычисления для всех остальных степеней вплоть до нулевой, в результате получим функцию c. Итак, если в сумме $f \oplus c$ двух функций n переменных степень функции f не превосходит k, а вес функции c не превосходит $2^{n-k-1}-1$, то приведенный метод позволяет по известной сумме $f \oplus c$ определить функции f и c.

14.4 Коды Боуза–Чоудхури–Хоквингема

В 1959 г. А. Хоквингем и независимо от него в 1960 г. Р. К. Боуз и Д. К. Рой-Чоудхури предложили простую и эффективную конструкцию, позволяющую строить коды с различными кодовыми расстояниями. Впоследствии эта конструкция была обобщена, и построенные на ее основе коды стали называться кодами Боуза–Чоудхури–Хоквингема или кратко БЧХ-кодами. Ниже рассматривается наиболее важный частный случай БЧХ-кодов — примитивные БЧХ-коды. Конструкция этих кодов описывается следующей теоремой.

Теорема 14.2. *Пусть $n = 2^m - 1$ и $\alpha_1, \ldots, \alpha_n$ — ненулевые элементы поля $GF(2^m)$. Тогда матрица*

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_n \\ \alpha_1^3 & \alpha_2^3 & \ldots & \alpha_n^3 \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_1^{2t-1} & \alpha_2^{2t-1} & \ldots & \alpha_n^{2t-1} \end{pmatrix}, \tag{14.6}$$

где α_j^i представлены в виде двоичных векторов высоты m, будет проверочной матрицей кода с расстоянием не меньшим $2t+1$. Код с проверочной матрицей (14.6) *называется примитивным БЧХ-кодом длины n с конструктивным расстоянием $2t+1$.*

Доказательство. В силу теоремы 13.4 достаточно показать, что любые $d-1$ столбцов матрицы (14.6) линейно независимы. Для этого установим следующее вспомогательное утверждение.

Лемма 14.1. *Пусть $n = 2^m - 1$ и $\alpha_1, \ldots, \alpha_n$ — ненулевые элементы поля $GF(2^m)$. Тогда в матрице*

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \ldots & \alpha_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_1^{d-1} & \alpha_2^{d-1} & \ldots & \alpha_n^{d-1} \end{pmatrix} \tag{14.7}$$

любые $d-1$ столбцов линейно независимы.

Доказательство. Допустим, что утверждение леммы не верно, и столбцы матрицы (14.7) с номерами $i_1, \ldots, i_{d-1}$ линейно зависимы. Тогда ранг матрицы

$$\begin{pmatrix} \alpha_{i_1} & \alpha_{i_2} & \ldots & \alpha_{i_{d-1}} \\ \alpha_{i_1}^2 & \alpha_{i_2}^2 & \ldots & \alpha_{i_{d-1}}^2 \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_{i_1}^{d-1} & \alpha_{i_2}^{d-1} & \ldots & \alpha_{i_{d-1}}^{d-1} \end{pmatrix}$$

не превосходит $d-2$. Поэтому строки этой матрицы линейно зависимы, т. е. найдется такой вектор $h = (h_1, \ldots, h_{d-1})$, что

$$h_1\alpha_{i_k} \oplus h_2\alpha_{i_k}^2 \oplus \cdots \oplus h_{d-1}\alpha_{i_k}^{d-1} = \alpha_{i_k}(h_1 \oplus h_2\alpha_{i_k} \oplus \cdots \oplus h_{d-1}\alpha_{i_k}^{d-2}) = 0$$

для $k = 1, \ldots, d-1$. Следовательно, многочлен $h_{d-1}x^{d-2} \oplus \cdots \oplus h_2x \oplus h_1$, степень которого меньше $d-1$, имеет не менее $d-1$ различных корней в поле $GF(2^m)$. Противоречие. Лемма доказана.

Теперь заметим, что в поле характеристики 2 для любого натурального l и любых $x_1, \ldots, x_l$ справедливо тождество $(x_1 \oplus \cdots \oplus x_l)^2 = x_1^2 \oplus \cdots \oplus x_l^2$. Поэтому суммы $\alpha_{i_1}^k \oplus \cdots \oplus \alpha_{i_{d-1}}^k$ и $\alpha_{i_1}^{2k} \oplus \cdots \oplus \alpha_{i_{d-1}}^{2k}$, составленные из элементов матриц (14.6) и (14.7), при любых $i_1, \ldots, i_{d-1}$ и любом k одновременно либо не равны, либо равны нулю, т. е. в этих матрицах столбцы с номерами $i_1, \ldots, i_{d-1}$ одновременно либо линейно зависимы, либо линейно независимы. Следовательно, в силу леммы 14.1, любые $d-1$ столбцов матрицы (14.6) линейно независимы. Теорема доказана.

Построим проверочную матрицу примитивного БЧХ-кода длины 15, исправляющего две ошибки. В качестве поля из 16 элементов возьмем поле $\mathbb{Z}_2[\alpha]$, где α — корень примитивного многочлена $x^4 \oplus x \oplus 1$. Так как α является порождающим элементом мультипликативной группы рассматриваемого поля, то его степени выглядят следующим образом:

$$\begin{array}{llll} \alpha^0=(0001), & \alpha^4=(0011), & \alpha^8=(0101), & \alpha^{12}=(1111),\\ \alpha^1=(0010), & \alpha^5=(0110), & \alpha^9=(1010), & \alpha^{13}=(1101),\\ \alpha^2=(0100), & \alpha^6=(1100), & \alpha^{10}=(0111), & \alpha^{14}=(1001),\\ \alpha^3=(1000), & \alpha^7=(1011), & \alpha^{11}=(1110). & \end{array} \tag{14.8}$$

Полагая в (14.6) $t=2$ и $\alpha_i=\alpha^{i-1}$, получим следующую матрицу

$$\begin{pmatrix} 0&0&0&1&0&0&1&1&0&1&0&1&1&1&1\\ 0&0&1&0&0&1&1&0&1&0&1&1&1&1&0\\ 0&1&0&0&1&1&0&1&0&1&1&1&1&0&0\\ 1&0&0&0&1&0&0&1&1&0&1&0&1&1&1\\ 0&1&1&1&1&0&1&1&1&1&0&1&1&1&1\\ 0&0&1&0&1&0&0&1&0&1&0&0&1&0&1\\ 0&0&0&1&1&0&0&0&1&1&0&0&0&1&1\\ 1&0&0&0&1&1&0&0&0&1&1&0&0&0&1 \end{pmatrix}, \tag{14.9}$$

которая и будет проверочной матрицей БЧХ-кода длины 15, исправляющего две ошибки.

14.5 Декодирование БЧХ-кодов

Теперь рассмотрим процедуру исправления ошибок в БЧХ-кодах. Пусть конструктивное расстояние БЧХ-кода длины $n=2^m-1$ равно $2t+1$. Допустим, что при передаче кодового слова $\boldsymbol{a}$ произошло $k\le t$ ошибок в позициях, соответствующих элементам $\alpha_{j_1},\dots,\alpha_{j_k}$ поля $GF(2^m)$, и в результате было получено слово $\boldsymbol{b}$. Для $i=1,\dots,k$ положим $X_i=\alpha_{j_i}$. Величины X_i называются локаторами ошибок. Нетрудно видеть, что синдром S принятого слова $\boldsymbol{b}$ выражается через локаторы ошибок следующим образом:

$$S=\begin{pmatrix} X_1\oplus\dots\oplus X_k\\ X_1^3\oplus\dots\oplus X_k^3\\ \dots\dots\dots\dots\\ X_1^{2t-1}\oplus\dots\oplus X_k^{2t-1}\end{pmatrix}.$$

Далее суммы $X_1^r\oplus\dots\oplus X_k^r$ будем обозначать символами S_r, при этом величины S_r будем называть компонентами синдрома не только для нечетных значений r, но также и для четных.

Введем многочлен

$$\Lambda(x)=(1\oplus xX_1)(1\oplus xX_2)\cdots(1\oplus xX_k)=\Lambda_kx^k\oplus\Lambda_{k-1}x^{k-1}\oplus\dots\oplus\Lambda_1x\oplus 1,$$

называемый многочленом *локаторов ошибок*. Если многочлен $\Lambda(x)$ известен, то вычислив его корни можно определить произошедшие ошибки. Так как переменная x может принимать не более чем n различных значений, то корни $\Lambda(x)$ можно найти последовательно вычисляя значения $\Lambda(x)$ на

всех элементах поля $GF(2^m)$. Для этого достаточно выполнить в общей сложности $\mathcal{O}(tn)$ действий над элементами поля $GF(2^m)$.

Опишем алгоритм нахождения коэффициентов многочлена $\Lambda(x)$. Для этого выполним ряд предварительных действий. Начнем с того, что для каждого $i=1,\dots,k$ и каждого $j=1,\dots,k$ умножим многочлен локаторов ошибок на X_i^{k+j} и подставим вместо переменной x его корень X_i^{-1}. В результате получим систему равенств

$$\Lambda_kX_i^j\oplus\Lambda_{k-1}X_i^{j+1}\oplus\dots\oplus\Lambda_1X_i^{k+j-1}\oplus X_i^{k+j}=0, \tag{14.10}$$

где $i=1,\dots,k$ и $j=1,\dots,k$. Суммируя при фиксированном j равенства из (14.10) по всем i от 1 до k, получим, что для каждого $j=1,\dots,k$ имеет место следующее равенство

$$\bigoplus_{i=1}^{k}(\Lambda_kX_i^j\oplus\Lambda_{k-1}X_i^{j+1}\oplus\dots\oplus\Lambda_1X_i^{k+j-1}\oplus X_i^{k+j})=$$
$$=\Lambda_k\bigoplus_{i=1}^{k}X_i^j\oplus\Lambda_{k-1}\bigoplus_{i=1}^{k}X_i^{j+1}\oplus\dots\oplus\Lambda_1\bigoplus_{i=1}^{k}X_i^{k+j-1}\oplus\bigoplus_{i=1}^{k}X_i^{k+j}=0.$$

Теперь заметим, что суммы, умножаемые в последнем равенстве на коэффициенты Λ_i, являются компонентами синдрома, и поэтому имеет место следующая система равенств:

$$\Lambda_kS_j\oplus\Lambda_{k-1}S_{j+1}\oplus\dots\oplus\Lambda_1S_{k+j-1}\oplus S_{k+j}=0,\quad j=1,\dots,k. \tag{14.11}$$

Из равенств (14.11) составим систему линейных уравнений относительно коэффициентов Λ_i. Нетрудно видеть, что в матричной форме рассматриваемая система имеет следующий вид:

$$\begin{pmatrix} S_1&S_2&S_3&\dots&S_{k-1}&S_k\\ S_2&S_3&S_4&\dots&S_k&S_{k+1}\\ S_3&S_4&S_5&\dots&S_{k+1}&S_{k+2}\\ \vdots&\vdots&\vdots&\ddots&\vdots&\vdots\\ S_k&S_{k+1}&S_{k+2}&\dots&S_{2k-2}&S_{2k-1}\end{pmatrix}\begin{pmatrix}\Lambda_k\\ \Lambda_{k-1}\\ \Lambda_{k-2}\\ \vdots\\ \Lambda_1\end{pmatrix}=\begin{pmatrix}S_{k+1}\\ S_{k+2}\\ S_{k+3}\\ \vdots\\ S_{2k}\end{pmatrix}.$$

Покажем, что если произошло ровно k ошибок, то матрица этой системы не вырождена, а если меньше чем k, то вырождена. Прямой проверкой легко убедиться в том, что

$$\begin{pmatrix} S_1&S_2&\dots&S_k\\ S_2&S_3&\dots&S_{k+1}\\ \vdots&\vdots&\ddots&\vdots\\ S_k&S_{k+1}&\dots&S_{2k-1}\end{pmatrix}=\begin{pmatrix} X_1&X_2&\dots&X_k\\ X_1^2&X_2^2&\dots&X_k^2\\ \vdots&\vdots&\ddots&\vdots\\ X_1^k&X_2^k&\dots&X_k^k\end{pmatrix}\begin{pmatrix}1&X_1&\dots&X_1^{k-1}\\ 1&X_2&\dots&X_2^{k-1}\\ \vdots&\vdots&\ddots&\vdots\\ 1&X_k&\dots&X_k^{k-1}\end{pmatrix}.$$

Повторяя рассуждения, приведенные в доказательстве леммы 14.1, нетрудно показать, что первый сомножитель в правой части последнего равенства

вырожден тогда и только тогда, когда среди величин X_i есть либо две одинаковые, либо одна из них равна нулю, а второй сомножитель вырожден тогда и только тогда, когда среди величин X_i есть две одинаковых. Поэтому, если произошло ровно k ошибок, то все величины X_i ненулевые и различные, и, следовательно, произведение двух невырожденных матриц также будет невырожденной матрицей. С другой стороны, если число ошибок меньше k, то в этом случае хотя бы одна из величин X_i равна нулю. Таким образом, первый сомножитель, а вместе с ним и все произведение, будут вырожденными матрицами.

Теперь все готово для описания алгоритма декодирования БЧХ-кода. Пусть конструктивное расстояние БЧХ-кода длины $n = 2^m - 1$ с проверочной матрицей (14.6) равно $2t + 1$. Тогда для декодирования слова $\boldsymbol{a}$ выполняем следующие действия.

1. Вычисляем компоненты синдрома $S_1, S_2, \ldots, S_{2t}$ и полагаем $k = t$.

2. Составляем матрицу

$$\mathbf{S}_k = \begin{pmatrix} S_1 & S_2 & S_3 & \ldots & S_{k-1} & S_k \\ S_2 & S_3 & S_4 & \ldots & S_k & S_{k+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ S_k & S_{k+1} & S_{k+2} & \ldots & S_{2k-2} & S_{2k-1} \end{pmatrix}$$

и определяем, вырождена это матрица или нет.

3. Если матрица $\mathbf{S}_k$ вырождена, то уменьшаем k на единицу и возвращаемся к п. 2. Если матрица $\mathbf{S}_k$ невырождена, то определяем, что число ошибок равно k и переходим к следующему пункту.

4. Решая матричное уравнение

$$\begin{pmatrix} S_1 & S_2 & S_3 & \ldots & S_{k-1} & S_k \\ S_2 & S_3 & S_4 & \ldots & S_k & S_{k+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ S_k & S_{k+1} & S_{k+2} & \ldots & S_{2k-2} & S_{2k-1} \end{pmatrix} \begin{pmatrix} \Lambda_k \\ \Lambda_{k-1} \\ \vdots \\ \Lambda_1 \end{pmatrix} = \begin{pmatrix} S_{k+1} \\ S_{k+2} \\ \vdots \\ S_{2k} \end{pmatrix},$$

находим коэффициенты Λ_i многочлена локаторов ошибок.

5. Вычисляя последовательно значения многочлена

$$\Lambda(x) = \Lambda_k x^k \oplus \Lambda_{k-1} x^{k-1} \oplus \cdots \oplus \Lambda_1 x \oplus 1$$

на всех элементах поля $GF(2^m)$, находим его корни $x_1, \ldots, x_k$, обращая которые, находим локаторы ошибок $X_1, \ldots, X_k$. Затем из равенств $X_i = \alpha_{j_i}$ определяем позиции ошибок j_i.

Приведенный алгоритм декодирования называется алгоритмом Питерсона–Горенстейна–Цирлера. Нетрудно видеть, что для выполнения этого алгоритма достаточно выполнить $\mathcal{O}(t^4 + nt)$ действий над элементами поля $GF(2^m)$. Рассмотрим работу алгоритма на примере исправления ошибок в слове (001111111111111) БЧХ-кодом с проверочной матрицей (14.9).

1. Умножая матрицу (14.9) на вектор (001111111111111), вычислим первую и третью компоненты синдрома S_1 и S_3:

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

Затем при помощи (14.8) выражаем эти компоненты в виде степеней порождающего элемента α: $S_1 = \alpha^4$ и $S_3 = \alpha^{14}$. Далее находим $S_2 = \alpha^8$ и $S_4 = \alpha^{16} = \alpha$.

2–3. Так как

$$\det \begin{pmatrix} S_1 & S_2 \\ S_2 & S_3 \end{pmatrix} = S_1 S_3 \oplus S_2^2 = \alpha^4 \alpha^{14} \oplus \alpha^8 \alpha^8 = \alpha^3 \oplus \alpha = \alpha^9 \neq 0,$$

то заключаем, что произошло две ошибки.

4. Решим матричное уравнение $\begin{pmatrix} S_1 & S_2 \\ S_2 & S_3 \end{pmatrix} \begin{pmatrix} \Lambda_2 \\ \Lambda_1 \end{pmatrix} = \begin{pmatrix} S_3 \\ S_4 \end{pmatrix}$, преобразуя соответствующую расширенную матрицу в единичную. Так как

$$\begin{pmatrix} \alpha^4 & \alpha^8 & \alpha^{14} \\ \alpha^8 & \alpha^{14} & \alpha \end{pmatrix} \sim \begin{pmatrix} 1 & \alpha^4 & \alpha^{10} \\ 1 & \alpha^6 & \alpha^8 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & \alpha^4 & \alpha^{10} \\ 0 & \alpha^4 \oplus \alpha^6 & \alpha^{10} \oplus \alpha^8 \end{pmatrix} \sim \begin{pmatrix} 1 & \alpha^4 & \alpha^{10} \\ 0 & \alpha^{12} & \alpha \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & \alpha^4 & \alpha^{10} \\ 0 & 1 & \alpha^4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & \alpha^{10} \oplus \alpha^8 \\ 0 & 1 & \alpha^4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & \alpha \\ 0 & 1 & \alpha^4 \end{pmatrix},$$

то $\Lambda_2 = \alpha$ и $\Lambda_1 = \alpha^4$.

5. Вычисляя значения многочлена $\Lambda(x) = \alpha x^2 \oplus \alpha^4 x \oplus 1$ на всех элементах поля $\mathbb{Z}_2[\alpha]$, найдем его корни $x_1 = 1$ и $x_2 = \alpha^{14}$. Обращая корни, находим локаторы ошибок $X_1 = 1$ и $X_2 = \alpha$, которые, очевидно, соответствуют первой и второй позициям.

Более быстрые алгоритмы декодирования БЧХ-кодов можно найти например в [4] и [25].

14.6 Задачи

14.1. Пусть x — натуральноее число, не превосходящее 15. Какое минимальное количество вопросов надо задать, чтобы определить x, если на каждый вопрос дается ответ "да" или "нет", и на один из вопросов может быть дан неверный ответ? Сформулировать вопросы, определяющие x.

14.2. Описать порождающие матрицы кода Хемминга длины $2^m - 1$ и расширенного кода Хемминга длины 2^m.

14.3. Найти сложность вычисления синдрома расширенного кода Хемминга длины 2^m.

14.4. Показать, что для всех n при $k = 1$ и $k = n - 2$ коды Рида–Маллера $\mathrm{PM}(n, k)$ являются максимальными линейными кодами с расстояниями 2^{n-1} и 4.

14.5. Оценить сложность декодирования кода Рида–Маллера длины 2^m порядка k.

14.6. Используя код Рида–Маллера длины 2^m порядка k, построить исправляющий $2^{m-k-1} - 1$ ошибок код длины $2^m - 1$.

14.7. Написать проверочную матрицу линейного $(16, 7)$-кода, исправляющего две и обнаруживающего три ошибки.

14.8. Написать проверочную матрицу БЧХ-кода длины 15 с конструктивным расстоянием 7 и найти ее ранг.

14.9. Написать проверочную матрицу БЧХ-кода длины 31 с конструктивным расстоянием 5 и найти ее ранг.

Лекция 15

Полиномиальные коды

Нетрудно видеть, что между линейным пространством двоичных последовательностей длины n и линейным пространством многочленов степени не выше $n-1$ над $\mathbb{Z}_2$ существует изоморфизм, ставящий в соответствие многочлену вектор его коэффициентов. Такой изоморфизм естественным образом приводит к понятию полиномиального кода. Код G называется *полиномиальным кодом* длины n с *порождающим многочленом* $h(x)$, если множество многочленов, соответствующих элементам кода G, состоит из всех тех многочленов степени не выше $n-1$, каждый из которых без остатка делится на многочлен $h(x)$.

15.1 Полиномиальные БЧХ-коды

Покажем, что примитивные БЧХ-коды являются полиномиальными кодами. Пусть $n = 2^m - 1$ и α — примитивный элемент поля $GF(2^m)$. Тогда матрица

$$\mathbf{H} = \begin{pmatrix} 1 & \alpha & \alpha^2 & \dots & \alpha^{n-1} \\ 1 & \alpha^3 & \alpha^{2\cdot 3} & \dots & \alpha^{(n-1)3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha^{2t-1} & \alpha^{2(2t-1)} & \dots & \alpha^{(n-1)(2t-1)} \end{pmatrix} \tag{15.1}$$

будет проверочной матрицей примитивного БЧХ-кода G с конструктивным расстоянием $2t+1$. Каждый элемент $\mathbf{g} = (g_0, \dots, g_{n-1})$ этого кода будем рассматривать как набор коэффициентов многочлена $g(x) = \bigoplus_{i=0}^{n-1} g_i x^i$. Так как

$$\begin{pmatrix} 1 & \alpha & \dots & \alpha^{n-1} \\ 1 & \alpha^3 & \dots & \alpha^{(n-1)3} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha^{2t-1} & \dots & \alpha^{(n-1)(2t-1)} \end{pmatrix} \begin{pmatrix} g_0 \\ g_1 \\ \vdots \\ g_{n-1} \end{pmatrix} = \begin{pmatrix} g(\alpha) \\ g(\alpha^3) \\ \vdots \\ g(\alpha^{2t-1}) \end{pmatrix},$$

то нетрудно видеть, что для любого многочлена $g(x)$, соответствующего элементу $\mathbf{g}$ рассматриваемого БЧХ-кода G, справедливы равенства

$$g(\alpha) = g(\alpha^3) = \dots = g(\alpha^{2t-1}) = 0. \tag{15.2}$$

Пусть $h_k(x)$ — минимальный многочлен элемента α^k, где $k = 1, 3, \ldots, 2t-1$. Тогда многочлен $h(x)$, являющийся наименьшим общим кратным минимальных многочленов $h_1, h_3, \ldots, h_{2t-1}$, будет делить многочлен, соответствующий любому элементу кода G. Верно и обратное — любой многочлен степени $n-1$, делящийся на $h(x)$, будет соответствовать элементу БЧХ-кода G с проверочной матрицей (15.1). Многочлен $h(x)$ называется *порождающим* многочленом рассматриваемого кода. Порождающий многочлен можно использовать для кодирования сообщений следующим образом. Передаваемое сообщение делится на блоки длины $n - \deg h(x)$. Затем каждый блок $\boldsymbol{a}$ рассматривается как набор коэффициентов многочлена $a(x)$ степени $n - \deg h(x) - 1$, и многочлен $a(x)$ умножается на порождающий многочлен $h(x)$. Произведение $g(x) = a(x)h(x)$ называется кодовым многочленом, и его коэффициенты образуют элемент кода. При декодировании полученное сообщение делится на блоки длины n. Каждый блок рассматривается как набор коэффициентов многочлена $b(x)$, являющегося суммой кодового многочлена $g(x)$ и многочлена ошибок $c(x)$, коэффициенты которого являются компонентами вектора ошибок $\mathbf{c}$. Так как $g(\alpha^k) = 0$ для $k = 1, 3, \ldots, 2t-1$, то

$$b(\alpha^k) = g(\alpha^k) \oplus c(\alpha^k) = c(\alpha^k) = S_k,$$

и далее декодирование производится в соответствии с приведенным выше алгоритмом Питерсона–Горинстейна–Цирлера.

Воспользуемся понятием порождающего многочлена и дадим новое определение примитивных БЧХ-кодов. Пусть $n = 2^m - 1$, α — примитивный элемент поля $GF(2^m)$, $h(x)$ — наименьшее общее кратное минимальных многочленов элементов $\alpha, \alpha^3, \ldots, \alpha^{2t-1}$ поля $GF(2^m)$. *Примитивным БЧХ-кодом* длины n с конструктивным расстоянием $2t+1$ будем называть полиномиальный код длины n с порождающим многочленом $h(x)$.

Построим порождающий многочлен БЧХ-кода длины 15, исправляющего три ошибки. В качестве поля из 16 элементов как и ранее возьмем поле $\mathbb{Z}_2[\alpha]$, где α — корень примитивного многочлена $x^4 \oplus x \oplus 1$. Напомним, что степени примитивного элемента α этого поля выглядят следующим образом:

$$\begin{array}{llll}
\alpha^0 = (0001), & \alpha^4 = (0011), & \alpha^8 = (0101), & \alpha^{12} = (1111),\\
\alpha^1 = (0010), & \alpha^5 = (0110), & \alpha^9 = (1010), & \alpha^{13} = (1101),\\
\alpha^2 = (0100), & \alpha^6 = (1100), & \alpha^{10} = (0111), & \alpha^{14} = (1001),\\
\alpha^3 = (1000), & \alpha^7 = (1011), & \alpha^{11} = (1110). &
\end{array}$$

Так как элемент α является корнем примитивного многочлена $x^4 \oplus x \oplus 1$, то для построения порождающего многочлена достаточно найти минимальные многочлены элементов α^3 и α^5 и умножить их произведение на многочлен $x^4 \oplus x \oplus 1$. Нетрудно видеть, что элементами сопряженными с α^3 будут элементы α^6, α^{12} и $\alpha^{24} = \alpha^9$. Поэтому минимальный многочлен для α^3 можно получить раскрыв скобки в произведении

$$(x \oplus \alpha^3)(x \oplus \alpha^6)(x \oplus \alpha^{12})(x \oplus \alpha^9).$$

Сначала раскроем скобки в двух первых сомножителях:

$$(x \oplus \alpha^3)(x \oplus \alpha^6) = x^2 \oplus (\alpha^6 \oplus \alpha^3)x \oplus \alpha^9 = x^2 \oplus \alpha^2 x \oplus \alpha^9.$$

Затем умножим полученный многочлен на третий сомножитель:

$$\begin{aligned}(x^2 \oplus \alpha^2 x \oplus \alpha^9)(x \oplus \alpha^{12}) &= x^3 \oplus (\alpha^{12} \oplus \alpha^2)x^2 \oplus (\alpha^{14} \oplus \alpha^9)x \oplus \alpha^{21} = \\ &= x^3 \oplus \alpha^7 x^2 \oplus \alpha^4 x \oplus \alpha^6.\end{aligned}$$

Наконец, последнее произведение дает окончательный результат:

$$\begin{aligned}(x^3 \oplus \alpha^7 x^2 \oplus \alpha^4 x \oplus \alpha^6)(x \oplus \alpha^9) &= \\ &= x^4 \oplus (\alpha^9 \oplus \alpha^7)x^3 \oplus (\alpha^{16} \oplus \alpha^4)x^2 \oplus (\alpha^{13} \oplus \alpha^6)x \oplus \alpha^{15} = \\ &= x^4 \oplus x^3 \oplus x^2 \oplus x \oplus 1.\end{aligned}$$

Теперь заметим, что $\alpha^{20} = \alpha^5$. Поэтому единственным элементом сопряженным с α^5 будет α^{10}. Тогда минимальный многочлен для α^5 находится следующим образом

$$(x \oplus \alpha^5)(x \oplus \alpha^{10}) = x^2 \oplus (\alpha^5 \oplus \alpha^{10})x \oplus \alpha^{15} = x^2 \oplus x \oplus 1.$$

Перемножая минимальные многочлены $x^4 \oplus x \oplus 1$, $x^4 \oplus x^3 \oplus x^2 \oplus x \oplus 1$ и $x^2 \oplus x \oplus 1$, получаем порождающий многочлен

$$x^{10} \oplus x^8 \oplus x^5 \oplus x^4 \oplus x^2 \oplus x \oplus 1$$

БЧХ-кода длины 15, исправляющего три ошибки.

15.2 Размерность примитивных БЧХ-кодов

Из теоремы 14.2 следует, что размерность примитивного БЧХ-кода длины $n = 2^m - 1$ с конструктивным расстоянием $2t+1$ не меньше, чем $n - t\log_2(n+1)$. Если $t \geq n/\log_2(n+1)$, то теорема 14.2 не гарантирует существование БЧХ-кода с расстоянием $2t+1$. В действительности при больших t степень наименьшего общего кратного минимальных многочленов $m_{\alpha^i}(x)$ первых t нечетных степеней порождающего элемента α мультипликативной группы поля становится меньше суммы степеней $m_{\alpha^i}(x)$, и как следствие, размерность кода становится больше разности $n - t\log_2(n+1)$. То, как это происходит рассмотрим на примере БЧХ-кодов длины 31.

Пусть α — примитивный элемент поля $GF(32)$. БЧХ-код с расстоянием 3 порождается минимальным многочленом элемента α, код с расстоянием 5 — наименьшим общим кратным минимальных многочленов элементов α и α^3 и т. д., вплоть до кода с расстоянием 15, который порождается наименьшим общим кратным минимальных многочленов элементов $\alpha, \alpha^3, \ldots, \alpha^{29}$. Для того, чтобы получить информацию об этих порождающих многочленах запишем все неравные нулю и единице элементы поля $GF(32)$ в таблицу,

помещая сопряженные элементы в одной строке. В первую строку таблицы поместим элементы, сопряженные с α, во вторую — с α^3 и т. д. В процессе заполнения очередной строки будем подчеркивать все появляющиеся в этой строке нечетные степени, исключая ту степень, с которой начинается заполнение. Нетрудно видеть, что каждая подчеркнутая степень $\underline{2k-1}$ означает, что далее в таблице не будет строки, начинающейся с элемента α^{2k-1}, так как все сопряженные с α^{2k-1} элементы находятся в той же строке, что и сам элемент α^{2k-1}. В результате получим следующее разбиение

$$\begin{array}{ccccc} \alpha & \alpha^2 & \alpha^4 & \alpha^8 & \alpha^{16} \\ \alpha^3 & \alpha^6 & \alpha^{12} & \alpha^{24} & \alpha^{\underline{17}} \\ \alpha^5 & \alpha^{10} & \alpha^{20} & \alpha^{\underline{9}} & \alpha^{18} \\ \alpha^7 & \alpha^{14} & \alpha^{28} & \alpha^{\underline{25}} & \alpha^{\underline{19}} \\ \alpha^{11} & \alpha^{22} & \alpha^{\underline{13}} & \alpha^{26} & \alpha^{\underline{21}} \\ \alpha^{15} & \alpha^{30} & \alpha^{\underline{29}} & \alpha^{\underline{27}} & \alpha^{\underline{23}} \end{array} \qquad (15.3)$$

не равных нулю и единице элементов поля $GF(32)$ на классы сопряженных элементов. Из этого разбиения, в частности, видно, что для кода с конструктивным расстоянием 9 его истинное кодовое расстояние будет не меньше 11, так как элемент α^9 сопряжен с элементом α^5 в силу чего эти элементы имеют общий минимальный многочлен.

Пусть $m_\beta(x)$ — минимальный многочлен элемента β, d_K — конструктивное расстояние кода, d — кодовое расстояние, $h(x)$ — порождающий многочлен кода. Из разбиения (15.3) нетрудно извлечь информацию о конструктивных расстояниях d_K, нижних оценках минимальных расстояний d, порождающих многочленах $h(x)$ и размерностях БЧХ-кодов длины 31. Эта информация представлена в таблице 15.1. Из таблицы видно, что существует состоящий из 64 элементов БЧХ-код длины 31 с кодовым расстоянием 15, т. е. с расстоянием, равным примерно половине длины кода. Далее покажем, что аналогичные БЧХ-коды существуют при всех значениях m. Точнее справедлива следующая теорема.

Таблица 15.1

d_K	$d \geq$	$h(x)$	dim
3	3	$m_\alpha(x)$	26
5	5	$m_\alpha(x) \cdot m_{\alpha^3}(x)$	21
7	7	$m_\alpha(x) \cdot m_{\alpha^3}(x) \cdot m_{\alpha^5}(x)$	16
9	11	$m_\alpha \cdot (x) m_{\alpha^3}(x) \cdot m_{\alpha^5}(x) \cdot m_{\alpha^7}(x)$	11
11	11	$m_\alpha(x) \cdot m_{\alpha^3}(x) \cdot m_{\alpha^5}(x) \cdot m_{\alpha^7}(x)$	11
13	15	$m_\alpha(x) \cdot m_{\alpha^3}(x) \cdot m_{\alpha^5}(x) \cdot m_{\alpha^7}(x) \cdot m_{\alpha^{11}}(x)$	6
15	15	$m_\alpha(x) \cdot m_{\alpha^3}(x) \cdot m_{\alpha^5}(x) \cdot m_{\alpha^7}(x) \cdot m_{\alpha^{11}}(x)$	6
17	31	$m_\alpha(x) \cdot m_{\alpha^3}(x) \cdot m_{\alpha^5}(x) \cdot m_{\alpha^7}(x) \cdot m_{\alpha^{11}}(x) \cdot m_{\alpha^{15}}(x)$	1

Теорема 15.1. *Размерность примитивного БЧХ-кода длины $n = 2^m - 1$ с конструктивным расстоянием $d = 2^{m-1} - 1$ равна $m + 1$. Минимальное расстояние этого кода совпадает с его конструктивным расстоянием.*

Доказательство. Пусть α — примитивный элемент поля $GF(2^m)$. Корнями порождающего многочлена $h(x)$ БЧХ-кода G должны быть элементы $\alpha, \alpha^3, \ldots, \alpha^{2t-1}$ и все сопряженные с ними. Пусть $k \in \{0, 1, 2, \ldots, 2^m - 1\}$ и $\boldsymbol{k}$ — набор длины m, составленный из разрядов двоичного представления k. В поле $GF(2^m)$ рассмотрим произвольный элемент α^k и сопряженные с ним элементы $\alpha^{k_1}, \ldots, \alpha^{k_s}$. Легко видеть, что наборы $\boldsymbol{k}_1, \ldots, \boldsymbol{k}_s$ будут циклическими сдвигами набора $\boldsymbol{k}$. Поэтому корнями порождающего многочлена $h(x)$ будут все такие элементы α^k, для которых соответствующий набор $\boldsymbol{k}$ является циклическим сдвигом хотя бы одного из наборов $\mathbf{1}, \mathbf{3}, \ldots, \boldsymbol{2t-1}$. Рассмотрим множество A, состоящее из всех двоичных ненулевых наборов длины m с не более чем $m-2$ единицами. Заметим, что для любого нечетного k, не превосходящего $2t - 1 = 2^m - 3$, соответствующий ему набор $\boldsymbol{k}$ принадлежит A вместе со всеми своими циклическими сдвигами. Верно и обратное — у каждого $\boldsymbol{k}$ из A хотя бы одному его сдвигу соответствует нечетное k, не превосходящее $2t - 1$. Поэтому степень $h(x)$ равна числу наборов в A, т. е.

$$\deg h(x) = \sum_{i=1}^{m-2} \binom{n}{i} = 2^m - m - 2.$$

Теперь первое утверждение теоремы следует из того, что размерность G равна разности его длины и степени порождающего многочлена.

Для доказательства равенства конструктивного и минимальных расстояний заметим, что порождающий многочлен $h(x)$ является делителем двучлена $x^{2^m-1} \oplus 1$ и при этом не делится на $x \oplus 1$. Поэтому найдется такой многочлен $g(x)$, что $h(x)g(x) = \bigoplus_{i=0}^{2^m-2} x^i$. Следовательно, набор из одних единиц принадлежит рассматриваемому БЧХ-коду, т. е. вместе с любым набором $\boldsymbol{a}$ коду принадлежит и его покомпонентное отрицание $\overline{\boldsymbol{a}}$. Таким образом, минимальное расстояние кода не превосходит половины его длины. Теорема доказана.

В двух следующих теоремах распространим утверждение теоремы 15.1 на примитивные БЧХ-коды, конструктивные кодовые расстояния которых равны $2^p - 1$ при различных значениях натурального параметра p. Для доказательства первой из этих теорем потребуется следующее известное утверждение об определителе матрицы Вандермонда

$$V_n = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \end{pmatrix}.$$

Лемма 15.1. *Для определителя матрицы Вандермонда V_n справедливо равенство*

$$\det V_n = \prod_{1 \leq j < i \leq n} (x_i - x_j). \qquad (15.4)$$

Доказательство. Лемму докажем индукцией по n. В основание индукции положим очевидный случай $n = 2$. Допустим, что утверждении леммы верно при $n \le m$. Покажем, что оно справедливо и при $n = m + 1$.

Прежде всего заметим, что если $x_i = x_j$, то матрица V_{m+1} будет содержать два одинаковых столбца, и, следовательно, ее определитель будет равен нулю. Поэтому в этом случае равенство (15.4) верно. Далее будем полагать, что все x_i различны.

В матрице V_{m+1} заменим x_{m+1} переменной x, после чего определитель разложим по последнему столбцу. Нетрудно видеть, что в этом случае определитель преобразованной матрицы будет многочленом степени m от x:

$$D(x) = \det \begin{pmatrix} 1 & 1 & \dots & 1 & 1 \\ x_1 & x_2 & \dots & x_m & x \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_1^m & x_2^m & \dots & x_m^m & x^m \end{pmatrix} = d_m x^m + \ldots + d_1 x + d_0,$$

где коэффициент d_m — определитель матрицы V_m. Многочлен $D(x)$ имеет ровно m корней — $x_1, \ldots, x_m$. Поэтому

$$D(x) = d_m (x - x_1)(x - x_2) \cdots (x - x_m).$$

Так как по предположению индукции $d_m = \prod_{1 \le j < i \le m} (x_i - x_j)$, то

$$\det V_{m+1} = D(x_{m+1}) = \prod_{1 \le j < i \le m} (x_i - x_j) \prod_{1 \le i \le 1} (x_{m+1} - x_i) = \prod_{1 \le j < i \le m+1} (x_i - x_j).$$

Лемма доказана.

Теорема 15.2. *Пусть $n = 2^m - 1$, $d = 2^p - 1$, $p = 1, \ldots, m-2$. Тогда для размерности $k(n,d)$ примитивного БЧХ-кода длины n с конструктивным расстоянием d справедливы равенства*

$$k(n,d) = \begin{cases} n - \frac{d-1}{2} \log_2(n+1), & \text{при } d+1 \le \sqrt{n+1}, \\ n \left(1 - \frac{d\varphi(n,d)}{2n}\right)^{\log_2 n} (1 + o(1)), & \text{при } d+1 > \sqrt{n+1}, \end{cases}$$

где $1 < \varphi(n,d) < 2$ и $\varphi(n,d) \to 1$ при $n/d \to \infty$.

Доказательство. Рассмотрим БЧХ-код G_1 длины $n = 2^m - 1$ с конструктивным расстоянием $d_1 = 2^p - 1$ и порождающим многочленом $h_1(x)$ и оценим его размерность, равную $2^m - 1 - \deg h_1(x)$. Корнями $h_1(x)$ будут все такие степени α^k примитивного элемента α, для которых соответствующий набор $\boldsymbol{k}$ является циклическим сдвигом хотя бы одного из наборов $\mathbf{1}, \mathbf{3}, \ldots, \boldsymbol{d_1 - 2}$. Представим $d_1 = 2^p - 1$ в виде двоичного m-разрядного числа. Легко видеть, что у этого числа старшие $l = m - p$ разрядов будут нулевыми. Также легко видеть, что ни один из циклических сдвигов наборов $\mathbf{1}, \mathbf{3}, \ldots, \boldsymbol{d_1 - 2}$ не совпадает ни с одним из циклических сдвигов набора $\boldsymbol{d_1}$. Поэтому $\deg h_1(x) = \deg h(x) - m$, где $h(x)$ — порождающий многочлен БЧХ-кода G длины n с конструктивным расстоянием $d = 2^p + 1$. Далее будем оценивать размерность кода G, зная которую, легко найдем и размерность кода G_1.

Заметим, что при $m \le 2l$ в любом ненулевом двоичном наборе длины m присутствует не более одного максимального блока из l или более стоящих рядом нулей. Поэтому все циклические сдвиги наборов $\mathbf{1}, \mathbf{3}, \ldots, \boldsymbol{d-2}$ различны, и, следовательно, $\deg h(x) = m \cdot \frac{d-1}{2}$. Очевидно, что в этом случае размерность рассматриваемого БЧХ-кода удовлетворяет равенству

$$k(n,d) = n - \frac{d-1}{2} \log_2(n+1). \tag{15.5}$$

Далее полагаем, что $m > 2l$. Легко видеть, что множество корней порождающего многочлена $h(x)$ совпадает с множеством таких степеней k элемента α, что либо в наборе $\boldsymbol{k}$ есть l стоящих рядом нулей, либо набор $\boldsymbol{k}$ можно представить в виде $0^s 1 \boldsymbol{a} 10^t$, где $1 \le s, t < l$, $s + t \ge l$ и в наборе $\boldsymbol{a}$ нет l стоящих рядом нулей. Пусть $N_{m,l}$ — мощность множества $\mathcal{N}_{m,l}$ двоичных наборов длины m, ни в одном из которых нет l стоящих рядом нулей. Так как число наборов вида $0^s 1 \boldsymbol{a} 10^t$ равно

$$(l-1)N_{m-l-2,l} + (l-2)N_{m-l-3,l} + \cdots + 2N_{m-2l+1,l} + N_{m-2l,l},$$

то $\deg h(x) = 2^m - 1 - N_{m,l} + \sum_{i=1}^{l-1} (l-i) N_{m-l-1-i,l}$, и для размерности $C_{m,l} = k(n,d)$ кода G при $m > 2l$ справедливо равенство

$$C_{m,l} = N_{m,l} - \sum_{i=1}^{l-1} (l-i) N_{m-l-1-i,l}. \tag{15.6}$$

Найдем $N_{m,l}$. Нетрудно видеть, что множество $\mathcal{N}_l$ двоичных наборов, ни в одном из которых нет l стоящих рядом нулей, порождается формулой

$$1^*((0 \cup 0^2 \cup \cdots \cup 0^{l-1})11^*)^*(\lambda \cup 0 \cup 0^2 \cup \cdots \cup 0^{l-1}).$$

Пусть вес $w(\boldsymbol{a})$ набора $\boldsymbol{a}$ равен его длине. Тогда для производящей функции $F^w_{\mathcal{N}_l}(x)$ множества $\mathcal{N}_l$ справедливо равенство

$$F^w_{\mathcal{N}_l}(x) = \frac{1}{1-x} \cdot \frac{1}{1 - \frac{x(x + x^2 + \cdots + x^{l-1})}{1-x}} \cdot (1 + x + x^2 + \cdots + x^{l-1}),$$

которое после несложных преобразований принимает следующий вид:

$$F^w_{\mathcal{N}_l}(x) = \frac{1 + x + x^2 + \cdots + x^{l-1}}{1 - x - x^2 - \cdots - x^l}.$$

Поэтому величина $N_{m,l}$ будет линейной комбинацией квазимногочленов, показатели которых будут корнями многочлена $g(y) = y^l - y^{l-1} - \cdots - y - 1$. При $l = 2$ этот многочлен превращается в трехчлен $y^2 - y - 1$ с корнями $(1+\sqrt{5})/2$ и $(1-\sqrt{5})/2$, и поэтому $N_{m,2} = c_0((1+\sqrt{5})/2)^m + c_1((1-\sqrt{5})/2)^m$,

где c_0 и c_1 — константы. Так как $(1-\sqrt{5})/2 < 1$, то $N_{m,2} \sim c_1((1+\sqrt{5})/2)^m$. Далее покажем, что при любом $l \geq 3$ все корни многочлена $g(y)$ различны, и только один из этих корней по модулю больше единицы.

Вместо многочлена $g(y) = y^l - y^{l-1} - \cdots - y - 1$ будем рассматривать многочлен $f(y) = y^{l+1} - 2y^l + 1$, получающийся из исходного умножением на $y-1$. Нетрудно видеть, что единственный ненулевой корень $f'(y)$ не будет корнем $f(y)$. Поэтому $f(y)$ не имеет кратных корней. Также легко видеть (см. рис. 15.1), что у $f(y)$ есть только один положительный корень больший единицы. Покажем, что любой неположительный (комплексный или отрицательнй) корень $f(y)$ по модулю меньше единицы. Пусть a — неположительный корень многочлена $f(y)$, $\boldsymbol{u}$ — вектор, представляющий на комплексной плоскости число a^{l+1}, $\boldsymbol{v}$ — вектор, представляющий на комплексной плоскости число $2a^l$, $\boldsymbol{w}$ — вектор, представляющий на комплексной плоскости действительную единицу. Тогда $\boldsymbol{u} + \boldsymbol{w} = \boldsymbol{v}$ и длины этих векторов связаны неравенством $\|\boldsymbol{u}\| + \|\boldsymbol{w}\| > \|\boldsymbol{v}\|$. Следовательно, $|a|^{l+1} > 2|a|^l - 1$. С другой стороны, a является корнем многочлена $y^l - y^{l-1} - \cdots - y - 1$. Рассмотрим векторы $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_l$, где $\boldsymbol{u}_i$ представляет на комплексной плоскости число a^i. Тогда $\boldsymbol{u}_l = \boldsymbol{u}_{l-1} + \cdots + \boldsymbol{u}_1 + \boldsymbol{w}$ и $\|\boldsymbol{u}_l\| < \|\boldsymbol{u}_{l-1}\| + \cdots + \|\boldsymbol{u}_1\| + \|\boldsymbol{w}\|$. При $|a| > 1$ последнее неравенство после умножения на $|a|-1$ превращается в неравенство $|a|^{l+1} < 2|a|^l - 1$, которое, очевидно, противоречит полученному выше неравенству $|a|^{l+1} > 2|a|^l - 1$. Таким образом, любой неположительный корень $f(y)$, а, следовательно, и корень $g(y)$ по модулю меньше единицы. Таким образом,

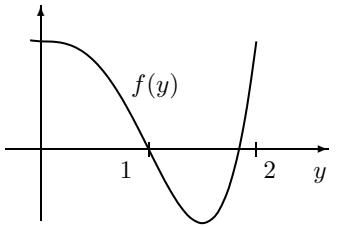


Рис. 15.1

$$N_{m,l} = c_0 y_0^m + c_1 y_1^m + \cdots + c_{l-1} y_{l-1}^m, \tag{15.7}$$

где c_i — константы, y_0 — единственный положительный корень многочлена $f(y)$ больший единицы, $|y_i| < 1$ при $i = 1, \ldots, l-1$.

Выразим константы c_i через корни y_i. Пусть $N_{0,l} = 1$. Так как $N_{m,l} = 2^l$ при $m < l$, то для определения c_i надо решить матричное уравнение

$$\begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \\ y_0 & y_1 & y_2 & \ldots & y_{l-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_0^{l-1} & y_1^{l-1} & y_2^{l-1} & \ldots & y_{l-1}^{l-1} \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{l-1} \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ \vdots \\ 2^{l-1} \end{pmatrix} \tag{15.8}$$

с матрицей Вандермонда в левой части. Уравнение (15.8) решим, найдя c_i при помощи формул Крамера и леммы 15.1. Затем сократим в числителе и знаменателе получившихся дробей одинаковые множители. Преобразуя оставшиеся множители при помощи формул

$$g(y) = \prod_{i=0}^{l-1}(y - y_i), \qquad g'(y) = \sum_{i=0}^{l-1}\prod_{j\neq i}(y - y_j)$$

и учитывая равенство $g(2) = 1 = 2y_i^l - y_i^{l+1}$, получим, что

$$c_i = \frac{\prod_{j\neq i}(2 - y_j)}{\prod_{j\neq i}(y_i - y_j)} = \frac{g(2)}{(2 - y_i)g'(y_i)} = \frac{y_i^l}{g'(y_i)}. \tag{15.9}$$

Теперь найдем значение $C_{m,l} = N_{m,l} - \sum_{j=1}^{l-1}(l-j)N_{m-l-1-j,l}$. Из (15.7) и (15.9) следует, что

$$\begin{aligned} C_{m,l} &= \sum_{i=1}^{l-1}\Big(c_i y_i^m - \sum_{j=1}^{l-1}(l-j)c_i y_i^{m-l-1-j}\Big) = \\ &= \sum_{i=1}^{l-1} c_i y_i^m \Big(1 - y_i^{-2l}\sum_{j=1}^{l-1}(l-j)y_i^{l-j-1}\Big) = \\ &= \sum_{i=1}^{l-1} y_i^m \,\frac{y_i^l - y_i^{-l}\sum_{j=1}^{l-1}(l-j)y_i^{l-j-1}}{g'(y_i)}. \end{aligned}$$

Так как $g'(y) = ly^{l-1} - \sum_{j=1}^{l-1}(l-j)y^{l-j-1}$, то

$$C_{m,l} = \sum_{i=1}^{l-1} y_i^m \,\frac{y_i^l - y_i^{-l}(ly_i^{l-1} - g'(y_i))}{g'(y_i)} = \sum_{i=1}^{l-1} y_i^m \left(\frac{y_i^l - ly_i^{-1}}{g'(y_i)} - y_i^{-l}\right). \tag{15.10}$$

Упростим дробь, стоящую в правой части равенства (15.10). Для этого заметим, что значения производных $g(y)$ и $f(y) = y^{l+1} - 2y^l + 1$ в корнях y_i многочлена $g(y)$ связаны равенством $g'(y_i) = f'(y_i)/(y_i - 1)$, и, кроме того,

$$y_i f'(y_i) = (l+1)y_i^{l+1} - 2ly_i^l = 2(l+1)y_i^l - (l+1) - 2ly_i^l = 2y_i^l - l - 1.$$

Следовательно, $g'(y_i) = (2y_i^l - l - 1)/y_i(y_i - 1)$. Подставив это выражение в дробь из (15.10), видим, что

$$\begin{aligned} \frac{y_i^l - ly_i^{-1}}{g'(y_i)} &= \frac{(y_i^l - ly_i^{-1})y_i(y_i - 1)}{2y_i^l - l - 1} = \\ &= \frac{(y_i^{l+1} - l)(y_i - 1)}{2y_i^l - l - 1} = \frac{(2y_i^l - l - 1)(y_i - 1)}{2y_i^l - l - 1} = y_i - 1. \end{aligned}$$

Таким образом,

$$C_{m,l} = \sum_{i=1}^{l-1} y_i^m(y_i - 1 - y_i^{-l}) = \sum_{i=1}^{l-1} y_i^m \cdot y_i^{-l}(y_i^{l+1} - y_i^l - 1) = \sum_{i=1}^{l-1} y_i^m.$$

Определим y_0 — единственный корень $g(y)$, модуль которого больше единицы. Для этого введем функции $\varphi_1(k,l)$ и $\varphi_2(k,l)$, положив

$$\varphi_1(1,l) = \Big(1 - \frac{1}{2^{l+1}}\Big)^{-l}, \qquad \varphi_1(k,l) = \Big(1 - \frac{1}{2^{l+1}} \cdot \varphi_1(k-1,l)\Big)^{-l},$$

$$\varphi_2(1,l)=\Big(1-\frac{1}{2^l}\Big)^{-l},\qquad \varphi_2(k,l)=\Big(1-\frac{1}{2^{l+1}}\cdot\varphi_2(k-1,l)\Big)^{-l}.$$

Покажем, что функция $\varphi_1(k,l)$ возрастает по k, а функция $\varphi_2(k,l)$ убывает по k при любом $l\ge 2$. Заметим, что

$$\varphi_1(2,l)=\Big(1-\frac{1}{2^{l+1}}\Big(1-\frac{1}{2^{l+1}}\Big)^{-l}\Big)^{-l}>\varphi_1(1,l)=\Big(1-\frac{1}{2^{l+1}}\Big)^{-l}\qquad(15.11)$$

так как неравенство

$$\frac{1}{2^{l+1}}\Big(1-\frac{1}{2^{l+1}}\Big)^{-l}>\frac{1}{2^{l+1}}$$

справедливо для величин, вычитаемых из единиц в левой и правой частях неравенства (15.11). Используем неравенство (15.11) как основание индукции для доказательства возрастания функции $\varphi_1(k,l)$ по k. Допустим, что $\varphi_1(k,l)>\varphi_1(k-1,l)$, тогда

$$\varphi_1(k+1,l)=\Big(1-\frac{1}{2^{l+1}}\cdot\varphi_1(k,l)\Big)^{-l}>\varphi_1(k,l)=\Big(1-\frac{1}{2^{l+1}}\cdot\varphi_1(k-1,l)\Big)^{-l}.$$

Таким образом $\varphi_1(k,l)$ возрастает по k.

Теперь покажем, что $\varphi_2(k,l)$ убывает по k. Сначала покажем, что

$$\frac{1}{2^{l+1}}\Big(1-\frac{1}{2^l}\Big)^{-l}<\frac{1}{2^l}.\qquad(15.12)$$

Нетрудно видеть, что (15.12) эквивалентно неравенству $\big(\frac{2^l}{2^l-1}\big)^l<2$. Так как $\binom{l}{i}<(2^l-1)$, то при $l\ge 3$

$$\Big(\frac{2^l}{2^l-1}\Big)^l=\Big(1+\frac{1}{2^l-1}\Big)^l=1+\frac{l}{2^l-1}+\sum_{i=2}^{l}\binom{l}{i}(2^l-1)^{-i}\le$$
$$\le 1+\frac{3}{7}+\sum_{i=2}^{l}(2^l-1)^{-i+1}\le 1+\frac{3}{7}+\sum_{i=2}^{l}7^{-i+1}\le\sum_{i=0}^{l}2^{-i}<2.$$

Неравенство (15.12) доказано. Из (15.12) легко следует, что

$$\varphi_2(2,l)=\Big(1-\frac{1}{2^{l+1}}\Big(1-\frac{1}{2^l}\Big)^{-l}\Big)^{-l}<\varphi_2(1,l)=\Big(1-\frac{1}{2^l}\Big)^{-l}<2.\qquad(15.13)$$

Допустим, что $\varphi_2(k,l)<\varphi_2(k-1,l)$. В этом случае

$$\varphi_2(k+1,l)=\Big(1-\frac{1}{2^{l+1}}\cdot\varphi_2(k,l)\Big)^{-l}<\varphi_2(k,l)=\Big(1-\frac{1}{2^{l+1}}\cdot\varphi_2(k-1,l)\Big)^{-l}.$$

Следовательно, $\varphi_2(k,l)$ убывает по k.

Положим

$$\psi_1(k,l)=2\Big(1-\frac{1}{2^{l+1}}\cdot\varphi_1(k,l)\Big),\qquad \psi_2(k,l)=2\Big(1-\frac{1}{2^{l+1}}\cdot\varphi_2(k,l)\Big).$$

Из монотонности функций $\varphi_i(k,l)$ следует, что функция $\psi_1(k,l)$ убывает по k, а функция $\psi_1(k,l)$ возрастает по k. Так как $0<\psi_i(k,l)<2$, то, очевидно, существуют пределы $\lim_{k\to\infty}\psi_1(k,l)=\psi_1(l)$ и $\lim_{k\to\infty}\psi_2(k,l)=\psi_2(l)$.

Покажем, что

$$f(\psi_2(k,l))<f(y_0)=0<f(\psi_1(k,l)).\qquad(15.14)$$

Сначала докажем правое неравенство. Для этого вычислим $f(\psi_1(k,l))$ и воспользуемся возрастанием $\varphi_1(k,l)$ по k. Нетрудно видеть, что

$$f(\psi_1(k,l))=\Big(2\Big(1-\frac{1}{2^{l+1}}\cdot\varphi_1(k,l)\Big)\Big)^{l+1}-2\cdot\Big(2\Big(1-\frac{1}{2^{l+1}}\cdot\varphi_1(k,l)\Big)\Big)^{l}+1=$$
$$=2^{l+1}\Big(1-\frac{1}{2^{l+1}}\cdot\varphi_1(k,l)\Big)^{l}\Big(1-\frac{1}{2^{l+1}}\cdot\varphi_1(k,l)-1\Big)+1=$$
$$=-\Big(1-\frac{1}{2^{l+1}}\cdot\varphi_1(k,l)\Big)^{l}\cdot\varphi_1(k,l)+1=-\frac{\varphi_1(k,l)}{\varphi_1(k+1,l)}+1>0.$$

Вычисляя $f(\psi_2(k,l))$ и используя убывание $\varphi_2(k,l)$ по k, видим, что

$$f(\psi_2(k,l))=\Big(2\Big(1-\frac{1}{2^{l+1}}\cdot\varphi_2(k,l)\Big)\Big)^{l+1}-2\cdot\Big(2\Big(1-\frac{1}{2^{l+1}}\cdot\varphi_2(k,2)\Big)\Big)^{l}+1=$$
$$=2^{l+1}\Big(1-\frac{1}{2^{l+1}}\cdot\varphi_2(k,l)\Big)^{l}\Big(1-\frac{1}{2^{l+1}}\cdot\varphi_2(k,l)-1\Big)+1=$$
$$=-\Big(1-\frac{1}{2^{l+1}}\cdot\varphi_2(k,l)\Big)^{l}\cdot\varphi_1(k,l)+1=-\frac{\varphi_2(k,l)}{\varphi_2(k+1,l)}+1<0.$$

Неравенства (15.14) доказаны. Из этих неравенств и возрастания $f(y)$ в окрестности y_0 (см. рис. 15.1) следует, что

$$\psi_2(k,l)\le\psi_2(l)\le y_0\le\psi_1(l)\le\psi_1(k',l)\qquad(15.15)$$

при любых k и k'.

Оценим разность $R(k,l)=\psi_1(k,l)-\psi_2(k,l)$. Так как

$$\varphi_1(k,l)=\Big(\frac{1}{2}\psi_1(k-1,l)\Big)^{-l},\qquad \varphi_2(k,l)=\Big(\frac{1}{2}\psi_2(k-1,l)\Big)^{-l},$$

то из (15.13) и (15.15) следует, что $\varphi_1(k,l))<\varphi_2(k',l)<\varphi_2(1,l)<2$ при любых k и k'. Поэтому

$$R(k+1,l)=\frac{1}{2^l}\Big(\varphi_2(k+1,l)-\varphi_1(k+1,l)\Big)=$$
$$=\psi_2(k,l)^{-l}-\psi_1(k,l)^{-l}=\frac{\psi_1(k,l)^l-\psi_2(k,l)^l}{\psi_2(k,l)^l\psi_1(k,l)^l}=$$
$$=(\psi_1(k,l)-\psi_2(k,l))\frac{\sum_{i=0}^{l-1}\psi_1(k,l)^i\psi_2(k,l)^{l-1-i}}{\psi_2(k,l)^l\psi_1(k,l)^l}=$$
$$=R(k,l)\frac{l\cdot\psi_1(k,l)^{l-1}\cdot\varphi_1(k,l)\varphi_2(k,l)}{2^{2l}}\le R(k,l)\frac{l}{2^{l-1}}\le R(k,l)\frac{3}{4}.$$

Таким образом разность $\psi_1(k,l) - \psi_2(k,l)$ убывает со скоростью геометрической прогрессии по k и, следовательно, стремится к нулю с ростом k. Поэтому $\psi(l) = \psi_1(l) = \psi_2(l)$. Из определения величин $\psi_1(k,l)$ и $\psi_2(k,l)$ следует, что с ростом k стремится к нулю и разность $\varphi_2(k,l) - \varphi_1(k,l)$, а так как $\varphi_i(k,l)$ монотонны по k, то $\lim_{k\to\infty}\varphi_1(k,l) = \lim_{k\to\infty}\varphi_2(k,l) = \varphi(l)$, и из (15.14) заключаем, что величина

$$\psi(l) = 2\left(1 - \frac{1}{2^{l+1}}\varphi(l)\right)$$

является искомым корнем y_0 многочлена $f(y)$, а, следовательно, и корнем многочлена $g(y)$.

Нетрудно видеть, что $C_{m,l} = 2^m\left(1 - \frac{1}{2^{l+1}}\varphi(l)\right)^m + \mathcal{O}(l)$, причем при больших m величина $\mathcal{O}(l)$ стремится к нулю, и $C_{m,l}$ становится ближайшим целым к $2^m\left(1 - \frac{1}{2^{l+1}}\varphi(l)\right)^m$. Так как $2^l \sim n/d$, то при $n\to\infty$ и $d-1 > \sqrt{n+1}$

$$C_{m,l} \sim 2^m\left(1 - \frac{1}{2^{l+1}}\varphi(l)\right)^m \sim n\left(1 - \frac{1}{2^{l+1}}\varphi(l)\right)^m \sim n\Big(1 - \frac{d\varphi(n,d)}{2n}\Big)^{\log_2 n}, \tag{15.16}$$

где $\varphi(n,d) = \varphi(l)$.

Равенства (15.5) и (15.16) доказаны для расстояний вида 2^p+1. Однако нетрудно видеть, что они остаются справедливыми и для расстояний вида 2^p-1, так как прибавление $\log_2(n+1)$ не изменяет вида формулы (15.5) и не нарушает асимптотики в (15.16). Теорема доказана.

Теорема 15.3. *Минимальное расстояние примитивного БЧХ-кода длины $n = 2^m - 1$ с конструктивным расстоянием $d = 2^p - 1$, где $p = 2,\dots,m-1$, совпадает с его конструктивным расстоянием.*

Доказательство. Пусть α — порождающий элемент поля $GF(2^m)$. Элементы $1, \alpha, \dots, \alpha^{m-1}$ образуют базис в поле $GF(2^m)$, поэтому каждый элемент β поля $GF(2^m)$ является линейной комбинацией $\bigoplus_{i=1}^m \beta_i\alpha^{i-1}$, где $\beta_i \in GF(2)$ — координаты β. Рассмотрим множество $\mathcal{F}_m$, состоящее из всех m-местных функций вида

$$f(x_1,\dots,x_m) = \bigoplus_{\beta\in GF(2^m)} f_\beta \cdot x_1^{\beta_1}\cdots x_m^{\beta_m},$$

где $f_\beta \in GF(2^m)$, $x_i \in GF(2)$, $x_i^1 = x$ и $x_i^0 = 1$. Далее значение функции f из $\mathcal{F}_m$ на наборе $\beta_1,\dots,\beta_m$ будем обозначать через $f(\beta)$, если $\beta = \bigoplus_{i=1}^m \beta_i\alpha^{i-1}$. Ненулевые элементы поля $GF(2^m)$ упорядочим по возрастанию их индексов по основанию α, и каждой функции f из $\mathcal{F}_m$ поставим в соответствие вектор

$$\boldsymbol{v}(f) = (f(\alpha^0), f(\alpha^1), \dots, f(\alpha^{2^m-2})) \tag{15.17}$$

ее значений на координатах ненулевых элементов поля $GF(2^m)$. Например, для линейной функции $h(x_1,\dots,x_m) = \bigoplus_{i=1}^m x_i\alpha^{i-1}$

$$\boldsymbol{v}(h) = (\alpha^0, \alpha^1, \dots, \alpha^{2^m-2}). \tag{15.18}$$

Если функция f имеет хотя бы одну фиктивную переменную и $f(0) = 0$, то

$$\bigoplus_{i=0}^{2^m-2} f(\alpha^i) = \bigoplus_{\beta\in GF(2^m)} f(\beta) = 0, \tag{15.19}$$

так как во второй сумме каждое значение встретится четное число раз. В частности, если $f(x_1,\dots,x_m) = x_{i_1}\cdots x_{i_k}$ — одночлен степени k, где $k < m$, то для f справедливо равенство (15.19). Поэтому нетрудно видеть, что (15.19) справедливо также и для любой суммы таких одночленов, т. е. для любой функции f из $\mathcal{F}_m$, степень которой меньше m и для которой $f(0) = 0$.

Положим $n = 2^m - 1$ и рассмотрим произведение

$$\begin{pmatrix} 1 & \alpha & \dots & \alpha^{n-1} \\ 1 & \alpha^3 & \dots & \alpha^{(n-1)3} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha^{2t-1} & \dots & \alpha^{(n-1)(2t-1)} \end{pmatrix} \begin{pmatrix} f_k(\alpha^0) \\ f_k(\alpha^1) \\ \vdots \\ f_k(\alpha^{n-1}) \end{pmatrix}$$

проверочной матрицы БЧХ-кода длины n с конструктивным расстоянием $2t+1$ и вектора значений функции $f_k(x_1,\dots,x_m) = (x_1\oplus 1)\cdots(x_k\oplus 1)$ степени k. Прежде всего заметим (см. (15.18)), что каждая строка

$$(\alpha^0, \alpha^r, \alpha^{2r}, \dots, \alpha^{(n-1)r})$$

проверочной матрицы рассматриваемого БЧХ-кода является вектором значений функции $\big(h(x_1,\dots,x_m)\big)^r = \big(\bigoplus_{i=1}^m x_i\alpha^{i-1}\big)^r$. Пусть $r = \sum_{j=1}^m r_j 2^{j-1}$. Тогда

$$\Big(\bigoplus_{i=1}^m x_i\alpha^{i-1}\Big)^r = \Big(\bigoplus_{i=1}^m x_i\alpha^{i-1}\Big)^{\sum_{j=1}^m r_j 2^{j-1}} = \prod_{j=1}^m \Big(\bigoplus_{i=1}^m x_i\alpha^{i-1}\Big)^{r_j 2^{j-1}} =$$
$$= \prod_{j=1}^m \Big(\bigoplus_{i=1}^m x_i(\alpha^{i-1})^{2^{j-1}}\Big)^{r_j} = \prod_{j=1}^m \Big(\bigoplus_{i=1}^m x_i\alpha_{ij}\Big)^{r_j},$$

где $\alpha_{ij} = (\alpha^{i-1})^{2^{j-1}} \in GF(2^m)$. Следовательно, $\deg(h(x_1,\dots,x_m))^r = \|\boldsymbol{r}\|$. Поэтому, если $\|\boldsymbol{r}\| + k < m$, то

$$\bigoplus_{i=0}^{n-1} \alpha^{ir} f_k(\alpha^i) = \bigoplus_{i=0}^{n-1} (h^r \cdot f_k)(\alpha^i) = 0,$$

так как $(h^r \cdot f_k)(0) = 0$.

Пусть $2t+1 = 2^p - 1$. Тогда $\|\boldsymbol{r}\| \le p-1$ для любого нечетного $r \le 2^p - 3$. Следовательно, при $k = m-p$ вектор значений $\boldsymbol{v}(f_k)$ функции f_k является элементом БЧХ-кода длины n с конструктивным расстоянием $2^p - 1$. Так как $\|\boldsymbol{v}(f_k)\| = 2^{m-k} - 1 = 2^p - 1$, то минимальное расстояние рассматриваемого БЧХ-кода равно $2^p - 1$. Теорема доказана.

15.3 Скорость примитивных БЧХ-кодов

Теорема 14.2 гарантирует существование примитивного БЧХ-кода длины $n = 2^m - 1$ с конструктивным расстоянием $2t + 1$ и скоростью

$$R = \frac{n - t\log_2(n+1)}{n} = 1 - \frac{t\log_2(n+1)}{n}. \qquad (15.20)$$

Сравним величину (15.20) с границами (13.13) скорости максимального линейного кода. Если $\log_2 t = o(\log_2 n)$, то

$$H\Big(\frac{t}{n}\Big) = -\frac{t}{n}\log_2\frac{t}{n} - \Big(1 - \frac{t}{n}\Big)\log_2\Big(1 - \frac{t}{n}\Big) = \\ = \frac{t}{n}\log_2 n\Big(1 - \mathcal{O}\Big(\frac{\log_2 t}{\log_2 n}\Big)\Big) \sim \frac{t}{n}\log_2 n.$$

Поэтому при $n \to \infty$ и $\log_2 t = o(\log_2 n)$ для скорости примитивных БЧХ-кодов длины n, исправляющих t ошибок, справедливо равенство

$$R = 1 - H\Big(\frac{t}{n}\Big)(1 + o(1)), \qquad (15.21)$$

правая часть которого с точностью до слагаемых вида $o(H(t/n))$ равна верхней оценке скорости максимального линейного кода из (13.8). Таким образом для малых расстояний БЧХ-коды являются асимптотически максимальными среди всех линейных кодов.

К сожалению, с ростом t ситуация меняется, и скорость БЧХ-кодов становится меньше не только верхней оценки в (13.13), но и нижней. Рассмотрим подробно случай больших относительных расстояний. Из теорем 15.2 и 15.3 следует, что для скорости R примитивного БЧХ-кода длины n с большим минимальным расстоянием $d = 2^p - 1$ справедливо асимптотическое равенство

$$R \sim \Big(1 - \frac{d\varphi(n,d)}{2n}\Big)^{\log_2 n}. \qquad (15.22)$$

Также нетрудно видеть, что если конструктивное расстояние d примитивного БЧХ-кода удовлетворяет неравенствам $2^{p-1} - 1 < d \le 2^p - 1$, то минимальное расстояние этого кода не превосходит $2^p - 1$. Следовательно, минимальное расстояние d_M любого примитивного БЧХ-кода не более чем в два раза больше его конструктивного расстояния d_K. Поэтому отсюда и из (15.22) следует, что если минимальное расстояние d_M по порядку величины растет быстрее чем $n/\log_2 n$, т. е. $d_M = \psi(n)n/\log_2 n$, где $\psi(n) \to \infty$ при $n \to \infty$, то

$$R \lesssim \Big(1 - \frac{d_M\varphi(n,d_K)}{4n}\Big)^{\log_2 n} \le \Big(1 - \frac{\psi(n)}{4\log_2 n}\Big)^{\log_2 n} = \\ = \Big(\Big(1 - \frac{\psi(n)}{4\log_2 n}\Big)^{4\log_2 n/\psi(n)}\Big)^{\psi(n)/4} \le 2^{-\psi(n)/4}.$$

Следовательно, если минимальное расстояние примитивного БЧХ-кода длины n растет быстрее чем $n/\log_2 n$, то его скорость стремится к нулю с ростом n. Поэтому при использовании БЧХ-кодов в двоичном симметричном канале при возрастании длины кода и стремлении к нулю вероятности неправильного декодирования также к нулю будет стремиться и скорость передачи информации.

15.4 Задачи

15.1. Показать, что не каждый линейный код является полиномиальным.

15.2. Пусть $\mathbf{g}$ — кодовый многочлен полиномиального кода. Описать алгоритм нахождения соответствующего $\mathbf{g}$ информационного многочлена $\boldsymbol{a}$.

15.3. Сформулировать и доказать аналог неравенства Варшамова–Гилберта для полиномиальных кодов.

15.4. Пусть $2^m - 1 = ab$, α — порождающий элемент поля $GF(2^m)$. Показать, что минимальное расстояние полиномиального кода, порожденного многочленом $\mathrm{НОК}(m_{\alpha^a}(x), m_{\alpha^{3a}}(x), \ldots, m_{\alpha^{(2t-1)a}}(x))$, не меньше чем $2t + 1$.

15.5. Код G называется *циклическим*, если вместе с каждым своим элементом $(g_1, g_2, g_3, \ldots, g_n)$ код G содержит и его циклический сдвиг $(g_2, g_3, \ldots, g_n, g_1)$. Показать, что любой примитивный БЧХ-код — циклический.

15.6. Написать проверочную матрицу и порождающий многочлен двоичного $(20, 11)$-кода, исправляющего две ошибки.

15.7. Показать, что при $n, d \to \infty$ и $d = o(n/\log_2 n)$ для скорости R БЧХ-кода длины n с расстоянием d справедливо следующее неравенство

$$1 - R = \mathcal{O}\Big(\frac{d\log_2 n}{2n}\Big).$$

15.8. Показать, что при $n, d \to \infty$ и $d = o(n)$ для скорости R БЧХ-кода длины n с расстоянием $d = 2^p - 1$ справедливо асимптотическое равенство

$$R \sim \mathbf{e}^{-\frac{d\log_2 n}{2n}}.$$

Лекция 16

Недвоичные коды

В этой лекции вводятся недвоичные коды, на которые переносится ряд утверждений и конструкций, рассмотренных в предыдущих лекциях. Затем эти коды используются в качестве составной части каскадных кодов, для которых существуют простые алгоритмы декодирования, и скорость которых может быть сколь угодно близка к пропускной способности двоичного симметричного канала.

16.1 Определения и свойства

Пусть $q = p^m$, где p — простое. Рассмотрим множество $\mathbb{F}_q^n$, состоящее из наборов длины n с компонентами из поля $GF(q)$. Расстоянием Хемминга $d(\boldsymbol{x}, \boldsymbol{y})$ между наборами $\boldsymbol{x}$ и $\boldsymbol{y}$ называется число разрядов, в которых эти наборы не совпадают. Весом $||\boldsymbol{x}||$ набора $\boldsymbol{x}$ из $\mathbb{F}_q^n$ называется число ненулевых разрядов этого набора. Нетрудно видеть, что

$$d(\boldsymbol{x}, \boldsymbol{y}) = ||\boldsymbol{x} - \boldsymbol{y}||. \tag{16.1}$$

Подмножество $G = \{\mathbf{g}_1, \ldots, \mathbf{g}_m\}$ множества $\mathbb{F}_q^n$ называется *q-ичным кодом длины n с кодовым расстоянием d*, если для любых двух его элементов $\mathbf{g}_i$ и $\mathbf{g}_j$ расстояние Хемминга между ними не меньше d. Как и в двоичном случае, говорят, что код G исправляет t независимых ошибок, если его кодовое расстояние не меньше чем $2t + 1$. Допустим, что под воздействием шума кодовый вектор $\mathbf{g}$ длины n превратился в вектор $\boldsymbol{v}$ такой же длины. В этом случае ненулевые компоненты вектора $\boldsymbol{c} = \boldsymbol{v} - \mathbf{g}$ указывают положение ошибок, а их величины называются значениями ошибок. Используя равенство (16.1), нетрудно показать, что для линейных q-ичных кодов справедливы теоремы, доказанные выше для двоичных кодов. В частности, имеют место следующие основные теоремы, доказательства которых полностью совпадают с доказательствами соответствующих теорем для двоичных кодов.

Теорема 16.1. *В каждом q-ичном линейном коде G кодовое расстояние d равно минимальному весу его ненулевого элемента*:

$$d = \min_{\mathbf{g} \neq \mathbf{0}, \mathbf{g} \in G} \|\mathbf{g}\|.$$

Теорема 16.2. *Для того, чтобы матрица $\mathbf{H}$ была проверочной матрицей q-ичного линейного кода с кодовым расстоянием не меньшим d необходимо и достаточно, чтобы любые $d-1$ столбцов матрицы $\mathbf{H}$ были линейно независимы.*

Как и в двоичном случае, при помощи теоремы 16.2 можно легко определить q-ичные коды, исправляющие одну ошибку. Для того, что матрица $\mathbf{H}$ с элементами из поля $GF(q)$ была проверочной матрицей кода, исправляющего одну ошибку, необходимо и достаточно, чтобы в этой матрице не было кратных столбцов. Любая матрица, у которой в каждом столбце первая ненулевая компонента равна единице и все столбцы различны, удовлетворяет этому условию. Например, матрица

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \\ 1 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \end{pmatrix} \tag{16.2}$$

будет проверочной матрицей троичного кода длины 13, исправляющего одну ошибку. Так как в $\{0, 1, 2\}^{13}$ шар радиуса один состоит из 27 наборов, а код с проверочной матрицей (16.2) состоит из 3^{10} элементов, то, очевидно, что этот код совершенный.

16.2 Недвоичные БЧХ-коды

Следующая теорема описывает конструкцию примитивных БЧХ-кодов над произвольным конечным полем.

Теорема 16.3. *Пусть $n = q^m - 1$ и $\alpha_1, \ldots, \alpha_n$ — ненулевые элементы поля $GF(q^m)$. Тогда матрица*

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_n^2 \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_1^{2t} & \alpha_2^{2t} & \cdots & \alpha_n^{2t} \end{pmatrix} \tag{16.3}$$

будет проверочной матрицей кода с расстоянием не меньшим $2t + 1$. Код с проверочной матрицей (16.3) *называется q-ичным примитивным БЧХ-кодом длины $q^m - 1$ с конструктивным расстоянием $2t + 1$*

Доказательство теоремы опустим, так как оно почти дословно совпадает с доказательством леммы 14.1.

Также как и двоичные БЧХ-коды, q-ичные БЧХ-коды можно определить при помощи корней порождающих многочленов. Пусть $n = q^m - 1$, α — примитивный элемент поля $GF(q^m)$, $h(x)$ — наименьшее общее кратное минимальных многочленов элементов $\alpha, \alpha^2, \ldots, \alpha^{2t}$ поля $GF(q^m)$. *Примитивным q-ичным БЧХ-кодом* длины n с конструктивным расстоянием $2t + 1$ будем называть полиномиальный код длины n с порождающим многочленом $h(x)$.

Найдем порождающий многочлен $h(x)$ троичного примитивного БЧХ-кода длины 26, исправляющего три ошибки. В качестве поля из 27 элементов возьмем поле $\mathbb{Z}_3[\alpha]$, где α — корень примитивного многочлена x^3+2x+1. Используя равенство $\alpha^3 = \alpha + 2$, легко находим все степени α:

$$\begin{array}{llll}
\alpha^0 = (001), & \alpha^7 = (122), & \alpha^{14} = (020), & \alpha^{21} = (101),\\
\alpha^1 = (010), & \alpha^8 = (202), & \alpha^{15} = (200), & \alpha^{22} = (022),\\
\alpha^2 = (100), & \alpha^9 = (011), & \alpha^{16} = (021), & \alpha^{23} = (220),\\
\alpha^3 = (012), & \alpha^{10} = (110), & \alpha^{17} = (210), & \alpha^{24} = (221),\\
\alpha^4 = (120), & \alpha^{11} = (112), & \alpha^{18} = (121), & \alpha^{25} = (201).\\
\alpha^5 = (212), & \alpha^{12} = (102), & \alpha^{19} = (222), & \\
\alpha^6 = (111), & \alpha^{13} = (002), & \alpha^{20} = (211), &
\end{array} \qquad (16.4)$$

Так как $x^3 + 2x + 1$ является минимальным многочленом для α и α^3, а минимальные многочлены α^2 и α^6 совпадают, то для нахождения $h(x)$ надо найти минимальные многочлены элеметов α^2, α^4 и α^5. Используя (16.4) находим:

$$\begin{aligned}
h_{\alpha^2}(x) &= (x-\alpha^2)(x-\alpha^6)(x-\alpha^{18}) =\\
&= x^3 - (\alpha^2+\alpha^6+\alpha^{18})x^2 + (\alpha^8+\alpha^{20}+\alpha^{24})x - \alpha^{26} =\\
&= x^3 - 2x^2 + x - 1 = x^3 + x^2 + x + 2;\\
h_{\alpha^4}(x) &= (x-\alpha^4)(x-\alpha^{12})(x-\alpha^{10}) =\\
&= x^3 - (\alpha^4+\alpha^{10}+\alpha^{12})x^2 + (\alpha^{14}+\alpha^{16}+\alpha^{22})x - \alpha^{26} =\\
&= x^3 - 2x^2 - 1 = x^3 + x^2 + 2;\\
h_{\alpha^5}(x) &= (x-\alpha^5)(x-\alpha^{15})(x-\alpha^{19}) =\\
&= x^3 - (\alpha^5+\alpha^{15}+\alpha^{19})x^2 + (\alpha^{20}+\alpha^{24}+\alpha^{28})x - \alpha^{13} =\\
&= x^3 - x^2 + x - 2 = x^3 + 2x^2 + x + 1.
\end{aligned}$$

Умножая x^3+2x+1 на найденные минимальные многочлены x^3+x^2+x+2, x^3+x^2+2 и x^3+2x^2+x+1, находим порождающий многочлен

$$h(x) = x^{12} + x^{11} + 2x^6 + x^3 + 2x^2 + 2x + 1 \qquad (16.5)$$

троичного примитивного БЧХ-кода длины 26, исправляющего три ошибки.

Рассмотрим процедуру исправления ошибок в q-ичных БЧХ-кодах. Допустим, что при передаче кодового слова $\boldsymbol{a}$ произошло k ошибок в позициях, соответствующих элементам $\alpha_{j_1}, \ldots, \alpha_{j_k}$ поля $GF(q^m)$. Для $i = 1, \ldots, k$ введем локаторы ошибок $X_i = \alpha_{j_i}$, при этом величину ошибки, соответствующей локатору X_i, обозначим через Y_i. Нетрудно видеть, что первые k компонент синдрома S принятого слова $\boldsymbol{b}$ выражаются через локаторы и

величины ошибок следующим образом:

$$\begin{pmatrix} X_1 & X_2 & \ldots & X_k \\ X_1^2 & X_2^2 & \ldots & X_k^2 \\ \vdots & \vdots & \ddots & \vdots \\ X_1^k & X_2^k & \ldots & X_k^k \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_k \end{pmatrix} = \begin{pmatrix} Y_1X_1 + \cdots + Y_kX_k \\ Y_1X_1^2 + \cdots + Y_kX_k^2 \\ \cdots\cdots\cdots\cdots \\ Y_1X_1^k + \cdots + Y_kX_k^k \end{pmatrix} = \begin{pmatrix} S_1 \\ S_2 \\ \cdots \\ S_k \end{pmatrix}. \qquad (16.6)$$

Следуя доказательству леммы 14.1, нетрудно показать, что матрица в левой части равенства (16.6) невырождена. Поэтому если известны локаторы ошибок X_i, то величины ошибок Y_i можно найти из (16.6).

Для нахождения локаторов ошибок применяется многочлен локаторов ошибок

$$\Lambda(x) = (1 - xX_1)(1 - xX_2)\ldots(1 - xX_k) = \Lambda_k x^k + \Lambda_{k-1}x^{k-1} + \cdots + \Lambda_1 x + 1,$$

коэффициенты которого можно найти так же как и в двоичном случае. Поэтому повторим рассуждения из раздела 14.4, учитывая при этом возможные различия между полями $GF(2^m)$ и $GF(q^m)$.

Сначала для каждого $i = 1, \ldots, k$ и каждого $j = 1, \ldots, k$ умножим многочлен локаторов ошибок на произведение $Y_iX_i^{k+j}$ и подставим вместо переменной x его корень X_i^{-1}. В результате получим систему равенств

$$\Lambda_k Y_i X_i^j + \Lambda_{k-1} Y_i X_i^{j+1} + \cdots + \Lambda_1 Y_i X_i^{k+j-1} + Y_i X_i^{k+j} = 0, \qquad (16.7)$$

где $i = 1, \ldots, k$ и $j = 1, \ldots, k$. Суммируя при фиксированном j равенства из (16.7) по всем i от 1 до k, получим, что для каждого $j = 1, \ldots, k$

$$\begin{aligned}
&\sum_{i=1}^{k} (\Lambda_k Y_i X_i^j + \Lambda_{k-1} Y_i X_i^{j+1} + \cdots + Y_i X_i^{k+j}) =\\
&\qquad = \Lambda_k \sum_{i=1}^{k} Y_i X_i^j + \Lambda_{k-1} \sum_{i=1}^{k} Y_i X_i^{j+1} + \cdots + \sum_{i=1}^{k} Y_i X_i^{k+j} = 0.
\end{aligned}$$

Так как суммы, умножаемые в последнем равенстве на коэффициенты Λ_i, являются компонентами синдрома, то

$$\Lambda_k S_j + \Lambda_{k-1} S_{j+1} + \cdots + \Lambda_1 S_{k+j-1} + S_{k+j} = 0, \quad j = 1, \ldots, k. \qquad (16.8)$$

Из равенств (16.8) составим систему уравнений относительно коэффициентов Λ_i. В матричной форме эта система имеет следующий вид:

$$\begin{pmatrix} S_1 & S_2 & S_3 & \ldots & S_{k-1} & S_k \\ S_2 & S_3 & S_4 & \ldots & S_k & S_{k+1} \\ S_3 & S_4 & S_5 & \ldots & S_{k+1} & S_{k+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ S_k & S_{k+1} & S_{k+2} & \ldots & S_{2k-2} & S_{2k-1} \end{pmatrix} \begin{pmatrix} \Lambda_k \\ \Lambda_{k-1} \\ \Lambda_{k-2} \\ \vdots \\ \Lambda_1 \end{pmatrix} = \begin{pmatrix} -S_{k+1} \\ -S_{k+2} \\ -S_{k+3} \\ \vdots \\ -S_{2k} \end{pmatrix}. \qquad (16.9)$$

Как и в двоичном случае матрица этой системы невырождена, если произошло ровно k ошибок, и вырождена, если число ошибок меньше, чем k. Для того, чтобы убедиться в этом достаточно представить матрицу системы (16.9) в виде произведения

$$\begin{pmatrix} S_1 & S_2 & \dots & S_k \\ S_2 & S_3 & \dots & S_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ S_k & S_{k+1} & \dots & S_{2k-1} \end{pmatrix} = \begin{pmatrix} Y_1X_1 & Y_2X_2 & \dots & Y_kX_k \\ Y_1X_1^2 & Y_2X_2^2 & \dots & Y_kX_k^2 \\ \vdots & \vdots & \ddots & \vdots \\ Y_1X_1^k & Y_2X_2^k & \dots & Y_kX_k^k \end{pmatrix} \begin{pmatrix} 1 & X_1 & \dots & X_1^{k-1} \\ 1 & X_2 & \dots & X_2^{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_k & \dots & X_k^{k-1} \end{pmatrix},$$

в котором левый множитель также можно представить в виде произведения

$$\begin{pmatrix} Y_1X_1 & Y_2X_2 & \dots & Y_kX_k \\ Y_1X_1^2 & Y_2X_2^2 & \dots & Y_kX_k^2 \\ \vdots & \vdots & \ddots & \vdots \\ Y_1X_1^k & Y_2X_2^k & \dots & Y_kX_k^k \end{pmatrix} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ X_1 & X_2 & \dots & X_k \\ \vdots & \vdots & \ddots & \vdots \\ X_1^{k-1} & X_2^{k-1} & \dots & X_k^{k-1} \end{pmatrix} \begin{pmatrix} Y_1X_1 & 0 & \dots & 0 \\ 0 & Y_2X_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & Y_kX_k \end{pmatrix}.$$

Таким образом, видим, что матрица системы (16.9) является произведением трех квадратных матриц, две из которых невырождены тогда и только тогда, когда различны все X_i (это легко следует из доказательства леммы 14.1), а третья — диагональная матрица, невырождена тогда и только тогда, когда все X_i и Y_i отличны от нуля.

Теперь приведем алгоритм декодирования БЧХ-кода. Пусть БЧХ-код длины $n = q^m - 1$ с проверочной матрицей (16.3) исправляет t ошибок. Тогда для декодирования слова $\boldsymbol{a}$ выполняем следующие действия.

1. Вычисляем компоненты синдрома $S_1, S_2, \dots, S_{2t}$ и полагаем $k = t$.

2. Составляем матрицу

$$\mathbf{S}_k = \begin{pmatrix} S_1 & S_2 & S_3 & \dots & S_{k-1} & S_k \\ S_2 & S_3 & S_4 & \dots & S_k & S_{k+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ S_k & S_{k+1} & S_{k+2} & \dots & S_{2k-2} & S_{2k-1} \end{pmatrix}$$

и определяем вырождена это матрица или нет.

3. Если матрица $\mathbf{S}_k$ вырождена, то уменьшаем k на единицу и возвращаемся к п. 2. Если матрица $\mathbf{S}_k$ невырождена, то определяем, что число ошибок равно k и переходим к следующему пункту.

4. Решая матричное уравнение

$$\begin{pmatrix} S_1 & S_2 & S_3 & \dots & S_{k-1} & S_k \\ S_2 & S_3 & S_4 & \dots & S_k & S_{k+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ S_k & S_{k+1} & S_{k+2} & \dots & S_{2k-2} & S_{2k-1} \end{pmatrix} \begin{pmatrix} \Lambda_k \\ \Lambda_{k-1} \\ \vdots \\ \Lambda_1 \end{pmatrix} = \begin{pmatrix} -S_{k+1} \\ -S_{k+2} \\ \vdots \\ -S_{2k} \end{pmatrix},$$

находим коэффициенты Λ_i многочлена локаторов ошибок.

5. Вычисляя последовательно значения многочлена

$$\Lambda(x) = \Lambda_k x^k + \Lambda_{k-1}x^{k-1} + \dots + \Lambda_1 x + 1$$

на всех элементах поля $GF(q^m)$, находим его корни $x_1, \dots, x_k$, обращая которые, находим локаторы ошибок $X_1, \dots, X_k$. Затем из равенств $X_i = \alpha_{j_i}$ определяем позиции ошибок j_i.

6. Решая матричное уравнение

$$\begin{pmatrix} X_1 & X_2 & \cdots & X_k \\ X_1^2 & X_2^2 & \dots & X_k^2 \\ \vdots & \vdots & \ddots & \vdots \\ X_1^k & X_2^k & \cdots & X_k^k \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_k \end{pmatrix} = \begin{pmatrix} S_1 \\ S_2 \\ \vdots \\ S_k \end{pmatrix},$$

находим величины ошибок $Y_1, \dots, Y_k$.

Рассмотрим работу алгоритма на примере исправления ошибок в слове (22222222222222222222222210) троичным БЧХ-кодом длины 26 с порождающим многочленом (16.5).

1. Вычисляя при помощи (16.4) значения многочлена

$$\begin{aligned} 2 + 2x + 2x^2 + 2x^3 + 2x^4 + 2x^5 + 2x^6 + 2x^7 + 2x^8 + \\ + 2x^9 + 2x^{10} + 2x^{11} + 2x^{12} + 2x^{13} + 2x^{14} + 2x^{15} + \\ + 2x^{16} + 2x^{17} + 2x^{18} + 2x^{19} + 2x^{20} + 2x^{21} + 2x^{22} + 2x^{23} + x^{24} \end{aligned}$$

на элементах $\alpha, \alpha^2, \alpha^4$ и α^5, найдем первую, вторую, четвертую и пятую компоненты синдрома:

$$S_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \alpha, \quad S_2 = \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix} = \alpha^8, \quad S_4 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = \alpha^{25}, \quad S_5 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \alpha^{10}.$$

Далее находим $S_3 = (S_1)^3 = \alpha^3$ и $S_6 = (S_2)^3 = \alpha^{24}$.

2–3. Положим $k = 3$. Так как

$$\begin{aligned} \det \begin{pmatrix} S_1 & S_2 & S_3 \\ S_2 & S_3 & S_4 \\ S_3 & S_4 & S_5 \end{pmatrix} = \det \begin{pmatrix} \alpha & \alpha^8 & \alpha^3 \\ \alpha^8 & \alpha^3 & \alpha^{25} \\ \alpha^3 & \alpha^{25} & \alpha^{10} \end{pmatrix} = \\ = \alpha^{14} + \alpha^{10} + \alpha^{10} - \alpha^9 - \alpha^0 - \alpha^{24} = 0, \end{aligned}$$

то заключаем, что произошло не более двух ошибок. Положим $k = 2$. Так как

$$\det \begin{pmatrix} S_1 & S_2 \\ S_2 & S_3 \end{pmatrix} = \det \begin{pmatrix} \alpha & \alpha^8 \\ \alpha^8 & \alpha^3 \end{pmatrix} = \alpha^4 - \alpha^{16} \neq 0,$$

то заключаем, что произошло две ошибки.

4. Решим матричное уравнение $\begin{pmatrix} S_1 & S_2 \\ S_2 & S_3 \end{pmatrix} \begin{pmatrix} \Lambda_2 \\ \Lambda_1 \end{pmatrix} = \begin{pmatrix} -S_3 \\ -S_4 \end{pmatrix}$, преобразуя соответствующую расширенную матрицу в единичную. Так как

$$\left(\begin{array}{cc|c} \alpha & \alpha^8 & -\alpha^3 \\ \alpha^8 & \alpha^3 & -\alpha^{25} \end{array}\right) \sim \left(\begin{array}{cc|c} \alpha & \alpha^8 & \alpha^{16} \\ \alpha^8 & \alpha^3 & \alpha^{12} \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & \alpha^7 & \alpha^{15} \\ 1 & \alpha^{21} & \alpha^4 \end{array}\right) \sim$$
$$\sim \left(\begin{array}{cc|c} 1 & \alpha^7 & \alpha^{15} \\ 0 & \alpha^{21}-\alpha^7 & \alpha^4-\alpha^{15} \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & \alpha^7 & \alpha^{15} \\ 0 & \alpha^3 & \alpha^{23} \end{array}\right)$$
$$\sim \left(\begin{array}{cc|c} 1 & \alpha^7 & \alpha^{15} \\ 0 & 1 & \alpha^{20} \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & \alpha^{15}-\alpha \\ 0 & 1 & \alpha^{20} \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & \alpha^{23} \\ 0 & 1 & \alpha^{20} \end{array}\right),$$

то $\Lambda_2 = \alpha^{23}$ и $\Lambda_1 = \alpha^{20}$.

5. Вычисляя значения многочлена $\Lambda(x) = \alpha^{23}x^2 + \alpha^{20}x + 1$ на всех ненулевых элементах поля $\mathbb{Z}_3[\alpha]$, найдем его корни $x_1 = \alpha$ и $x_2 = \alpha^2$. Обращая корни, находим локаторы ошибок $X_1 = \alpha^{25}$ и $X_2 = \alpha^{24}$, которые, очевидно, соответствуют предпоследней и последней позициям.

6. Решим матричное уравнение $\begin{pmatrix} X_1 & X_2 \\ X_1^2 & X_2^2 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} S_1 \\ S_2 \end{pmatrix}$ для определения значений ошибок Y_1 и Y_2:

$$\left(\begin{array}{cc|c} \alpha^{25} & \alpha^{24} & \alpha \\ \alpha^{24} & \alpha^{22} & \alpha^8 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & \alpha^{25} & \alpha^2 \\ 1 & \alpha^{24} & \alpha^{10} \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & \alpha^{25} & \alpha^2 \\ 0 & \alpha^{24}-\alpha^{25} & \alpha^{10}-\alpha^2 \end{array}\right) \sim$$
$$\sim \left(\begin{array}{cc|c} 1 & \alpha^{25} & \alpha^2 \\ 0 & \alpha^{14} & \alpha \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & \alpha^{25} & \alpha^2 \\ 0 & 1 & \alpha^{13} \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & \alpha^2-\alpha^{12} \\ 0 & 1 & 2 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \end{array}\right).$$

Следовательно, $Y_1 = 1$ и $Y_2 = 2$, и в векторе ошибок первые 24 компоненты равны нулю, 25 компонента равна двум, а двадцать шестая — единице. Вычитая из полученного вектора вектор ошибок, находим, что все компоненты переданного слова равны двойкам.

16.3 Коды Рида–Соломона

Важный частный случай q-ичных БЧХ-кодов длины $q^m - 1$ представляют коды с $m = 1$. Такие коды называются кодами *Рида–Соломона*, так как они впервые были описаны И. Ридом и Г. Соломоном в публикации 1960 г. После появления в том же 1960 г. работы Н. Цирлера и Д. Горенстейна, содержащей обобщение конструкции БЧХ-кодов на недвоичный случай, оказалось, что коды Рида–Соломона являются частным случаем БЧХ-кодов.

Легко видеть, что размерность k q-ичного кода Рида–Соломона длины $n = q - 1$ с конструктивным расстоянием d равна $n - d + 1$, т. е. длина, размерность и конструктивное расстояние кодов Рида–Соломона связаны равенством

$$d = n - k + 1. \tag{16.10}$$

В тоже время известна следующая теорема, неравенство которой называется границей Синглтона.

Теорема 16.4. *Для любого линейного (n, k)-кода с минимальным расстоянием d, справедливо неравенство $d \le n - k + 1$.*

Доказательство. Пусть $\mathbf{H}$ — проверочная матрица линейного (n, k)-кода с минимальным расстоянием d. Ранг этой матрицы равен $n - k$, и при этом любые ее $d - 1$ столбцов линейно независимы. Следовательно, $n - k \ge d - 1$. Теорема доказана.

Из сравнения границы Синглтона и равенства (16.10) следует, что конструктивное расстояние кодов Рида–Соломона совпадает с их минимальным расстоянием, а сами коды являются максимальными.

Найдем порождающий многочлен кода Рида–Соломона длины 15, исправляющего две ошибки. В качестве поля из 16 элементов, как и ранее (см. стр. 216 и 222), возьмем поле $\mathbb{Z}_2[\alpha]$, где α — корень примитивного многочлена $x^4 \oplus x \oplus 1$. Напомним, что степени примитивного элемента α этого поля выглядят следующим образом:

$$\begin{array}{llll} \alpha^0 = (0001), & \alpha^4 = (0011), & \alpha^8 = (0101), & \alpha^{12} = (1111), \\ \alpha^1 = (0010), & \alpha^5 = (0110), & \alpha^9 = (1010), & \alpha^{13} = (1101), \\ \alpha^2 = (0100), & \alpha^6 = (1100), & \alpha^{10} = (0111), & \alpha^{14} = (1001), \\ \alpha^3 = (1000), & \alpha^7 = (1011), & \alpha^{11} = (1110). & \end{array}$$

В этом случае многочлен

$$\begin{aligned} h(x) &= (x \oplus \alpha)(x \oplus \alpha^2)(x \oplus \alpha^3)(x \oplus \alpha^4) = \\ &= x^4 \oplus (\alpha \oplus \alpha^2 \oplus \alpha^3 \oplus \alpha^4)x^3 \oplus \\ &\quad \oplus (\alpha^3 \oplus \alpha^4 \oplus \alpha^6 \oplus \alpha^7)x^2 \oplus (\alpha^6 \oplus \alpha^7 \oplus \alpha^8 \oplus \alpha^9)x \oplus \alpha^{10} = \\ &= x^4 \oplus \alpha^{13}x^3 \oplus \alpha^6 x^2 \oplus \alpha^3 x \oplus \alpha^{10} \end{aligned}$$

будет искомым порождающим многочленом.

16.4 Каскадные коды

В 1965 г. Г. Форни предложил метод комбинирования двух кодов для построения нового кода, длина и скорость которого равны произведению длин и скоростей исходных кодов, а кратность ошибок, исправляемых новым кодом, не меньше произведения кратностей ошибок, исправляемых исходными кодами. При этом главное достоинство метода Форни заключается в том, что декодирование нового кода сводится к последовательному применению алгоритмов декодирования исходных кодов.

Пусть G_1 — двоичный линейный (n_1, k_1)-код, исправляющий t_1 ошибок. Любой элемент $\mathbf{g}_1$ этого кода является линейным преобразованием $\mathbf{G}_1(\boldsymbol{a}_1)$ двоичного вектора $\boldsymbol{a}_1$ длины k_1. Пусть $q = 2^{k_1}$ и G_2 — q-ичный линейный (n_2, k_2)-код, исправляющий t_2 ошибок. Любой элемент $\mathbf{g}_2 = (b_1, \ldots, b_{n_2})$ кода G_2 является вектором длины n_2, каждый элемент b_i которого можно

представить в виде двоичного вектора $\boldsymbol{b}_i$ длины k_1. *Каскадным двоичным* (n_1n_2, k_1k_2)*-кодом* H с внутренним кодом G_1 и внешним кодом G_2 называется множество $\{\boldsymbol{h}_j\}$, где любой элемент $\boldsymbol{h}$ получен из соответствующего элемента $\mathbf{g}_2 = (b_1, \ldots, b_{n_2})$ кода G_2 подстановкой вектора $\mathbf{G}_1(\boldsymbol{b}_i)$ вместо элемента b_i, т. е. $\boldsymbol{h} = (\mathbf{G}_1(\boldsymbol{b}_1), \ldots, \mathbf{G}_1(\boldsymbol{b}_{n_2}))$. Нетрудно видеть, что код H состоит из $2^{k_1k_2}$ элементов, и его скорость R равна произведению скоростей R_1 и R_2 кодов G_1 и G_2.

Кодирование кодом H выполняется следующим образом. Информационная двоичная последовательность делится на блоки по k_1k_2 символов. Затем каждый блок длины k_1k_2 делится на блоки длины k_1, после чего блоки длины k_1 рассматриваются как элементы поля $GF(2^{k_1})$, и k_2 таких блоков кодируются внешним кодом G_2 в состоящие из элементов поля $GF(2^{k_1})$ блоки длины n_2. Наконец, каждый элемент $GF(2^{k_1})$ в таком блоке рассматривается как двоичный вектор длины k_1 и кодируется внутренним кодом G_1 в двоичный вектор длины n_1. В результате исходный двоичный блок длины k_1k_2 преобразуется в двоичный блок длины n_1n_2.

Декодирование выполняется в обратном порядке. Подлежащая декодированию двоичная последовательность делится на блоки по n_1n_2 символов. Затем каждый блок длины n_1n_2 делится на n_2 блоков длины n_1, после чего эти блоки рассматриваются как элементы кода G_1 и декодируются в блоки длины k_1, которые, в свою очередь, далее рассматриваются как элементы поля $GF(2^{k_1})$. Наконец блок из n_2 таких элементов декодируется внешним кодом G_2 в блок длины k_2, каждый элемент которого рассматривается как двоичная последовательность длины k_1. В результате декодируемый двоичный блок длины n_1n_2 преобразуется в двоичный блок длины k_1k_2. Нетрудно видеть, что набор длины n_1n_2 будет декодирован неправильно только в том случае, когда в этом наборе найдется не меньше чем t_2+1 блоков длины n_1, в каждом из которых произойдет не меньше чем t_1+1 ошибок. Таким образом, рассмотренный алгоритм декодирования позволяет исправлять все ошибки, кратность которых не превосходит $t_1t_2 + t_1 + t_2$.

Пусть ε — произвольно малая положительная постоянная. Рассмотрим линейный код G_1 длины n, у которого вероятность неправильного декодирования не превосходит ε, а скорость не меньше $1-H(p)-\varepsilon$. Существование такого кода следует из прямой теоремы Шеннона о кодировании в двоичном симметричном канале. Будем полагать, что код G_1 состоит из 2^m элементов. Пусть G_2 — код Рида–Соломона длины $N = 2^m-1$, исправляющий $2\varepsilon N$ ошибок. Очевидно, что скорость G_2 не меньше $1-4\varepsilon$. Каскадный код G с внутренним кодом G_1 и внешним кодом G_2 будет иметь длину nN, скорость не меньше чем

$$(1-H(p)-\varepsilon)(1-4\varepsilon) = 1-H(p)-\varepsilon(5-H(p)-4\varepsilon) > 1-H(p)-5\varepsilon,$$

а из (13.16) следует, что при достаточно больших n и N для вероятности P неправильного декодирования кода G имеет место неравенство

$$P \le \sum_{i \ge 4\varepsilon} \binom{N}{i} \varepsilon^i (1-\varepsilon)^{N-i} \le N 2^{-\gamma N} \le \varepsilon.$$

При этом нетрудно видеть (см. (13.11) и алгоритм декодирования БЧХ-кодов), что сложность декодирования кода G есть $(nN)^{\mathcal{O}(1)}$.

Таким образом, каскадные коды позволяют передавать информацию по двоичному симметричному каналу с скоростью, сколь угодно близкой к максимально возможной для данного канала, сколь угодно малой вероятностью неправильного декодирования и полиномиальной, относительно длины кода, сложностью декодирования, т. е. являются хорошими в смысле п. 13.5.

16.5 Задачи

16.1. Показать, что в линейном q-ичном коде множества исправляемых ошибок всех элементов совпадают.

16.2. Пусть q — степень простого числа. Совершенный q-ичный код с расстоянием 3 называется q-ичным кодом Хемминга. Описать проверочную матрицу q-ичного кода Хемминга. При каких n существуют q-ичные коды Хемминга длины n?

16.3. Показать, что если числа n, m и d удовлетворяют неравенству

$$q^{n-m} > \sum_{i=0}^{d-2} \binom{n-1}{i} (q-1)^i,$$

то существует линейный q-ичный (n, m)-код с расстоянием d.

16.4. Оценить мощность и написать проверочную матрицу троичного кода длины 11, 12 и 13 с минимальным расстоянием 3.

16.5. Показать, что в q-ичном БЧХ-коде длины $n = q^m - 1$ с расстоянием $d < \sqrt{n}$ содержится ровно $q^{n-m(d-1-\lfloor (d-1)/q \rfloor)}$ элементов.

16.6. Написать проверочную матрицу и порождающий многочлен троичного БЧХ-кода длины 26, исправляющего две ошибки.

16.7. Найти какой-либо многочлен $g(x) \in \mathbb{Z}_5[x]$, являющийся порождающим многочленом БЧХ-кода длины 24, исправляющего две онибки.

16.8. Говорят, что код G длины n исправляет пакеты ошибок длины s, если множество ошибок этого кода состоит из всех тех наборов длины n, ненулевые компоненты которых расположены в s идущих друг за другом разрядах, т. е. в каждом таком наборе номера первой и последней ненулевой компонент отличаются не более чем на $s-1$. Показать, что существует двоичный $(155, 135)$-код, исправляющий пакеты ошибок длины 6.

Дополнение А

Конечные поля

В этом дополнении приводятся используемые в различных разделах дискретной математики сведения о конечных полях, их структуре и свойствах.

А.1 Циклические группы

Рассматривая множество $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ с операциями сложения и умножения по модулю n, элементы этого множества будем заключать в квадратные скобки и отмечать нижним индексом n, т. е. если $k, m \in \mathbb{Z}$, то $[k]_n, [m]_n \in \mathbb{Z}_n$, и арифметические операции над $[k]_n$ и $[m]_n$ выполняются по модулю n.

Теорема А.1. *Любая конечная циклическая группа порядка n изоморфна группе $(\mathbb{Z}_n, +)$.*

Доказательство. Пусть конечная циклическая группа G порядка n порождается элементом g. Рассмотрим отображение $f : g^k \to k$. Тогда для любых элементов $g_1 = g^k$ и $g_2 = g^m$ группы G имеем:

$$f(g_1 g_2) = f(g^k g^m) = f(g^{k+m}) = [k+m]_m = \\ = [k]_n + [m]_n = f(g^k) + f(g^m) = f(g_1) + f(g_2).$$

Таким образом f — изоморфизм. Теорема доказана.

Теорема А.2. *Пусть m делит n. В любой конечной циклической группе порядка n существует единственная подгруппа порядка m. Все подгруппы циклической группы циклические.*

Доказательство. В силу теоремы А.1, достаточно показать, что доказываемое свойство справедливо для $(\mathbb{Z}_n, +)$. Прежде всего отметим, что по крайней мере одна подгруппа порядка m существует. Эта подгруппа образована числами, делящимися на $d = n/m$: $0, d, \ldots, (m-1)d$, порождается элементом d, и поэтому является циклической. Отсутствие других подгрупп следует из того, что для каждого элемента a такой подгруппы должно выполняться равенство $m \cdot a = 0 \pmod n$, т. е. каждый ее элемент должен

делится на d. В $\mathbb{Z}_n$ таких чисел ровно m, и все они принадлежат первой подгруппе. Теорема доказана.

Теорема А.3. *В любой циклической группе порядка n существует ровно $\varphi(n)$ порождающих элементов.*

Доказательство. Снова воспользуемся теоремой А.1. Покажем, что в $(\mathbb{Z}_n, +)$ порождающими элементами будут все числа взаимнопростые с n и только они. Известно, что взаимная простота чисел n и a эквивалентна существованию таких целых k и m, что $kn + ma = 1$. Из этого равенства следует, что

$$m \cdot a = 1 \pmod n. \qquad (А.1)$$

Поэтому единица, порождающий элемент в $(\mathbb{Z}_n, +)$, является m-й степенью элемента a. Следовательно, a — порождающий элемент в $(\mathbb{Z}_n, +)$. С другой стороны, если a — порождающий элемент, то единица должна быть степенью a, т. е. должно существовать целое m, для которого справедливо (А.1), откуда в свою очередь следует взаимная простота n и a. Теорема доказана.

Простым следствием двух предыдущих теорем является следующее утверждение.

Теорема А.4. *Пусть m делит n. В любой конечной циклической группе порядка n существует $\varphi(m)$ элементов порядка m.*

Доказательство. Каждый элемент порядка m порождает подгруппу такого же порядка. Так как в циклической группе есть только одна подгруппа порядка m, то все элементы порядка m принадлежат одной и той же подгруппе и являются ее порождающими элементами. Теорема доказана.

А.2 Кольца

Множество $\mathbb{K}$ с определенными на нем бинарными операциями $+$ и $*$ называется *кольцом*, если:

(1) множество $\mathbb{K}$ с операцией $+$ является абелевой группой;

(2) множество $\mathbb{K}$ с операцией $*$ является полугруппой;

(3) для всех a, b и c из множества $\mathbb{K}$ выполняются законы дистрибутивности:

$$a * (b + c) = a * b + a * c, \qquad (b + c) * a = b * a + c * a.$$

Группа $(\mathbb{K}, +)$ называется аддитивной группой кольца, а ее единичный элемент — нулевым элементом кольца. Нулевой элемент кольца обозначается символом 0.

Кольцо $\mathbb{K}$ называется *кольцом с единицей*, если существует такой элемент 1, что $a * 1 = e * 1$ для любого $a \in \mathbb{K}$.

Кольцо $\mathbb{K}$ называется *коммутативным*, если операция $*$ коммутативна.

Кольцо $\mathbb{K}$ называется *кольцом без делителей нуля*, если из равенства $a * b = 0$ следует либо $a = 0$, либо $b = 0$.

Лемма А.1. *В кольце без делителей нуля выполняется закон сокращения: если $c \neq 0$ и $ac = bc$, то $a = b$.*

Доказательство. Из равенства $ac = bc$ следует, что $(a - b)c = 0$, и так как $c \neq 0$ и в кольце нет делителей нуля, то $a - b = 0$, т. е. $a = b$. Лемма доказана.

Коммутативное кольцо с единицей $1 \neq 0$, в котором каждый ненулевой элемент обратим, называется *полем*. Множество ненулевых элементов поля образует коммутативную группу, которая называется *мультипликативной группой поля*, а ее единичный элемент называется единичным элементом (единицей) поля. Нетрудно видеть, что полями являются множества $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}_p$ при простом p с обычными для этих множеств операциями сложения и умножения.

Далее сумму k одинаковых элементов x поля $\mathbb{F}$ будем обозначать через $k \cdot x$. *Характеристикой поля* $\mathbb{F}$ называется минимальное натуральное число m, для которого $m \cdot 1 = 0$, а если такого числа нет, то говорят, что характеристика поля равна нулю. Поле называется конечным, если оно состоит из конечного числа элементов. Нетрудно видеть, что характеристика любого конечного поля является простым числом.

Подмножество $\mathbb{F}'$ элементов поля $\mathbb{F}$ называется *подполем* поля $\mathbb{F}$, если $\mathbb{F}'$ является полем. Если $\mathbb{F}'$ — подполе поля $\mathbb{F}$, то $\mathbb{F}$ называется *расширением* поля $\mathbb{F}'$. Поле называется *простым*, если оно не имеет собственного подполя. Очевидно, что при любом простом p поле $\mathbb{Z}_p$ будет простым. Также нетрудно видеть, что в любом конечном поле существует простое подполе, порожденное его нулевым и единичным элементами.

А.3 Кольцо многочленов

Пусть $\mathbb{F}$ — поле. Многочлен

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_i x^i + \cdots + a_1 x + a_0,$$

где $a_0, a_1, \ldots, a_n \in \mathbb{F}$, назовем *многочленом над полем* $\mathbb{F}$ степени n относительно переменной x если $a_n \neq 0$. Степень многочлена $p(x)$ обозначим через $\deg p(x)$. Если коэффициент при старшей степени многочлена равен единице, то многочлен называется *нормированным*. Множество, состоящее из всех многочленов любой конечной степени над $\mathbb{F}$, обозначим через $\mathbb{F}[x]$. Без доказательства приведем следующее очевидное утверждение.

Теорема А.5. *Пусть $\mathbb{F}$ — поле. Тогда $\mathbb{F}[x]$ — коммутативное кольцо с единицей и без делителей нуля.*

Любой многочлен $c(x)$, который делит многочлены $a(x)$ и $b(x)$, называется общим делителем этих многочленов. Нормированный многочлен наибольшей степени среди общих делителей многочленов $a(x)$ и $b(x)$ называется их наибольшим общим делителем и обозначается символом $(a(x), b(x))$.

Многочлены, не имеющие общих делителей, называются взаимнопростыми. Легко видеть, что $(a(x), b(x)) = 1$ для любых взаимнопростых многочленов $a(x)$ и $b(x)$.

Теорема А.6. *Пусть $\mathbb{F}$ — поле, $a(x)$ и $b(x)$ — многочлены над $\mathbb{F}$. Ненулевой нормированный многочлен минимальной степени $d(x)$, удовлетворяющий равенству*

$$d(x) = a(x)s(x) + b(x)t(x), \tag{А.2}$$

где $s(x)$ и $t(x)$ — многочлены над $\mathbb{F}$, является наибольшим общим делителем многочленов $a(x)$ и $b(x)$.

Доказательство. Покажем, что $a(x)$ делится на $d(x)$. Представим $a(x)$ в виде $a(x) = p(x)d(x) + q(x)$, где $\deg q(x) < \deg d(x)$. Тогда

$$\begin{aligned} q(x) &= a(x) + p(x)d(x) = a(x) + p(x)(a(x)s(x) + b(x)t(x)) = \\ &= a(x)(1 + p(x)s(x)) + b(x)(p(x)t(x)) = a(x)s'(x) + b(x)t'(x), \end{aligned}$$

где $s'(x) = 1 + p(x)s(x)$ и $t'(x) = p(x)t(x)$. Так как $d(x)$ — ненулевой нормированный многочлен минимальной степени, удовлетворяющий (А.2), то $q(x) = 0$ и, следовательно, $a(x)$ делится на $d(x)$. Аналогичным образом доказывается делимость на $d(x)$ числа $b(x)$. Следовательно, $d(x)$ является общим делителем многочленов $a(x)$ и $b(x)$. Теперь покажем, что $d(x)$ будет наибольшим общим делителем. Действительно, если $a(x)$ и $b(x)$ делятся на $d(x)d'(x)$, где $\deg d'(x) > 1$, то в (А.2) правая часть будет делиться на $d(x)d'(x)$, а левая не будет. Следовательно, $d(x) = (a(x), b(x))$. Теорема доказана.

Многочлен $p(x)$ называется *неприводимым* над полем $\mathbb{F}$, если он не имеет делителей ненулевой степени над этим полем. Из определения неприводимого многочлена легко следует, что любой многочлен можно представить в виде произведения неприводимых многочленов, например, в $\mathbb{Z}_2[x]$ имеет место равенство

$$x^4 + x = x \cdot (x + 1)(x^2 + x + 1).$$

Лемма А.2. *Если произведение $a(x)b(x)$ многочленов $a(x)$ и $b(x)$ над полем $\mathbb{F}$ делится на неприводимый многочлен $p(x)$, то хотя бы один из сомножителей также делится на $p(x)$.*

Доказательство. Допустим, что утверждение леммы не верно. Тогда ни один из сомножителей не делится на $p(x)$. Следовательно, $(a(x), p(x)) = 1$ и $(b(x), p(x)) = 1$, и из теоремы А.6 следует, что найдутся такие многочлены $s_1(x), s_2(x)$ и $t_1(x), t_2(x)$, что

$$1 = s_1(x)a(x) + s_2(x)p(x), \qquad 1 = t_1(x)b(x) + t_2(x)p(x).$$

Перемножив эти равенства, получим, что

$$1 = (s_1(x)a(x) + s_2(x)p(x)) \cdot (t_1(x)b(x) + t_2(x)p(x)) =$$

$$= (s_1(x)t_1(x)) \cdot a(x)b(x) + \\ + (s_2(x)t_1(x)b(x) + s_1(x)t_2(x)a(x) + s_2(x)t_2(x)p(x)) \cdot p(x),$$

т. е., в силу теоремы А.6 многочлены $a(x)b(x)$ и $p(x)$ взаимнопросты. Противоречие. Лемма доказана.

Лемма А.3. *Если произведение $a_1(x)a_2(x)\cdots a_n(x)$ многочленов $a_i(x)$ над полем $\mathbb{F}$ делится на неприводимый многочлен $p(x)$, то хотя бы один из сомножителей также делится на $p(x)$.*

Доказательство. Лемму докажем индукцией по числу сомножителей. Случай $n = 2$ доказан в предыдущей лемме. Допустим, утверждение леммы справедливо для любых произведений, содержащих не более k сомножителей. Тогда, если произведение $(a_1(x)a_2(x)\cdots a_k(x))a_{k+1}(x)$ делится на $p(x)$, то из леммы А.2 следует, что либо $a_1(x)a_2(x)\cdots a_k(x)$, либо $a_{k+1}(x)$ делится на $p(x)$. В первом случае утверждение леммы следует из предположения индукции. Во втором случае справедливость леммы очевидна. Лемма доказана.

Теорема А.7. *Каждый многочлен над полем $\mathbb{F}$ единственным образом раскладывается в произведение неприводимых многочленов.*

Доказательство. Теорему докажем методом математической индукцией. В основание индукции положим многочлены степени 1, единственность разложения которых очевидна. Предположим, что каждый многочлен, степень которого не превосходит n, разлагается на неприводимые множители единственным образом. Покажем, что из этого предположения следует единственность разложения любого многочлена степени $n+1$. Действительно, если некоторый многочлен $t(x)$ степени $n+1$ имеет два различных разложения в произведение неприводимых сомножителей, то

$$t(x) = p_1^{k_1}(x)p_2^{k_2}(x)\cdots p_n^{k_n}(x) = q_1^{s_1}(x)q_2^{s_2}(x)\cdots q_m^{s_m}(x).$$

Покажем, что $p_1(x)$ совпадает с одним из многочленов $q_i(x)$. Так как неприводимый многочлен $p_1(x)$ делит произведение

$$\underbrace{q_1(x)\cdots q_1(x)}_{s_1 \text{ раз}}\underbrace{q_2(x)\cdots q_2(x)}_{s_2 \text{ раз}}\cdot\ldots\ldots\cdot\underbrace{q_m(x)\cdots q_m(x)}_{s_m \text{ раз}} \qquad (A.3)$$

то в силу леммы А.3 многочлен $p_1(x)$ также делит один из его сомножителей. В (А.3) все сомножители — неприводимые многочлены. Следовательно, $p_1(x)$ совпадает с одним из них. Без ограничения общности будем полагать, что $p_1(x) = q_1(x)$. Сокращая левую и правую части равенства (А.3) на $p_1(x)$, получим

$$p_1^{k_1-1}(x)p_2^{k_2}(x)\cdots p_n^{k_n}(x) = q_1^{s_1-1}(x)q_2^{s_2}(x)\cdots q_m^{s_m}(x). \qquad (A.4)$$

По предположению индукции многочлен $t(x)/p_1(x)$ раскладывается на неприводимые множители единственным образом. Поэтому в (А.4) $n = m$ и для каждого $i \in \{1,\ldots,n\}$ справедливы равенства $p_i(x) = q_i(x)$ и $k_i = s_i$. Теорема доказана.

Лемма А.4. *Многочлен $f(x)$ над полем $\mathbb{F}$ делится на двучлен $x-\alpha$, где $\alpha \in \mathbb{F}$, тогда и только тогда, когда $f(\alpha) = 0$.*

Доказательство. Если многочлен $f(x)$ делится на $x-\alpha$, то $f(x) = h(x)(x-\alpha)$. Тогда $f(\alpha) = h(\alpha)(\alpha-\alpha) = 0$. С другой стороны, $f(x) = h(x)(x-\alpha)+\beta$, и если $f(\alpha) = 0$, то легко видеть, что $\beta = 0$. Следовательно, $f(x)$ делится на $x-\alpha$. Лемма доказана.

Теорема А.8. *Многочлен f степени n над полем $\mathbb{F}$ имеет в поле $\mathbb{F}$ не более n корней.*

Доказательство. Теорему докажем индукцией по степени многочлена. Очевидно, что ненулевой многочлен нулевой степени не имеет корней. Этот случай ($\deg f = 0$) положим в основание индукции. Допустим, что утверждение теоремы справедливо для всех многочленов степени не более k. Пусть $\deg f = k+1$. Если многочлен $f(x)$ в поле $\mathbb{F}$ не имеет корней, то утверждение теоремы очевидно. Пусть α — корень многочлена $f(x)$. Тогда в силу леммы А.4 многочлен $f(x)$ делится на $x-\alpha$, т. е. $f(x) = h(x)(x-\alpha)$, где многочлен k-й степени $h(x)$ по предположению индукции имеет в $\mathbb{F}$ не более k корней. Следовательно, f имеет в поле $\mathbb{F}$ не более $k+1$ корней. Теорема доказана.

А.4 Поле многочленов

Множество многочленов из $\mathbb{Z}_p[x]$ степени не выше $n-1$ с операциями сложения и умножения по модулю многочлена $h(x)$ степени n называется факторкольцом $\mathbb{Z}_p[x]/h(x)$. Нетрудно видеть, что $\mathbb{Z}_p[x]/h(x)$ является коммутативным кольцом с единицей. Более того, справедлива следующая теорема.

Теорема А.9. *Пусть p — простое, $h(x)$ — многочлен из $\mathbb{Z}_p[x]$ степени n неприводимый над полем $\mathbb{Z}_p$. Тогда фактор-кольцо $\mathbb{Z}_p[x]/h(x)$ является полем из p^n элементов.*

Доказательство. Кольцо $\mathbb{Z}_p[x]/h(x)$ является коммутативным кольцом с единицей и состоит из p^n элементов, поэтому для доказательства теоремы достаточно показать, что любой ненулевой элемент этого кольца имеет обратный по умножению. Так как $(f(x), h(x)) = 1$ для любого $f(x)$ из $\mathbb{Z}_p[x]/h(x)$, то в силу теоремы А.6 найдутся такие многочлены $s(x)$ и $t(x)$, что $1 = s(x)f(x) + t(x)h(x)$. Поэтому $f(x)s(x) = 1 \pmod{h(x)}$, т. е. $f(x)$ имеет обратный элемент в $\mathbb{Z}_p[x]/h(x)$. Теорема доказана.

Теорема А.10. *Для любого простого p и любого натурального n существует конечное поле из p^n элементов. Любое конечное поле характеристики p состоит из p^n элементов.*

Доказательство. Существование поля из p^n элементов для простого p и натурального n легко следует из теоремы А.9 и существования в $\mathbb{Z}_p[x]$

неприводимого над $\mathbb{Z}_p$ многочлена любой степени n. В свою очередь, существование такого многочлена является простым следствием теоремы 3.2 о числе неприводимых многочленов, доказанной на стр. 43.

В любом конечном поле $\mathbb{F}$ характеристики p существует простое подполе $\mathbb{F}'$ из p элементов. Нетрудно видеть, что поле $\mathbb{F}$ будет линейным конечномерным пространством над своим простым подполем $\mathbb{F}'$, и поэтому число элементов в поле $\mathbb{F}$ будет натуральной степенью числа элементов в его простом подполе $\mathbb{F}'$. Теорема доказана.

Теорема А.9 позволяет задавать конечные поля в явном виде и выполнять в этих полях операции сложения и умножения. Для задания поля из p^n элементов выберем в $\mathbb{Z}_p[x]$ неприводимый над $\mathbb{Z}_p$ многочлен n-й степени $h(x)$. Элементы поля будем представлять в виде упорядоченных наборов $(a_0,\dots,a_{n-1})$ длины n с элементами из $\mathbb{Z}_p$. При этом операция сложения выполняется покомпонентно, а для умножения элементов $a=(a_0,\dots,a_{n-1})$ и $b=(b_0,\dots,b_{n-1})$ надо вычислить произведение $r(x)=r_0+\dots+r_{n-1}x^{n-1}$ многочленов $a(x)=a_0+\dots+a_{n-1}x^{n-1}$ и $b(x)=b_0+\dots+b_{n-1}x^{n-1}$ по модулю многочлена $h(x)$ и в качестве произведения ab взять набор $(r_0,\dots,r_{n-1})$ коэффициентов многочлена $r(x)$.

В качестве примера рассмотрим поле $\mathbb{F}$ из 27 элементов. Для этого понадобится неприводимый нормированный многочлен третьей степени из $\mathbb{Z}_3[x]$. Так как любой приводимый многочлен третьей степени над $\mathbb{Z}_3$ имеет корень в $\mathbb{Z}_3$, то многочлен $x(x+1)(x+2)+1=x^3+2x+1$ будет неприводимым над $\mathbb{Z}_3$. Элементами поля $\mathbb{F}$ будут наборы вида (a_0,a_1,a_2), где $a_i\in\mathbb{Z}_3$. Найдем сумму и произведение элементов $(1,0,1)$ и $(0,1,2)$. Легко видеть, что $(1,1,1)+(0,2,2)=(1,0,0)$. Далее, так как

$$\begin{aligned}(1+x+x^2)(2x+2x^2) &= 2x+x^2+x^3+2x^4 = \\ &= (1+2x+x^3)(1+2x)+(2+x),\end{aligned}$$

то $(1,1,1)\cdot(0,2,2)=(2,1,0)$.

А.5 Структура конечного поля

Теорема А.11. *Мультипликативная группа конечного поля циклическая.*

Доказательство. Рассмотрим поле $\mathbb{F}_q$, состоящее из q элементов. Пусть $q-1=p_1^{k_1}\cdots p_r^{k_r}$ — разложение числа $q-1$ на простые множители. При каждом i многочлен $x^{(q-1)/p_i}-1$ имеет в мультипликативной группе поля $\mathbb{F}_q$ не более $(q-1)/p_i$ корней, поэтому для каждого i в этой группе найдется элемент g_i такой, что

$$g_i^{(q-1)/p_i}\neq 1. \tag{A.5}$$

Положим $h_i=g_i^{(q-1)/p_i^{k_i}}$, $h=\prod_{i=1}^r h_i$ и покажем, что h будет порождающим элементом в $\mathbb{F}_q^*$. Очевидно, что для этого достаточно доказать, что

$h^{(q-1)/p_i}\neq 1$ для каждого i. Сделаем это только для $i=1$, так как легко видеть, что доказательства для разных значений i аналогичны.

Так как группа $\mathbb{F}_q^*$ абелева, то $h^{(q-1)/p_1}=\prod_{i=1}^r h_i^{(q-1)/p_1}$. Далее заметим, что произведение $((q-1)/p_i^{k_i})\cdot(q-1)/p_1$ кратно $q-1$, если $i\neq 1$. Поэтому

$$h_i^{(q-1)/p_1}=1 \quad \text{при} \quad i\neq 1. \tag{A.6}$$

Теперь найдем порядок элемента h_1. Так как $h_1^{p_1^{k_1}}=1$, то порядок элемента h_1 равен p_1^s, где s не превосходит k_1. Но из неравенства (A.5) следует, что $h_1^{p_1^{k_1-1}}=g_i^{(q-1)/p_i}\neq 1$. Поэтому порядок элемента h_1 равен $p_1^{k_1}$. Учитывая, что $(q-1)/p_1$ не делится на $p_1^{k_1}$, заключаем, что

$$h_1^{(q-1)/p_1}\neq 1. \tag{A.7}$$

Из (A.6) и (A.7) следует, что $h^{(q-1)/p_1}\neq 1$. Теорема доказана.

Порождающий элемент мультипликативной группы конечного поля называется примитивным элементом этого поля. Из теоремы А.3 следует, что в поле из q элементов существует ровно $\varphi(q-1)$ примитивных элементов. Пусть α — примитивный элемент поля $\mathbb{F}_q$ и β — ненулевой элемент этого поля. Минимальное i, для которого $\beta=\alpha^i$, называется индексом элемента β по основанию α.

Лемма А.5. *Пусть $g(x)$ — произвольный многочлен над $\mathbb{Z}_p$. Тогда*

$$(g(x))^p=g(x^p).$$

Доказательство. Так как p простое, то в правой части равенства

$$(a+b)^p=a^p+\binom{p}{1}a^{p-1}b+\dots+\binom{p}{k}a^{p-k}b^k+\dots+\binom{p}{p-1}ab^{p-1}+b^p$$

все коэффициенты $\binom{p}{k}$ при $k\neq 0,p$ делятся на p. Поэтому в поле характеристики p

$$(a+b)^p=a^p+b^p \tag{A.8}$$

для любых a и b. Используя равенство (A.8) в качестве основания индукции, индукцией по числу слагаемых нетрудно показать, что

$$(a_1+\dots+a_n)^p=a_1^p+\dots+a_n^p \tag{A.9}$$

для любого n и любых $a_1,\dots,a_n$. Поэтому для любого многочлена $g(x)$

$$\begin{aligned}(g(x))^p &= (g_nx^n+\dots+g_1x+g_0)^p = \\ &= (g_nx^p)^n+\dots+(g_1x^p+(g_0)^p = g_n(x^p)^n+\dots+g_1x^p+g_0=g(x^p).\end{aligned}$$

Лемма доказана.

Формальной производной многочлена $f(x) = a_n x^n + \cdots + x_1 x + a_0$ называется многочлен $f'(x) = n a_n x^{n-1} + \cdots + a_1$. Заметим, что преобразование многочлена в его формальную производную является линейным преобразованием. Используя это свойство, нетрудно показать, что для формальной производной произведения двух многочленов справедливо равенство

$$(f(x)h(x))' = f'(x)h(x) + f(x)h'(x).$$

Теорема А.12.

$$x^{p^n} - x = \prod h(x), \tag{A.10}$$

где произведение берется по всем нормированным неприводимым над $\mathbb{Z}_p$ многочленам из $\mathbb{Z}_p[x]$, степени которых делят n.

Доказательство. Пусть $h(x)$ — неприводимый многочлен степени m. Так как $\mathbb{Z}_p[x]/h(x)$ является полем, которое состоит из p^m элементов, то в мультипликативной группе этого поля порядок любого элемента, в том числе и одночлена x, является делителем числа $p^m - 1$, и, следовательно,

$$x^{p^m} = x \,(\mathrm{mod}\, h(x)). \tag{A.11}$$

Представим n в виде $n = km + r$, где $0 \le r < m$. Тогда, учитывая равенство (A.11), имеем

$$x^{p^n} = x^{p^{km+r}} = \left(x^{p^{km}}\right)^{p^r} = x^{p^r} \ (\mathrm{mod}\, h(x)). \tag{A.12}$$

Если n делится на m, т. е. $r = 0$, то в силу предыдущего равенства

$$x^{p^n} - x = 0 \ (\mathrm{mod}\,(h(x)),$$

и, следовательно, двучлен $x^{p^n} - x$ делится на многочлен $h(x)$.

Теперь предположим, что двучлен $x^{p^n} - x$ делится на $h(x)$ и n не делится на m, т. е. $r > 0$. В этом случае

$$x^{p^n} - x = 0 \ (\mathrm{mod}\,(h(x)),$$

и из $(A.12)$ следует равенство

$$x^{p^r} = x \ (\mathrm{mod}\,(h(x)).$$

Далее заметим, что в силу леммы А.5 для любого элемента g поля $\mathbb{Z}_p[x]/h(x)$ из предыдущего равенства следует, что

$$(g(x))^{p^r} = g(x^{p^r}) = g(x) \ (\mathrm{mod}\,(h(x)).$$

Таким образом порядок любого элемента поля $\mathbb{Z}_p[x]/h(x)$ не превосходит $p^r - 1$. Так как $r < m$, то в этом поле нет порождающего элемента, что противоречит теореме А.11.

Поэтому двучлен $x^{p^n} - x$ является произведением всех неприводимых многочленов, степени которых делят n. Пусть $h(x)$ — один из таких многочленов. Покажем, что $x^{p^n} - x$ не делится на $(h(x))^2$. Для этого рассмотрим формальную производную произведения квадрата $h(x)$ и произвольного многочлена $g(x)$. Легко видеть, что

$$\begin{aligned}\left(h(x)^2 g(x)\right)' &= 2h(x)h'(x)g(x) + h(x)^2 g'(x) = \\ &= h(x)(2h'(x)g(x) + h(x)g'(x)).\end{aligned}$$

Поэтому производная произведения $(h(x))^2 g(x)$ либо делится на $h(x)$, либо равна нулю (если $2h'(x)g(x) + h(x)g'(x) = 0$). Дифференцируя $x^{p^n} - x$, легко находим, что производная этого двучлена равна минус единице. Таким образом, в разложении $x^{p^n} - x$ на неприводимые многочлены нет кратных множителей. Следовательно,

$$x^{p^n} - x = \prod h(x), \tag{A.13}$$

где произведение берется по всем неприводимым многочленам, степени которых делят n. Теорема доказана.

Число неприводимых многочленов степени m обозначим через $P(m)$. Из теоремы А.12 легко извлекается следующее утверждение — лемма А.6, доказанное на стр. 41 с использованием метода производящих функций.

Лемма А.6. *Для последовательности $P(n)$ справедливо рекуррентное равенство*

$$p^n = \sum_{m \mid n} m P(m).$$

Два поля $\mathbb{F}$ и $\mathbb{F}'$ называются изоморфными, если существует такое взаимнооднозначное отображение φ поля $\mathbb{F}$ в поле $\mathbb{F}'$, что $\varphi(a+b) = \varphi(a) + \varphi(b)$ и $\varphi(ab) = \varphi(a)\varphi(b)$ для любых элементов a и b поля $\mathbb{F}$.

Лемма А.7. *Пусть p — простое. Любое конечное поле $\mathbb{F}$ из p элементов изоморфно полю $\mathbb{Z}_p$.*

Доказательство. Пусть $\mathbf{1}$ — единица поля $\mathbb{F}$. Любой элемент поля $\mathbb{F}$ можно представить в виде $k\mathbf{1}$ — суммы k единиц этого поля. Очевидно, что $n\mathbf{1} = [n]_p \mathbf{1}$ при любом n из $\mathbb{N}$. Тогда для отображения φ из $\mathbb{F}$ в $\mathbb{Z}_p$, заданного равенством $\varphi(n\mathbf{1}) = n$, и для любых k и m из $\{0, 1, \ldots, p-1\}$ справедливо равенство

$$\begin{aligned}\varphi(k\mathbf{1} + m\mathbf{1}) = \varphi((k+m)\mathbf{1}) = \varphi([k+m]_p\mathbf{1}) = \\ = [k+m]_p = [k]_p + [m]_p = \varphi(k\mathbf{1}) + \varphi(m\mathbf{1}).\end{aligned}$$

Аналогичное равенство имеет место для умножения:

$$\begin{aligned}\varphi(k\mathbf{1} \cdot m\mathbf{1}) = \varphi((k \cdot m)\mathbf{1}) = \varphi([k \cdot m]_p\mathbf{1}) = \\ = [k \cdot m]_p = [k]_p \cdot [m]_p = \varphi(k\mathbf{1}) \cdot \varphi(m\mathbf{1}).\end{aligned}$$

Следовательно, φ — изоморфизм. Лемма доказана.

Пусть p — простое, и $\mathbb{F}_p$ — поле из p элементов. Из леммы А.7 легко следует, что $\mathbb{F}_p[x] \cong \mathbb{Z}_p[x]$. Пусть $\varphi : \mathbb{Z}_p[x] \to \mathbb{F}_p[x]$ — изоморфизм. Этот изоморфизм будем использовать для определения значений многочленов из $\mathbb{Z}_p[x]$ на элементах поля $\mathbb{F}_q$ с простым подполем $\mathbb{F}_p$ полагая, что для $f(x) \in \mathbb{Z}_p[x]$ и $\alpha \in \mathbb{F}_q$ значение $f(\alpha)$ многочлена f на элементе α равно $(\varphi(f))(\alpha)$.

Каждый элемент поля $\mathbb{F}$ из p^n элементов является корнем двучлена $x^{p^n} - x$, который в силу теоремы А.12 разлагается в произведение всех нормированных неприводимых над $\mathbb{Z}_p$ многочленов, степени которых делят n. Так как $x^{p^n} - x$ не имеет кратных множителей, то каждый элемент поля $\mathbb{F}$ будет простым корнем одного из этих неприводимых над $\mathbb{Z}_p$ многочленов.

Теорема А.13. *В любом поле $\mathbb{F}$ из p^n элементов для любого m, являющегося делителем n, существует единственное подполе из p^m элементов. Других подполей в поле $\mathbb{F}$ нет.*

Доказательство. Если $\mathbb{F}'$ — подполе поля $\mathbb{F}$, то $\mathbb{F}$ будет линейным пространством над своим подполем $\mathbb{F}'$. В этом случае найдется натуральное k такое, что $p^n = |\mathbb{F}'|^k$. Очевидно, что последнее равенство возможно только в том случае, когда $n = km$.

Теперь покажем, что если m делит n, то в $\mathbb{F}$ существует подполе из p^m элементов. Рассмотрим множество N элементов поля $\mathbb{F}$, являющихся корнями всех неприводимых многочленов, степени которых равны m или делят m. Из теоремы А.12 следует, что каждый из этих элементов является корнем двучлена $x^{p^m} - x$. Покажем, что сумма и произведение любых двух элементов из N также принадлежат этому множеству. Если $a^{p^m} - a = 0$ и $b^{p^m} - b = 0$, то в силу равенства (А.8)

$$(a+b)^{p^m} - (a+b) = (a^{p^m} + b^{p^m}) - (a+b) = (a^{p^m} - a) + (b^{p^m} - b) = 0,$$

т. е. сумма $a+b$ принадлежит N. Аналогичная цепочка равенств имеет место и для произведения ab:

$$(ab)^{p^m} - ab = a^{p^m} \cdot b^{p^m} - ab = ab - ab = 0.$$

Таким образом множество N замкнуто относительно сложения и умножения. Также нетрудно видеть, что если a является корнем двучлена $x^{p^m} - x$, то и его обратные элементы по сложению и умножению будут корнями этого двучлена. Следовательно, N является полем.

Наконец заметим, что мультипликативная группа подполя будет подгруппой мультипликативной группы поля. Поэтому единственность подполя из p^m элементов в поле из p^n элементов легко следует из теоремы А.2. Теорема доказана.

Пусть α — элемент поля из p^n элементов. Элементы $\alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^{k-1}}$, где $\alpha^{p^k} = \alpha$, называются сопряженными с α. Так как в поле из p^n элементов $\alpha^{p^n} = \alpha$, то α имеет не более $n-1$ сопряженных элементов.

Теорема А.14. *Если α — корень неприводимого над $\mathbb{Z}_p$ многочлена $h(x)$ степени k, то корнями этого многочлена являются все элементы, сопряженные с α и других корней многочлен $h(x)$ не имеет.*

Доказательство. Так как $(h(x))^p = h(x^p)$, то, очевидно, что любой элемент сопряженный с α, будет корнем многочлена $h(x)$. Поэтому для доказательства теоремы достаточно показать, что все элементы $\alpha, \alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^{k-1}}$ различны. Допустим, что это не так и $\alpha^{p^m} = \alpha$ при $m < k$. Тогда α будет корнем многочлена $x^{p^m} - x$ и, в силу теоремы А.12, будет также корнем неприводимого многочлена $g(x)$ степени t, где t является делителем m. Разделив неприводимый многочлен $h(x)$ на $g(x)$, получим, что $h(x) = g(x)f(x) + r(x)$, где $r(x) \neq 0$ и $\deg r(x) \leq t \leq m < k$. Так как $h(\alpha) = 0$ и $g(\alpha) = 0$, то, очевидно, что и $r(\alpha) = 0$. Вместе с α корнями многочлена $r(x)$ должны быть и все элементы, сопряженные с α, т. е. число корней многочлена $r(x)$ не меньше k, что, очевидно, невозможно, так как его степень строго меньше k. Таким образом, сделанное предположение ложно и $m = k$. Теорема доказана.

Пусть $\mathbb{F}$ — поле из p^n элементов, α — элемент этого поля. Нормированный многочлен $f(x)$ из $\mathbb{Z}_p[x]$ называется *минимальным многочленом* элемента α, если степень многочлена f минимальна среди всех многочленов из $\mathbb{Z}_p[x]$, для которых α является корнем. Очевидно, что для любого α существует единственный минимальный многочлен, этот многочлен является неприводимым над $\mathbb{Z}_p[x]$, и его корнями в $\mathbb{F}$ являются вместе с α все элементы, сопряженные с α.

Неприводимый над $\mathbb{Z}_p$ многочлен $h(x)$ из $\mathbb{Z}_p[x]$ называется *примитивным многочленом*, если в поле $\mathbb{Z}_p[x]/h(x)$ одночлен x является порождающим элементом мультипликативной группы. Из теорем А.14 и А.4 следует, что в $\mathbb{Z}_p[x]$ существует ровно $\varphi(p^n - 1)/n$ нормированных примитивных многочленов степени n.

Теорема А.15. *Пусть поле $\mathbb{F}$ состоит из p^n элементов, а его элемент α — корень неприводимого многочлена степени n. Тогда любой элемент поля $\mathbb{F}$ является линейной комбинацией элементов $1, \alpha, \ldots, \alpha^{n-1}$ с коэффициентами из простого подполя поля $\mathbb{F}$.*

Доказательство. Прежде всего заметим, что существует ровно p^n различных линейных комбинаций, и каждая линейная комбинация является элементом поля $\mathbb{F}$. Более того, среди этих линейных комбинаций нет двух, равных одному и тому же элементу поля $\mathbb{F}$, так как в противном случае разность двух равных линейных комбинаций будет ненулевой линейной комбинацией, равной нулевому элементу поля. Используя лемму А.7, нетрудно показать, что существование такой линейной комбинации равносильно тому, что α будет корнем многочлена степени не больше $n-1$, что, очевидно, противоречит тому, что α является корнем неприводимого многочлена степени n. Теорема доказана.

Теорема А.15 показывает, что поле $\mathbb{F}$ из p^n элементов с простым подполем $\mathbb{F}_p$ можно представить как $\mathbb{F}_p[\alpha]$, где α — корень неприводимого над $\mathbb{Z}_p$

многочлена степени n. Такое представление приводит к следующему утверждению.

Теорема А.16. *Любые два конечных поля, состоящие из одного и того же числа элементов, изоморфны.*

Доказательство. Пусть p — простое, $h(x)$ — нормированный неприводимый над $\mathbb{Z}_p$ многочлен степени n из $\mathbb{Z}_p[x]$. Покажем, что произвольное поле из p^n элементов изоморфно полю $\mathbb{Z}_p[x]/h(x)$.

В силу леммы А.7 далее без ограничения общности будем полагать, что простое подполе поля $\mathbb{F}$ совпадает с полем $\mathbb{Z}_p$. В поле $\mathbb{F}$ выберем элемент α, являющийся корнем многочлена $h(x)$. Покажем, что отображение $\varphi : \sum_{i=0}^{n-1} a_i\alpha^i \to \sum_{i=0}^{n-1} a_i x^i$ поля $\mathbb{F}$ в поле $\mathbb{Z}_p[x]/h(x)$ будет искомым изоморфизмом. Очевидно, что φ взаимнооднозначно и линейно, и в силу своей линейности сохраняет операцию сложения. Поэтому для доказательства теоремы достаточно показать, что f сохраняет умножение. Каждый элемент $a = \sum_{i=0}^{n-1} a_i\alpha^i$ поля $\mathbb{F}$ будем рассматривать как многочлен $a(\alpha)$ относительно α. В силу законов дистрибутивности умножение элементов $a = \sum_{i=0}^{n-1} a_i\alpha^i$ и $b = \sum_{i=0}^{n-1} b_i\alpha^i$ поля $\mathbb{F}$ производится так же, как и умножение многочленов. Поэтому справедливо равенство

$$ab = a(\alpha)b(\alpha) = c(\alpha)h(\alpha) + r(\alpha),$$

где $r(\alpha)$ — остаток от деления многочлена $a(\alpha)b(\alpha)$ на многочлен $h(\alpha)$. Так как $h(\alpha) = 0$, то $ab = r(\alpha)$. Поэтому

$$\varphi(a(\alpha))\varphi(b(\alpha)) = a(x)b(x) = r(x) = \varphi(r(\alpha)),$$

т. е. φ действительно сохраняет умножение. Теорема доказана.

Таким образом, любые два поля из p^n элементов изоморфны, поэтому для всех таких полей, называемых полями Галуа, часто используется единое обозначение $GF(p^n)$.

Литература

[1] *Андреев А. Е.* О сложности реализации частичных булевых функций схемами из функциональных элементов. — Дискретная математика. 1989. Вып. 4. С. 36–45.

[2] *Ансель Ж.* О числе монотонных булевых функций n переменных. — В кн.: Кибернетический сборник. Новая серия. Вып.5. — М.: Мир, 1968, с. 53–63.

[3] *Берлекэмп Э.* Алгебраическая теория кодирования. — М.: Мир, 1971.

[4] *Блейхут Р.* Теория и практика кодов, контролирующих ошибки. — М.: Мир, 1986.

[5] *де Брейн Н. Дж.* Теория перечисления Пойа. — В кн.: Прикладная комбинаторная математика. — М.: Мир, 1968, с. 61–106.

[6] *Гаврилов Г. П., Сапоженко А. А.* Задачи и упражнения по курсу дискретной математики. — 2-е изд. М.: Наука, 1992.

[7] *Галлагер Р.* Теория информации и надежная связь. — М.: Советское радио, 1974.

[8] *Грэхем Р., Кнут Д., Паташник О.* Конкретная математика. — М.: Мир, 1988.

[9] *Гульден Я., Джексон Д.* Перечислительная комбинаторика. — М.: Наука, 1990.

[10] *Ежов И. И., Скороход А. В., Ядренко М. И.* Элементы комбинаторики. — М.: Наука, 1977.

[11] *Емеличев В. А. и др.* Лекции по теории графов. — М.: Наука, 1990.

[12] *Забалуев Р. Н.* О средней сложности булевых функций, заданных полиномами Жегалкина. — Дискретн. анализ и исслед. опер., Серия 1. 2004. Вып. 3. С. 3–15.

[13] *Кострикин А. И.* Введение в алгебру. — М.: Наука, 1977.

[14] *Кричевский Р. Е.* Сжатие и поиск информации. — М.: Радио и связь, 1989.

[15] *Левенштейн В. И.* Введение в теорию функций k-значной логики. — В кн.: Дискретная математика и математические вопросы кибернетики, т. 1, под ред. С. В. Яблонского и О. Б. Лупанова. — М.: Наука, 1974, с. 207–302.

[16] *Лидл Р., Нидеррайтер Г.* Конечные поля. — М.: Мир, 1988.

[17] *Ложкин С. А.* Лекции по основам кибернетики. — М.: Изд. отд. ВМК МГУ, 2004.

[18] *Лупанов О. Б.* О синтезе некоторых классов управляющих систем. — В кн.: Проблемы кибернетики. Вып. 10. — М.: Физматгиз, 1963, с. 63–97.

[19] *Лупанов О. Б.* Об одном подходе к синтезу управляющих систем — принципе локального кодирования. — В кн.: Проблемы кибернетики. Вып. 14. — М.: Физматгиз, 1965, с. 31–110.

[20] *Лупанов О. Б.* Асимптотические оценки сложности управляющих систем. — М.: МГУ, 1984.

[21] *Мак-Вильямс Ф. Дж., Слоэн Н. Дж. А.* Теория кодов, исправляющих ошибки. — М.: Связь, 1979.

[22] *Манн Г. Б.* О числе информационных символов в кодах Боуза–Чоудхури. — В кн.: Кибернетический сборник. Вып. 8. — М.: Мир, 1966, с. 33–41.

[23] *Нечипорук Э. И.* О топологических принципах самокорректирования. — В кн.: Проблемы кибернетики. Вып. 21. — М.: Наука, 1969, с. 5–102.

[24] *Нигматуллин Р. Г.* Сложность булевых функций. — М.: Наука, 1991.

[25] *Питерсон У., Уэлдон Э.* Коды, исправляющие ошибки. — М.: Мир, 1976.

[26] *Редфилд Г. Дж.* Теория распределений, приведенных по группе. — В кн.: Перечислительные залачи комбинаторного анализа. — М.: Мир, 1979, с. 9–35.

[27] *Редькин Н. П.* О реализации монотонных функций контактными схемами. — В кн.: Проблемы кибернетики. Вып. 35. — М.: Наука, 1979, с. 87–110.

[28] *Сэвидж Дж. Э.* Сложность вычислений. — М.: Факториал, 1998.

[29] *Угольников А. Б.* О реализации монотонных функций схемами из функциональных элементов. — В кн.: Проблемы кибернетики. Вып. 31. — М.: Наука, 1976, с. 167–185.

[30] *Уилсон Р.* Введение в теорию графов. — М.: Мир, 1977.

[31] *Феллер В.* Введение в теорию вероятностей и ее приложения. т. 1. — М.: Мир, 1984.

[32] *Фихтенгольц Г. М.* Основы математического анализа. т. 1–2. — М.: Физматгиз, 1960.

[33] *Харари Ф.* Теория графов. — М.: Мир, 1973.

[34] *Холл М.* Комбинаторика. — М.: Мир, 1970.

[35] *Храпченко В. М.* Об асимптотической оценке времени сложения параллельного сумматора. — В кн.: Проблемы кибернетики. Вып. 19. — М.: Наука, 1967, с. 107–122.

[36] *Чашкин А. В.* О сложности булевых матриц, графов и соответствующих им булевых функций. — Дискретная математика, 1994, №2, с. 43–73.

[37] *Чашкин А. В.* О среднем времени вычисления значений булевых функций. — Дискретный анализ и исследование операций. 1997. Вып. 1. С. 60–78.

[38] *Чашкин А. В.* О сложности реализации булевых функций формулами. — Дискретный анализ и исследование операций. 2005. Вып. 2. С. 56–72.

[39] *Шенхаге А., Штрассен В.* Быстрое умножение больших чисел. — В кн.: Кибернетический сборник. Новая серия. Вып. 10. — М.: Мир, 1973, с. 87–98.

[40] *Шоломов Л. А.* О реализации недоопределенных булевых функций схемами из функциональных элементов. — В кн.: Проблемы кибернетики. Вып. 21. — М.: Наука, 1969, с. 215–226.

[41] *Яблонский С. В.* Введение в теорию функций k-значной логики. — В кн.: Дискретная математика и математические вопросы кибернетики, т. 1, под ред. С. В. Яблонского и О. Б. Лупанова. — М.: Наука, 1974, с. 9–66.

[42] *Яблонский С. В.* Введение в дискретную математику. — 3-е изд. М.: Высш. школа, 2001.

[43] *Ajtai M., Komlos Ja., Szemeredi E.* Sorting in $O(n \log n)$ parallel steps. — Combinatorica. 1983. V. 3. № 1. P. 1–19.